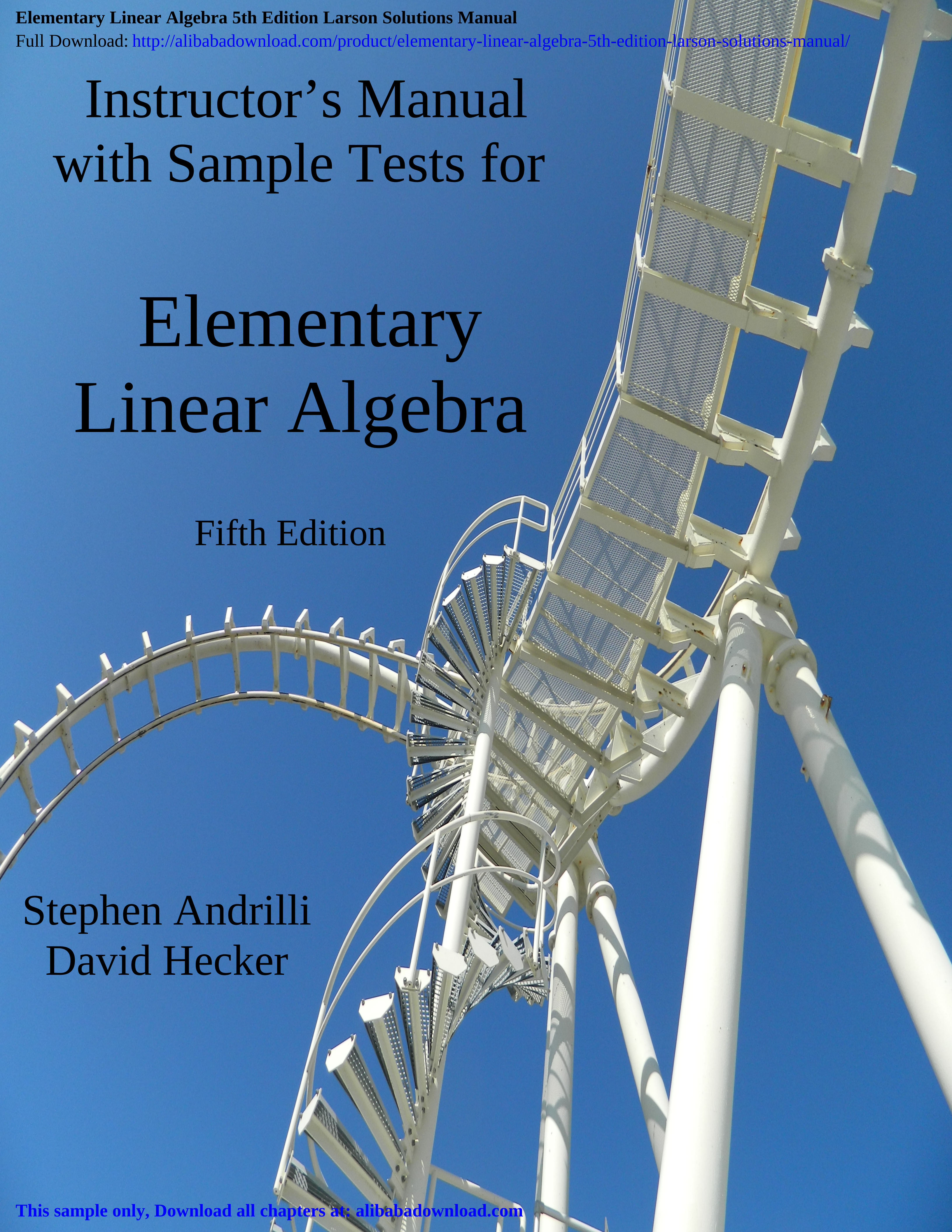


Instructor's Manual with Sample Tests for

Elementary Linear Algebra

Fifth Edition

Stephen Andrilli
David Hecker

Dedication

To all the instructors who have used the various editions of our book over the years

Table of Contents

Preface

This *Instructor's Manual with Sample Tests* is designed to accompany *Elementary Linear Algebra*, *5th edition*, by Stephen Andrilli and David Hecker.

This manual contains answers for all the computational exercises in the textbook and detailed solutions for virtually all of the problems that ask students for proofs. The exceptions are typically those exercises that ask students to verify a particular computation, or that ask for a proof for which detailed hints have already been supplied in the textbook. A few proofs that are extremely trivial in nature have also been omitted.

This manual also contains sample Chapter Tests for the material in Chapters 1 through 7, as well as answer keys for these tests.

Additional information regarding the textbook, this manual, the *Student Solutions Manual*, and linear algebra in general can be found at the web site for the textbook, where you found this manual.

Thank you for using our textbook.

Stephen Andrilli

David Hecker

August 2015

Answers to Exercises

Chapter 1

Section 1.1

(1) (a) $[9, -4]$, distance $= \sqrt{97}$
(b) $[-6, 1, 1]$, distance $= \sqrt{38}$
(c) $[-1, -1, 2, -3, -4]$, distance $= \sqrt{31}$

(2) (a) $(3, 4, 2)$ (see Figure 1)
(b) $(0, 5, 3)$ (see Figure 2)
(c) $(1, -2, 0)$ (see Figure 3)
(d) $(3, 0, 0)$ (see Figure 4)

(3) (a) $(7, -13)$ (b) $(6, 4, -8)$ (c) $(-1, 3, -1, 4, 6)$

(4) (a) $\left(\frac{16}{3}, -\frac{13}{3}, 8\right)$ (b) $(-\frac{20}{3}, -1, -6, -1)$

(5) (a) $\left[\frac{3}{\sqrt{70}}, -\frac{5}{\sqrt{70}}, \frac{6}{\sqrt{70}}\right]$; shorter, since length of original vector is > 1
(b) $\left[-\frac{6}{7}, \frac{2}{7}, 0, -\frac{3}{7}\right]$; shorter, since length of original vector is > 1
(c) $[0.6, -0.8]$; neither, since given vector is a unit vector
(d) $\left[\frac{1}{\sqrt{11}}, -\frac{2}{\sqrt{11}}, -\frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{2}{\sqrt{11}}\right]$; longer, since length of original vector $=\frac{\sqrt{11}}{5} < 1$

(6) (a) Parallel (b) Parallel (c) Not parallel (d) Not parallel

(7) (a) $[-6, 12, 15]$
(b) $[10, 6, -12]$
(c) $[-7, 1, 11]$
(d) $[-9, -2, 4]$
(e) $[-10, -32, -1]$
(f) $[-35, 3, 20]$

(8) (a) $\mathbf{x} + \mathbf{y} = [1, 1]$, $\mathbf{x} - \mathbf{y} = [-3, 9]$, $\mathbf{y} - \mathbf{x} = [3, -9]$ (see Figure 5)
(b) $\mathbf{x} + \mathbf{y} = [3, -5]$, $\mathbf{x} - \mathbf{y} = [17, 1]$, $\mathbf{y} - \mathbf{x} = [-17, -1]$ (see Figure 6)
(c) $\mathbf{x} + \mathbf{y} = [1, 8, -5]$, $\mathbf{x} - \mathbf{y} = [3, 2, -1]$, $\mathbf{y} - \mathbf{x} = [-3, -2, 1]$ (see Figure 7)
(d) $\mathbf{x} + \mathbf{y} = [-2, -4, 4]$, $\mathbf{x} - \mathbf{y} = [4, 0, 6]$, $\mathbf{y} - \mathbf{x} = [-4, 0, -6]$ (see Figure 8)

(9) With $A = (7, -3, 6)$, $B = (11, -5, 3)$, and $C = (10, -7, 8)$, the length of side AB = length of side AC $= \sqrt{29}$. The triangle is isosceles, but not equilateral, since the length of side BC is $\sqrt{30}$.

(10) (a) $[10, -10]$ (b) $[-5\sqrt{3}, -15]$ (c) $[0, 0] = \mathbf{0}$

(11) See Figures 1.7 and 1.8 in Section 1.1 of the textbook.

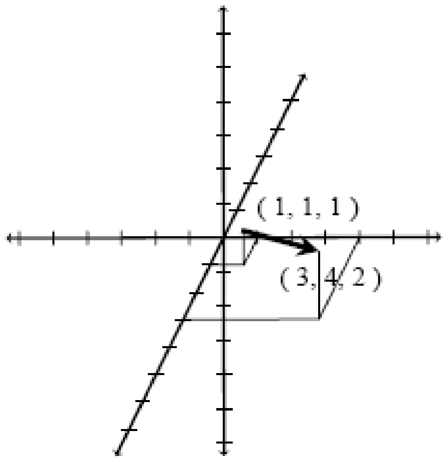


Figure 1

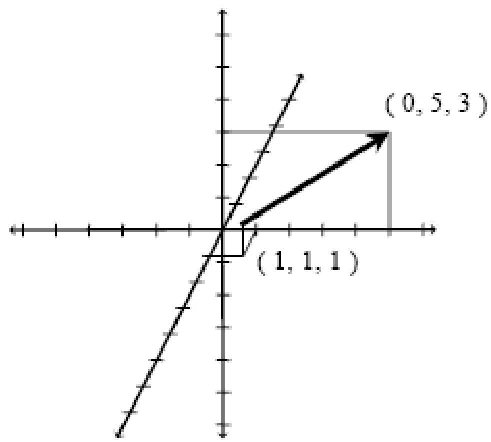


Figure 2

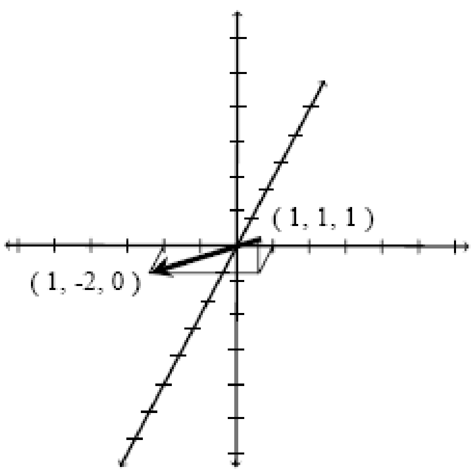


Figure 3

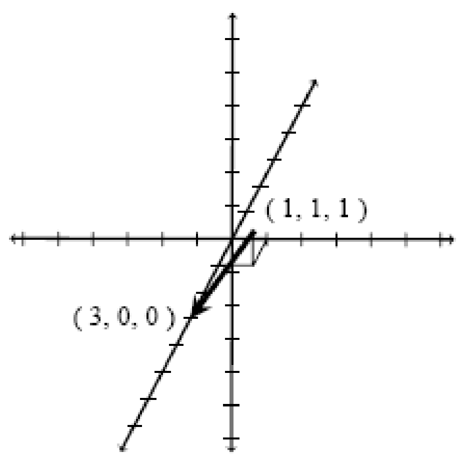
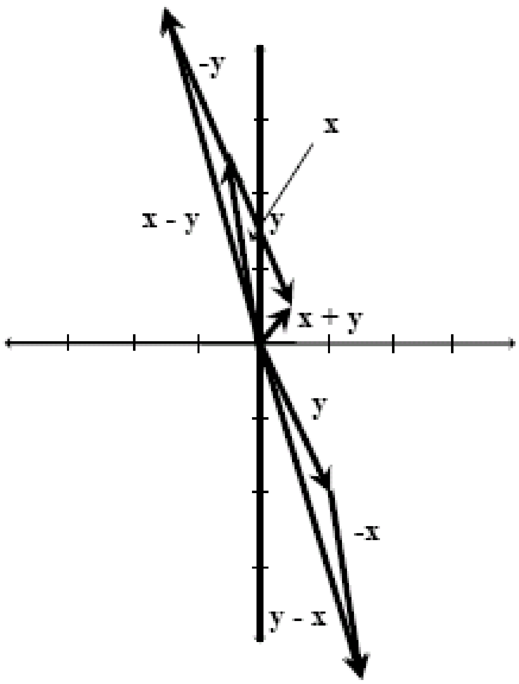
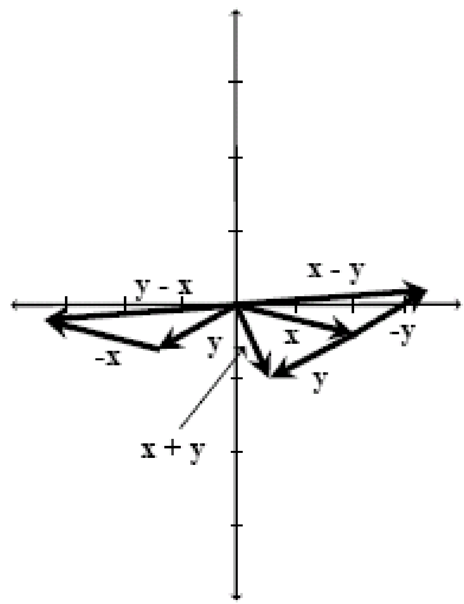


Figure 4



Scale: Each tickmark = 2 units

Figure 5



Scale: Each tickmark = 5 units

Figure 6

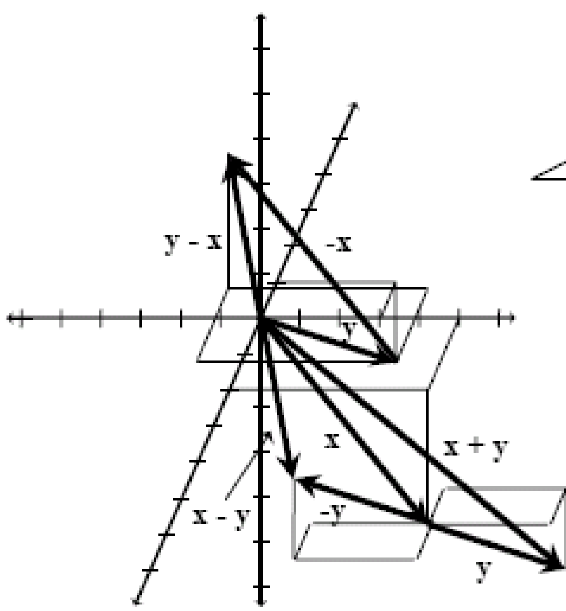


Figure 7

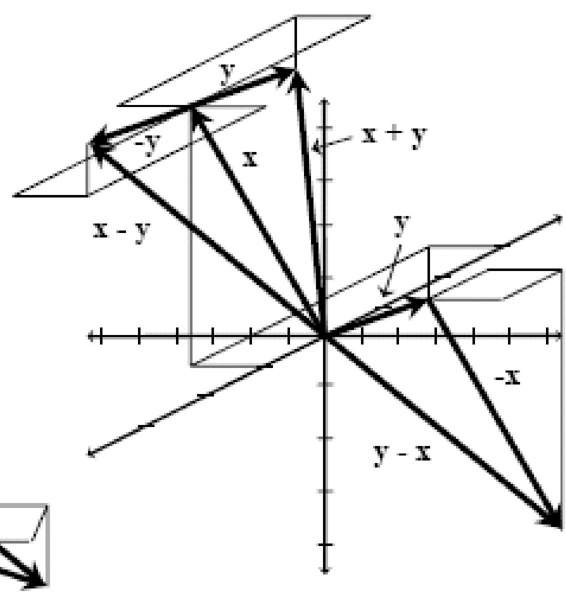


Figure 8

(12) See Figure 9. Both represent the same diagonal vector by the associative law of addition for vectors.

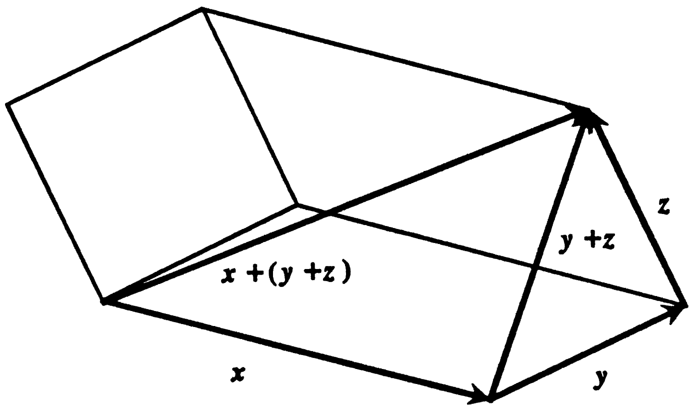


Figure 9: $\mathbf{x} + (\mathbf{y} + \mathbf{z})$

(13) $[0.5 - 0.6\sqrt{2}, -0.4\sqrt{2}] \approx [-0.3485, -0.5657]$

(14) Net velocity $= [-2\sqrt{2}, -3 + 2\sqrt{2}]$, resultant speed ≈ 2.83 km/hr

(15) Net velocity $= \left[-\frac{3\sqrt{2}}{2}, \frac{8-3\sqrt{2}}{2}\right]$; speed ≈ 2.83 km/hr

(16) $[-8 - \sqrt{2}, -\sqrt{2}]$

(17) Acceleration $= \frac{1}{20}[\frac{12}{13}, -\frac{344}{65}, \frac{392}{65}] \approx [0.0462, -0.2646, 0.3015]$

(18) Acceleration $= \left[\frac{13}{2}, 0, \frac{4}{3}\right]$

(19) $\frac{180}{\sqrt{14}}[-2, 3, 1] \approx [-96.22, 144.32, 48.11]$

(20) $\mathbf{a} = [\frac{-mg}{1+\sqrt{3}}, \frac{mg}{1+\sqrt{3}}]$; $\mathbf{b} = [\frac{mg}{1+\sqrt{3}}, \frac{mg\sqrt{3}}{1+\sqrt{3}}]$

(21) Let $\mathbf{a} = [a_1, \ldots, a_n]$.

(a) $\|\mathbf{a}\|^2 = a_1^2 + \cdots + a_n^2$ is a sum of squares, which must be nonnegative. But then $||\mathbf{a}|| \geq 0$ because the square root of a nonnegative real number is a nonnegative real number.

(b) If $||\mathbf{a}|| = 0$, then $\|\mathbf{a}\|^2 = a_1^2 + \cdots + a_n^2 = 0$, which is only possible if every $a_i = 0$. Thus, $\mathbf{a} = \mathbf{0}$.

(22) In each part, suppose that $\mathbf{x} = [x_1, \ldots, x_n]$, $\mathbf{y} = [y_1, \ldots, y_n]$, and $\mathbf{z} = [z_1, \ldots, z_n]$.

(a) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = [x_1, \ldots, x_n] + [(y_1 + z_1), \ldots, (y_n + z_n)] = [(x_1 + (y_1 + z_1)), \ldots, (x_n + (y_n + z_n))]$
$= [((x_1 + y_1) + z_1), \ldots, ((x_n + y_n) + z_n)] = [(x_1 + y_1), \ldots, (x_n + y_n)] + [z_1, \ldots, z_n] = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$

(b) $\mathbf{x} + (-\mathbf{x}) = [x_1, \ldots, x_n] + [-x_1, \ldots, -x_n] = [(x_1 + (-x_1)), \ldots, (x_n + (-x_n))] = [0, \ldots, 0]$. Also, $(-\mathbf{x}) + \mathbf{x} = \mathbf{x} + (-\mathbf{x})$ (by part (1) of Theorem 1.3) $= \mathbf{0}$, by the above.

(c) $c(\mathbf{x}+\mathbf{y}) = c([(x_1+y_1), \ldots, (x_n+y_n)]) = [c(x_1+y_1), \ldots, c(x_n+y_n)] = [(cx_1+cy_1), \ldots, (cx_n+cy_n)] = [cx_1, \ldots, cx_n] + [cy_1, \ldots, cy_n] = c\mathbf{x} + c\mathbf{y}$

(d) $(cd)\mathbf{x} = [((cd)x_1), \ldots, ((cd)x_n)] = [(c(dx_1)), \ldots, (c(dx_n))] = c[(dx_1), \ldots, (dx_n)]$
$= c(d[x_1, \ldots, x_n]) = c(d\mathbf{x})$

(23) If $c = 0$, done. Otherwise, $(\frac{1}{c})(c\mathbf{x}) = \frac{1}{c}(\mathbf{0}) \Longrightarrow (\frac{1}{c} \cdot c)\mathbf{x} = \mathbf{0}$ (by part (7) of Theorem 1.3) $\Longrightarrow \mathbf{x} = \mathbf{0}$. Thus either $c = 0$ or $\mathbf{x} = \mathbf{0}$.

(24) $c_1\mathbf{x} = c_2\mathbf{x} \Longrightarrow c_1\mathbf{x} - c_2\mathbf{x} = \mathbf{0} \Longrightarrow (c_1 - c_2)\mathbf{x} = \mathbf{0} \Longrightarrow (c_1 - c_2) = 0$ or $\mathbf{x} = \mathbf{0}$, by Theorem 1.4. But since $c_1 \neq c_2$, $(c_1 - c_2) \neq 0$. Hence, $\mathbf{x} = \mathbf{0}$.

(25) (a) F (b) T (c) T (d) F (e) T (f) F (g) F (h) F

Section 1.2

(1) (a) $\arccos(-\frac{27}{5\sqrt{37}})$, or about 152.6°, or 2.66 radians

(b) $\arccos(\frac{46}{\sqrt{74}\sqrt{29}})$, or about 6.8°, or 0.12 radians

(c) $\arccos(0)$, which is 90°, or $\frac{\pi}{2}$ radians

(d) $\arccos(-\frac{435}{\sqrt{2175}\sqrt{87}}) = \arccos(-1)$, which is 180°, or π radians (since $\mathbf{x} = -5\mathbf{y}$)

(2) The vector from A_1 to A_2 is $[2, -7, -3]$, and the vector from A_1 to A_3 is $[5, 4, -6]$. These vectors are orthogonal.

(3) (a) $[a, b] \cdot [-b, a] = a(-b) + ba = 0$. Similarly, $[a, -b] \cdot [b, a] = 0$.

(b) A vector in the direction of the line $ax + by + c = 0$ is $[b, -a]$, while a vector in the direction of $bx - ay + d = 0$ is $[a, b]$.

(4) (a) 15 joules (b) $\frac{1040\sqrt{5}}{9} \approx 258.4$ joules (c) $-\frac{189\sqrt{15}}{5} \approx -146.4$ joules

(5) Note that $\mathbf{y}\cdot\mathbf{z}$ is a scalar, so $\mathbf{x}\cdot(\mathbf{y}\cdot\mathbf{z})$ is not defined.

(6) In all parts, let $\mathbf{x} = [x_1, x_2, \ldots, x_n]$, $\mathbf{y} = [y_1, y_2, \ldots, y_n]$, and $\mathbf{z} = [z_1, z_2, \ldots, z_n]$.

(a) $\mathbf{x} \cdot \mathbf{y} = [x_1, x_2, \ldots, x_n] \cdot [y_1, y_2, \ldots, y_n] = x_1y_1 + \cdots + x_ny_n = y_1x_1 + \cdots + y_nx_n$
$= [y_1, y_2, \ldots, y_n] \cdot [x_1, x_2, \ldots, x_n] = \mathbf{y} \cdot \mathbf{x}$

(b) $\mathbf{x} \cdot \mathbf{x} = [x_1, x_2, \ldots, x_n] \cdot [x_1, x_2, \ldots, x_n] = x_1x_1 + \cdots + x_nx_n = x_1^2 + \cdots + x_n^2$. Now $x_1^2 + \cdots + x_n^2$ is a sum of squares, each of which must be nonnegative. Hence, the sum is also nonnegative, and so its square root is defined. Thus, $0 \leq \mathbf{x} \cdot \mathbf{x} = x_1^2 + \cdots + x_n^2 = \left(\sqrt{x_1^2 + \cdots + x_n^2}\right)^2 = \|\mathbf{x}\|^2$.

(c) Suppose $\mathbf{x} \cdot \mathbf{x} = 0$. From part (b), $0 = \mathbf{x} \cdot \mathbf{x} = x_1^2 + \cdots + x_n^2 \geq x_i^2$, for each i, since all terms in the sum are nonnegative. Hence, $0 \geq x_i^2$ for each i. But $x_i^2 \geq 0$, because it is a square. Hence each $x_i = 0$. Therefore, $\mathbf{x} = \mathbf{0}$.

(d) $c(\mathbf{x} \cdot \mathbf{y}) = c\left([x_1, x_2, \ldots, x_n] \cdot [y_1, y_2, \ldots, y_n]\right) = c\left(x_1y_1 + \cdots + x_ny_n\right)$
$= cx_1y_1 + \cdots + cx_ny_n = [cx_1, cx_2, \ldots, cx_n] \cdot [y_1, y_2, \ldots, y_n] = (c\mathbf{x}) \cdot \mathbf{y}$.
Next, $c(\mathbf{x} \cdot \mathbf{y}) = c(\mathbf{y} \cdot \mathbf{x})$ (by part (a)) $= (c\mathbf{y}) \cdot \mathbf{x}$ (from above) $= \mathbf{x} \cdot (c\mathbf{y})$, by part (a).

(e) $(\mathbf{x}+\mathbf{y})\cdot\mathbf{z} = ([x_1, x_2, \ldots, x_n] + [y_1, y_2, \ldots, y_n])\cdot[z_1, z_2, \ldots, z_n]$
$= [x_1+y_1, x_2+y_2, \ldots, x_n+y_n]\cdot[z_1, z_2, \ldots, z_n]$
$= (x_1+y_1)z_1 + (x_2+y_2)z_2 + \cdots + (x_n+y_n)z_n$
$= (x_1z_1 + x_2z_2 + \cdots + x_nz_n) + (y_1z_1 + y_2z_2 + \cdots + y_nz_n)$.
Also, $(\mathbf{x}\cdot\mathbf{z}) + (\mathbf{y}\cdot\mathbf{z}) = ([x_1, x_2, \ldots, x_n]\cdot[z_1, z_2, \ldots, z_n]) + ([y_1, y_2, \ldots, y_n]\cdot[z_1, z_2, \ldots, z_n])$
$= (x_1z_1 + x_2z_2 + \cdots + x_nz_n) + (y_1z_1 + y_2z_2 + \cdots + y_nz_n)$.
Hence, $(\mathbf{x}+\mathbf{y})\cdot\mathbf{z} = (\mathbf{x}\cdot\mathbf{z}) + (\mathbf{y}\cdot\mathbf{z})$.

(7) No; consider $\mathbf{x} = [1, 0]$, $\mathbf{y} = [0, 1]$, and $\mathbf{z} = [1, 1]$.

(8) A method similar to the first part of the proof of Lemma 1.6 in the textbook yields:
$\|\mathbf{a}-\mathbf{b}\|^2 \geq 0 \Longrightarrow (\mathbf{a}\cdot\mathbf{a}) - (\mathbf{b}\cdot\mathbf{a}) - (\mathbf{a}\cdot\mathbf{b}) + (\mathbf{b}\cdot\mathbf{b}) \geq 0 \Longrightarrow 1 - 2(\mathbf{a}\cdot\mathbf{b}) + 1 \geq 0 \Longrightarrow \mathbf{a}\cdot\mathbf{b} \leq 1$.

(9) Note that $(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) = (\mathbf{x}\cdot\mathbf{x}) + (\mathbf{y}\cdot\mathbf{x}) - (\mathbf{x}\cdot\mathbf{y}) - (\mathbf{y}\cdot\mathbf{y}) = \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$. Hence, $(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) = 0$ implies $\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2$, which means $\|\mathbf{x}\| = \|\mathbf{y}\|$ (since both are nonnegative).

(10) Note that $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2$, while $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2$.
Hence, $\frac{1}{2}(\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2) = \frac{1}{2}(2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

(11) (a) From the first equation in the solution to Exercise 10 above, $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ implies $2(\mathbf{x}\cdot\mathbf{y}) = 0$, which means $\mathbf{x}\cdot\mathbf{y} = 0$.
(b) From the first equation in the solution to Exercise 10 above, $\mathbf{x}\cdot\mathbf{y} = 0$ implies $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

(12) Note that $\|\mathbf{x}+\mathbf{y}+\mathbf{z}\|^2 = \|(\mathbf{x}+\mathbf{y})+\mathbf{z}\|^2$
$= \|\mathbf{x}+\mathbf{y}\|^2 + 2((\mathbf{x}+\mathbf{y})\cdot\mathbf{z}) + \|\mathbf{z}\|^2$
$= \|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2 + 2(\mathbf{x}\cdot\mathbf{z}) + 2(\mathbf{y}\cdot\mathbf{z}) + \|\mathbf{z}\|^2$
$= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \|\mathbf{z}\|^2$, since $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are mutually orthogonal.

(13) From the first two equations in the solution for Exercise 10 above,
$\frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) = \frac{1}{4}(4(\mathbf{x}\cdot\mathbf{y})) = \mathbf{x}\cdot\mathbf{y}$.

(14) Since $\mathbf{x}$ is orthogonal to both $\mathbf{y}$ and $\mathbf{z}$, we have $\mathbf{x}\cdot(c_1\mathbf{y} + c_2\mathbf{z}) = c_1(\mathbf{x}\cdot\mathbf{y}) + c_2(\mathbf{x}\cdot\mathbf{z}) = c_1(0) + c_2(0) = 0$.

(15) Suppose $\mathbf{y} = c\mathbf{x}$, for some $c \neq 0$. Then, $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}\cdot(c\mathbf{x}) = c\,(\mathbf{x}\cdot\mathbf{x}) = c\,\|\mathbf{x}\|^2 = \|\mathbf{x}\|\,(c\,\|\mathbf{x}\|) = \|\mathbf{x}\|\,(\pm|c|\,\|\mathbf{x}\|)$
$= \pm\,\|\mathbf{x}\|\,\|c\mathbf{x}\| = \pm\,\|\mathbf{x}\|\,\|\mathbf{y}\|$.

(16) (a) Length $= \sqrt{3}s$ (b) angle $= \arccos(\frac{\sqrt{3}}{3}) \approx 54.7°$, or 0.955 radians

(17) (a) $\mathbf{proj_a b} = [-\frac{3}{5}, -\frac{3}{10}, -\frac{3}{2}]$; $\mathbf{b} - \mathbf{proj_a b} = [\frac{8}{5}, \frac{43}{10}, -\frac{3}{2}]$; $(\mathbf{b} - \mathbf{proj_a b})\cdot\mathbf{a} = 0$
(b) $\mathbf{proj_a b} = [-\frac{6}{5}, 1, \frac{2}{5}]$; $\mathbf{b} - \mathbf{proj_a b} = [-\frac{14}{5}, -4, \frac{8}{5}]$; $(\mathbf{b} - \mathbf{proj_a b})\cdot\mathbf{a} = 0$
(c) $\mathbf{proj_a b} = [\frac{1}{6}, 0, -\frac{1}{6}, \frac{1}{3}]$; $\mathbf{b} - \mathbf{proj_a b} = [\frac{17}{6}, -1, \frac{1}{6}, -\frac{4}{3}]$; $(\mathbf{b} - \mathbf{proj_a b})\cdot\mathbf{a} = 0$
(d) $\mathbf{proj_a b} = [-1, \frac{3}{2}, -2, -\frac{3}{2}]$; $\mathbf{b} - \mathbf{proj_a b} = [6, -\frac{1}{2}, -6, \frac{7}{2}]$; $(\mathbf{b} - \mathbf{proj_a b})\cdot\mathbf{a} = 0$

(18) (a) $\mathbf{0}$ (zero vector). The dropped perpendicular travels along $\mathbf{b}$ to the common initial point of $\mathbf{a}$ and $\mathbf{b}$.
(b) The vector $\mathbf{b}$. The terminal point of $\mathbf{b}$ lies on the line through $\mathbf{a}$, so the dropped perpendicular has length zero.

(19) $a\mathbf{i}$, $b\mathbf{j}$, $c\mathbf{k}$

(20) (a) Parallel: $[\frac{20}{29}, -\frac{30}{29}, \frac{40}{29}]$, orthogonal: $[-\frac{194}{29}, \frac{88}{29}, \frac{163}{29}]$
(b) Parallel: $[-\frac{1}{2}, 1 - \frac{1}{2}]$, orthogonal: $[-\frac{11}{2}, 1, \frac{15}{2}]$
(c) Parallel: $[\frac{60}{49}, -\frac{40}{49}, \frac{120}{49}]$, orthogonal: $[-\frac{354}{49}, \frac{138}{49}, \frac{223}{49}]$

(21) From the lower triangle in the figure, we have $(\mathbf{proj_r x}) + (\mathbf{proj_r x} - \mathbf{x})$ = reflection of $\mathbf{x}$ (see Figure 10).

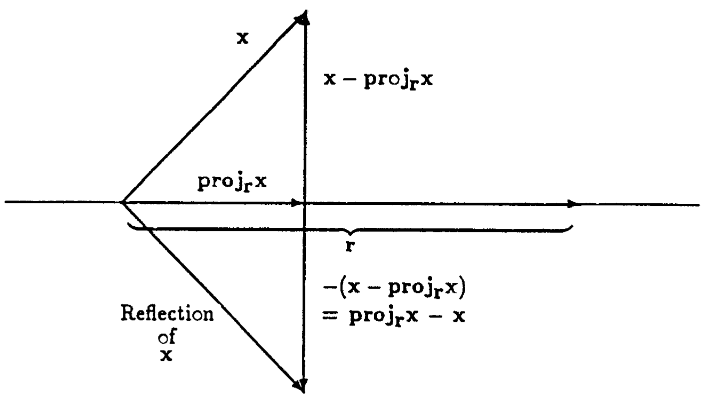


Figure 10: Reflection of a vector $\mathbf{x}$ through a line.

(22) For the case $\|\mathbf{x}\| \le \|\mathbf{y}\|$: $|\,\|\mathbf{x}\| - \|\mathbf{y}\|\,| = \|\mathbf{y}\| - \|\mathbf{x}\| = \|\mathbf{x}+\mathbf{y}+(-\mathbf{x})\| - \|\mathbf{x}\| \le \|\mathbf{x}+\mathbf{y}\| + \|-\mathbf{x}\| - \|\mathbf{x}\|$ (by the Triangle Inequality) $= \|\mathbf{x}+\mathbf{y}\| + \|\mathbf{x}\| - \|\mathbf{x}\| = \|\mathbf{x}+\mathbf{y}\|$.

The case $\|\mathbf{x}\| \ge \|\mathbf{y}\|$ is done similarly, with the roles of $\mathbf{x}$ and $\mathbf{y}$ reversed.

(23) (a) Note that $\mathbf{proj_x y} = [\frac{8}{5}, -\frac{6}{5}, 2] = \frac{2}{5}\mathbf{x}$, and $\mathbf{y} - \mathbf{proj_x y} = [\frac{7}{5}, -\frac{24}{5}, -4]$ is orthogonal to $\mathbf{x}$.

Let $\mathbf{w} = \mathbf{y} - \mathbf{proj_x y}$. Then since $\mathbf{y} = \mathbf{proj_x y} + (\mathbf{y} - \mathbf{proj_x y})$, we have $\mathbf{y} = \frac{2}{5}\mathbf{x} + \mathbf{w}$, where $\mathbf{w}$ is orthogonal to $\mathbf{x}$.

(b) Let $c = (\mathbf{x}\cdot\mathbf{y})/(\|\mathbf{x}\|^2)$ (so that $c\mathbf{x} = \mathbf{proj_x y}$), and let $\mathbf{w} = \mathbf{y} - \mathbf{proj_x y}$, which is orthogonal to $\mathbf{x}$ by the argument before Theorem 1.11. Then $\mathbf{y} = c\mathbf{x} + \mathbf{w}$, where $\mathbf{w}$ is orthogonal to $\mathbf{x}$.

(c) Suppose $c\mathbf{x}+\mathbf{w} = d\mathbf{x}+\mathbf{v}$. Then $(d-c)\mathbf{x} = \mathbf{w}-\mathbf{v}$, and $(d-c)\mathbf{x}\cdot(\mathbf{w}-\mathbf{v}) = (\mathbf{w}-\mathbf{v})\cdot(\mathbf{w}-\mathbf{v}) = \|\mathbf{w}-\mathbf{v}\|^2$. But $(d-c)\mathbf{x}\cdot(\mathbf{w}-\mathbf{v}) = (d-c)(\mathbf{x}\cdot\mathbf{w}) - (d-c)(\mathbf{x}\cdot\mathbf{v}) = 0$, since $\mathbf{v}$ and $\mathbf{w}$ are orthogonal to $\mathbf{x}$. Hence $\|\mathbf{w}-\mathbf{v}\|^2 = 0 \Longrightarrow \mathbf{w}-\mathbf{v} = \mathbf{0} \Longrightarrow \mathbf{w} = \mathbf{v}$. Then, $c\mathbf{x} = d\mathbf{x} \Longrightarrow c = d$, from Theorem 1.4, since $\mathbf{x}$ is nonzero.

(24) If θ is the angle between $\mathbf{x}$ and $\mathbf{y}$, and ϕ is the angle between $\mathbf{proj_x y}$ and $\mathbf{proj_y x}$, then

$$\cos\phi = \frac{\mathbf{proj_x y}\cdot\mathbf{proj_y x}}{\|\mathbf{proj_x y}\|\,\|\mathbf{proj_y x}\|} = \frac{\left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x}\cdot\left(\frac{\mathbf{y}\cdot\mathbf{x}}{\|\mathbf{y}\|^2}\right)\mathbf{y}}{\left\|\left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x}\right\|\;\left\|\left(\frac{\mathbf{y}\cdot\mathbf{x}}{\|\mathbf{y}\|^2}\right)\mathbf{y}\right\|} = \frac{\left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\left(\frac{\mathbf{y}\cdot\mathbf{x}}{\|\mathbf{y}\|^2}\right)(\mathbf{x}\cdot\mathbf{y})}{\left(\frac{|\mathbf{x}\cdot\mathbf{y}|}{\|\mathbf{x}\|^2}\right)\|\mathbf{x}\|\;\left(\frac{|\mathbf{y}\cdot\mathbf{x}|}{\|\mathbf{y}\|^2}\right)\|\mathbf{y}\|}$$

$$= \frac{\left(\frac{(\mathbf{x}\cdot\mathbf{y})^3}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}\right)}{\left(\frac{(\mathbf{x}\cdot\mathbf{y})^2}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)} = \frac{(\mathbf{x}\cdot\mathbf{y})}{\|\mathbf{x}\|\,\|\mathbf{y}\|} = \cos\theta.$$

Hence $\phi = \theta$.

(25) (a) T (b) T (c) F (d) F (e) T (f) F

Section 1.3

(1) (a) We have $\|4\mathbf{x} + 7\mathbf{y}\| \leq \|4\mathbf{x}\| + \|7\mathbf{y}\| = 4\|\mathbf{x}\| + 7\|\mathbf{y}\| \leq 7\|\mathbf{x}\| + 7\|\mathbf{y}\| = 7(\|\mathbf{x}\| + \|\mathbf{y}\|)$.

(b) Let $m = \max\{|c|, |d|\}$. Then $\|c\mathbf{x} \pm d\mathbf{y}\| \leq m(\|\mathbf{x}\| + \|\mathbf{y}\|)$.

(2) (a) Note that $6j - 5 = 3(2(j-1)) + 1$. Let $k = 2(j-1)$.

(b) Consider the number 4.

(3) Note that since $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$, $\mathbf{proj_x y} = \mathbf{0}$ iff $(\mathbf{x}\cdot\mathbf{y})/(\|\mathbf{x}\|^2) = 0$ iff $\mathbf{x}\cdot\mathbf{y} = 0$ iff $\mathbf{y}\cdot\mathbf{x} = 0$ iff $(\mathbf{y}\cdot\mathbf{x})/(\|\mathbf{y}\|^2) = 0$ iff $\mathbf{proj_y x} = \mathbf{0}$.

(4) If $\mathbf{y} = c\mathbf{x}$ (for $c > 0$), then $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}+c\mathbf{x}\| = (1+c)\|\mathbf{x}\| = \|\mathbf{x}\| + c\|\mathbf{x}\| = \|\mathbf{x}\| + \|c\mathbf{x}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$. On the other hand, if $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$, then $\|\mathbf{x} + \mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$. Now $\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y})\cdot(\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2$, while $(\|\mathbf{x}\| + \|\mathbf{y}\|)^2 = \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2$. Hence $\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|$. By Result 4, $\mathbf{y} = c\mathbf{x}$, for some $c > 0$.

(5) (a) Suppose $\mathbf{y} \neq \mathbf{0}$. We must show that $\mathbf{x}$ is not orthogonal to $\mathbf{y}$. Now $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2$, so $\|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2$. Hence $\|\mathbf{y}\|^2 = -2(\mathbf{x}\cdot\mathbf{y})$. Since $\mathbf{y} \neq \mathbf{0}$, we have $\|\mathbf{y}\|^2 \neq 0$, and so $\mathbf{x}\cdot\mathbf{y} \neq 0$.

(b) Suppose $\mathbf{x}$ is not a unit vector. We must show that $\mathbf{x}\cdot\mathbf{y} \neq 1$.
Now $\mathbf{proj_x y} = \mathbf{x} \Longrightarrow ((\mathbf{x}\cdot\mathbf{y})/(\|\mathbf{x}\|^2))\mathbf{x} = 1\mathbf{x} \Longrightarrow (\mathbf{x}\cdot\mathbf{y})/(\|\mathbf{x}\|^2) = 1 \Longrightarrow \mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|^2$.
But then $\|\mathbf{x}\| \neq 1 \Longrightarrow \|\mathbf{x}\|^2 \neq 1 \Longrightarrow \mathbf{x}\cdot\mathbf{y} \neq 1$.

(6) (a) Consider $\mathbf{x} = [1, 0, 0]$ and $\mathbf{y} = [1, 1, 0]$.

(b) If $\mathbf{x} \neq \mathbf{y}$, then $\mathbf{x} \cdot \mathbf{y} \neq \|\mathbf{x}\|^2$.

(c) Yes

(7) See the answer for Exercise 11(a) in Section 1.2.

(8) If $\|\mathbf{x}\| > \|\mathbf{y}\|$, then $\|\mathbf{x}\|^2 > \|\mathbf{y}\|^2$, and so $\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 > 0$. But then $(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) > 0$, and so the cosine of the angle between $(\mathbf{x}+\mathbf{y})$ and $(\mathbf{x}-\mathbf{y})$ is positive. Thus the angle between $(\mathbf{x}+\mathbf{y})$ and $(\mathbf{x}-\mathbf{y})$ is acute.

(9) (a) Contrapositive: If $\mathbf{x} = \mathbf{0}$, then $\mathbf{x}$ is not a unit vector.
Converse: If $\mathbf{x}$ is nonzero, then $\mathbf{x}$ is a unit vector.
Inverse: If $\mathbf{x}$ is not a unit vector, then $\mathbf{x} = \mathbf{0}$.

(b) (Let $\mathbf{x}$ and $\mathbf{y}$ be nonzero vectors.)
Contrapositive: If $\mathbf{y} \neq \mathbf{proj_x y}$, then $\mathbf{x}$ is not parallel to $\mathbf{y}$.
Converse: If $\mathbf{y} = \mathbf{proj_x y}$, then $\mathbf{x} \parallel \mathbf{y}$.
Inverse: If $\mathbf{x}$ is not parallel to $\mathbf{y}$, then $\mathbf{y} \neq \mathbf{proj_x y}$.

(c) (Let $\mathbf{x}$, $\mathbf{y}$ be nonzero vectors.)
Contrapositive: If $\mathbf{proj_y x} \neq \mathbf{0}$, then $\mathbf{proj_x y} \neq \mathbf{0}$.
Converse: If $\mathbf{proj_y x} = \mathbf{0}$, then $\mathbf{proj_x y} = \mathbf{0}$.
Inverse: If $\mathbf{proj_x y} \neq \mathbf{0}$, then $\mathbf{proj_y x} \neq \mathbf{0}$.

(10) (a) Converse: Let $\mathbf{x}$ and $\mathbf{y}$ be nonzero vectors in $\mathbb{R}^n$. If $\|\mathbf{x} + \mathbf{y}\| > \|\mathbf{y}\|$, then $\mathbf{x}\cdot\mathbf{y} \geq 0$.

(b) Let $\mathbf{x} = [2, -1]$ and $\mathbf{y} = [0, 2]$.

(11) (a) Converse: Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be vectors in $\mathbb{R}^n$. If $\mathbf{y} = \mathbf{z}$, then $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{z}$. The converse is obviously true, but the original statement is false in general, with counterexample $\mathbf{x} = [1, 1]$, $\mathbf{y} = [1, -1]$, and $\mathbf{z} = [-1, 1]$.

(b) Converse: Let $\mathbf{x}$ and $\mathbf{y}$ be vectors in $\mathbb{R}^n$. If $\|\mathbf{x} + \mathbf{y}\| \geq \|\mathbf{y}\|$, then $\mathbf{x} \cdot \mathbf{y} = 0$. The original statement is true, but the converse is false in general. Proof of the original statement follows from $\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + 2(\mathbf{x} \cdot \mathbf{y}) + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \geq \|\mathbf{y}\|^2$.
Counterexample to converse: let $\mathbf{x} = [1, 0]$, $\mathbf{y} = [1, 1]$.

(c) Converse: Let $\mathbf{x}$, $\mathbf{y}$ be vectors in $\mathbb{R}^n$, with $n > 1$. If $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$, then $\mathbf{x}\cdot\mathbf{y} = 0$. The converse is obviously true, but the original statement is false in general, with counterexample $\mathbf{x} = [1, -1]$ and $\mathbf{y} = [1, 1]$.

(12) Suppose $\mathbf{x} \perp \mathbf{y}$ and n is odd. Then $\mathbf{x} \cdot \mathbf{y} = 0$. Now $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$. But each product $x_i y_i$ equals either 1 or -1. If exactly k of these products equal 1, then $\mathbf{x} \cdot \mathbf{y} = k - (n - k) = -n + 2k$. Hence $-n + 2k = 0$, and so $n = 2k$, contradicting n odd.

(13) Suppose that $[6, 5]$, $[-2, 3]$, and $[x_1, x_2]$ are mutually orthogonal, with $[x_1, x_2] \neq [0, 0]$. Then $6x_1+5x_2 = 0$ and $-2x_1 + 3x_2 = 0$. Multiplying the latter equation by 3, we obtain $-6x_1 + 9x_2 = 0$. Adding this to the first equation gives $14x_2 = 0$, which means $x_2 = 0$, and hence $x_1 = 0$. Thus, $[x_1, x_2] = [0, 0]$, a contradiction.

(14) Assume that $||\mathbf{x}|| = 1$. We know that $||\mathbf{proj_x y}|| \neq \mathbf{x}\cdot\mathbf{y}$. However, $\mathbf{proj_x y} = \left((\mathbf{x}\cdot\mathbf{y})/||\mathbf{x}||^2\right)\mathbf{x}$, so $||\mathbf{proj_x y}|| = |\mathbf{x}\cdot\mathbf{y}|/||\mathbf{x}||^2 = |\mathbf{x}\cdot\mathbf{y}|$. But $\mathbf{x} \cdot \mathbf{y} > 0$ because the angle between $\mathbf{x}$ and $\mathbf{y}$ is acute, so $||\mathbf{proj_x y}|| = \mathbf{x}\cdot\mathbf{y}$, a contradiction. Thus, $||\mathbf{x}|| \neq 1$.

(15) Base Step ($n = 1$): $\mathbf{x}_1 = \mathbf{x}_1$.
Inductive Step: Assume $\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{n-1} + \mathbf{x}_n = \mathbf{x}_n + \mathbf{x}_{n-1} + \cdots + \mathbf{x}_2 + \mathbf{x}_1$, for some $n \geq 1$.
Prove: $\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{n-1} + \mathbf{x}_n + \mathbf{x}_{n+1} = \mathbf{x}_{n+1} + \mathbf{x}_n + \mathbf{x}_{n-1} + \cdots + \mathbf{x}_2 + \mathbf{x}_1$.
But,

$$\begin{aligned}
\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{n-1} + \mathbf{x}_n + \mathbf{x}_{n+1} &= (\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{n-1} + \mathbf{x}_n) + \mathbf{x}_{n+1} \\
&= \mathbf{x}_{n+1} + (\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_{n-1} + \mathbf{x}_n) \\
&= \mathbf{x}_{n+1} + (\mathbf{x}_n + \mathbf{x}_{n-1} + \cdots + \mathbf{x}_2 + \mathbf{x}_1) \\
&\qquad \text{(by the inductive hypothesis)} \\
&= \mathbf{x}_{n+1} + \mathbf{x}_n + \mathbf{x}_{n-1} + \cdots + \mathbf{x}_2 + \mathbf{x}_1.
\end{aligned}$$

(16) Base Step ($m = 1$): $\|\mathbf{x}_1\| \leq \|\mathbf{x}_1\|$.
Inductive Step: Assume $\|\mathbf{x}_1 + \cdots + \mathbf{x}_m\| \leq \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_m\|$, for some $m \geq 1$.
Prove: $\|\mathbf{x}_1 + \cdots + \mathbf{x}_m + \mathbf{x}_{m+1}\| \leq \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_m\| + \|\mathbf{x}_{m+1}\|$.
But, by the Triangle Inequality,

$$\begin{aligned}
\|(\mathbf{x}_1 + \cdots + \mathbf{x}_m) + \mathbf{x}_{m+1}\| &\leq \|\mathbf{x}_1 + \cdots + \mathbf{x}_m\| + \|\mathbf{x}_{m+1}\| \\
&\leq \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_m\| + \|\mathbf{x}_{m+1}\|,
\end{aligned}$$

by the inductive hypothesis.

(17) Base Step ($k = 1$): $\|\mathbf{x}_1\|^2 = \|\mathbf{x}_1\|^2$.
Inductive Step: Assume $\|\mathbf{x}_1 + \cdots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \cdots + \|\mathbf{x}_k\|^2$.

Prove: $\|\mathbf{x}_1 + \cdots + \mathbf{x}_k + \mathbf{x}_{k+1}\|^2 = \|\mathbf{x}_1\|^2 + \cdots + \|\mathbf{x}_k\|^2 + \|\mathbf{x}_{k+1}\|^2$.
We have

$$\begin{aligned} \|(\mathbf{x}_1 + \cdots + \mathbf{x}_k) + \mathbf{x}_{k+1}\|^2 &= \|\mathbf{x}_1 + \cdots + \mathbf{x}_k\|^2 + 2((\mathbf{x}_1 + \cdots + \mathbf{x}_k)\cdot\mathbf{x}_{k+1}) + \|\mathbf{x}_{k+1}\|^2 \\ &= \|\mathbf{x}_1 + \cdots + \mathbf{x}_k\|^2 + \|\mathbf{x}_{k+1}\|^2 \\ &\qquad \text{(since } \mathbf{x}_{k+1} \text{is orthogonal to all of } \mathbf{x}_1, \ldots, \mathbf{x}_k) \\ &= \|\mathbf{x}_1\|^2 + \cdots + \|\mathbf{x}_k\|^2 + \|\mathbf{x}_{k+1}\|^2, \end{aligned}$$

by the inductive hypothesis.

(18) Base Step ($k = 1$): We must show $(a_1\mathbf{x}_1)\cdot\mathbf{y} \leq |a_1|\,\|\mathbf{y}\|$.
But,

$$\begin{aligned} (a_1\mathbf{x}_1)\cdot\mathbf{y} &\leq |(a_1\mathbf{x}_1)\cdot\mathbf{y}| \leq \|a_1\mathbf{x}_1\|\,\|\mathbf{y}\| \\ &\qquad \text{(by the Cauchy-Schwarz Inequality)} \\ &= |a_1|\,\|\mathbf{x}_1\|\,\|\mathbf{y}\| = |a_1|\,\|\mathbf{y}\|, \end{aligned}$$

since $\mathbf{x}_1$ is a unit vector.
Inductive Step: Assume $(a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k)\cdot\mathbf{y} \leq (|a_1| + \cdots + |a_k|)\|\mathbf{y}\|$, for some $k \geq 1$.
Prove: $(a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k + a_{k+1}\mathbf{x}_{k+1})\cdot\mathbf{y} \leq (|a_1| + \cdots + |a_k| + |a_{k+1}|)\|\mathbf{y}\|$.
We have

$$\begin{aligned} ((a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k) + a_{k+1}\mathbf{x}_{k+1})\cdot\mathbf{y} &= (a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k)\cdot\mathbf{y} + (a_{k+1}\mathbf{x}_{k+1})\cdot\mathbf{y} \\ &\leq (|a_1| + \cdots + |a_k|)\|\mathbf{y}\| + (a_{k+1}\mathbf{x}_{k+1})\cdot\mathbf{y} \\ &\qquad \text{(by the inductive hypothesis)} \\ &\leq (|a_1| + \cdots + |a_k|)\|\mathbf{y}\| + |a_{k+1}|\,||\mathbf{y}|| \\ &\qquad \text{(by an argument similar to the Base Step)} \\ &= (|a_1| + \cdots + |a_k| + |a_{k+1}|)\|\mathbf{y}\|. \end{aligned}$$

(19) Step 1 cannot be reversed, because y could equal $\pm(x^2 + 2)$.
Step 2 cannot be reversed, because y^2 could equal $x^4 + 4x^2 + c$.
Step 4 cannot be reversed, because in general y does not have to equal $x^2 + 2$.
Step 6 cannot be reversed, since $\frac{dy}{dx}$ could equal $2x + c$.

All other steps remain true when reversed.

(20) (a) For every unit vector $\mathbf{x}$ in $\mathbb{R}^3$, $\mathbf{x} \cdot [1, -2, 3] \neq 0$.
(b) $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{x}\cdot\mathbf{y} \leq 0$, for some vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.
(c) $\mathbf{x} = \mathbf{0}$ or $\|\mathbf{x} + \mathbf{y}\| \neq \|\mathbf{y}\|$, for all vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.
(d) There is some vector $\mathbf{x} \in \mathbb{R}^n$ for which $\mathbf{x}\cdot\mathbf{x} \leq 0$.
(e) There is an $\mathbf{x} \in \mathbb{R}^3$ such that for every nonzero $\mathbf{y} \in \mathbb{R}^3$, $\mathbf{x} \cdot \mathbf{y} \neq 0$.
(f) For every $\mathbf{x} \in \mathbb{R}^4$, there is some $\mathbf{y} \in \mathbb{R}^4$ such that $\mathbf{x}\cdot\mathbf{y} \neq 0$.

(21) (a) Contrapositive: If $\mathbf{x} \neq \mathbf{0}$ and $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{y}\|$, then $\mathbf{x} \cdot \mathbf{y} \neq 0$.
Converse: If $\mathbf{x} = \mathbf{0}$ or $\|\mathbf{x} - \mathbf{y}\| > \|\mathbf{y}\|$, then $\mathbf{x} \cdot \mathbf{y} = 0$.
Inverse: If $\mathbf{x} \cdot \mathbf{y} \neq 0$, then $\mathbf{x} \neq \mathbf{0}$ and $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{y}\|$.

(b) Contrapositive: If $\|\mathbf{x}-\mathbf{y}\| \leq \|\mathbf{y}\|$, then either $\mathbf{x}=\mathbf{0}$ or $\mathbf{x}\cdot\mathbf{y} \neq 0$.
Converse: If $\|\mathbf{x}-\mathbf{y}\| > \|\mathbf{y}\|$, then $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{x}\cdot\mathbf{y} = 0$.
Inverse: If $\mathbf{x}=\mathbf{0}$ or $\mathbf{x}\cdot\mathbf{y} \neq 0$, then $\|\mathbf{x}-\mathbf{y}\| \leq \|\mathbf{y}\|$.

(22) Suppose $\mathbf{x} \neq \mathbf{0}$. We must prove $\mathbf{x}\cdot\mathbf{y} \neq 0$ for some vector $\mathbf{y} \in \mathbb{R}^n$. Let $\mathbf{y} = \mathbf{x}$.

(23) Let $\mathbf{x} = [1,1]$, $\mathbf{y} = [1,-1]$.

(24) Let $\mathbf{y} = [1,-2,2]$. Then since $\mathbf{x}\cdot\mathbf{y} \geq 0$, Result 3 implies that $\|\mathbf{x}+\mathbf{y}\| > \|\mathbf{y}\| = 3$.

(25) (a) F (b) T (c) T (d) F (e) F (f) F (g) F (h) T (i) F

Section 1.4

(1) (a) $\begin{bmatrix} 2 & 1 & 3 \\ 2 & 7 & -5 \\ 9 & 0 & -1 \end{bmatrix}$

(b) Impossible

(c) $\begin{bmatrix} -16 & 8 & 12 \\ 0 & 20 & -4 \\ 24 & 4 & -8 \end{bmatrix}$

(d) $\begin{bmatrix} -32 & 8 & 6 \\ -8 & 2 & 14 \\ 0 & 6 & -8 \end{bmatrix}$

(e) Impossible

(f) $\begin{bmatrix} -7 & 0 & 8 \\ -1 & 3 & 1 \\ 9 & 9 & -5 \end{bmatrix}$

(g) $\begin{bmatrix} -23 & 14 & -9 \\ -5 & 8 & 8 \\ -9 & -18 & 1 \end{bmatrix}$

(h) Impossible

(i) $\begin{bmatrix} -1 & 1 & 12 \\ -1 & 5 & 8 \\ 8 & -3 & -4 \end{bmatrix}$

(j) $\begin{bmatrix} -1 & 1 & 12 \\ -1 & 5 & 8 \\ 8 & -3 & -4 \end{bmatrix}$

(k) $\begin{bmatrix} -12 & 8 & -6 \\ 10 & -8 & 26 \end{bmatrix}$

(l) Impossible

(m) Impossible

(n) $\begin{bmatrix} 13 & -6 & 2 \\ 3 & -3 & -5 \\ 3 & 5 & 1 \end{bmatrix}$

(2) Square: $\mathbf{B}, \mathbf{C}, \mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{J}, \mathbf{K}, \mathbf{L}, \mathbf{M}, \mathbf{N}, \mathbf{P}, \mathbf{Q}$
Diagonal: $\mathbf{B}, \mathbf{G}, \mathbf{N}$
Upper triangular: $\mathbf{B}, \mathbf{G}, \mathbf{L}, \mathbf{N}$
Lower triangular: $\mathbf{B}, \mathbf{G}, \mathbf{M}, \mathbf{N}, \mathbf{Q}$
Symmetric: $\mathbf{B}, \mathbf{F}, \mathbf{G}, \mathbf{J}, \mathbf{N}, \mathbf{P}$
Skew-symmetric: $\mathbf{H}$ (but not $\mathbf{E}, \mathbf{C}, \mathbf{K}$)
Transposes: $\mathbf{A}^T = \begin{bmatrix} -1 & 0 & 6 \\ 4 & 1 & 0 \end{bmatrix}$, $\mathbf{B}^T = \mathbf{B}$, $\mathbf{C}^T = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$, and so on

(3) (a) $\begin{bmatrix} 3 & -\frac{1}{2} & \frac{5}{2} \\ -\frac{1}{2} & 2 & 1 \\ \frac{5}{2} & 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & 0 & 4 \\ -\frac{3}{2} & -4 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & \frac{3}{2} & 0 \\ \frac{3}{2} & 3 & -1 \\ 0 & -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{3}{2} & -4 \\ \frac{3}{2} & 0 & 0 \\ 4 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} + \begin{bmatrix} 0 & 3 & 4 & -1 \\ -3 & 0 & -1 & 2 \\ -4 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} -3 & 7 & -2 & -1 \\ 7 & 4 & 3 & -6 \\ -2 & 3 & 5 & -8 \\ -1 & -6 & -8 & -5 \end{bmatrix} + \begin{bmatrix} 0 & -4 & 7 & -3 \\ 4 & 0 & 2 & 5 \\ -7 & -2 & 0 & -6 \\ 3 & -5 & 6 & 0 \end{bmatrix}$

(4) If $\mathbf{A}^T = \mathbf{B}^T$, then $\left(\mathbf{A}^T\right)^T = \left(\mathbf{B}^T\right)^T$, implying $\mathbf{A} = \mathbf{B}$, by part (1) of Theorem 1.13.

(5) (a) If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}^T$ is $n \times m$. If $\mathbf{A} = \mathbf{A}^T$, then $m = n$.

(b) If $\mathbf{A}$ is a diagonal matrix and if $i \neq j$, then $a_{ij} = 0 = a_{ji}$.

(c) Follows directly from part (b), since $\mathbf{I}_n$ is diagonal.

(d) The matrix must be a square zero matrix; that is, $\mathbf{O}_n$, for some n.

(6) (a) If $i \neq j$, then $a_{ij} + b_{ij} = 0 + 0 = 0$.

(b) Use the fact that $a_{ij} + b_{ij} = a_{ji} + b_{ji}$.

(7) Use induction on n. Base Step ($n = 1$): Obvious. Inductive Step: Assume $\mathbf{A}_1, \ldots, \mathbf{A}_{k+1}$ and $\mathbf{B} = \sum_{i=1}^{k} \mathbf{A}_i$ are upper triangular. Prove that $\mathbf{D} = \sum_{i=1}^{k+1} \mathbf{A}_i$ is upper triangular. Let $\mathbf{C} = \mathbf{A}_{k+1}$. Then $\mathbf{D} = \mathbf{B} + \mathbf{C}$. Hence, $d_{ij} = b_{ij} + c_{ij} = 0 + 0 = 0$ if $i > j$ (by the inductive hypothesis). Hence $\mathbf{D}$ is upper triangular.

(8) (a) Let $\mathbf{B} = \mathbf{A}^T$. Then $b_{ij} = a_{ji} = a_{ij} = b_{ji}$. Let $\mathbf{D} = c\mathbf{A}$. Then $d_{ij} = ca_{ij} = ca_{ji} = d_{ji}$.

(b) Let $\mathbf{B} = \mathbf{A}^T$. Then $b_{ij} = a_{ji} = -a_{ij} = -b_{ji}$.
Let $\mathbf{D} = c\mathbf{A}$. Then $d_{ij} = ca_{ij} = c(-a_{ji}) = -ca_{ji} = -d_{ji}$.

(9) (a) Part (4): Let $\mathbf{B} = \mathbf{A} + (-\mathbf{A})$ $(= (-\mathbf{A}) + \mathbf{A}$, by part (1)). Then $b_{ij} = a_{ij} + (-a_{ij}) = 0$.

(b) Part (5): Let $\mathbf{D} = c(\mathbf{A} + \mathbf{B})$ and let $\mathbf{E} = c\mathbf{A} + c\mathbf{B}$. Then, $d_{ij} = c(a_{ij} + b_{ij}) = ca_{ij} + cb_{ij} = e_{ij}$.

(c) Part (7): Let $\mathbf{B} = (cd)\mathbf{A}$ and let $\mathbf{E} = c(d\mathbf{A})$. Then $b_{ij} = (cd)a_{ij} = c(da_{ij}) = e_{ij}$.

(10) (a) Part (1): $(\mathbf{A}^T)^T$ and $\mathbf{A}$ are both $m \times n$ matrices. The (i, j) entry of $(\mathbf{A}^T)^T = (j, i)$ entry of $\mathbf{A}^T$ $= (i, j)$ entry of $\mathbf{A}$. Thus, $(\mathbf{A}^T)^T = \mathbf{A}$.

(b) Part (2), Subtraction Case: $(\mathbf{A} - \mathbf{B})^T$ and $\mathbf{A}^T - \mathbf{B}^T$ are both $n \times m$ matrices. The (i, j) entry of $(\mathbf{A} - \mathbf{B})^T = (j, i)$ entry of $\mathbf{A} - \mathbf{B} = (j, i)$ entry of $\mathbf{A} - (j,i)$ entry of $\mathbf{B} = (i, j)$ entry of $\mathbf{A}^T$ $-$ (i,j) entry of $\mathbf{B}^T = (i, j)$ entry of $\mathbf{A}^T - \mathbf{B}^T$. Thus, $(\mathbf{A} - \mathbf{B})^T = \mathbf{A}^T - \mathbf{B}^T$.

(c) Part (3): $(c\mathbf{A})^T$ and $c(\mathbf{A}^T)$ are both $n \times m$ matrices. The (i, j) entry of $(c\mathbf{A})^T = (j, i)$ entry of $c\mathbf{A} = c((j, i)$ entry of $\mathbf{A}) = c((i, j)$ entry of $\mathbf{A}^T) = (i, j)$ entry of $c(\mathbf{A}^T)$. Thus, $(c\mathbf{A})^T = c(\mathbf{A}^T)$.

(11) Assume $c \neq 0$. We must show $\mathbf{A} = \mathbf{O}_{mn}$. But for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$, $ca_{ij} = 0$ with $c \neq 0$. Hence, all $a_{ij} = 0$.

(12) (a) $\mathbf{S} + \mathbf{V} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) = \frac{1}{2}(2\mathbf{A}) = \mathbf{A}$;
$\mathbf{S}^T = (\frac{1}{2}(\mathbf{A} + \mathbf{A}^T))^T = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)^T = \frac{1}{2}(\mathbf{A}^T + (\mathbf{A}^T)^T) = \frac{1}{2}(\mathbf{A}^T + \mathbf{A}) = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) = \mathbf{S}$;
$\mathbf{V}^T = (\frac{1}{2}(\mathbf{A} - \mathbf{A}^T))^T = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T)^T = \frac{1}{2}(\mathbf{A}^T - (\mathbf{A}^T)^T) = \frac{1}{2}(\mathbf{A}^T - \mathbf{A}) = -\frac{1}{2}(\mathbf{A} - \mathbf{A}^T) = -\mathbf{V}$

(b) $\mathbf{S}_1 - \mathbf{V}_1 = \mathbf{S}_2 - \mathbf{V}_2$

(c) Follows immediately from part (b).

(d) Part (a) shows $\mathbf{A}$ can be decomposed as the sum of a symmetric matrix and a skew-symmetric matrix, while parts (b) and (c) show that the decomposition for $\mathbf{A}$ is unique.

(13) (a) Trace ($\mathbf{B}$) = 1, trace ($\mathbf{C}$) = 0, trace ($\mathbf{E}$) = −6, trace ($\mathbf{F}$) = 2, trace ($\mathbf{G}$) = 18, trace ($\mathbf{H}$) = 0, trace ($\mathbf{J}$) = 1, trace ($\mathbf{K}$) = 4, trace ($\mathbf{L}$) = 3, trace ($\mathbf{M}$) = 0, trace ($\mathbf{N}$) = 3, trace ($\mathbf{P}$) = 0, trace ($\mathbf{Q}$) = 1

(b) Part (i): Let $\mathbf{D} = \mathbf{A}+\mathbf{B}$. Then $\text{trace}(\mathbf{D}) = \sum_{i=1}^n d_{ii} = \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = \text{trace}(\mathbf{A})+\text{trace}(\mathbf{B})$.
Part (ii): Let $\mathbf{B} = c\mathbf{A}$. Then $\text{trace}(\mathbf{B}) = \sum_{i=1}^n b_{ii} = \sum_{i=1}^n ca_{ii} = c\sum_{i=1}^n a_{ii} = c(\text{trace}(\mathbf{A}))$.
Part (iii): Let $\mathbf{B} = \mathbf{A}^T$. Then $\text{trace}(\mathbf{B}) = \sum_{i=1}^n b_{ii} = \sum_{i=1}^n a_{ii}$ (since $b_{ii} = a_{ii}$ for all i) = trace($\mathbf{A}$).

(c) Not necessarily: consider the matrices $\mathbf{L}$ and $\mathbf{N}$ in Exercise 2. (Note: If $n = 1$, the statement is true.)

(14) (a) F (b) T (c) F (d) T (e) T

Section 1.5

(1) (a) Impossible

(b) $\begin{bmatrix} 34 & -24 \\ 42 & 49 \\ 8 & -22 \end{bmatrix}$

(c) Impossible

(d) $\begin{bmatrix} 73 & -34 \\ 77 & -25 \\ 19 & -14 \end{bmatrix}$

(e) $[-38]$

(f) $\begin{bmatrix} -24 & 48 & -16 \\ 3 & -6 & 2 \\ -12 & 24 & -8 \end{bmatrix}$

(g) Impossible

(h) $[56, -8]$

(i) Impossible

(j) Impossible

(k) $\begin{bmatrix} 22 & 9 & -6 \\ 9 & 73 & 18 \\ 2 & -6 & 4 \end{bmatrix}$

(l) $\begin{bmatrix} 5 & 3 & 2 & 5 \\ 4 & 1 & 3 & 1 \\ 1 & 1 & 0 & 2 \\ 4 & 1 & 3 & 1 \end{bmatrix}$

(m) $[226981]$

(n) $\begin{bmatrix} 146 & 5 & -603 \\ 154 & 27 & -560 \\ 38 & -9 & -193 \end{bmatrix}$

(o) Impossible

(2) (a) No (b) Yes (c) No (d) Yes (e) No

(3) (a) $[15, -13, -8]$

(b) $\begin{bmatrix} 11 \\ 6 \\ 3 \end{bmatrix}$

(c) $[4]$

(d) $[2, 8, -2, 12]$

(4) (a) Valid, by Theorem 1.16, part (1)

(b) Invalid

(c) Valid, by Theorem 1.16, part (1)

(d) Valid, by Theorem 1.16, part (2)

(e) Valid, by Theorem 1.18

(f) Invalid

(g) Valid, by Theorem 1.16, part (3)

(h) Valid, by Theorem 1.16, part (2)

(i) Invalid

(j) Valid, by Theorem 1.16, part (3), and Theorem 1.18

(5)

	Salary	Fringe Benefits
Outlet 1	\$367500	\$78000
Outlet 2	\$225000	\$48000
Outlet 3	\$765000	\$162000
Outlet 4	\$360000	\$76500

(6)

	Tickets	Food	Souvenirs
June	\$1151300	\$3056900	\$2194400
July	\$1300700	\$3456700	\$2482400
August	\$981100	\$2615900	\$1905100

(7)

	Field 1	Field 2	Field 3
Nitrogen	1.00	0.45	0.65
Phosphate	0.90	0.35	0.75
Potash	0.95	0.35	0.85

(in tons)

(8)

	Chip 1	Chip 2	Chip 3	Chip 4
Rocket 1	2131	1569	1839	2750
Rocket 2	2122	1559	1811	2694
Rocket 3	2842	2102	2428	3618
Rocket 4	2456	1821	2097	3124

(9) (a) One example: $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$

(b) One example: $\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) Consider $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

(10) (a) Third row, fourth column entry of $\mathbf{AB}$

(b) Fourth row, first column entry of $\mathbf{AB}$

(c) Third row, second column entry of $\mathbf{BA}$

(d) Second row, fifth column entry of $\mathbf{BA}$

(11) (a) $\sum_{k=1}^{n} a_{3k} b_{k2}$

(b) $\sum_{k=1}^{m} b_{4k} a_{k1}$

(12) (a) $[-27, 43, -56]$

(b) $\begin{bmatrix} 56 \\ -57 \\ 18 \end{bmatrix}$

(13) (a) $[-61, 15, 20, 9]$

(b) $\begin{bmatrix} 43 \\ -41 \\ -12 \end{bmatrix}$

(14) (a) Let $\mathbf{B} = \mathbf{Ai}$. Then $\mathbf{B}$ is $m \times 1$ and $b_{k1} = \sum_{j=1}^{m} a_{kj} i_{j1} = (a_{k1})(1) + (a_{k2})(0) + (a_{k3})(0) = a_{k1}$.

(b) $\mathbf{Ae}_i = i$th column of $\mathbf{A}$.

(c) By part (b), each column of $\mathbf{A}$ is easily seen to be the zero vector by letting $\mathbf{x}$ equal each of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in turn.

(15) (a) Proof of Part (2): The (i, j) entry of $\mathbf{A}(\mathbf{B} + \mathbf{C})$

$=$ (ith row of $\mathbf{A}$)·(jth column of $(\mathbf{B} + \mathbf{C})$)

$=$ (ith row of $\mathbf{A}$)·(jth column of $\mathbf{B}$ + jth column of $\mathbf{C}$)

$=$ (ith row of $\mathbf{A}$)·(jth column of $\mathbf{B}$)
+ (ith row of $\mathbf{A}$)·(jth column of $\mathbf{C}$)

$=$ ((i, j) entry of $\mathbf{AB}$) + ((i, j) entry of $\mathbf{AC}$)

$=$ (i, j) entry of $(\mathbf{AB} + \mathbf{AC})$.

(b) Proof of Part (3): The (i, j) entry of $(\mathbf{A} + \mathbf{B})\mathbf{C}$
$= (i\text{th row of } (\mathbf{A} + \mathbf{B}))\cdot(j\text{th column of } \mathbf{C})$
$= (i\text{th row of } \mathbf{A} + i\text{th row of } \mathbf{B})\cdot(j\text{th column of } \mathbf{C})$
$= (i\text{th row of } \mathbf{A})\cdot(j\text{th column of } \mathbf{C}) + (i\text{th row of } \mathbf{B})\cdot(j\text{th column of } \mathbf{C})$
$= ((i, j) \text{ entry of } \mathbf{AC}) + ((i, j) \text{ entry of } \mathbf{BC})$
$= (i, j) \text{ entry of } (\mathbf{AC} + \mathbf{BC})$.

(c) For the first equation in Part (4), the (i, j) entry of $c(\mathbf{AB})$
$= c((i\text{th row of } \mathbf{A})\cdot(j\text{th column of } \mathbf{B}))$
$= (c(i\text{th row of } \mathbf{A}))\cdot(j\text{th column of } \mathbf{B})$
$= (i\text{th row of } c\mathbf{A})\cdot(j\text{th column of } \mathbf{B})$
$= (i, j) \text{ entry of } (c\mathbf{A})\mathbf{B}$.

For the second equation in Part (4), the (i, j) entry of $(c\mathbf{A})\mathbf{B}$
$= (i\text{th row of } c\mathbf{A})\cdot(j\text{th column of } \mathbf{B})$
$= (c(i\text{th row of } \mathbf{A}))\cdot(j\text{th column of } \mathbf{B})$
$= (i\text{th row of } \mathbf{A})\cdot(c(j\text{th column of } \mathbf{B}))$.
$= (i\text{th row of } \mathbf{A})\cdot(j\text{th column of } c\mathbf{B})$
$= (i, j) \text{ entry of } \mathbf{A}(c\mathbf{B})$.

(16) Let $\mathbf{B} = \mathbf{AO}_{np}$. Clearly $\mathbf{B}$ is an $m \times p$ matrix and $b_{ij} = \sum_{k=1}^{n} a_{ik}0 = 0$.

(17) Proof that $\mathbf{AI}_n = \mathbf{A}$: Let $\mathbf{B} = \mathbf{AI}_n$. Then $\mathbf{B}$ is clearly an $m \times n$ matrix and $b_{jl} = \sum_{k=1}^{n} a_{jk}i_{kl} = a_{jl}$, since $i_{kl} = 0$ unless $k = l$, in which case $i_{kk} = 1$. The proof that $\mathbf{I}_m\mathbf{A} = \mathbf{A}$ is similar.

(18) (a) We need to show $c_{ij} = 0$ if $i \neq j$. Now, if $i \neq j$, both factors in each term of the formula for c_{ij} in the formal definition of matrix multiplication are zero, except possibly for the terms $a_{ii}b_{ij}$ and $a_{ij}b_{jj}$. But since the factors b_{ij} and a_{ij} also equal zero, all terms in the formula for c_{ij} equal zero.

(b) Assume $i > j$. Consider the term $a_{ik}b_{kj}$ in the formula for c_{ij}. If $i > k$, then $a_{ik} = 0$. If $i \leq k$, then $j < k$, so $b_{kj} = 0$. Hence all terms in the formula for c_{ij} equal zero.

(c) Let $\mathbf{L}_1$ and $\mathbf{L}_2$ be lower triangular matrices. $\mathbf{U}_1 = \mathbf{L}_1^T$ and $\mathbf{U}_2 = \mathbf{L}_2^T$ are then upper triangular, and $\mathbf{L}_1\mathbf{L}_2 = \mathbf{U}_1^T\mathbf{U}_2^T = (\mathbf{U}_2\mathbf{U}_1)^T$ (by Theorem 1.18). But by part (b), $\mathbf{U}_2\mathbf{U}_1$ is upper triangular. So $\mathbf{L}_1\mathbf{L}_2$ is lower triangular.

(19) Base Step: Clearly, $(c\mathbf{A})^1 = c^1\mathbf{A}^1$.
Inductive Step: Assume $(c\mathbf{A})^n = c^n\mathbf{A}^n$, and prove $(c\mathbf{A})^{n+1} = c^{n+1}\mathbf{A}^{n+1}$. Now, $(c\mathbf{A})^{n+1} = (c\mathbf{A})^n(c\mathbf{A})$ $= (c^n\mathbf{A}^n)(c\mathbf{A})$ (by the inductive hypothesis) $= c^n c\mathbf{A}^n\mathbf{A}$ (by part (4) of Theorem 1.16) $= c^{n+1}\mathbf{A}^{n+1}$.

(20) (a) Proof of Part (1): Base Step: $\mathbf{A}^{s+0} = \mathbf{A}^s = \mathbf{A}^s\mathbf{I} = \mathbf{A}^s\mathbf{A}^0$.
Inductive Step: Assume $\mathbf{A}^{s+t} = \mathbf{A}^s\mathbf{A}^t$ for some $t \geq 0$. We must prove $\mathbf{A}^{s+(t+1)} = \mathbf{A}^s\mathbf{A}^{t+1}$. But $\mathbf{A}^{s+(t+1)} = \mathbf{A}^{(s+t)+1} = \mathbf{A}^{s+t}\mathbf{A} = (\mathbf{A}^s\mathbf{A}^t)\mathbf{A}$ (by the inductive hypothesis) $= \mathbf{A}^s(\mathbf{A}^t\mathbf{A}) = \mathbf{A}^s\mathbf{A}^{t+1}$.

(b) Proof of Part (2): (We only need prove $(\mathbf{A}^s)^t = \mathbf{A}^{st}$.) Base Step: $(\mathbf{A}^s)^0 = \mathbf{I} = \mathbf{A}^0 = \mathbf{A}^{s0}$.
Inductive Step: Assume $(\mathbf{A}^s)^t = \mathbf{A}^{st}$ for some integer $t \geq 0$. We must prove $(\mathbf{A}^s)^{t+1} = \mathbf{A}^{s(t+1)}$. But $(\mathbf{A}^s)^{t+1} = (\mathbf{A}^s)^t\mathbf{A}^s = \mathbf{A}^{st}\mathbf{A}^s$ (by the inductive hypothesis) $= \mathbf{A}^{st+s}$ (by part (1)) $= \mathbf{A}^{s(t+1)}$.

(21) $\mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l}$ (by part (1) of Theorem 1.17) $= \mathbf{A}^{l+k} = \mathbf{A}^l\mathbf{A}^k$ (by part (1) of Theorem 1.17).

(22) (a) $\mathbf{A}(c_1\mathbf{B}_1 + c_2\mathbf{B}_2 + \cdots + c_k\mathbf{B}_k) = c_1\mathbf{AB}_1 + c_2\mathbf{AB}_2 + \cdots + c_k\mathbf{AB}_k$ (by parts (2) and (4) of Theorem 1.16) $= c_1\mathbf{B}_1\mathbf{A} + c_2\mathbf{B}_2\mathbf{A} + \cdots + c_k\mathbf{B}_k\mathbf{A}$ (since $\mathbf{A}$ commutes with $\mathbf{B}_1$, $\mathbf{B}_2$, …, $\mathbf{B}_k$) $= (c_1\mathbf{B}_1 + c_2\mathbf{B}_2 + \cdots + c_k\mathbf{B}_k)\mathbf{A}$ (by part (3) of Theorem 1.16).

(b) By Exercise 21, $\mathbf{A}^j$ commutes with $\mathbf{A}^i$, for $i = 0$ to k. Therefore, by part (a), $\mathbf{A}^j$ commutes with $c_o\mathbf{I}_n + c_1\mathbf{A} + c_2\mathbf{A}^2 + \cdots + c_k\mathbf{A}^k$.

(c) By part (b), $(c_o\mathbf{I}_n + c_1\mathbf{A} + c_2\mathbf{A}^2 + \cdots + c_k\mathbf{A}^k)$ commutes with $\mathbf{A}^j$ for every nonnegative integer j. Therefore, by part (a), $(c_o\mathbf{I}_n + c_1\mathbf{A} + c_2\mathbf{A}^2 + \cdots + c_k\mathbf{A}^k)$ commutes with $(d_o\mathbf{I}_n + d_1\mathbf{A} + d_2\mathbf{A}^2 + \cdots + d_m\mathbf{A}^m)$.

(23) (a) If $\mathbf{A}$ is an $m \times n$ matrix, and $\mathbf{B}$ is a $p \times r$ matrix, then the fact that $\mathbf{AB}$ exists means $n = p$, and the fact that $\mathbf{BA}$ exists means $m = r$. Then $\mathbf{AB}$ is an $m \times m$ matrix, while $\mathbf{BA}$ is an $n \times n$ matrix. If $\mathbf{AB} = \mathbf{BA}$, then $m = n$.

(b) Note that by the Distributive Law, $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2$.

(24) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{A}(\mathbf{CB}) = (\mathbf{AC})\mathbf{B} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB})$.

(25) Use the fact that $\mathbf{A}^T\mathbf{B}^T = (\mathbf{BA})^T$, while $\mathbf{B}^T\mathbf{A}^T = (\mathbf{AB})^T$.

(26) $(\mathbf{AA}^T)^T = (\mathbf{A}^T)^T\mathbf{A}^T$ (by Theorem 1.18) $= \mathbf{AA}^T$. (Similarly for $\mathbf{A}^T\mathbf{A}$.)

(27) (a) If $\mathbf{A}, \mathbf{B}$ are both skew-symmetric, then $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T = (-\mathbf{B})(-\mathbf{A}) = (-1)(-1)\mathbf{BA} = \mathbf{BA}$. The symmetric case is similar.

(b) Use the fact that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T = \mathbf{BA}$, since $\mathbf{A}, \mathbf{B}$ are symmetric.

(28) (a) The (i, i) entry of $\mathbf{AA}^T$ = (ith row of $\mathbf{A}$)$\cdot$(ith column of $\mathbf{A}^T$) = (ith row of $\mathbf{A}$)$\cdot$(ith row of $\mathbf{A}$) = sum of the squares of the entries in the ith row of $\mathbf{A}$. Hence, trace($\mathbf{AA}^T$) is the sum of the squares of the entries from all rows of $\mathbf{A}$.

(c) $\text{Trace}(\mathbf{AB}) = \sum_{i=1}^n(\sum_{k=1}^n a_{ik}b_{ki}) = \sum_{i=1}^n\sum_{k=1}^n b_{ki}a_{ik} = \sum_{k=1}^n\sum_{i=1}^n b_{ki}a_{ik}$.
Reversing the roles of the dummy variables k and i gives $\sum_{i=1}^n\sum_{k=1}^n b_{ik}a_{ki}$, which is equal to trace($\mathbf{BA}$).

(29) (a) Consider any matrix of the form $\begin{bmatrix} 1 & 0 \\ x & 0 \end{bmatrix}$.

(c) $(\mathbf{I}_n - \mathbf{A})^2 = \mathbf{I}_n^2 - \mathbf{I}_n\mathbf{A} - \mathbf{A}\mathbf{I}_n + \mathbf{A}^2 = \mathbf{I}_n - \mathbf{A} - \mathbf{A} + \mathbf{A} = \mathbf{I}_n - \mathbf{A}$.

(d) $\begin{bmatrix} 2 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$

(e) $\mathbf{A}^2 = (\mathbf{A})\mathbf{A} = (\mathbf{AB})\mathbf{A} = \mathbf{A}(\mathbf{BA}) = \mathbf{AB} = \mathbf{A}$.

(30) (a) (ith row of $\mathbf{A}$) $\cdot$ (jth column of $\mathbf{B}$) = (i, j)th entry of $\mathbf{O}_{mp} = 0$.

(b) Consider $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

(c) Let $\mathbf{C} = \begin{bmatrix} 2 & -4 \\ 0 & 2 \\ 2 & 0 \end{bmatrix}$.

(31) A 2×2 matrix $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ that commutes with every other 2×2 matrix must have the form $\mathbf{A} = c\mathbf{I}_2$. For, if $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $\mathbf{AB} = \begin{bmatrix} 0 & a_{11} \\ 0 & a_{21} \end{bmatrix}$, which must equal $\mathbf{BA} = \begin{bmatrix} a_{21} & a_{22} \\ 0 & 0 \end{bmatrix}$. Hence,

$a_{21} = 0$ and $a_{11} = a_{22}$. Let $c = a_{11} = a_{22}$. Then $\mathbf{A} = \begin{bmatrix} c & a_{12} \\ 0 & c \end{bmatrix}$. Also, if $\mathbf{D} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, then $\mathbf{AD} = \begin{bmatrix} a_{12} & 0 \\ c & 0 \end{bmatrix}$ must equal $\mathbf{DA} = \begin{bmatrix} 0 & 0 \\ c & a_{12} \end{bmatrix}$, which gives $a_{12} = 0$, and so $\mathbf{A} = c\mathbf{I}_2$.

Finally, note that $c\mathbf{I}_2$ actually does commute with every 2×2 matrix $\mathbf{M}$, since $(c\mathbf{I}_2)\mathbf{M} = c(\mathbf{I}_2\mathbf{M}) = c\mathbf{M} = c(\mathbf{MI}_2) = \mathbf{M}(c\mathbf{I}_2)$.

(32) (a) T (b) T (c) T (d) F (e) F (f) F (g) F

Chapter 1 Review Exercises

(1) Yes. Vectors corresponding to adjacent sides are orthogonal. Vectors corresponding to opposite sides are parallel, with one pair having slope $\frac{3}{5}$and the other pair having slope $-\frac{5}{3}$.

(2) $\mathbf{u} = \left[\frac{5}{\sqrt{394}}, -\frac{12}{\sqrt{394}}, \frac{15}{\sqrt{394}}\right] \approx [0.2481, -0.5955, 0.7444]$; slightly longer.

(3) Net velocity $= \left[4\sqrt{2} - 5, -4\sqrt{2}\right] \approx [0.6569, -5.6569]$; speed ≈ 5.6947 mi/hr.

(4) $\mathbf{a} = [-10, 9, 10]$ m/sec^2

(5) $|\mathbf{x} \cdot \mathbf{y}| = 74 \leq \|\mathbf{x}\| \ \ \|\mathbf{y}\| \approx 90.9$

(6) $\theta \approx 136°$

(7) $\mathbf{proj_a b} = \left[\frac{114}{25}, -\frac{38}{25}, \frac{19}{25}, \frac{57}{25}\right] = [4.56, -1.52, 0.76, 2.28]$;
$\mathbf{b} - \mathbf{proj_a b} = \left[-\frac{14}{25}, -\frac{62}{25}, \frac{56}{25}, -\frac{32}{25}\right] = [-0.56, -2.48, 2.24, -1.28]$;
$\mathbf{a} \cdot (\mathbf{b} - \mathbf{proj_a b}) = 0$.

(8) -1782 joules

(9) We must prove that $(\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = 0 \Longrightarrow \|\mathbf{x}\| = \|\mathbf{y}\|$. But,
$(\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) = 0 \Longrightarrow \mathbf{x} \cdot \mathbf{x} - \mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{x} - \mathbf{y} \cdot \mathbf{y} = 0$
$\Longrightarrow \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 = 0 \Longrightarrow \|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 \Longrightarrow \|\mathbf{x}\| = \|\mathbf{y}\|$.

(10) First, $\mathbf{x} \neq \mathbf{0}$, or else $\mathbf{proj_x y}$ is not defined. Also, $\mathbf{y} \neq \mathbf{0}$, since that would imply $\mathbf{proj_x y} = \mathbf{y}$. Now, assume $\mathbf{x} \parallel \mathbf{y}$. Then, there is a scalar $c \neq 0$ such that $\mathbf{y} = c\mathbf{x}$. Hence,

$$\mathbf{proj_x y} = \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x} = \left(\frac{\mathbf{x} \cdot c\mathbf{x}}{\|\mathbf{x}\|^2}\right)\mathbf{x} = \left(\frac{c\,\|\mathbf{x}\|^2}{\|\mathbf{x}\|^2}\right)\mathbf{x} = c\mathbf{x} = \mathbf{y},$$

contradicting the assumption that $\mathbf{y} \neq \mathbf{proj_x y}$.

(11) (a) $3\mathbf{A} - 4\mathbf{C}^T = \begin{bmatrix} 3 & 2 & 13 \\ -11 & -19 & 0 \end{bmatrix}$; $\mathbf{AB} = \begin{bmatrix} 15 & -21 & -4 \\ 22 & -30 & 11 \end{bmatrix}$; $\mathbf{BA}$ is not defined;

$\mathbf{AC} = \begin{bmatrix} 23 & 14 \\ -5 & 23 \end{bmatrix}$; $\mathbf{CA} = \begin{bmatrix} 30 & -11 & 17 \\ 2 & 0 & 18 \\ -11 & 5 & 16 \end{bmatrix}$; $\mathbf{A}^3$ is not defined;

$\mathbf{B}^3 = \begin{bmatrix} 97 & -128 & 24 \\ -284 & 375 & -92 \\ 268 & -354 & 93 \end{bmatrix}$

(b) Third row of $\mathbf{BC} = [5, 8]$.

(12) $\mathbf{S} = \begin{bmatrix} 4 & -\frac{1}{2} & \frac{11}{2} \\ -\frac{1}{2} & 7 & -1 \\ \frac{11}{2} & -1 & -2 \end{bmatrix}$; $\mathbf{V} = \begin{bmatrix} 0 & -\frac{5}{2} & -\frac{1}{2} \\ \frac{5}{2} & 0 & -2 \\ \frac{1}{2} & 2 & 0 \end{bmatrix}$

(13) Now, $(3(\mathbf{A}-\mathbf{B})^T)^T = 3\left((\mathbf{A}-\mathbf{B})^T\right)^T$ (by part (3) of Theorem 1.13) $= 3(\mathbf{A}-\mathbf{B})$ (by part (1) of Theorem 1.13). Also, $-(3(\mathbf{A}-\mathbf{B})^T) = 3(-1)(\mathbf{A}^T - \mathbf{B}^T)$ (parts (2) and (3) of Theorem 1.13) $=$ $3(-1)((-\mathbf{A})-(-\mathbf{B}))$ (definition of skew-symmetric) $= 3(\mathbf{A}-\mathbf{B})$. Hence, $(3(\mathbf{A}-\mathbf{B})^T)^T = -(3(\mathbf{A}-\mathbf{B})^T)$, and so $3(\mathbf{A}-\mathbf{B})^T$ is skew-symmetric.

(14) Let $\mathbf{C} = \mathbf{A}+\mathbf{B}$. Now, $c_{ij} = a_{ij} + b_{ij}$. But for $i < j$, $a_{ij} = b_{ij} = 0$. Hence, for $i < j$, $c_{ij} = 0$. Thus, $\mathbf{C}$ is lower triangular.

(15)

	Price	Shipping Cost
Company I	\$168500	\$24200
Company II	\$202500	\$29100
Company III	\$155000	\$22200

(16) Take the transpose of both sides of $\mathbf{A}^T\mathbf{B}^T = \mathbf{B}^T\mathbf{A}^T$ to get $\mathbf{BA} = \mathbf{AB}$.
Then, $(\mathbf{AB})^2 = (\mathbf{AB})(\mathbf{AB}) = \mathbf{A}(\mathbf{BA})\mathbf{B} = \mathbf{A}(\mathbf{AB})\mathbf{B} = \mathbf{A}^2\mathbf{B}^2$.

(17) Negation: For every square matrix $\mathbf{A}$, $\mathbf{A}^2 = \mathbf{A}$. Counterexample: $\mathbf{A} = [-1]$.

(18) If $\mathbf{A} \neq \mathbf{O}_{22}$, then some row of $\mathbf{A}$, say the ith row, is nonzero. Apply Result 5 in Section 1.3 with $\mathbf{x} = (i\text{th row of } \mathbf{A})$.

(19) Base Step ($k = 2$): Suppose $\mathbf{A}$ and $\mathbf{B}$ are upper triangular $n \times n$ matrices, and let $\mathbf{C} = \mathbf{AB}$. Then $a_{ij} = b_{ij} = 0$, for $i > j$. Hence, for $i > j$,

$$c_{ij} = \sum_{m=1}^{n} a_{im}b_{mj} = \sum_{m=1}^{i-1} 0 \cdot b_{mj} + a_{ii}b_{ij} + \sum_{m=i+1}^{n} a_{im} \cdot 0 = a_{ii}(0) = 0.$$

Thus, $\mathbf{C}$ is upper triangular.
Inductive Step: Let $\mathbf{A}_1, \ldots, \mathbf{A}_{k+1}$ be upper triangular matrices. Then, the product $\mathbf{C} = \mathbf{A}_1 \cdots \mathbf{A}_k$ is upper triangular by the Inductive Hypothesis, and so the product $\mathbf{A}_1 \cdots \mathbf{A}_{k+1} = \mathbf{C}\mathbf{A}_{k+1}$ is upper triangular by the Base Step.

(20) (a) F (c) F (e) F (g) F (i) F (k) T (m) T (o) F (q) F
(b) T (d) F (f) T (h) F (j) T (l) T (n) F (p) F (r) T

Chapter 2

Section 2.1

(1) (a) Consistent; $\{(-2, 3, 6)\}$

(b) Consistent; $\{(5, -4, 2)\}$

(c) Inconsistent; $\{\}$

(d) Consistent; $\{(c + 7, 2c - 3, c, 6) \mid c \in \mathbb{R}\}$; $(7, -3, 0, 6)$, $(8, -1, 1, 6)$, $(9, 1, 2, 6)$

(e) Consistent; $\{(2b - d - 4, b, 2d + 5, d, 2) \mid b, d \in \mathbb{R}\}$; $(-4, 0, 5, 0, 2)$, $(-2, 1, 5, 0, 2)$, $(-5, 0, 7, 1, 2)$

(f) Consistent; $\{(b + 3c - 2, b, c, 8) \mid b, c \in \mathbb{R}\}$; $(-2, 0, 0, 8)$, $(-1, 1, 0, 8)$, $(1, 0, 1, 8)$

(g) Consistent; $\{(6, -1, 3)\}$

(h) Inconsistent; $\{\}$

(2) (a) $\{(3c + 13e + 46, c + e + 13, c, -2e + 5, e) \mid c, e \in \mathbb{R}\}$

(b) $\{(3b - 12d + 50e - 13, b, 2d - 8e + 3, d, e, -2) \mid b, d, e \in \mathbb{R}\}$

(c) $\{(-20c + 9d - 153f - 68, 7c - 2d + 37f + 15, c, d, 4f + 2, f) \mid c, d, f \in \mathbb{R}\}$

(d) $\{\}$

(3) 52 nickels, 64 dimes, 32 quarters

(4) $y = 2x^2 - x + 5$

(5) $y = -3x^3 + 4x^2 - 5x + 6$

(6) $x^2 + y^2 - 6x - 8y = 0$, or $(x - 3)^2 + (y - 4)^2 = 25$

(7) In each part, $R(\mathbf{AB}) = (R(\mathbf{A}))\mathbf{B}$, which equals the given matrix.

(a) $\begin{bmatrix} 26 & 15 & -6 \\ 6 & 4 & 1 \\ 0 & -6 & 12 \\ 10 & 4 & -14 \end{bmatrix}$ (b) $\begin{bmatrix} 26 & 15 & -6 \\ 10 & 4 & -14 \\ 18 & 6 & 15 \\ 6 & 4 & 1 \end{bmatrix}$

(8) (a) To save space, we write $\langle \mathbf{C} \rangle_i$ for the ith row of a matrix $\mathbf{C}$.

For the Type (I) operation $R : \langle i \rangle \leftarrow c\langle i \rangle$: Now, $\langle R(\mathbf{AB}) \rangle_i = c\langle \mathbf{AB} \rangle_i = c\langle \mathbf{A} \rangle_i \mathbf{B}$ (by the hint in the text) $= \langle R(\mathbf{A}) \rangle_i \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_i$. But, if $k \neq i$, $\langle R(\mathbf{AB}) \rangle_k = \langle \mathbf{AB} \rangle_k = \langle \mathbf{A} \rangle_k \mathbf{B}$ (by the hint in the text) $= \langle R(\mathbf{A}) \rangle_k \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_k$.

For the Type (II) operation $R : \langle i \rangle \leftarrow c\langle j \rangle + \langle i \rangle$: Now, $\langle R(\mathbf{AB}) \rangle_i = c\langle \mathbf{AB} \rangle_j + \langle \mathbf{AB} \rangle_i = c\langle \mathbf{A} \rangle_j \mathbf{B} + \langle \mathbf{A} \rangle_i \mathbf{B}$ (by the hint in the text) $= (c\langle \mathbf{A} \rangle_j + \langle \mathbf{A} \rangle_i)\mathbf{B} = \langle R(\mathbf{A}) \rangle_i \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_i$. But, if $k \neq i$, $\langle R(\mathbf{AB}) \rangle_k = \langle \mathbf{AB} \rangle_k = \langle \mathbf{A} \rangle_k \mathbf{B}$ (by the hint in the text) $= \langle R(\mathbf{A}) \rangle_k \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_k$.

For the Type (III) operation $R : \langle i \rangle \longleftrightarrow \langle j \rangle$: Now, $\langle R(\mathbf{AB}) \rangle_i = \langle \mathbf{AB} \rangle_j = \langle \mathbf{A} \rangle_j \mathbf{B}$ (by the hint in the text) $= \langle R(\mathbf{A}) \rangle_i \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_i$. Similarly, $\langle R(\mathbf{AB}) \rangle_j = \langle \mathbf{AB} \rangle_i = \langle \mathbf{A} \rangle_i \mathbf{B}$ (by the hint in the text) $= \langle R(\mathbf{A}) \rangle_j \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_j$. And, if $k \neq i$ and $k \neq j$, $\langle R(\mathbf{AB}) \rangle_k = \langle \mathbf{AB} \rangle_k = \langle \mathbf{A} \rangle_k \mathbf{B}$ (by the hint in the text) $= \langle R(\mathbf{A}) \rangle_k \mathbf{B} = \langle R(\mathbf{A})\mathbf{B} \rangle_k$.

(b) Use induction on n, the number of row operations used.
Base Step: The case $n = 1$ is part (1) of Theorem 2.1, and is proven in part (a) of this exercise.
Inductive Step: Assume that

$$R_n(\cdots(R_2(R_1(\mathbf{AB})))\cdots) = R_n(\cdots(R_2(R_1(\mathbf{A})))\cdots)\mathbf{B}$$

and prove that

$$R_{n+1}(R_n(\cdots(R_2(R_1(\mathbf{AB})))\cdots)) = R_{n+1}(R_n(\cdots(R_2(R_1(\mathbf{A})))\cdots))\mathbf{B}.$$

Now,

$$R_{n+1}(R_n(\cdots(R_2(R_1(\mathbf{AB})))\cdots)) = R_{n+1}(R_n(\cdots(R_2(R_1(\mathbf{A})))\cdots)\mathbf{B})$$

(by the inductive hypothesis) $= R_{n+1}(R_n(\cdots(R_2(R_1(\mathbf{A})))\cdots))\mathbf{B}$ (by part (a)).

(9) Multiplying a row by zero changes all of its entries to zero, essentially erasing all of the information in the row.

(10) Suppose that $\mathbf{A}$, $\mathbf{B}$, $\mathbf{X}_1$, and $\mathbf{X}_2$ are as given in the problem.

(a) Let c be a scalar. Then,

$$\mathbf{A}(\mathbf{X}_1 + c(\mathbf{X}_2 - \mathbf{X}_1)) = \mathbf{AX}_1 + c\mathbf{A}(\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{B} + c\mathbf{AX}_2 - c\mathbf{AX}_1 = \mathbf{B} + c\mathbf{B} - c\mathbf{B} = \mathbf{B}.$$

Hence, $\mathbf{X}_1 + c(\mathbf{X}_2 - \mathbf{X}_1)$ is a solution to $\mathbf{AX} = \mathbf{B}$.

(b) Suppose $\mathbf{X}_1 + c(\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{X}_1 + d(\mathbf{X}_2 - \mathbf{X}_1)$. Then,

$$\begin{aligned} & \mathbf{X}_1 + c(\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{X}_1 + d(\mathbf{X}_2 - \mathbf{X}_1) \\ \Longrightarrow\ & c(\mathbf{X}_2 - \mathbf{X}_1) = d(\mathbf{X}_2 - \mathbf{X}_1) \\ \Longrightarrow\ & (c-d)(\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{0} \\ \Longrightarrow\ & (c-d) = 0 \text{ or } (\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{0}. \end{aligned}$$

However, $\mathbf{X}_2 \neq \mathbf{X}_1$, and so $c = d$.

(c) Parts (a) and (b) show that, for each different real number c, $\mathbf{X}_1 + c(\mathbf{X}_2 - \mathbf{X}_1)$ is a different solution to $\mathbf{AX} = \mathbf{B}$. Therefore, since there are an infinite number of real numbers, $\mathbf{AX} = \mathbf{B}$ has an infinite number of solutions.

(11) (a) T (b) F (c) F (d) F (e) T (f) T (g) F

Section 2.2

(1) Matrices in (a), (b), (c), (d), and (f) are not in reduced row echelon form.
Matrix in (a) fails condition 2 of the definition.
Matrix in (b) fails condition 4 of the definition.
Matrix in (c) fails condition 1 of the definition.
Matrix in (d) fails conditions 1, 2, and 3 of the definition.
Matrix in (f) fails condition 3 of the definition.

(2) (a) $\left[\begin{array}{ccc|c} 1 & 4 & 0 & -13 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(b) $\mathbf{I}_4$

(c) $\left[\begin{array}{ccccc} 1 & -2 & 0 & 11 & -23 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(d) $\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

(e) $\left[\begin{array}{ccccc|c} 1 & -2 & 0 & 2 & -1 & 1 \\ 0 & 0 & 1 & -1 & 3 & 2 \end{array}\right]$

(f) $\mathbf{I}_5$

(3) (a) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 6 \end{array}\right]$; $(-2, 3, 6)$

(e) $\left[\begin{array}{ccccc|c} 1 & -2 & 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & -2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]$; $\{(2b-d-4, b, 2d+5, d, 2) \,|\, b, d \in \mathbb{R}\}$

(g) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$; $\{(6, -1, 3)\}$

(4) (a) Solution set $= \{(c-2d, -3d, c, d) \,|\, c, d \in \mathbb{R}\}$; one particular solution $= (-3, -6, 1, 2)$

(b) Solution set $= \{(e, -2d-e, -3d, d, e) \,|\, d, e \in \mathbb{R}\}$; one particular solution $= (1, 1, 3, -1, 1)$

(c) Solution set $= \{(-4b+2d-f, b, -3d+2f, d, -2f, f) \,|\, b, d, f \in \mathbb{R}\}$; one particular solution $= (-3, 1, 0, 2, -6, 3)$

(5) (a) $\{(2c, -4c, c) \,|\, c \in \mathbb{R}\} = \{c(2, -4, 1) \,|\, c \in \mathbb{R}\}$

(b) $\{(0, 0, 0)\}$

(c) $\{(0, 0, 0, 0)\}$

(d) $\{(3d, -2d, d, d) \,|\, d \in \mathbb{R}\}$
$= \{d(3, -2, 1, 1) \,|\, d \in \mathbb{R}\}$

(6) (a) $a = 2,\ b = 15,\ c = 12,\ d = 6$

(b) $a = 2,\ b = 25,\ c = 16,\ d = 18$

(c) $a = 4,\ b = 2,\ c = 4,\ d = 1,\ e = 4$

(d) $a = 3,\ b = 11,\ c = 2,\ d = 3,\ e = 2,\ f = 6$

(7) (a) $A = 3,\ B = 4,\ C = -2$

(b) $A = 4,\ B = -3,\ C = 1,\ D = 0$

(8) Solution for system $\mathbf{AX} = \mathbf{B}_1$: $(6, -51, 21)$; solution for system $\mathbf{AX} = \mathbf{B}_2$: $(\frac{35}{3}, -98, \frac{79}{2})$.

(9) Solution for system $\mathbf{AX} = \mathbf{B}_1$: $(-1, -6, 14, 9)$; solution for system $\mathbf{AX} = \mathbf{B}_2$: $(-10, 56, 8, \frac{10}{3})$.

(10) (a) R_1: (III): $\langle 1 \rangle \leftrightarrow \langle 2 \rangle$
R_2: (I): $\langle 1 \rangle \leftarrow \frac{1}{2} \langle 1 \rangle$
R_3: (II): $\langle 2 \rangle \leftarrow \frac{1}{3} \langle 2 \rangle$
R_4: (II): $\langle 1 \rangle \leftarrow -2 \langle 2 \rangle + \langle 1 \rangle$;

(b) $\mathbf{AB} = \begin{bmatrix} 57 & -21 \\ 56 & -20 \end{bmatrix}$; $R_4(R_3(R_2(R_1(\mathbf{AB})))) = R_4(R_3(R_2(R_1(\mathbf{A}))))\mathbf{B} = \begin{bmatrix} -10 & 4 \\ 19 & -7 \end{bmatrix}$.

(11) (a) $\mathbf{A}(\mathbf{X}_1 + \mathbf{X}_2) = \mathbf{AX}_1 + \mathbf{AX}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$; $\mathbf{A}(c\mathbf{X}_1) = c(\mathbf{AX}_1) = c\mathbf{0} = \mathbf{0}$.

(b) Any nonhomogeneous system with two equations and two unknowns that has a unique solution will serve as a counterexample. For instance, consider

$$\begin{cases} x + y = 1 \\ x - y = 1 \end{cases}.$$

This system has a unique solution: $(1, 0)$. Let (s_1, s_2) and (t_1, t_2) both equal $(1, 0)$. Then the sum of solutions is not a solution in this case. Also, if $c \neq 1$, then the scalar multiple of a solution by c is not a solution.

(c) $\mathbf{A}(\mathbf{X}_1 + \mathbf{X}_2) = \mathbf{A}\mathbf{X}_1 + \mathbf{A}\mathbf{X}_2 = \mathbf{B} + \mathbf{0} = \mathbf{B}$.

(d) Let $\mathbf{X}_1$ be the unique solution to $\mathbf{AX} = \mathbf{B}$. Suppose $\mathbf{AX} = \mathbf{0}$ has a nontrivial solution $\mathbf{X}_2$. Then, by (c), $\mathbf{X}_1 + \mathbf{X}_2$ is a solution to $\mathbf{AX} = \mathbf{B}$, and $\mathbf{X}_1 \neq \mathbf{X}_1 + \mathbf{X}_2$ since $\mathbf{X}_2 \neq \mathbf{0}$. This contradicts the uniqueness of $\mathbf{X}_1$.

(12) If $a \neq 0$, then pivoting in the first column of $\left[\begin{array}{cc|c} a & b & 0 \\ c & d & 0 \end{array}\right]$ yields $\left[\begin{array}{cc|c} 1 & \frac{b}{a} & 0 \\ 0 & d - c\left(\frac{b}{a}\right) & 0 \end{array}\right]$. By Theorem 2.2, there is a nontrivial solution if and only if the $(2, 2)$ entry of this matrix is zero, which occurs if and only if $ad - bc = 0$.

If $a = 0$ and $c \neq 0$, then $\left[\begin{array}{cc|c} a & b & 0 \\ c & d & 0 \end{array}\right] = \left[\begin{array}{cc|c} 0 & b & 0 \\ c & d & 0 \end{array}\right]$. Swapping the first and second rows and then pivoting in the first column yields $\left[\begin{array}{cc|c} 1 & \frac{d}{c} & 0 \\ 0 & b & 0 \end{array}\right]$. By Theorem 2.2, there is a nontrivial solution if and only if the $(2, 2)$ entry of this matrix (that is, b) equals zero. But b is zero if and only if $ad - bc = 0$, since $a = 0$ and $c \neq 0$.

Finally, if both a and c equal 0, then $ad - bc = 0$ and $(1, 0)$ is a nontrivial solution.

(13) (a) The contrapositive is: If $\mathbf{AX} = \mathbf{0}$ has only the trivial solution, then $\mathbf{A}^2\mathbf{X} = \mathbf{0}$ has only the trivial solution.

Let $\mathbf{X}_1$ be a solution to $\mathbf{A}^2\mathbf{X} = \mathbf{0}$. Then $\mathbf{0} = \mathbf{A}^2\mathbf{X}_1 = \mathbf{A}(\mathbf{A}\mathbf{X}_1)$. Thus $\mathbf{A}\mathbf{X}_1$ is a solution to $\mathbf{AX} = \mathbf{0}$. Hence $\mathbf{A}\mathbf{X}_1 = \mathbf{0}$ by the premise. Thus, $\mathbf{X}_1 = \mathbf{0}$, using the premise again.

(b) The contrapositive is: Let t be a positive integer. If $\mathbf{AX} = \mathbf{0}$ has only the trivial solution, then $\mathbf{A}^t\mathbf{X} = \mathbf{0}$ has only the trivial solution.

Proceed by induction. The statement is clearly true when $t = 1$, completing the Base Step.

Inductive Step: Assume that if $\mathbf{AX} = \mathbf{0}$ has only the trivial solution, then $\mathbf{A}^t\mathbf{X} = \mathbf{0}$ has only the trivial solution. We must prove that if $\mathbf{AX} = \mathbf{0}$ has only the trivial solution, then $\mathbf{A}^{t+1}\mathbf{X} = \mathbf{0}$ has only the trivial solution.

Let $\mathbf{X}_1$ be a solution to $\mathbf{A}^{t+1}\mathbf{X} = \mathbf{0}$. Then $\mathbf{A}(\mathbf{A}^t\mathbf{X}_1) = \mathbf{0}$. Thus $\mathbf{A}^t\mathbf{X}_1 = \mathbf{0}$, since $\mathbf{A}^t\mathbf{X}_1$ is a solution to $\mathbf{AX} = \mathbf{0}$. But then $\mathbf{X}_1 = \mathbf{0}$ by the inductive hypothesis.

(14) (a) T (b) T (c) F (d) T (e) F (f) F

Section 2.3

(1) (a) A row operation of Type (I) converts $\mathbf{A}$ to $\mathbf{B}$: $\langle 2 \rangle \leftarrow -5\,\langle 2 \rangle$.

(b) A row operation of Type (III) converts $\mathbf{A}$ to $\mathbf{B}$: $\langle 1 \rangle \leftrightarrow \langle 3 \rangle$.

(c) A row operation of Type (II) converts $\mathbf{A}$ to $\mathbf{B}$: $\langle 2 \rangle \leftarrow \langle 3 \rangle + \langle 2 \rangle$.

(2) (a) $\mathbf{B} = \mathbf{I}_3$. The sequence of row operations converting $\mathbf{A}$ to $\mathbf{B}$ is:

(I): $\langle 1 \rangle \leftarrow \frac{1}{4} \langle 1 \rangle$
(II): $\langle 2 \rangle \leftarrow 2 \langle 1 \rangle + \langle 2 \rangle$
(II): $\langle 3 \rangle \leftarrow -3 \langle 1 \rangle + \langle 3 \rangle$
(III): $\langle 2 \rangle \leftrightarrow \langle 3 \rangle$
(II): $\langle 1 \rangle \leftarrow 5 \langle 3 \rangle + \langle 1 \rangle$

(b) The sequence of row operations converting $\mathbf{B}$ to $\mathbf{A}$ is:

(II): $\langle 1 \rangle \leftarrow -5 \langle 3 \rangle + \langle 1 \rangle$
(III): $\langle 2 \rangle \leftrightarrow \langle 3 \rangle$
(II): $\langle 3 \rangle \leftarrow 3 \langle 1 \rangle + \langle 3 \rangle$
(II): $\langle 2 \rangle \leftarrow -2 \langle 1 \rangle + \langle 2 \rangle$
(I): $\langle 1 \rangle \leftarrow 4 \langle 1 \rangle$

(3) (a) Let $\mathbf{C}$ be the common reduced row echelon form matrix for $\mathbf{A}$ and $\mathbf{B}$. Then, $\mathbf{A}$ and $\mathbf{B}$ are both row equivalent to $\mathbf{C}$. Also, by part (1) of Theorem 2.4, $\mathbf{C}$ is row equivalent to $\mathbf{B}$. But, since $\mathbf{A}$ is row equivalent to $\mathbf{C}$, and $\mathbf{C}$ is row equivalent to $\mathbf{B}$, part (2) of Theorem 2.4 asserts that $\mathbf{A}$ is row equivalent to $\mathbf{B}$.

(b) The common reduced row echelon form is $\mathbf{I}_3$.

(c) The sequence of row operations is:

(II): $\langle 3 \rangle \leftarrow 2 \langle 2 \rangle + \langle 3 \rangle$
(I): $\langle 3 \rangle \leftarrow -1 \langle 3 \rangle$
(II): $\langle 1 \rangle \leftarrow -9 \langle 3 \rangle + \langle 1 \rangle$
(II): $\langle 2 \rangle \leftarrow 3 \langle 3 \rangle + \langle 2 \rangle$
(II): $\langle 3 \rangle \leftarrow -\frac{9}{5} \langle 2 \rangle + \langle 3 \rangle$
(II): $\langle 1 \rangle \leftarrow -\frac{3}{5} \langle 2 \rangle + \langle 1 \rangle$
(I): $\langle 2 \rangle \leftarrow -\frac{1}{5} \langle 2 \rangle$
(II): $\langle 3 \rangle \leftarrow -3 \langle 1 \rangle + \langle 3 \rangle$
(II): $\langle 2 \rangle \leftarrow -2 \langle 1 \rangle + \langle 2 \rangle$
(I): $\langle 1 \rangle \leftarrow -5 \langle 1 \rangle$

(4) (a) Assume $\mathbf{B}$ is row equivalent to $\mathbf{A}$. Let $\mathbf{C}$ be the reduced row echelon form matrix for $\mathbf{A}$. Then $\mathbf{A}$ is row equivalent to $\mathbf{C}$. But, since $\mathbf{B}$ is row equivalent to $\mathbf{A}$, and $\mathbf{A}$ is row equivalent to $\mathbf{C}$, part (2) of Theorem 2.4 asserts that $\mathbf{A}$ is row equivalent to $\mathbf{C}$. Now, by Theorem 2.6, the reduced row echelon form matrix for $\mathbf{B}$ is unique, so it must be $\mathbf{C}$. Thus $\mathbf{A}$ and $\mathbf{C}$ have the same reduced row echelon form matrix.

(b) The reduced row echelon form matrices for $\mathbf{A}$ and $\mathbf{B}$ are, respectively,

$$\begin{bmatrix} 1 & 0 & 0 & -2 & 3 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(5) (a) 2 (b) 1 (c) 2 (d) 3 (e) 3 (f) 2

(6) (a) Rank = 3. Theorem 2.7 predicts that there is only the trivial solution. Solution set = $\{(0,0,0)\}$

(b) Rank = 2. Theorem 2.7 predicts that nontrivial solutions exist.
Solution set = $\{(3c, -4c, c) \mid c \in \mathbb{R}\}$

(7) In the following answers, the asterisk represents any real entry:

(a) Smallest rank = 1 $\left[\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ Largest rank = 4 $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$

(b) Smallest rank = 1 $\left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$ Largest rank = 3 $\left[\begin{array}{cccc|c} 1 & 0 & 0 & * & * \\ 0 & 1 & 0 & * & * \\ 0 & 0 & 1 & * & * \end{array}\right]$

(c) Smallest rank = 2 $\left[\begin{array}{cccc|c} 1 & * & * & * & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$ Largest rank = 3 $\left[\begin{array}{cccc|c} 1 & 0 & * & * & 0 \\ 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$

(d) Smallest rank = 1 $\left[\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ Largest rank = 3 $\left[\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(8) (a) $\mathbf{x} = -\frac{21}{11}\mathbf{a}_1 + \frac{6}{11}\mathbf{a}_2$

(b) $\mathbf{x} = 3\mathbf{a}_1 - 4\mathbf{a}_2 + \mathbf{a}_3$

(c) Not possible

(d) $\mathbf{x} = \frac{1}{2}\mathbf{a}_1 - \frac{1}{2}\mathbf{a}_2 + \frac{1}{2}\mathbf{a}_3$

(e) The answer is not unique; one possible answer is $\mathbf{x} = -3\mathbf{a}_1 + 2\mathbf{a}_2 + 0\mathbf{a}_3$.

(f) Not possible (g) $\mathbf{x} = 2\mathbf{a}_1 - \mathbf{a}_2 - \mathbf{a}_3$ (h) Not possible

(9) (a) Yes: 5(row 1) − 3(row 2) − 1(row 3)

(b) Not in row space

(c) Not in row space

(d) Yes: −3(row 1) + 1(row 2)

(e) Yes, but the linear combination of the rows is not unique; one possible expression for the given vector is
−3(row 1) + 1(row 2) + 0(row 3).

(10) (a) $[13, -23, 60] = -2\mathbf{q}_1 + \mathbf{q}_2 + 3\mathbf{q}_3$

(b) $\mathbf{q}_1 = 3\mathbf{r}_1 - \mathbf{r}_2 - 2\mathbf{r}_3$; $\mathbf{q}_2 = 2\mathbf{r}_1 + 2\mathbf{r}_2 - 5\mathbf{r}_3$; $\mathbf{q}_3 = \mathbf{r}_1 - 6\mathbf{r}_2 + 4\mathbf{r}_3$

(c) $[13, -23, 60] = -\mathbf{r}_1 - 14\mathbf{r}_2 + 11\mathbf{r}_3$

(11) (a) (i) $\mathbf{B} = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$;

(ii) $[1, 0, -1, 2] = -\frac{7}{8}[0, 4, 12, 8] + \frac{1}{2}[2, 7, 19, 18] + 0[1, 2, 5, 6]$;
$[0, 1, 3, 2] = \frac{1}{4}[0, 4, 12, 8] + 0[2, 7, 19, 18] + 0[1, 2, 5, 6]$

(iii) $[0, 4, 12, 8] = 0[1, 0, -1, 2] + 4[0, 1, 3, 2]$;
$[2, 17, 19, 18] = 2[1, 0, -1, 2] + 7[0, 1, 3, 2]$;
$[1, 2, 5, 6] = 1[1, 0, -1, 2] + 2[0, 1, 3, 2]$

(b) (i) $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;

(ii) $[1, 2, 3, 0, -1] = -\frac{5}{3}[1, 2, 3, -4, -21] - \frac{4}{3}[-2, -4, -6, 5, 27] + 0[13, 26, 39, 5, 12] + 0[2, 4, 6, -1, -7]$;
$[0, 0, 0, 1, 5] = -\frac{2}{3}[1, 2, 3, -4, -21] - \frac{1}{3}[-2, -4, -6, 5, 27] + 0[13, 26, 39, 5, 12] + 0[2, 4, 6, -1, -7]$

(iii) $[1, 2, 3, -4, -21] = 1[1, 2, 3, 0, -1] - 4[0, 0, 0, 1, 5]$;
$[-2, -4, -6, 5, 27] = -2[1, 2, 3, 0, -1] + 5[0, 0, 0, 1, 5]$;
$[13, 26, 39, 5, 12] = 13[1, 2, 3, 0, -1] + 5[0, 0, 0, 1, 5]$;
$[2, 4, 6, -1, -7] = 2[1, 2, 3, 0, -1] - 1[0, 0, 0, 1, 5]$

(12) Suppose that all main diagonal entries of $\mathbf{A}$ are nonzero. Then, for each i, perform the row operation $\langle i \rangle \leftarrow (1/a_{ii})\langle i \rangle$ on the matrix $\mathbf{A}$. This will convert $\mathbf{A}$ into $\mathbf{I}_n$. We prove the converse by contrapositive. Suppose some diagonal entry a_{ii} equals 0. Then the ith column of $\mathbf{A}$ has all zero entries. No step in the row reduction process will alter this column of zeroes, and so the unique reduced row echelon form for the matrix must contain at least one column of zeroes, and so cannot equal $\mathbf{I}_n$.

(13) (a) Suppose we are performing row operations on an $m \times n$ matrix $\mathbf{A}$. Throughout this part, we will write $\langle \mathbf{B} \rangle_i$ for the ith row of a matrix $\mathbf{B}$.

For the Type (I) operation $R : \langle i \rangle \leftarrow c\langle i \rangle$: Now R^{-1} is $\langle i \rangle \leftarrow \frac{1}{c}\langle i \rangle$. Clearly, R and R^{-1} change only the ith row of $\mathbf{A}$. We want to show that $R^{-1}R$ leaves $\langle \mathbf{A} \rangle_i$ unchanged. But $\langle R^{-1}(R(\mathbf{A})) \rangle_i = \frac{1}{c}\langle R(\mathbf{A}) \rangle_i = \frac{1}{c}(c\langle \mathbf{A} \rangle_i) = \langle \mathbf{A} \rangle_i$.

For the Type (II) operation $R : \langle i \rangle \leftarrow c\langle j \rangle + \langle i \rangle$: Now R^{-1} is $\langle i \rangle \leftarrow -c\langle j \rangle + \langle i \rangle$. Again, R and R^{-1} change only the ith row of $\mathbf{A}$, and we need to show that $R^{-1}R$ leaves $\langle \mathbf{A} \rangle_i$ unchanged. But $\langle R^{-1}(R(\mathbf{A})) \rangle_i = -c\langle R(\mathbf{A}) \rangle_j + \langle R(\mathbf{A}) \rangle_i = -c\langle \mathbf{A} \rangle_j + \langle R(\mathbf{A}) \rangle_i = -c\langle \mathbf{A} \rangle_j + c\langle \mathbf{A} \rangle_j + \langle \mathbf{A} \rangle_i = \langle \mathbf{A} \rangle_i$.

For the Type (III) operation $R : \langle i \rangle \leftrightarrow \langle j \rangle$: Now, $R^{-1} = R$. Also, R changes only the ith and jth rows of $\mathbf{A}$, and these get swapped. Obviously, a second application of R swaps them back to where they were, proving that R is indeed its own inverse.

(b) Suppose $\mathbf{C}$ is row equivalent to $\mathbf{D}$, and $\mathbf{D}$ is row equivalent to $\mathbf{E}$. Then by the definition of row equivalence, there is a sequence $R_1, R_2, ..., R_m$ of row operations converting $\mathbf{C}$ to $\mathbf{D}$, and a sequence $S_1, S_2, ..., S_n$ of row operations converting $\mathbf{D}$ to $\mathbf{E}$. But then the combined sequence $R_1, R_2, ..., R_m, S_1, S_2, ..., S_n$ of row operations converts $\mathbf{C}$ to $\mathbf{E}$, and so $\mathbf{C}$ is row equivalent to $\mathbf{E}$.

(c) An approach similar to that used for Type (II) operations in the abridged proof of Theorem 2.5 in the text works just as easily for Type (I) and Type (III) operations. However, here is a different approach: Suppose R is a row operation, and let $\mathbf{X}$ satisfy $\mathbf{AX} = \mathbf{B}$. Multiplying both sides of this matrix equation by the matrix $R(\mathbf{I})$ yields $R(\mathbf{I})\mathbf{AX} = R(\mathbf{I})\mathbf{B}$, implying $R(\mathbf{IA})\mathbf{X} = R(\mathbf{IB})$, by Theorem 2.1. Thus, $R(\mathbf{A})\mathbf{X} = R(\mathbf{B})$, showing that $\mathbf{X}$ is a solution to the new linear system obtained from $\mathbf{AX} = \mathbf{B}$ after the row operation R is performed.

(14) The zero vector is a solution to $\mathbf{AX} = \mathbf{0}$, but it is not a solution for $\mathbf{AX} = \mathbf{B}$.

(15) Consider the systems

$$\begin{cases} x + y = 1 \\ x + y = 0 \end{cases} \quad \text{and} \quad \begin{cases} x - y = 1 \\ x - y = 2 \end{cases}.$$

The reduced row echelon matrices for these inconsistent systems are, respectively,

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right].$$

Thus, the original augmented matrices are not row equivalent, since their reduced row echelon forms are different.

(16) (a) Plug each of the 5 points in turn into the equation for the conic. This will give a homogeneous system of 5 equations in the 6 variables a, b, c, d, e, and f. This system has a nontrivial solution, by Corollary 2.2.

(b) Yes. In this case, there will be even fewer equations, so Corollary 2.3 again applies.

(17) Because $\mathbf{A}$ and $\mathbf{B}$ are row equivalent, $\mathbf{A} = R_n(\cdots(R_2(R_1(\mathbf{B})))\cdots)$ for some row operations $R_1, \ldots, R_n$. Now, if $\mathbf{D}$ is the unique reduced row echelon form matrix to which $\mathbf{A}$ is row equivalent, then for some additional row operations $R_{n+1}, \ldots, R_{n+k}$,

$$\begin{aligned} \mathbf{D} &= R_{n+k}(\cdots(R_{n+2}(R_{n+1}(\mathbf{A})))\cdots) \\ &= R_{n+k}(\cdots(R_{n+2}(R_{n+1}(R_n(\cdots(R_2(R_1(\mathbf{B})))\cdots))))\cdots), \end{aligned}$$

showing that $\mathbf{B}$ also has $\mathbf{D}$ as its reduced echelon form matrix. Therefore, by the definition of rank, $\text{rank}(\mathbf{B})$ = the number of nonzero rows in $\mathbf{D} = \text{rank}(\mathbf{A})$.

(18) (a) $R_k(\cdots(R_1(\mathbf{A}))\cdots)$ and $\mathbf{A}$ are clearly row equivalent. Use Exercise 17.

(b) Since the row reduction process has no effect on rows having all zeroes, at least k of the m rows in the reduced row echelon form of $\mathbf{A}$ are rows of all zeroes. Thus, $\text{rank}(\mathbf{A}) \leq m - k$.

(c) Since $\mathbf{A}$ is in reduced row echelon form and has k rows of zeroes, $\text{rank}(\mathbf{A}) = m - k$. But $\mathbf{AB}$ has at least k rows of zeroes, so $\text{rank}(\mathbf{AB}) \leq m - k$ by part (b).

(d) Let $\mathbf{A} = R_k(\cdots(R_2(R_1(\mathbf{D})))\cdots)$, where $\mathbf{D}$ is in reduced row echelon form. Then $\text{rank}(\mathbf{AB}) = \text{rank}(R_k(\cdots(R_2(R_1(\mathbf{D})))\cdots)\mathbf{B}) = \text{rank}(R_k(\cdots(R_2(R_1(\mathbf{DB})))\cdots))$ (by part (2) of Theorem 2.1) $= \text{rank}(\mathbf{DB})$ (by part (a)) $\leq \text{rank}(\mathbf{D})$ (by part (c)) $= \text{rank}(\mathbf{A})$ (by definition of rank).

(19) As in the abridged proof of Theorem 2.9 in the text, let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ represent the rows of $\mathbf{A}$, and let $\mathbf{b}_1, \ldots, \mathbf{b}_m$ represent the rows of $\mathbf{B}$.

For the Type (I) operation $R : \langle i \rangle \leftarrow c\langle i \rangle$: Now $\mathbf{b}_i = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + c\mathbf{a}_i + 0\mathbf{a}_{i+1} + \cdots + 0\mathbf{a}_m$, and, for $k \neq i$, $\mathbf{b}_k = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_k + 0\mathbf{a}_{k+1} + \cdots + 0\mathbf{a}_m$. Hence, each row of $\mathbf{B}$ is a linear combination of the rows of $\mathbf{A}$, implying it is in the row space of $\mathbf{A}$.

For the Type (II) operation $R : \langle i \rangle \leftarrow c\langle j \rangle + \langle i \rangle$: Now $\mathbf{b}_i = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + c\mathbf{a}_j + 0\mathbf{a}_{j+1} + \cdots + \mathbf{a}_i + 0\mathbf{a}_{i+1} + \ldots + 0\mathbf{a}_m$, where our notation assumes $i > j$. (An analogous argument works for $i < j$.) And, for $k \neq i$, $\mathbf{b}_k = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_k + 0\mathbf{a}_{k+1} + \cdots + 0\mathbf{a}_m$. Hence, each row of $\mathbf{B}$ is a linear combination of the rows of $\mathbf{A}$, implying it is in the row space of $\mathbf{A}$.

For the Type (III) operation $R : \langle i \rangle \leftrightarrow \langle j \rangle$: Now, $\mathbf{b}_i = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_j + 0\mathbf{a}_{j+1} + \cdots + 0\mathbf{a}_m$, $\mathbf{b}_j = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_i + 0\mathbf{a}_{i+1} + \cdots + 0\mathbf{a}_m$, and, for $k \neq i$, $k \neq j$, $\mathbf{b}_k = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_k + 0\mathbf{a}_{k+1} + \cdots + 0\mathbf{a}_m$. Hence, each row of $\mathbf{B}$ is a linear combination of the rows of $\mathbf{A}$, implying it is in the row space of $\mathbf{A}$.

(20) Let k be the number of matrices *between* $\mathbf{A}$ and $\mathbf{B}$ when performing row operations to get from $\mathbf{A}$ to $\mathbf{B}$. Use a proof by induction on k.
Base Step: If $k = 0$, then there are no intermediary matrices, and Exercise 19 shows that the row space of $\mathbf{B}$ is contained in the row space of $\mathbf{A}$.
Inductive Step: Given the chain

$$\mathbf{A} \to \mathbf{D}_1 \to \mathbf{D}_2 \to \cdots \to \mathbf{D}_k \to \mathbf{D}_{k+1} \to \mathbf{B},$$

we must show that the row space of $\mathbf{B}$ is contained in the row space of $\mathbf{A}$. The inductive hypothesis shows that the row space of $\mathbf{D}_{k+1}$ is in the row space of $\mathbf{A}$, since there are only k matrices between $\mathbf{A}$ and $\mathbf{D}_{k+1}$ in the chain. Thus, each row of $\mathbf{D}_{k+1}$ can be expressed as a linear combination of the rows of $\mathbf{A}$. But by Exercise 19, each row of $\mathbf{B}$ can be expressed as a linear combination of the rows of $\mathbf{D}_{k+1}$. Hence, by Lemma 2.8, each row of $\mathbf{B}$ can be expressed as a linear combination of the rows of $\mathbf{A}$, and therefore is in the row space of $\mathbf{A}$. By Lemma 2.8 again, the row space of $\mathbf{B}$ is contained in the row space of $\mathbf{A}$.

(21) Let x_{ij} represent the jth coordinate of $\mathbf{x}_i$. The corresponding homogeneous system in variables $a_1, \ldots, a_{n+1}$ is

$$\begin{cases} a_1x_{11} + a_2x_{21} + \cdots + a_{n+1}x_{n+1,1} = 0 \\ \quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad\quad \vdots \qquad\quad \vdots \\ a_1x_{1n} + a_2x_{2n} + \cdots + a_{n+1}x_{n+1,n} = 0 \end{cases},$$

which has a nontrivial solution for $a_1, \ldots, a_{n+1}$, by Corollary 2.3.

(22) (a) T (b) T (c) F (d) F (e) F (f) T

Section 2.4

(1) The product of each given pair of matrices equals $\mathbf{I}$.

(2) (a) Rank = 2; nonsingular (b) Rank = 2; singular (c) Rank = 3; nonsingular (d) Rank = 4; nonsingular (e) Rank = 3; singular

(3) No inverse exists for (b), (e) and (f).

(a) $\begin{bmatrix} \frac{1}{10} & \frac{1}{15} \\ \frac{3}{10} & -\frac{2}{15} \end{bmatrix}$ (c) $\begin{bmatrix} -\frac{2}{21} & -\frac{5}{84} \\ \frac{1}{7} & -\frac{1}{28} \end{bmatrix}$ (d) $\begin{bmatrix} \frac{3}{11} & \frac{2}{11} \\ \frac{4}{11} & -\frac{1}{11} \end{bmatrix}$

(4) No inverse exists for (b) and (e).

(a) $\begin{bmatrix} 1 & 3 & 2 \\ -1 & 0 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} \frac{3}{2} & 0 & \frac{1}{2} \\ -3 & \frac{1}{2} & -\frac{1}{2} \\ -\frac{8}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix}$

(d) $\begin{bmatrix} 1 & -1 & 1 & 2 \\ -7 & 5 & -10 & -19 \\ -2 & 1 & -2 & -4 \\ 3 & -2 & 4 & 8 \end{bmatrix}$

(f) $\begin{bmatrix} 4 & -1 & 3 & -6 \\ -3 & 1 & -5 & 10 \\ 10 & -2 & 0 & 1 \\ 1 & 0 & -3 & 6 \end{bmatrix}$

(5) (a) $\begin{bmatrix} \frac{1}{a_{11}} & 0 \\ 0 & \frac{1}{a_{22}} \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{a_{11}} & 0 & 0 \\ 0 & \frac{1}{a_{22}} & 0 \\ 0 & 0 & \frac{1}{a_{33}} \end{bmatrix}$ (c) $\begin{bmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{a_{nn}} \end{bmatrix}$

(6) (a) The general inverse is $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. When $\theta = \frac{\pi}{6}$, the inverse is $\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$.

When $\theta = \frac{\pi}{4}$, the inverse is $\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$. When $\theta = \frac{\pi}{2}$, the inverse is $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

(b) The general inverse is $\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$. When $\theta = \frac{\pi}{6}$, the inverse is $\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

When $\theta = \frac{\pi}{4}$, the inverse is $\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$. When $\theta = \frac{\pi}{2}$, the inverse is $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(7) (a) Inverse = $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{7}{3} & \frac{5}{3} \end{bmatrix}$; solution set = $\{(3, -5)\}$

(b) Inverse = $\begin{bmatrix} 1 & 0 & -3 \\ \frac{8}{5} & \frac{2}{5} & -\frac{17}{5} \\ \frac{1}{5} & -\frac{1}{5} & -\frac{4}{5} \end{bmatrix}$; solution set = $\{(-2, 4, -3)\}$

(c) Inverse = $\begin{bmatrix} 1 & -13 & -15 & 5 \\ -3 & 3 & 0 & -7 \\ -1 & 2 & 1 & -3 \\ 0 & -4 & -5 & 1 \end{bmatrix}$; solution set = $\{(5, -8, 2, -1)\}$

(8) (a) Consider $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. (b) Consider $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. (c) $\mathbf{A} = \mathbf{A}^{-1}$ if $\mathbf{A}$ is involutory.

(9) (a) $\mathbf{A} = \mathbf{I}_2$, $\mathbf{B} = -\mathbf{I}_2$, $\mathbf{A} + \mathbf{B} = \mathbf{O}_2$

(b) $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(c) If $\mathbf{A} = \mathbf{B} = \mathbf{I}_2$, then $\mathbf{A} + \mathbf{B} = 2\mathbf{I}_2$, $\mathbf{A}^{-1} = \mathbf{B}^{-1} = \mathbf{I}_2$, $\mathbf{A}^{-1} + \mathbf{B}^{-1} = 2\mathbf{I}_2$, and $(\mathbf{A} + \mathbf{B})^{-1} = \frac{1}{2}\mathbf{I}_2$, so $\mathbf{A}^{-1} + \mathbf{B}^{-1} \neq (\mathbf{A} + \mathbf{B})^{-1}$.

(10) (a) $\mathbf{B}$ must be the zero matrix.

(b) No. $\mathbf{AB} = \mathbf{I}_n$ implies $\mathbf{A}^{-1}$ exists (and equals $\mathbf{B}$). Multiply both sides of $\mathbf{AC} = \mathbf{O}_n$ on the left by $\mathbf{A}^{-1}$.

(11) $\ldots, \mathbf{A}^{-9}, \mathbf{A}^{-5}, \mathbf{A}^{-1}, \mathbf{A}^{3}, \mathbf{A}^{7}, \mathbf{A}^{11}, \ldots$

(12) $\mathbf{B}^{-1}\mathbf{A}$ is the inverse of $\mathbf{A}^{-1}\mathbf{B}$.

(13) $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$ (by Theorem 2.12, part (4)) $= \mathbf{A}^{-1}$.

(14) (a) Suppose first that the matrix contains a column of zeroes. No row operations can alter a column of zeroes, so the unique reduced row echelon form for the matrix must also contain a column of zeroes, and thus cannot equal $\mathbf{I}_n$. Hence the matrix has rank less than n. But by Theorem 2.15, an $n \times n$ matrix is nonsingular if and only if its rank equals n. Thus, the matrix must be singular.

Suppose a matrix contains a row of all zeroes. We assume that it is nonsingular and derive a contradiction. If the matrix is nonsingular, then part (4) of Theorem 2.12 says that its transpose is also nonsingular. But the transpose contains a column of all zeroes. This contradicts the result obtained in the first paragraph.

(b) Such a matrix will contain a column of all zeroes. Use part (a).

(c) When pivoting in the ith column of the matrix during row reduction, the (i, i) entry will be nonzero, allowing a pivot in that position. Then, since all entries in that column below the main diagonal are already zero, none of the rows below the ith row are changed in that column. Hence, none of the entries below the ith row are changed when that row is used as the pivot row. Thus, none of the nonzero entries on the main diagonal are affected by the row reduction steps on previous columns (and the matrix stays in upper triangular form throughout the process). Therefore the matrix row reduces to $\mathbf{I}_n$.

(d) If $\mathbf{A}$ is lower triangular with no zeroes on the main diagonal, then $\mathbf{A}^T$ is upper triangular with no zeroes on the main diagonal. By part (c), $\mathbf{A}^T$ is nonsingular. Hence $\mathbf{A} = (\mathbf{A}^T)^T$ is nonsingular by part (4) of Theorem 2.12.

(e) Note that when row reducing the ith column of $\mathbf{A}$, all of the following occur:

1. The pivot a_{ii} is nonzero, so we use a Type (I) operation to change it to a 1. This changes only row i.
2. No nonzero targets appear below row i, so all Type (II) row operations only change rows above row i.
3. Because of (2), no entries are changed below the main diagonal, and no main diagonal entries are changed by Type (II) operations.

When row reducing $[\mathbf{A}|\mathbf{I}_n]$, we use the exact same row operations we use to reduce $\mathbf{A}$. Since $\mathbf{I}_n$ is also upper triangular, fact (3) above shows that all the zeroes below the main diagonal of $\mathbf{I}_n$ remain zero when the row operations are applied. Thus, the result of the row operations, namely $\mathbf{A}^{-1}$, is upper triangular.

(15) (a) Part (1): Since $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$, we must have $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
Part (2): For $k > 0$, to show $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$, we must show that $\mathbf{A}^k(\mathbf{A}^{-1})^k = \mathbf{I}_n$. We proceed by induction on k.
Base Step: For $k = 1$, clearly $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$.
Inductive Step: Assume $\mathbf{A}^k(\mathbf{A}^{-1})^k = \mathbf{I}_n$. Prove $\mathbf{A}^{k+1}(\mathbf{A}^{-1})^{k+1} = \mathbf{I}_n$.
Now, $\mathbf{A}^{k+1}(\mathbf{A}^{-1})^{k+1} = \mathbf{A}\mathbf{A}^k(\mathbf{A}^{-1})^k\mathbf{A}^{-1} = \mathbf{A}\mathbf{I_n}\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$. This concludes the proof for $k > 0$.
We now show $\mathbf{A}^k(\mathbf{A}^{-1})^k = \mathbf{I}_n$ for $k \leq 0$.
For $k = 0$, clearly $\mathbf{A}^0(\mathbf{A}^{-1})^0 = \mathbf{I}_n\mathbf{I}_n = \mathbf{I}_n$. The case $k = -1$ is covered by part (1) of the theorem.
For $k \leq -2$, $(\mathbf{A}^k)^{-1} = ((\mathbf{A}^{-1})^{-k})^{-1}$ (by definition) $= ((\mathbf{A}^{-k})^{-1})^{-1}$ (by the $k > 0$ case) $= \mathbf{A}^{-k}$ (by part (1)).

(b) To show $(\mathbf{A}_1 \cdots \mathbf{A}_m)^{-1} = \mathbf{A}_m^{-1} \cdots \mathbf{A}_1^{-1}$, we must prove that $(\mathbf{A}_1 \cdots \mathbf{A}_m)(\mathbf{A}_m^{-1} \cdots \mathbf{A}_1^{-1}) = \mathbf{I}_n$. Use induction on n.
Base Step: For $n = 1$, clearly $\mathbf{A}_1\mathbf{A}_1^{-1} = \mathbf{I}_n$.
Inductive Step: Assume that $(\mathbf{A}_1 \cdots \mathbf{A}_m)(\mathbf{A}_m^{-1} \cdots \mathbf{A}_1^{-1}) = \mathbf{I}_n$.
Prove that $(\mathbf{A}_1 \cdots \mathbf{A}_{m+1})(\mathbf{A}_{m+1}^{-1} \cdots \mathbf{A}_1^{-1}) = \mathbf{I}_n$.
Now, $(\mathbf{A}_1 \cdots \mathbf{A}_{m+1})(\mathbf{A}_{m+1}^{-1} \cdots \mathbf{A}_1^{-1}) = (\mathbf{A}_1 \cdots \mathbf{A}_m)\mathbf{A}_{m+1}\mathbf{A}_{m+1}^{-1}(\mathbf{A}_m^{-1} \cdots \mathbf{A}_1^{-1})$
$= (\mathbf{A}_1 \cdots \mathbf{A}_m)\mathbf{I}_n(\mathbf{A}_m^{-1} \cdots \mathbf{A}_1^{-1}) = (\mathbf{A}_1 \cdots \mathbf{A}_m)(\mathbf{A}_m^{-1} \cdots \mathbf{A}_1^{-1}) = \mathbf{I}_n$.

(16) We must prove that $(c\mathbf{A})(\frac{1}{c}\mathbf{A}^{-1}) = \mathbf{I}$. But, $(c\mathbf{A})(\frac{1}{c}\mathbf{A}^{-1}) = c\frac{1}{c}\mathbf{A}\mathbf{A}^{-1} = 1\mathbf{I} = \mathbf{I}$.

(17) (a) Let $p = -s$, $q = -t$. Then $p, q > 0$. Now, $\mathbf{A}^{s+t} = \mathbf{A}^{-(p+q)} = (\mathbf{A}^{-1})^{p+q} = (\mathbf{A}^{-1})^p(\mathbf{A}^{-1})^q$ (by Theorem 1.17) $= \mathbf{A}^{-p}\mathbf{A}^{-q} = \mathbf{A}^s\mathbf{A}^t$.

(b) Let $q = -t$. Then $(\mathbf{A}^s)^t = (\mathbf{A}^s)^{-q} = ((\mathbf{A}^s)^{-1})^q = ((\mathbf{A}^{-1})^s)^q$ (by Theorem 2.12, part (2)) $= (\mathbf{A}^{-1})^{sq}$ (by Theorem 1.17) $= \mathbf{A}^{-sq}$ (by Theorem 2.12, part (2)) $= \mathbf{A}^{s(-q)} = \mathbf{A}^{st}$. Similarly, $(\mathbf{A}^s)^t = ((\mathbf{A}^{-1})^s)^q$ (as before) $= ((\mathbf{A}^{-1})^q)^s$ (by Theorem 1.17) $= (\mathbf{A}^{-q})^s = (\mathbf{A}^t)^s$.

(18) First assume $\mathbf{AB} = \mathbf{BA}$. Then $(\mathbf{AB})^2 = \mathbf{ABAB} = \mathbf{A}(\mathbf{BA})\mathbf{B} = \mathbf{A}(\mathbf{AB})\mathbf{B} = \mathbf{A}^2\mathbf{B}^2$.
Conversely, if $(\mathbf{AB})^2 = \mathbf{A}^2\mathbf{B}^2$, then $\mathbf{ABAB} = \mathbf{AABB} \Longrightarrow \mathbf{A}^{-1}\mathbf{ABAB}\mathbf{B}^{-1} = \mathbf{A}^{-1}\mathbf{AABB}\mathbf{B}^{-1} \Longrightarrow \mathbf{BA} = \mathbf{AB}$.

(19) If $(\mathbf{AB})^q = \mathbf{A}^q\mathbf{B}^q$ for all $q \geq 2$, use $q = 2$ and the proof in Exercise 18 to show $\mathbf{BA} = \mathbf{AB}$.
Conversely, we need to show that $\mathbf{BA} = \mathbf{AB} \Longrightarrow (\mathbf{AB})^q = \mathbf{A}^q\mathbf{B}^q$, for all $q \geq 2$.
First, we prove that $\mathbf{BA} = \mathbf{AB} \Longrightarrow \mathbf{AB}^q = \mathbf{B}^q\mathbf{A}$, for all $q \geq 2$. We use induction on q.
Base Step ($q = 2$): $\mathbf{AB}^2 = \mathbf{A}(\mathbf{BB}) = (\mathbf{AB})\mathbf{B} = (\mathbf{BA})\mathbf{B} = \mathbf{B}(\mathbf{AB}) = \mathbf{B}(\mathbf{BA}) = (\mathbf{BB})\mathbf{A} = \mathbf{B}^2\mathbf{A}$.
Inductive Step: $\mathbf{AB}^{q+1} = \mathbf{A}(\mathbf{B}^q\mathbf{B}) = (\mathbf{AB}^q)\mathbf{B} = (\mathbf{B}^q\mathbf{A})\mathbf{B}$ (by the inductive hypothesis) $= \mathbf{B}^q(\mathbf{AB})$ $= \mathbf{B}^q(\mathbf{BA}) = (\mathbf{B}^q\mathbf{B})\mathbf{A} = \mathbf{B}^{q+1}\mathbf{A}$.
Now we use this "lemma" ($\mathbf{BA} = \mathbf{AB} \Longrightarrow \mathbf{AB}^q = \mathbf{B}^q\mathbf{A}$ for all $q \geq 2$) to prove $\mathbf{BA} = \mathbf{AB} \Longrightarrow (\mathbf{AB})^q = \mathbf{A}^q\mathbf{B}^q$ for all $q \geq 2$. Again, we proceed by induction on q.
Base Step ($q = 2$): $(\mathbf{AB})^2 = (\mathbf{AB})(\mathbf{AB}) = \mathbf{A}(\mathbf{BA})\mathbf{B} = \mathbf{A}(\mathbf{AB})\mathbf{B} = \mathbf{A}^2\mathbf{B}^2$.
Inductive Step: $(\mathbf{AB})^{q+1} = (\mathbf{AB})^q(\mathbf{AB}) = (\mathbf{A}^q\mathbf{B}^q)(\mathbf{AB})$ (by the inductive hypothesis) $= \mathbf{A}^q(\mathbf{B}^q\mathbf{A})\mathbf{B}$ $= \mathbf{A}^q(\mathbf{AB}^q)\mathbf{B}$ (by the lemma) $= (\mathbf{A}^q\mathbf{A})(\mathbf{B}^q\mathbf{B}) = \mathbf{A}^{q+1}\mathbf{B}^{q+1}$.

(20) Base Step ($k = 0$): $\mathbf{I}_n = (\mathbf{A}^1 - \mathbf{I}_n)(\mathbf{A} - \mathbf{I}_n)^{-1}$.
Inductive Step: Assume $\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^k = (\mathbf{A}^{k+1} - \mathbf{I}_n)(\mathbf{A} - \mathbf{I}_n)^{-1}$, for some k.
Prove $\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^k + \mathbf{A}^{k+1} = (\mathbf{A}^{k+2} - \mathbf{I}_n)(\mathbf{A} - \mathbf{I}_n)^{-1}$.
Now, $\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^k + \mathbf{A}^{k+1} = (\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^k) + \mathbf{A}^{k+1}$
$= (\mathbf{A}^{k+1} - \mathbf{I}_n)(\mathbf{A} - \mathbf{I}_n)^{-1} + \mathbf{A}^{k+1}(\mathbf{A} - \mathbf{I}_n)(\mathbf{A} - \mathbf{I}_n)^{-1}$ (where the first term is obtained from the inductive hypothesis) $= ((\mathbf{A}^{k+1} - \mathbf{I}_n) + \mathbf{A}^{k+1}(\mathbf{A} - \mathbf{I}_n))(\mathbf{A} - \mathbf{I}_n)^{-1} = (\mathbf{A}^{k+2} - \mathbf{I}_n)(\mathbf{A} - \mathbf{I}_n)^{-1}$.

(21) Suppose $\mathbf{A}$ is an $n \times k$ matrix and $\mathbf{B}$ is a $k \times n$ matrix. Suppose, further, that $\mathbf{AB} = \mathbf{I}_n$ and $n > k$. By Corollary 2.3, the homogeneous system having $\mathbf{B}$ as its matrix of coefficients has a nontrivial solution $\mathbf{X}$. That is, there is a nonzero vector $\mathbf{X}$ such that $\mathbf{BX} = \mathbf{0}$. But then $\mathbf{X} = \mathbf{I}_n\mathbf{X} = (\mathbf{AB})\mathbf{X} = \mathbf{A}(\mathbf{BX})$ $= \mathbf{A0} = \mathbf{0}$, a contradiction.

(22) (a) F (b) T (c) T (d) F (e) F (f) T

Chapter 2 Review Exercises

(1) (a) staircase pattern of pivots after Gauss-Jordan Method: $\left[\begin{array}{ccc|c} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right]$;

complete solution set $= \{(-6, 8, -5)\}$

(b) staircase pattern of pivots after Gauss-Jordan Method: $\left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{19}{6} & 0 \\ 0 & 1 & 0 & \frac{1}{6} & 0 \\ 0 & 0 & 1 & \frac{7}{6} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$; no solutions

(c) staircase pattern of pivots after Gauss-Jordan Method: $\left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -1 & -5 \\ 0 & 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right]$;

complete solution set $= \{[-5 - c + e, 1 - 2c - e, c, 1 + 2e, e] \mid c, e \in \mathbb{R}\}$

(2) $y = -2x^3 + 5x^2 - 6x + 3$

(3) (a) No. Entries above pivots need to be zero. (b) No. Rows 3 and 4 should be switched.

(4) $a = 4$, $b = 7$, $c = 4$, $d = 6$

(5) (i) $x_1 = -308$, $x_2 = -18$, $x_3 = -641$, $x_4 = 108$ (ii) $x_1 = -29$, $x_2 = -19$, $x_3 = -88$, $x_4 = 36$

(6) Corollary 2.3 applies since there are more variables than equations.

(7) (a) (I): $\langle 3 \rangle \leftarrow -6 \langle 3 \rangle$ (b) (II): $\langle 2 \rangle \leftarrow 3 \langle 4 \rangle + \langle 2 \rangle$ (c) (III): $\langle 2 \rangle \leftrightarrow \langle 3 \rangle$

(8) (a) $\text{rank}(\mathbf{A}) = 2$, $\text{rank}(\mathbf{B}) = 4$, $\text{rank}(\mathbf{C}) = 3$
(b) $\mathbf{AX} = \mathbf{0}$ and $\mathbf{CX} = \mathbf{0}$: infinite number of solutions; $\mathbf{BX} = \mathbf{0}$: one solution

(9) The reduced row echelon form matrices for $\mathbf{A}$ and $\mathbf{B}$ are both $\begin{bmatrix} 1 & 0 & -3 & 0 & 2 \\ 0 & 1 & 2 & 0 & -4 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$. Therefore, $\mathbf{A}$ and $\mathbf{B}$ are row equivalent by an argument similar to that in Exercise 3(b) of Section 2.3.

(10) (a) Yes. $[-34, 29, -21] = 5[2, -3, -1] + 2[5, -2, 1] - 6[9, -8, 3]$

(b) Yes. $[-34, 29, -21]$ is a linear combination of the rows of the matrix.

(11) $\mathbf{A}^{-1} = \begin{bmatrix} -\frac{2}{9} & \frac{1}{9} \\ -\frac{1}{6} & \frac{1}{3} \end{bmatrix}$

(12) (a) Nonsingular. $\mathbf{A}^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{5}{2} & 2 \\ -1 & 6 & 4 \\ 1 & -5 & -3 \end{bmatrix}$ (b) Singular

(13) This is true by the Inverse Method, which is justified in Section 2.4.

(14) No, by part (1) of Theorem 2.16.

(15) The inverse of the coefficient matrix is $\begin{bmatrix} 3 & -1 & -4 \\ 2 & -1 & -3 \\ 1 & -2 & -2 \end{bmatrix}$.

The solution set is $x_1 = -27$, $x_2 = -21$, $x_3 = -1$.

(16) (a) Because $\mathbf{B}$ is nonsingular, $\mathbf{B}$ is row equivalent to $\mathbf{I}_m$ (see Exercise 13). Thus, there is a sequence of row operations, $R_1, \ldots, R_k$ such that $R_1(\cdots(R_k(\mathbf{B}))\cdots) = \mathbf{I}_m$. Hence, by part (2) of Theorem 2.1,

$$R_1(\cdots(R_k(\mathbf{BA}))\cdots) = R_1(\cdots(R_k(\mathbf{B}))\cdots)\mathbf{A} = \mathbf{I}_m\mathbf{A} = \mathbf{A}.$$

Therefore, $\mathbf{BA}$ is row equivalent to $\mathbf{A}$. Thus, by Exercise 17 in Section 2.3, $\mathbf{BA}$ and $\mathbf{A}$ have the same rank.

(b) By part (d) of Exercise 18 in Section 2.3, $\text{rank}(\mathbf{AC}) \leq \text{rank}(\mathbf{A})$.
Similarly, $\text{rank}(\mathbf{A}) = \text{rank}((\mathbf{AC})\mathbf{C}^{-1}) \leq \text{rank}(\mathbf{AC})$. Hence, $\text{rank}(\mathbf{AC}) = \text{rank}(\mathbf{A})$.

(17)

(a) F	(d) T	(g) T	(j) T	(m) T	(p) T	(s) T
(b) F	(e) F	(h) T	(k) F	(n) T	(q) T	
(c) F	(f) T	(i) F	(l) F	(o) F	(r) T	

Chapter 3

Section 3.1

(1) (a) -17 (b) 6 (c) 0 (d) 1 (e) -108 (f) 156 (g) -40 (h) -60 (i) 0 (j) -3

(2) (a) $\begin{vmatrix} 4 & 3 \\ -2 & 4 \end{vmatrix} = 22$ (b) $\begin{vmatrix} 0 & 2 & -3 \\ 1 & 4 & 2 \\ 4 & -1 & 1 \end{vmatrix} = 65$ (c) $\begin{vmatrix} -3 & 0 & 5 \\ 2 & -1 & 4 \\ 6 & 4 & 0 \end{vmatrix} = 118$

(3) (a) $(-1)^{2+2}\begin{vmatrix} 4 & -3 \\ 9 & -7 \end{vmatrix} = -1$

(b) $(-1)^{2+3}\begin{vmatrix} -9 & 6 \\ 4 & 3 \end{vmatrix} = 51$

(c) $(-1)^{4+3}\begin{vmatrix} -5 & 2 & 13 \\ -8 & 2 & 22 \\ -6 & -3 & -16 \end{vmatrix} = 222$

(d) $(-1)^{1+2}\begin{vmatrix} x-4 & x-3 \\ x-1 & x+2 \end{vmatrix} = -2x+11$

(4) Same answers as Exercise 1.

(5) (a) 0 (b) -251 (c) -60 (d) 352

(6) $$a_{11}a_{22}a_{33}a_{44} + a_{11}a_{23}a_{34}a_{42} + a_{11}a_{24}a_{32}a_{43} + a_{12}a_{21}a_{34}a_{43} + a_{12}a_{23}a_{31}a_{44} + a_{12}a_{24}a_{33}a_{41}$$
$$+ a_{13}a_{21}a_{32}a_{44} + a_{13}a_{22}a_{34}a_{41} + a_{13}a_{24}a_{31}a_{42} + a_{14}a_{21}a_{33}a_{42} + a_{14}a_{22}a_{31}a_{43} + a_{14}a_{23}a_{32}a_{41}$$
$$- a_{11}a_{22}a_{34}a_{43} - a_{11}a_{23}a_{32}a_{44} - a_{11}a_{24}a_{33}a_{42} - a_{11}a_{21}a_{33}a_{44} - a_{12}a_{23}a_{34}a_{41} - a_{12}a_{24}a_{31}a_{43}$$
$$- a_{13}a_{21}a_{34}a_{42} - a_{13}a_{22}a_{31}a_{44} - a_{13}a_{24}a_{32}a_{41} - a_{14}a_{21}a_{32}a_{43} - a_{14}a_{22}a_{33}a_{41} - a_{14}a_{23}a_{31}a_{42}$$

(7) Let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, and let $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(8) (a) Perform the basketweaving method.

(b) $$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{a} = (a_2b_3 - a_3b_2)a_1 + (a_3b_1 - a_1b_3)a_2 + (a_1b_2 - a_2b_1)a_3$$
$$= a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_1a_2b_3 + a_1a_3b_2 - a_2a_3b_1 = 0.$$
Similarly, $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{b} = 0$.

(9) (a) 7 (b) 18 (c) 12 (d) 0

(10) Let $\mathbf{x} = [x_1, x_2]$ and $\mathbf{y} = [y_1, y_2]$. Then

$$\mathbf{proj_x y} = \left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x} = \frac{1}{\|\mathbf{x}\|^2}[(x_1y_1 + x_2y_2)x_1, (x_1y_1 + x_2y_2)x_2].$$

Hence,

$$\begin{aligned}
\mathbf{y} - \mathbf{proj_x y} &= \frac{1}{\|\mathbf{x}\|^2}(\|\mathbf{x}\|^2\mathbf{y}) - \mathbf{proj_x y} \\
&= \frac{1}{\|\mathbf{x}\|^2}[(x_1^2 + x_2^2)y_1, (x_1^2 + x_2^2)y_2] - \frac{1}{\|\mathbf{x}\|^2}[(x_1y_1 + x_2y_2)x_1, (x_1y_1 + x_2y_2)x_2] \\
&= \frac{1}{\|\mathbf{x}\|^2}[x_1^2y_1 + x_2^2y_1 - x_1^2y_1 - x_1x_2y_2, x_1^2y_2 + x_2^2y_2 - x_1x_2y_1 - x_2^2y_2]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\|\mathbf{x}\|^2}[x_2^2y_1 - x_1x_2y_2, x_1^2y_2 - x_1x_2y_1] \\
&= \frac{1}{\|\mathbf{x}\|^2}[x_2(x_2y_1 - x_1y_2), x_1(x_1y_2 - x_2y_1)] \\
&= \frac{x_1y_2 - x_2y_1}{\|\mathbf{x}\|^2}[-x_2, x_1].
\end{aligned}$$

Thus,

$$\begin{aligned}
\|\mathbf{x}\|\,\|\mathbf{y} - \mathbf{proj_x y}\| &= \|\mathbf{x}\|\frac{|x_1y_2 - x_2y_1|}{\|\mathbf{x}\|^2}\sqrt{x_2^2 + x_1^2} = |x_1y_2 - x_2y_1| \\
&= \text{absolute value of } \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.
\end{aligned}$$

(11) (a) 18 (b) 7 (c) 63 (d) 8

(12) First, notice from the definition of $\mathbf{x} \times \mathbf{y}$ in Exercise 8, that

$$\|\mathbf{x} \times \mathbf{y}\| = \sqrt{(x_2y_3 - x_3y_2)^2 + (x_1y_3 - x_3y_1)^2 + (x_1y_2 - x_2y_1)^2}.$$

We will verify that $\|\mathbf{x} \times \mathbf{y}\|$ is equal to the area of the parallelogram determined by $\mathbf{x}$ and $\mathbf{y}$. Now, from the solution to Exercise 10 above, the area of this parallelogram is equal to $\|\mathbf{x}\|\,\|\mathbf{y} - \mathbf{proj_x y}\|$. One can verify $\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\|\,\|\mathbf{y} - \mathbf{proj_x y}\|$ by a tedious, brute force, argument. (Algebraically expand and simplify $\|\mathbf{x}\|^2\,\|\mathbf{y} - \mathbf{proj_x y}\|^2$ to get $(x_2y_3 - x_3y_2)^2 + (x_1y_3 - x_3y_1)^2 + (x_1y_2 - x_2y_1)^2$.) An alternate approach is the following: Note that

$$\begin{aligned}
\|\mathbf{x} \times \mathbf{y}\|^2 &= (x_2y_3 - x_3y_2)^2 + (x_1y_3 - x_3y_1)^2 + (x_1y_2 - x_2y_1)^2 \\
&= x_2^2y_3^2 - 2x_2x_3y_2y_3 + x_3^2y_2^2 + x_1^2y_3^2 - 2x_1x_3y_1y_3 + x_3^2y_1^2 + x_1^2y_2^2 - 2x_1x_2y_1y_2 + x_2^2y_1^2.
\end{aligned}$$

Using some algebraic manipulation, this can be expressed as

$$\|\mathbf{x} \times \mathbf{y}\|^2 = (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2.$$

Therefore,

$$\begin{aligned}
\|\mathbf{x} \times \mathbf{y}\|^2 &= \|\mathbf{x}\|^2\,\|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2 \\
&= \frac{\|\mathbf{x}\|^2}{\|\mathbf{x}\|^2}\left(\|\mathbf{x}\|^2\,\|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2\right) \\
&= \|\mathbf{x}\|^2\left(\frac{\|\mathbf{x}\|^2\,\|\mathbf{y}\|^2}{\|\mathbf{x}\|^2} - \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{x}\|^2}\right) \\
&= \|\mathbf{x}\|^2\left(\frac{\|\mathbf{x}\|^2\,(\mathbf{y} \cdot \mathbf{y})}{\|\mathbf{x}\|^2} - 2\left(\frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{x}\|^2}\right) + \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{x}\|^2}\right) \\
&= \|\mathbf{x}\|^2\left((\mathbf{y} \cdot \mathbf{y}) - 2\left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|^2}\right)(\mathbf{x} \cdot \mathbf{y}) + \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|^2}\right)^2(\mathbf{x} \cdot \mathbf{x})\right)
\end{aligned}$$

$$\begin{aligned} &= \|\mathbf{x}\|^2 \left(\mathbf{y} - \frac{(\mathbf{x}\cdot\mathbf{y})\,\mathbf{x}}{\|\mathbf{x}\|^2}\right) \cdot \left(\mathbf{y} - \frac{(\mathbf{x}\cdot\mathbf{y})\,\mathbf{x}}{\|\mathbf{x}\|^2}\right) \\ &= \|\mathbf{x}\|^2 \left\| \mathbf{y} - \left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x} \right\|^2 \\ &= \|\mathbf{x}\|^2 \|\mathbf{y} - \mathbf{proj_x y}\|^2. \end{aligned}$$

Hence, $\|\mathbf{x}\times\mathbf{y}\| = \|\mathbf{x}\|\,\|\mathbf{y} - \mathbf{proj_x y}\|$, the area of the parallelogram determined by $\mathbf{x}$ and $\mathbf{y}$.

Now, we determine the volume of the parallelepiped. As in the hint, let $\mathbf{h} = \mathbf{proj}_{(\mathbf{x}\times\mathbf{y})}\mathbf{z}$, the perpendicular from $\mathbf{z}$ to the parallelogram determined by $\mathbf{x}$ and $\mathbf{y}$. Since the area of this parallelogram equals $\|\mathbf{x}\times\mathbf{y}\|$, the volume of the parallelepiped equals

$$\|\mathbf{h}\|\,\|\mathbf{x}\times\mathbf{y}\| = \|\mathbf{proj}_{(\mathbf{x}\times\mathbf{y})}\mathbf{z}\|\,\|\mathbf{x}\times\mathbf{y}\| = \left|\frac{\mathbf{z}\cdot(\mathbf{x}\times\mathbf{y})}{\|\mathbf{x}\times\mathbf{y}\|}\right| \|\mathbf{x}\times\mathbf{y}\| = |\mathbf{z}\cdot(\mathbf{x}\times\mathbf{y})| = |(\mathbf{x}\times\mathbf{y})\cdot\mathbf{z}|\,.$$

But from the definition in Exercise 8,

$$\begin{aligned} |(\mathbf{x}\times\mathbf{y})\cdot\mathbf{z}| &= |\ (x_2y_3 - x_3y_2)z_1 + (x_3y_1 - x_1y_3)z_2 + (x_1y_2 - x_2y_1)z_3\ | \\ &= |\ x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_3y_2z_1 - x_1y_3z_2 - x_2y_1z_3\ | \\ &= \text{absolute value of } \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}. \end{aligned}$$

(13) (a) Base Step: If $n = 1$, then $\mathbf{A} = [a_{11}]$, so $c\mathbf{A} = [ca_{11}]$, and $|c\mathbf{A}| = ca_{11} = c|\mathbf{A}|$.
Inductive Step: Assume that if $\mathbf{A}$ is an $n\times n$ matrix and c is a scalar, then $|c\mathbf{A}| = c^n|\mathbf{A}|$. Prove that if $\mathbf{A}$ is an $(n+1)\times(n+1)$ matrix, and c is a scalar, then $|c\mathbf{A}| = c^{n+1}|\mathbf{A}|$. Let $\mathbf{B} = c\mathbf{A}$. Then

$$|\mathbf{B}| = b_{n+1,1}\mathcal{B}_{n+1,1} + b_{n+1,2}\mathcal{B}_{n+1,2} + \cdots + b_{n+1,n}\mathcal{B}_{n+1,n} + b_{n+1,n+1}\mathcal{B}_{n+1,n+1}.$$

Each $\mathcal{B}_{n+1,i} = (-1)^{n+1+i}|\mathbf{B}_{n+1,i}| = c^n(-1)^{n+1+i}|\mathbf{A}_{n+1,i}|$ (by the inductive hypothesis, since $\mathcal{B}_{n+1,i}$ is an $n\times n$ matrix) $= c^n\mathbf{A}_{n+1,i}$. Thus,

$$\begin{aligned} |c\mathbf{A}| &= |\mathbf{B}| = ca_{n+1,1}(c^n\mathcal{A}_{n+1,1}) + ca_{n+1,2}(c^n\mathcal{A}_{n+1,2}) + \cdots + ca_{n+1,n}(c^n\mathcal{A}_{n+1,n}) \\ &\qquad + ca_{n+1,n+1}(c^n\mathcal{A}_{n+1,n+1}) \\ &= c^{n+1}(a_{n+1,1}\mathcal{A}_{n+1,1} + a_{n+1,2}\mathcal{A}_{n+1,2} + \cdots + a_{n+1,n}\mathcal{A}_{n+1,n} + a_{n+1,n+1}\mathcal{A}_{n+1,n+1}) \\ &= c^{n+1}|\mathbf{A}|. \end{aligned}$$

(b) $|2\mathbf{A}| = 2^3|\mathbf{A}|$ (since $\mathbf{A}$ is a 3×3 matrix) $= 8|\mathbf{A}|$.

(14)

$$\begin{vmatrix} x & -1 & 0 & 0 \\ 0 & x & -1 & 0 \\ 0 & 0 & x & -1 \\ a_0 & a_1 & a_2 & a_3+x \end{vmatrix} = a_0(-1)^{4+1}\begin{vmatrix} -1 & 0 & 0 \\ x & -1 & 0 \\ 0 & x & -1 \end{vmatrix} + a_1(-1)^{4+2}\begin{vmatrix} x & 0 & 0 \\ 0 & -1 & 0 \\ 0 & x & -1 \end{vmatrix}$$
$$+ a_2(-1)^{4+3}\begin{vmatrix} x & -1 & 0 \\ 0 & x & 0 \\ 0 & 0 & -1 \end{vmatrix} + a_3(-1)^{4+4}\begin{vmatrix} x & -1 & 0 \\ 0 & x & -1 \\ 0 & 0 & x \end{vmatrix}.$$

The four 3×3 determinants in the previous equation can be calculated using "basketweaving" as -1, x, $-x^2$, and x^3, respectively. Therefore, the original 4×4 determinant equals

$$(-a_0)(-1) + (a_1)(x) + (-a_2)(-x^2) + (a_3)(x^3) = a_0 + a_1 x + a_2 x^2 + a_3 x^3.$$

(15) (a) $x = -5$ or $x = 2$ (b) $x = -2$ or $x = -1$ (c) $x = 3$, $x = 1$, or $x = 2$

(16) (a) Use "basketweaving" and factor. (b) 20

(17) (a) Area of $T = \frac{\sqrt{3}s^2}{4}$.
(b) Suppose one side of T extends from (a, b) to (c, d), where (a, b) and (c, d) are lattice points. Then $s^2 = (c-a)^2 + (d-b)^2$, an integer. Hence, $\frac{s^2}{4}$ is rational, and the area of $T = \frac{\sqrt{3}s^2}{4}$ is irrational.
(c) Suppose the vertices of T are lattice points $X = (a, b)$, $Y = (c, d)$, $Z = (e, f)$. Then side XY is expressed using vector $[c-a,\ d-b]$, and side XZ is expressed using vector $[e-a,\ f-b]$. Hence, the area of $T = \frac{1}{2}$(area of the parallelogram formed by $[c-a,\ d-b]$ and $[e-a,\ f-b]$) $= \frac{1}{2}\begin{vmatrix} c-a & d-b \\ e-a & f-b \end{vmatrix}$.
(d) $\frac{1}{2}\begin{vmatrix} c-a & d-b \\ e-a & f-b \end{vmatrix} = \frac{1}{2}((c-a)(f-b) - (d-b)(e-a))$, which is $\frac{1}{2}$ times the difference of two products of integers, and hence, is rational.

(18) (a) F (b) T (c) F (d) F (e) T

Section 3.2

(1) (a) (II): $\langle 1\rangle \leftarrow -3\,\langle 2\rangle + \langle 1\rangle$; determinant $= 1$
(b) (III): $\langle 2\rangle \leftrightarrow \langle 3\rangle$; determinant $= -1$
(c) (I): $\langle 3\rangle \leftarrow -4\,\langle 3\rangle$; determinant $= -4$
(d) (II): $\langle 2\rangle \leftarrow 2\,\langle 1\rangle + \langle 2\rangle$; determinant $= 1$
(e) (I): $\langle 1\rangle \leftarrow \frac{1}{2}\,\langle 1\rangle$; determinant $= \frac{1}{2}$
(f) (III): $\langle 1\rangle \leftrightarrow \langle 2\rangle$; determinant $= -1$

(2) (a) 30 (b) -3 (c) -4 (d) 3 (e) 35 (f) 20

(3) (a) Determinant $= -2$; nonsingular
(b) Determinant $= 1$; nonsingular
(c) Determinant $= -79$; nonsingular
(d) Determinant $= 0$; singular

(4) (a) Determinant $= -1$; the system has only the trivial solution.
(b) Determinant $= 0$; the system has nontrivial solutions. (One nontrivial solution is $(1, 7, 3)$).
(c) Determinant $= -42$; the system has only the trivial solution.

(5) Use Theorem 3.2.

(6) Use row operations to reverse the order of the rows of $\mathbf{A}$. This can be done using 3 Type (III) operations, an odd number. (These operations are $\langle 1\rangle \leftrightarrow \langle 6\rangle$, $\langle 2\rangle \leftrightarrow \langle 5\rangle$ and $\langle 3\rangle \leftrightarrow \langle 4\rangle$.) Hence, by part (3) of Theorem 3.3, $|\mathbf{A}| = -a_{16}a_{25}a_{34}a_{43}a_{52}a_{61}$.

(7) If $|\mathbf{A}| \neq 0$, $\mathbf{A}^{-1}$ exists. Hence $\mathbf{AB} = \mathbf{AC} \Longrightarrow \mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{AC} \Longrightarrow \mathbf{B} = \mathbf{C}$.

(8) (a) Base Step: When $n = 1$, $\mathbf{A} = [a_{11}]$. Then $\mathbf{B} = R(\mathbf{A})$ (with $i = 1$) $= [ca_{11}]$. Hence, $|\mathbf{B}| = ca_{11} = c|\mathbf{A}|$.

(b) Inductive Hypothesis: Assume that if $\mathbf{A}$ is an $n \times n$ matrix, and R is the row operation (I): $\langle i \rangle \leftarrow c\,\langle i \rangle$, and $\mathbf{B} = R(\mathbf{A})$, then $|\mathbf{B}| = c|\mathbf{A}|$.

Note: In parts (c) and (d), we must prove that if $\mathbf{A}$ is an $(n+1) \times (n+1)$ matrix and R is the row operation (I): $\langle i \rangle \leftarrow c\,\langle i \rangle$, and $\mathbf{B} = R(\mathbf{A})$, then $|\mathbf{B}| = c|\mathbf{A}|$.

(c) Inductive Step when $i \neq n$:

$$\begin{aligned} |\mathbf{B}| &= b_{n+1,1}\mathcal{B}_{n+1,1} + \cdots + b_{n+1,n}\mathcal{B}_{n+1,n} + b_{n+1,n+1}\mathcal{B}_{n+1,n+1} \\ &= b_{n+1,1}(-1)^{n+1+1}|\mathbf{B}_{n+1,1}| + \cdots + b_{n+1,n}(-1)^{n+1+n}|\mathbf{B}_{n+1,n}| \\ &\qquad + b_{n+1,n+1}(-1)^{n+1+n+1}|\mathbf{B}_{n+1,n+1}|. \end{aligned}$$

But since $|\mathbf{B}_{n+1,i}|$ is the determinant of the $n \times n$ matrix $\mathbf{A}_{n+1,i}$ after the row operation (I): $\langle i \rangle \leftarrow c\,\langle i \rangle$ has been performed (since $i \neq n$), the inductive hypothesis shows that each $|\mathbf{B}_{n+1,i}| = c|\mathbf{A}_{n+1,i}|$. Since each $b_{n+1,i} = a_{n+1,i}$, we have

$$\begin{aligned} |\mathbf{B}| &= ca_{n+1,1}(-1)^{n+1+1}|\mathbf{A}_{n+1,1}| + \cdots + ca_{n+1,n}(-1)^{n+1+n}|\mathbf{A}_{n+1,n}| \\ &\qquad + ca_{n+1,n+1}(-1)^{n+1+n+1}|\mathbf{A}_{n+1,n+1}| \\ &= c|\mathbf{A}|. \end{aligned}$$

(d) Inductive Step when $i = n$: Again,

$$|\mathbf{B}| = b_{n+1,1}\mathcal{B}_{n+1,1} + \cdots + b_{n+1,n}\mathcal{B}_{n+1,n} + b_{n+1,n+1}\mathcal{B}_{n+1,n+1}.$$

Since $i = n$, each $b_{n+1,i} = ca_{n+1,i}$, while each $\mathcal{B}_{n+1,i} = \mathcal{A}_{n+1,i}$ (since the first n rows of $\mathbf{A}$ are left unchanged by the row operation in this case). Hence,

$$|\mathbf{B}| = c(a_{n+1,1}\mathcal{A}_{n+1,1}) + \cdots + c(a_{n+1,n}\mathcal{A}_{n+1,n}) + c(a_{n+1,n+1}\mathcal{A}_{n+1,n+1}) = c|\mathbf{A}|.$$

(9) (a) In order to add a multiple of one row to another, we need at least two rows. Hence $n = 2$ here.

(b) Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then the row operation (II): $\langle 1 \rangle \leftarrow c\langle 2 \rangle + \langle 1 \rangle$ produces

$$\mathbf{B} = \begin{bmatrix} ca_{21} + a_{11} & ca_{22} + a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

and $|\mathbf{B}| = (ca_{21} + a_{11})a_{22} - (ca_{22} + a_{12})a_{21}$, which reduces to $a_{11}a_{22} - a_{12}a_{21} = |\mathbf{A}|$.

(c) Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then the row operation (II): $\langle 2 \rangle \leftarrow c\langle 1 \rangle + \langle 2 \rangle$ produces

$$\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} \\ ca_{11} + a_{21} & ca_{12} + a_{22} \end{bmatrix},$$

and $|\mathbf{B}| = a_{11}(ca_{12} + a_{22}) - a_{12}(ca_{11} + a_{21})$, which reduces to $a_{11}a_{22} - a_{12}a_{21} = |\mathbf{A}|$.

(10) (a) Inductive Hypothesis: Assume that if $\mathbf{A}$ is an $(n-1) \times (n-1)$ matrix and R is a Type (II) row operation, then $|R(\mathbf{A})| = |\mathbf{A}|$.
To complete the Inductive Step, we must prove that if $\mathbf{A}$ is an $n \times n$ matrix, and R is a Type (II) row operation, then $|R(\mathbf{A})| = |\mathbf{A}|$.
In what follows, let R be the Type (II) row operation $\langle i \rangle \leftarrow c\,\langle j \rangle + \langle i \rangle$ (so, $i \neq j$).

(b) Inductive Step when $i \neq n$, $j \neq n$: Let $\mathbf{B} = R(\mathbf{A})$. Then $|\mathbf{B}| = b_{n1}\mathcal{B}_{n1} + \cdots + b_{nn}\mathcal{B}_{nn}$. But since the nth row of $\mathbf{A}$ is not affected by R, each $b_{ni} = a_{ni}$. Also, each $\mathcal{B}_{ni} = (-1)^{n+i}|\mathbf{B}_{ni}|$, and $|\mathbf{B}_{ni}|$ is the determinant of the $(n-1) \times (n-1)$ matrix $R(\mathbf{A}_{ni})$. But by the inductive hypothesis, each $|\mathbf{B}_{ni}| = |\mathbf{A}_{ni}|$. Hence, $|\mathbf{B}| = |\mathbf{A}|$.

(c) Since $i \neq j$, we can assume $j \neq n$. Let $i = n$, and let $k \neq n$.
The nth row of $R_1(R_2(R_1(\mathbf{A})))$
$= k$th row of $R_2(R_1(\mathbf{A}))$
$= k$th row of $R_1(\mathbf{A}) + c(j$th row of $R_1(\mathbf{A}))$
$= n$th row of $\mathbf{A} + c(j$th row of $\mathbf{A})$
$= n$th row of $R(\mathbf{A})$.
The kth row of $R_1(R_2(R_1(\mathbf{A})))$
$= n$th row of $R_2(R_1(\mathbf{A}))$
$= n$th row of $R_1(\mathbf{A})$ (since $k \neq n$)
$= k$th row of $\mathbf{A} = k$th row of $R(\mathbf{A})$.
The lth row (where $l \neq k, n$) of $R_1(R_2(R_1(\mathbf{A})))$
$= l$th row of $R_2(R_1(\mathbf{A}))$ (since $l \neq k, n$)
$= l$th row of $R_1(\mathbf{A})$ (since $l \neq k$)
$= l$th row of $\mathbf{A}$ (since $l \neq k, n$)
$= l$th row of $R(\mathbf{A})$ (since $l \neq n = i$).

(d) Inductive Step when $i = n$: From part (c), we have:
$|R(\mathbf{A})| = |R_1(R_2(R_1(\mathbf{A})))| = -|R_2(R_1(\mathbf{A}))|$ (by part (3) of Theorem 3.3)
$= -|R_1(\mathbf{A})|$ (by part (b), since $k, j \neq n = i$)
$= -(-|\mathbf{A}|)$ (by part (3) of Theorem 3.3) $= |\mathbf{A}|$.

(e) Since $i \neq j$, we can assume $i \neq n$. Now,
the ith row of $R_1(R_3(R_1(\mathbf{A})))$
$= i$th row of $R_3(R_1(\mathbf{A}))$
$= i$th row of $R_1(\mathbf{A}) + c(k$th row of $R_1(\mathbf{A}))$
$= i$th row of $\mathbf{A} + c(n$th row of $\mathbf{A})$ (since $i \neq k$)
$= i$th row of $R(\mathbf{A})$ (since $j = n$).
The kth row of $R_1(R_3(R_1(\mathbf{A})))$
$= n$th row of $R_3(R_1(\mathbf{A}))$
$= n$th row of $R_1(\mathbf{A})$ (since $i \neq n$)
$= k$th row of $\mathbf{A} = k$th row of $R(\mathbf{A})$ (since $i \neq k$).
The nth row of $R_1(R_3(R_1(\mathbf{A})))$
$= k$th row of $R_3(R_1(\mathbf{A}))$
$= k$th row of $R_1(\mathbf{A})$ (since $k \neq i$) $= n$th row of $\mathbf{A}$
$= n$th row of $R(\mathbf{A})$ (since $i \neq n$).
The lth row (where $l \neq i, k, n$) of $R_1(R_3(R_1(\mathbf{A})))$
$= l$th row of $R_3(R_1(\mathbf{A}))$
$= l$th row of $R_1(\mathbf{A})$
$= l$th row of $\mathbf{A}$
$= l$th row of $R(\mathbf{A})$.

(f) Inductive Step when $j = n$:
$|R(\mathbf{A})| = |R_1(R_3(R_1(\mathbf{A})))| = -|R_3(R_1(\mathbf{A}))|$ (by part (3) of Theorem 3.3)
$= -|R_1(\mathbf{A})|$ (by part (b))
$= -(-|\mathbf{A}|)$ (by part (3) of Theorem 3.3) $= |\mathbf{A}|$.

(11) (a) Suppose the ith row of $\mathbf{A}$ is a row of all zeroes. Let R be the row operation (I): $\langle i\rangle \leftarrow 2\langle i\rangle$. Then $|R(\mathbf{A})| = 2|\mathbf{A}|$ by part (1) of Theorem 3.3. But also $|R(\mathbf{A})| = |\mathbf{A}|$ since R has no effect on $\mathbf{A}$. Hence, $2|\mathbf{A}| = |\mathbf{A}|$, so $|\mathbf{A}| = 0$.

(b) If $\mathbf{A}$ contains a row of zeroes, $\text{rank}(\mathbf{A}) < n$ (since the reduced row echelon form of $\mathbf{A}$ also contains at least one row of zeroes). Hence by Corollary 3.6, $|\mathbf{A}| = 0$.

(12) (a) Let rows i, j of $\mathbf{A}$ be identical. Let R be the row operation (III): $\langle i\rangle \leftrightarrow \langle j\rangle$. Then, $R(\mathbf{A}) = \mathbf{A}$, and so $|R(\mathbf{A})| = |\mathbf{A}|$. But by part (3) of Theorem 3.3, $|R(\mathbf{A})| = -|\mathbf{A}|$. Hence, $|\mathbf{A}| = -|\mathbf{A}|$, and so $|\mathbf{A}| = 0$.

(b) If two rows of $\mathbf{A}$ are identical, then subtracting one of these rows from the other shows that $\mathbf{A}$ is row equivalent to a matrix with a row of zeroes. Use part (b) of Exercise 11.

(13) (a) Suppose the ith row of $\mathbf{A} = c$(jth row of $\mathbf{A}$). Using the row operation (II): $\langle i\rangle \leftarrow -c\langle j\rangle + \langle i\rangle$ shows that $\mathbf{A}$ is row equivalent to a matrix with a row of zeroes. Use part (a) of Exercise 11.

(b) Use the given hint, Theorem 2.7 and Corollary 3.6.

(14) (a) We present two different proofs of this result.

First proof: Let $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}$. If $|\mathbf{B}| = 0$, then, when row reducing $\mathbf{A}$ into upper triangular form, one of the first m columns will not be a pivot column. Similarly, if $|\mathbf{B}| \neq 0$ and $|\mathbf{D}| = 0$, one of the last $(n-m)$ columns will not be a pivot column. In either case, some column will not contain a pivot, and so $\text{rank}(\mathbf{A}) < n$. Hence, in both cases, $|\mathbf{A}| = 0 = |\mathbf{B}||\mathbf{D}|$.

Now suppose that both $|\mathbf{B}|$ and $|\mathbf{D}|$ are nonzero. Then the row operations for the first m pivots used to put $\mathbf{A}$ into upper triangular form will be the same as the row operations for the first m pivots used to put $\mathbf{B}$ into upper triangular form (putting 1's along the entire main diagonal of $\mathbf{B}$). These operations convert $\mathbf{A}$ into the matrix $\begin{bmatrix} \mathbf{U} & \mathbf{G} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}$, where $\mathbf{U}$ is an $m \times m$ upper triangular matrix with 1's on the main diagonal, and $\mathbf{G}$ is some $m \times (n-m)$ matrix. Hence, at this point in the computation of $|\mathbf{B}|$, the multiplicative factor P that we use to calculate the determinant (as in Example 5 in the text) equals $1/|\mathbf{B}|$.

The remaining $(n-m)$ pivots to put $\mathbf{A}$ into upper triangular form will involve the same row operations needed to put $\mathbf{D}$ into upper triangular form (with 1's all along the main diagonal). Clearly, this means that the multiplicative factor P must be multiplied by $1/|\mathbf{D}|$, and so the new value of P equals $1/(|\mathbf{B}||\mathbf{D}|)$. Since the final matrix is in upper triangular form with 1's all along the main diagonal, its determinant is 1. Multiplying this by the reciprocal of the final value of P (as in Example 5) yields $|\mathbf{A}| = |\mathbf{B}||\mathbf{D}|$, completing the first proof.

Second proof: Let $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}$, where $\mathbf{B}$ is an $m \times m$ submatrix, $\mathbf{C}$ is an $m \times (n-m)$ submatrix, $\mathbf{D}$ is an $(n-m) \times (n-m)$ submatrix, and $\mathbf{O}$ is an $(n-m) \times m$ zero submatrix. Let $k = n - m$, so that $\mathbf{D}$ is a $k \times k$ matrix. We will prove $|\mathbf{A}| = |\mathbf{B}|\,|\mathbf{D}|$ by induction on k.

Base Step ($k = 1$):

$$|\mathbf{A}| = a_{n1}\mathcal{A}_{n1} + \cdots + a_{nn}\mathcal{A}_{nn} = 0\mathcal{A}_{n1} + \cdots + 0\mathcal{A}_{n(n-1)} + d_{11}\,|\mathbf{B}| = |\mathbf{D}|\,|\mathbf{B}|\,.$$

Inductive Step: Assume $|\mathbf{A}| = |\mathbf{B}|\,|\mathbf{D}|$ whenever $\mathbf{D}$ is a $(k-1)\times(k-1)$ matrix. We must prove $|\mathbf{A}| = |\mathbf{B}|\,|\mathbf{D}|$ when $\mathbf{D}$ is a $k \times k$ matrix. Now, $|\mathbf{A}| =$

$$a_{n1}\mathcal{A}_{n1}+\cdots+a_{nn}\mathcal{A}_{nn} = 0\mathcal{A}_{n1}+\cdots+0\mathcal{A}_{n(n-k)}+d_{k1}\,(-1)^{2n-k+1}\left|\mathbf{A}_{n(n-k+1)}\right|+\cdots+d_{kk}\,(-1)^{2n}\,|\mathbf{A}_{nn}|\,.$$

But $\mathbf{A}_{n(n-k+i)} = \begin{bmatrix} \mathbf{B} & \mathbf{G}_i \\ \mathbf{O} & \mathbf{D}_{ki} \end{bmatrix}$, where $\mathbf{G}_i$ is the matrix $\mathbf{C}$ with its ith column removed. Hence, by the inductive hypothesis $|\mathbf{A}| = d_{k1}\,(-1)^{2n-k+1}\,|\mathbf{B}|\,|\mathbf{D}_{k1}|+\cdots+d_{kk}\,(-1)^{2n}\,|\mathbf{B}|\,|\mathbf{D}_{kk}|$. Now, $k+i$ can be expressed as $(2n-k+i)+2(k-n)$, and so has the same parity (even or odd) as $2n-k+i$, and hence $(-1)^{2n-k+i} = (-1)^{k+i}$. Therefore, $|\mathbf{A}| = d_{k1}\,(-1)^{k+1}\,|\mathbf{B}|\,|\mathbf{D}_{k1}| + \cdots + d_{kk}\,(-1)^{2k}\,|\mathbf{B}|\,|\mathbf{D}_{kk}|$ $= |\mathbf{B}| \left(d_{k1}\,(-1)^{k+1}\,|\mathbf{D}_{k1}| + \cdots + d_{kk}\,(-1)^{2k}\,|\mathbf{D}_{kk}|\right) = |\mathbf{B}|\,|\mathbf{D}|$, completing the second proof.

(b) $(18-18)(20-3) = 0(17) = 0$

(15) Follow the hint in the text. If $\mathbf{A}$ is row equivalent to $\mathbf{B}$ in reduced row echelon form, follow the method in Example 5 in the text to find determinants using row reduction by maintaining a multiplicative factor P throughout the process. But, since the same rules work for f with regard to row operations as for the determinant, the same factor P will be produced for f. If $\mathbf{B} = \mathbf{I}_n$, then $f(\mathbf{B}) = 1 = |\mathbf{B}|$. Since the factor P is the same for f and the determinant, $f(\mathbf{A}) = (1/P)f(\mathbf{B}) = (1/P)|\mathbf{B}| = |\mathbf{A}|$. If, instead, $\mathbf{B} \neq \mathbf{I}_n$, then the nth row of $\mathbf{B}$ will be all zeroes. Perform the operation (I): $\langle n\rangle \leftarrow 2\langle n\rangle$ on $\mathbf{B}$, which yields $\mathbf{B}$. Then, $f(\mathbf{B}) = 2f(\mathbf{B})$, implying $f(\mathbf{B}) = 0$. Hence, $f(\mathbf{A}) = (1/P)0 = 0$.

(16) By Corollary 3.6 and part (1) of Theorem 2.7, the homogeneous system $\mathbf{AX} = \mathbf{0}$ has nontrivial solutions. Let $\mathbf{B}$ be any $n \times n$ matrix such that every column of $\mathbf{B}$ is a nontrivial solution for $\mathbf{AX} = \mathbf{0}$. Then the ith column of $\mathbf{AB} = \mathbf{A}$(ith column of $\mathbf{B}$) $= \mathbf{0}$ for every i. Hence, $\mathbf{AB} = \mathbf{O}_n$.

(17) (a) F (b) T (c) F (d) F (e) F (f) T

Section 3.3

(1) (a) $a_{31}(-1)^{3+1}|\mathbf{A}_{31}| + a_{32}(-1)^{3+2}|\mathbf{A}_{32}| + a_{33}(-1)^{3+3}|\mathbf{A}_{33}| + a_{34}(-1)^{3+4}|\mathbf{A}_{34}|$

(b) $a_{11}(-1)^{1+1}|\mathbf{A}_{11}| + a_{12}(-1)^{1+2}|\mathbf{A}_{12}| + a_{13}(-1)^{1+3}|\mathbf{A}_{13}| + a_{14}(-1)^{1+4}|\mathbf{A}_{14}|$

(c) $a_{14}(-1)^{1+4}|\mathbf{A}_{14}| + a_{24}(-1)^{2+4}|\mathbf{A}_{24}| + a_{34}(-1)^{3+4}|\mathbf{A}_{34}| + a_{44}(-1)^{4+4}|\mathbf{A}_{44}|$

(d) $a_{11}(-1)^{1+1}|\mathbf{A}_{11}| + a_{21}(-1)^{2+1}|\mathbf{A}_{21}| + a_{31}(-1)^{3+1}|\mathbf{A}_{31}| + a_{41}(-1)^{4+1}|\mathbf{A}_{41}|$

(2) (a) -76 (b) 465 (c) 102 (d) 410

(3) (a) $\{(-4, 3, -7)\}$ (b) $\{(5, -1, -3)\}$ (c) $\{(6, 4, -5)\}$ (d) $\{(4, -1, -3, 6)\}$

(4) (a) Suppose R is the Type (I) operation $\langle i\rangle \leftarrow c\,\langle i\rangle$. Then $R(\mathbf{I}_n)$ is the diagonal matrix having 1's in every entry on the main diagonal, except for the (i, i) entry, which has the value c. But, since $R(\mathbf{I}_n)$ is a diagonal matrix, it is symmetric.

(b) Suppose R is the Type (III) operation $\langle i\rangle \leftrightarrow \langle j\rangle$. Then the only nonzero entries off the main diagonal for $R(\mathbf{I}_n)$ are the (i, j) entry and the (j, i) entry, which are both 1's. Hence, all entries across the main diagonal from each other in $R(\mathbf{I}_n)$ are equal, and so $R(\mathbf{I}_n)$ is symmetric.

(c) Suppose R is the Type (II) operation $\langle i\rangle \leftarrow c\langle j\rangle + \langle i\rangle$ and S is the Type (II) operation $\langle j\rangle \leftarrow c\langle i\rangle + \langle j\rangle$. Then $R(\mathbf{I}_n)$ is the $n \times n$ matrix having 1's on the main diagonal and 0's everywhere else, except for the (i, j) entry, which has the value c. Similarly, $S(\mathbf{I}_n)$ has 1's all along the main diagonal and 0's everywhere else, except for the (j, i) entry, which has the value c. However, this exactly describes the matrix $(R(\mathbf{I}_n))^T$.

(5) (a) If $\mathbf{A}$ is nonsingular, $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ by part (4) of Theorem 2.12. For the converse, $\mathbf{A}^{-1} = ((\mathbf{A}^T)^{-1})^T$.

(b) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}| = |\mathbf{B}||\mathbf{A}| = |\mathbf{BA}|$.

(6) (a) Use the fact that $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

(b) $|\mathbf{AB}| = |-\mathbf{BA}| \implies |\mathbf{AB}| = (-1)^n|\mathbf{BA}|$ (by Corollary 3.4) $\implies |\mathbf{A}||\mathbf{B}| = (-1)|\mathbf{B}||\mathbf{A}|$ (since n is odd). Hence, either $|\mathbf{A}| = 0$ or $|\mathbf{B}| = 0$.

(7) (a) $|\mathbf{AA}^T| = |\mathbf{A}||\mathbf{A}^T| = |\mathbf{A}||\mathbf{A}| = |\mathbf{A}|^2 \geq 0$.

(b) $|\mathbf{AB}^T| = |\mathbf{A}||\mathbf{B}^T| = |\mathbf{A}^T||\mathbf{B}|$

(8) (a) $\mathbf{A}^T = -\mathbf{A}$, so $|\mathbf{A}| = |\mathbf{A}^T| = |-\mathbf{A}| = (-1)^n|\mathbf{A}|$ (by Corollary 3.4) $= (-1)|\mathbf{A}|$ (since n is odd). Hence $|\mathbf{A}| = 0$.

(b) Consider $\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

(9) (a) $\mathbf{I}_n^T = \mathbf{I}_n = \mathbf{I}_n^{-1}$.

(b) Consider $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(c) $|\mathbf{A}^T| = |\mathbf{A}^{-1}| \implies |\mathbf{A}| = 1/|\mathbf{A}|$ $\implies |\mathbf{A}|^2 = 1 \implies |\mathbf{A}| = \pm 1$.

(10) Note that

$$\mathbf{B} = \begin{bmatrix} 9 & 0 & -3 \\ 3 & 2 & -1 \\ -6 & 0 & 1 \end{bmatrix}$$

has a determinant equal to -18. Hence, if $\mathbf{B} = \mathbf{A}^2$, then $|\mathbf{A}|^2$ is negative, a contradiction.

(11) (a) Base Step: $k = 2$: $|\mathbf{A}_1\mathbf{A}_2| = |\mathbf{A}_1||\mathbf{A}_2|$ by Theorem 3.7.
Inductive Step: Assume $|\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k| = |\mathbf{A}_1||\mathbf{A}_2|\cdots|\mathbf{A}_k|$, for any $n \times n$ matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$. Prove $|\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k\mathbf{A}_{k+1}| = |\mathbf{A}_1||\mathbf{A}_2|\cdots|\mathbf{A}_k||\mathbf{A}_{k+1}|$, for any $n \times n$ matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k, \mathbf{A}_{k+1}$. But, $|\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k\mathbf{A}_{k+1}| = |(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)\mathbf{A}_{k+1}| = |\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k||\mathbf{A}_{k+1}|$ (by Theorem 3.7) $= |\mathbf{A}_1||\mathbf{A}_2|\cdots|\mathbf{A}_k||\mathbf{A}_{k+1}|$ by the inductive hypothesis.

(b) Use part (a) with $\mathbf{A}_i = \mathbf{A}$, for $1 \leq i \leq k$. Or, for an induction proof:
Base Step ($k = 1$): Obvious.
Inductive Step: $|\mathbf{A}^{k+1}| = |\mathbf{A}^k\mathbf{A}| = |\mathbf{A}^k||\mathbf{A}| = |\mathbf{A}|^k|\mathbf{A}| = |\mathbf{A}|^{k+1}$.

(c) Use part (b): $|\mathbf{A}|^k = |\mathbf{A}^k| = |\mathbf{O}_n| = 0$, and so $|\mathbf{A}| = 0$. Or, for an induction proof:
Base Step ($k = 1$): Obvious.
Inductive Step: Assume.$\mathbf{A}^{k+1} = \mathbf{O}_n$. If $\mathbf{A}$ is singular, then $|\mathbf{A}| = 0$, and we are done. If $\mathbf{A}$ is nonsingular, $\mathbf{A}^{k+1} = \mathbf{O}_n \implies \mathbf{A}^{k+1}\mathbf{A}^{-1} = \mathbf{O}_n\mathbf{A}^{-1} \implies \mathbf{A}^k = \mathbf{O}_n \implies |\mathbf{A}| = 0$ by the inductive hypothesis, which implies $\mathbf{A}$ is singular, a contradiction.

(12) (a) $|\mathbf{A}^n| = |\mathbf{A}|^n$. If $|\mathbf{A}| = 0$, then $|\mathbf{A}|^n = 0$ is not prime. Similarly, if $|\mathbf{A}| = \pm 1$, then $|\mathbf{A}|^n = \pm 1$ and is not prime. Otherwise, letting $t = |\mathbf{A}|$, we have $|t| > 1$, and $|\mathbf{A}|^n$ has $|t|$ as a positive proper divisor, for $n \geq 2$.

(b) $|\mathbf{A}|^n = |\mathbf{A}^n| = |\mathbf{I}| = 1$. Since $|\mathbf{A}|$ is an integer, $|\mathbf{A}| = \pm 1$. If n is odd, then $|\mathbf{A}| \neq -1$, so $|\mathbf{A}| = 1$.

(13) (a) Let $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$. Suppose $\mathbf{P}$ is an $n \times n$ matrix. Then $\mathbf{P}^{-1}$ is also an $n \times n$ matrix. Hence $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is not defined unless $\mathbf{A}$ is also an $n \times n$ matrix, and thus, the product $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is also $n \times n$.

(b) For example, consider

$$\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \mathbf{A} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -4 \\ 16 & 11 \end{bmatrix}, \quad \text{and,}$$
$$\mathbf{B} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}^{-1} \mathbf{A} \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 10 & -12 \\ 4 & -5 \end{bmatrix}.$$

(c) $\mathbf{A} = \mathbf{I}_n\mathbf{A}\mathbf{I}_n = \mathbf{I}_n^{-1}\mathbf{A}\mathbf{I}_n$.

(d) If $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, then $\mathbf{P}\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{P}^{-1} = \mathbf{P}\mathbf{B}\,\mathbf{P}^{-1} \Longrightarrow \mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$. Hence $(\mathbf{P}^{-1})^{-1}\mathbf{B}\mathbf{P}^{-1} = \mathbf{A}$.

(e) $\mathbf{A}$ similar to $\mathbf{B}$ implies there is a matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$. $\mathbf{B}$ similar to $\mathbf{C}$ implies there is a matrix $\mathbf{Q}$ such that $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}$. Hence, $\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \mathbf{P}^{-1}(\mathbf{Q}^{-1}\mathbf{C}\mathbf{Q})\mathbf{P} = (\mathbf{P}^{-1}\mathbf{Q}^{-1})\mathbf{C}(\mathbf{Q}\mathbf{P}) = (\mathbf{Q}\mathbf{P})^{-1}\mathbf{C}(\mathbf{Q}\mathbf{P})$, and so $\mathbf{A}$ is similar to $\mathbf{C}$.

(f) $\mathbf{A}$ similar to $\mathbf{I}_n$ implies there is a matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{P}^{-1}\mathbf{I}_n\mathbf{P}$. Hence, $\mathbf{A} = \mathbf{P}^{-1}\mathbf{P} = \mathbf{I}_n$.

(g) $|\mathbf{B}| = |\mathbf{P}^{-1}\mathbf{A}\mathbf{P}| = |\mathbf{P}^{-1}||\mathbf{A}||\mathbf{P}| = |\mathbf{P}^{-1}||\mathbf{P}||\mathbf{A}| = (1/|\mathbf{P}|)|\mathbf{P}||\mathbf{A}| = |\mathbf{A}|$.

(14) (a) Let R be the Type (III) row operation $\langle k\rangle \leftrightarrow \langle k-1\rangle$, let $\mathbf{A}$ be an $n \times n$ matrix and let $\mathbf{B} = R(\mathbf{A})$. Then, by part (3) of Theorem 3.3, $|\mathbf{B}| = (-1)|\mathbf{A}|$. Next, notice that the submatrix $\mathbf{A}_{(k-1)j} = \mathbf{B}_{kj}$ because the $(k-1)$st row of $\mathbf{A}$ becomes the kth row of $\mathbf{B}$, implying the same row is being eliminated from both matrices, and since all other rows maintain their original relative positions (notice the same column is being eliminated in both cases). Hence,

$$\begin{aligned}
&a_{(k-1)1}\mathcal{A}_{(k-1)1} + a_{(k-1)2}\mathcal{A}_{(k-1)2} + \cdots + a_{(k-1)n}\mathcal{A}_{(k-1)n} \\
&= b_{k1}(-1)^{(k-1)+1}|\mathbf{A}_{(k-1)1}| + b_{k2}(-1)^{(k-1)+2}|\mathbf{A}_{(k-1)2}| + \cdots + b_{kn}(-1)^{(k-1)+n}|\mathbf{A}_{(k-1)n}| \\
&\qquad \text{(because } a_{(k-1)j} = b_{kj} \text{ for } 1 \leq j \leq n) \\
&= b_{k1}(-1)^{k}|\mathbf{B}_{k1}| + b_{k2}(-1)^{k+1}|\mathbf{B}_{k2}| + \cdots + b_{kn}(-1)^{k+n-1}|\mathbf{B}_{kn}| \\
&= (-1)(b_{k1}(-1)^{k+1}|\mathbf{B}_{k1}| + b_{k2}(-1)^{k+2}|\mathbf{B}_{k2}| + \cdots + b_{kk}(-1)^{k+n}|\mathbf{B}_{kn}|) \\
&= (-1)|\mathbf{B}| \text{ (by applying part (1) of Theorem 3.11 along the } k\text{th row of } \mathbf{B}, \\
&\qquad \text{since we have assumed this result is valid for } i = k) \\
&= (-1)|R(\mathbf{A})| \\
&= (-1)(-1)|\mathbf{A}| \\
&= |\mathbf{A}|, \text{ finishing the proof.}
\end{aligned}$$

(b) The definition of the determinant is the Base Step for an induction proof on the row number, counting down from n to 1. Part (a) is the Inductive Step.

(15) (a) Case 1: Assume $1 \leq k < j$ and $1 \leq i < m$. Then the (i,k) entry of $(\mathbf{A}_{jm})^T$
$= (k,i)$ entry of $\mathbf{A}_{jm} = (k,i)$ entry of $\mathbf{A} = (i,k)$ entry of $\mathbf{A}^T = (i,k)$ entry of $(\mathbf{A}^T)_{mj}$.
Case 2: Assume $j \leq k < n$ and $1 \leq i < m$. Then the (i,k) entry of $(\mathbf{A}_{jm})^T$
$= (k,i)$ entry of $\mathbf{A}_{jm} = (k+1,i)$ entry of $\mathbf{A} = (i,k+1)$ entry of $\mathbf{A}^T = (i,k)$ entry of $(\mathbf{A}^T)_{mj}$.
Case 3: Assume $1 \leq k < j$ and $m \leq i < n$. Then the (i,k) entry of $(\mathbf{A}_{jm})^T$

$= (k,i)$ entry of $\mathbf{A}_{jm} = (k,i+1)$ entry of $\mathbf{A} = (i+1,k)$ entry of $\mathbf{A}^T = (i,k)$ entry of $(\mathbf{A}^T)_{mj}$.

Case 4: Assume $j \le k < n$ and $m \le i < n$. Then the (i,k) entry of $(\mathbf{A}_{jm})^T$

$= (k,i)$ entry of $\mathbf{A}_{jm} = (k+1,i+1)$ entry of $\mathbf{A} = (i+1,k+1)$ entry of $\mathbf{A}^T$

$= (i,k)$ entry of $(\mathbf{A}^T)_{mj}$.

Hence, the corresponding entries of $(\mathbf{A}_{jm})^T$ and $(\mathbf{A}^T)_{mj}$ are all equal, proving that the matrices themselves are equal.

(b) $a_{1j}\mathcal{A}_{1j} + a_{2j}\mathcal{A}_{2j} + \cdots + a_{nj}\mathcal{A}_{nj} = a_{1j}(-1)^{1+j}|\mathbf{A}_{1j}| + a_{2j}(-1)^{2+j}|\mathbf{A}_{2j}| + \cdots + a_{nj}(-1)^{n+j}|\mathbf{A}_{nj}|$

$= a_{1j}(-1)^{1+j}\left|(\mathbf{A}_{1j})^T\right| + a_{2j}(-1)^{2+j}\left|(\mathbf{A}_{2j})^T\right| + \cdots + a_{nj}(-1)^{n+j}\left|(\mathbf{A}_{nj})^T\right|$ (by Theorem 3.10)

$= a_{1j}(-1)^{j+1}\left|\left(\mathbf{A}^T\right)_{j1}\right| + a_{2j}(-1)^{j+2}\left|\left(\mathbf{A}^T\right)_{j2}\right| + \cdots + a_{nj}(-1)^{j+n}\left|\left(\mathbf{A}^T\right)_{jn}\right|$ (by part (a))

$= \left|\mathbf{A}^T\right|$ (using part (1) of Theorem 3.11 along the jth row of $\mathbf{A}^T$) $= |\mathbf{A}|$, (by Theorem 3.10).

(16) Let $\mathbf{A}$, i, j, and $\mathbf{B}$ be as given in the exercise and its hint. Then by Exercise 12 in Section 3.2, $|\mathbf{B}| = 0$, since its ith and jth rows are the same. Also, since every row of $\mathbf{A}$ equals the corresponding row of $\mathbf{B}$, with the exception of the jth row, the submatrices $\mathbf{A}_{jk}$ and $\mathbf{B}_{jk}$ are equal for $1 \le k \le n$. Hence, $\mathcal{A}_{jk} = \mathcal{B}_{jk}$ for $1 \le k \le n$. Now, computing the determinant of $\mathbf{B}$ using a cofactor expansion along the jth row (part (1) of Theorem 3.11) yields $0 = |\mathbf{B}| = b_{j1}\mathcal{B}_{j1} + b_{j2}\mathcal{B}_{j2} + \cdots + b_{jn}\mathcal{B}_{jn} = a_{i1}\mathcal{B}_{j1} + a_{i2}\mathcal{B}_{j2} + \cdots + a_{in}\mathcal{B}_{jn}$ (because the jth row of $\mathbf{B}$ equals the ith row of $\mathbf{A}$) $= a_{i1}\mathcal{A}_{j1} + a_{i2}\mathcal{A}_{j2} + \cdots + a_{in}\mathcal{A}_{jn}$, completing the proof.

(17) For $i \ne j$, the (i,j) entry of $\mathbf{A}\mathbf{B}^T$ is $a_{i1}\mathcal{A}_{j1} + a_{i2}\mathcal{A}_{j2} + \cdots + a_{in}\mathcal{A}_{jn}$. This equals 0 by Exercise 16. The (i,i) entry of $\mathbf{A}\mathbf{B}^T$ is $a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in}$, which equals $|\mathbf{A}|$, by part (1) of Theorem 3.11. Hence, $\mathbf{A}\mathbf{B}^T = (|\mathbf{A}|)\,\mathbf{I}_n$.

(18) Assume in what follows that $\mathbf{A}$ is an $n \times n$ matrix. Let $\mathbf{A}_i$ be the ith matrix (as defined in Theorem 3.12) for the matrix $\mathbf{A}$.

(a) Let $\mathbf{C} = R([\mathbf{A}|\mathbf{B}])$. We must show that for each i, $1 \le i \le n$, the matrix $\mathbf{C}_i$ (as defined in Theorem 3.12) for $\mathbf{C}$ is identical to $R(\mathbf{A}_i)$.

First we consider the columns of $\mathbf{C}_i$ other than the ith column. If $1 \le j \le n$ with $j \ne i$, then the jth column of $\mathbf{C}_i$ is the same as the jth column of $\mathbf{C} = R([\mathbf{A}|\mathbf{B}])$. But since the jth column of $[\mathbf{A}|\mathbf{B}]$ is identical to the jth column of $\mathbf{A}_i$, it follows that the jth column of $R([\mathbf{A}|\mathbf{B}])$ is identical to the jth column of $R(\mathbf{A}_i)$. Therefore, for $1 \le j \le n$ with $j \ne i$, the jth column of $\mathbf{C}_i$ is the same as the jth column of $R(\mathbf{A}_i)$.

Finally, we consider the ith column of $\mathbf{C}_i$. Now, the ith column of $\mathbf{C}_i$ is identical to the last column of $R([\mathbf{A}|\mathbf{B}]) = [R(\mathbf{A})|R(\mathbf{B})]$, which is $R(\mathbf{B})$. But since the ith column of $\mathbf{A}_i$ equals $\mathbf{B}$, it follows that the ith column of $R(\mathbf{A}_i) = R(\mathbf{B})$. Thus, the ith column of $\mathbf{C}_i$ is identical to the ith column of $R(\mathbf{A}_i)$.

Therefore, since $\mathbf{C}_i$ and $R(\mathbf{A}_i)$ agree in every column, we have $\mathbf{C}_i = R(\mathbf{A}_i)$.

(b) Consider each type of operation in turn. First, if R is the Type (I) operation $\langle k\rangle \leftarrow c\,\langle k\rangle$ for some $c \ne 0$, then by part (1) of Theorem 3.3,

$$\frac{|R(\mathbf{A}_i)|}{|R(\mathbf{A})|} = \frac{c|\mathbf{A}_i|}{c|\mathbf{A}|} = \frac{|\mathbf{A}_i|}{|\mathbf{A}|}.$$

If R is any Type (II) operation, then by part (2) of Theorem 3.3,

$$\frac{|R(\mathbf{A}_i)|}{|R(\mathbf{A})|} = \frac{|\mathbf{A}_i|}{|\mathbf{A}|}.$$

If R is any Type (III) operation, then

$$\frac{|R(\mathbf{A}_i)|}{|R(\mathbf{A})|} = \frac{(-1)|\mathbf{A}_i|}{(-1)|\mathbf{A}|} = \frac{|\mathbf{A}_i|}{|\mathbf{A}|}.$$

(c) In the special case when $n = 1$, we have $\mathbf{A} = [1]$. Let $\mathbf{B} = [b]$. Hence, $\mathbf{A}_1$ (as defined in Theorem 3.12) $= \mathbf{B} = [b]$. Then the formula in Theorem 3.12 gives

$$x_1 = \frac{|\mathbf{A}_1|}{|\mathbf{A}|} = \frac{b}{1} = b,$$

which is the correct solution to the equation $\mathbf{AX} = \mathbf{B}$ in this case. For the remainder of the proof, we assume $n > 1$.

If $\mathbf{A} = \mathbf{I}_n$, then the solution to the system is $\mathbf{X} = \mathbf{B}$. Therefore we must show that, for each i, $1 \le i \le n$, the formula given in Theorem 3.12 yields $x_i = b_i$ (the ith entry of $\mathbf{B}$). First, note that $|\mathbf{A}| = |\mathbf{I}_n| = 1$. Now, the ith matrix $\mathbf{A}_i$ (as defined in Theorem 3.12) for $\mathbf{A}$ is identical to $\mathbf{I}_n$ except in its ith column, which equals $\mathbf{B}$. Therefore, the ith row of $\mathbf{A}_i$ has b_i as its ith entry, and zeroes elsewhere. Thus, a cofactor expansion along the ith row of $\mathbf{A}_i$ yields

$$|\mathbf{A}_i| = b_i(-1)^{i+i}|\mathbf{I}_{n-1}| = b_i.$$

Hence, for each i, the formula in Theorem 3.12 produces

$$x_i = \frac{|\mathbf{A}_i|}{|\mathbf{A}|} = \frac{b_i}{1} = b_i,$$

completing the proof.

(d) Because $\mathbf{A}$ is nonsingular, $[\mathbf{A}|\mathbf{B}]$ row reduces to $[\mathbf{I}_n|\mathbf{X}]$, where $\mathbf{X}$ is the unique solution to the system. Let $[\mathbf{C}|\mathbf{D}]$ represent any intermediate augmented matrix during this row reduction process. Now, by part (a) and repeated use of part (b), the ratio

$$\frac{|\mathbf{C}_i|}{|\mathbf{C}|} \text{ is identical to the ratio } \frac{|\mathbf{A}_i|}{|\mathbf{A}|}$$

obtained from the original augmented matrix, for each i, $1 \le i \le n$. But part (c) proves that for the final augmented matrix, $[\mathbf{I}_n|\mathbf{X}]$, this common ratio gives the correct solution for x_i, for each i, $1 \le i \le n$. Since all of the systems corresponding to these intermediate matrices have the same unique solution, the formula in Theorem 3.12 gives the correct solution for the original system, $[\mathbf{A}|\mathbf{B}]$, as well. Thus, Cramer's Rule is validated.

(19) Suppose $\mathbf{A}$ is an $n \times n$ matrix with $|\mathbf{A}| = 0$. Then, by Theorem 3.10, $\left|\mathbf{A}^T\right| = 0$. Thus, by Exercise 16 in Section 3.2, there is an $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{A}^T\mathbf{C} = \mathbf{O}_n$. Taking the transpose of both sides of this equation yields $\mathbf{C}^T\mathbf{A} = \mathbf{O}_n^T = \mathbf{O}_n$. Letting $\mathbf{B} = \mathbf{C}^T$ completes the proof.

(20) (a) T (b) T (c) F (d) T (e) T (f) T

Section 3.4

(1) (a) $x^2 - 7x + 14$ (b) $x^3 - 6x^2 + 3x + 10$ (c) $x^3 - 8x^2 + 21x - 18$ (d) $x^3 - 8x^2 + 7x - 5$ (e) $x^4 - 3x^3 - 4x^2 + 12x$

(2) (a) $E_2 = \{b[1,1]\}$ (b) $E_2 = \{c[-1,1,0]\}$ (c) $E_{-1} = \{b[1,2,0] + c[0,0,1]\}$

(3) In the answers for this exercise, b, c, and d represent arbitrary scalars.

(a) $\lambda = 1$, $E_1 = \{a[1,0]\}$, algebraic multiplicity of $\lambda = 2$

(b) $\lambda_1 = 2$, $E_2 = \{b[1,0]\}$, algebraic multiplicity of $\lambda_1 = 1$;
$\lambda_2 = 3$, $E_3 = \{b[1,-1]\}$, algebraic multiplicity of $\lambda_2 = 1$

(c) $\lambda_1 = 1$, $E_1 = \{a[1,0,0]\}$, algebraic multiplicity of $\lambda_1 = 1$;
$\lambda_2 = 2$, $E_2 = \{b[0,1,0]\}$, algebraic multiplicity of $\lambda_2 = 1$;
$\lambda_3 = -5$, $E_{-5} = \{c[-\frac{1}{6}, \frac{3}{7}, 1]\} = \{c[-7,18,42]\}$, algebraic multiplicity of $\lambda_3 = 1$

(d) $\lambda_1 = 1$, $E_1 = \{b[3,1]\}$, algebraic multiplicity of $\lambda_1 = 1$;
$\lambda_2 = -1$, $E_{-1} = \{b[7,3]\}$, algebraic multiplicity of $\lambda_2 = 1$

(e) $\lambda_1 = 0$, $E_0 = \{c[1,3,2]\}$, algebraic multiplicity of $\lambda_1 = 1$;
$\lambda_2 = 2$, $E_2 = \{b[0,1,0] + c[1,0,1]\}$, algebraic multiplicity of $\lambda_2 = 2$

(f) $\lambda_1 = 13$, $E_{13} = \{c[4,1,3]\}$, algebraic multiplicity of $\lambda_1 = 1$;
$\lambda_2 = -13$, $E_{-13} = \{b[1,-4,0] + c[3,0,-4]\}$, algebraic multiplicity of $\lambda_2 = 2$

(g) $\lambda_1 = 1$, $E_1 = \{d[2,0,1,0]\}$, algebraic multiplicity of $\lambda_1 = 1$;
$\lambda_2 = -1$, $E_{-1} = \{d[0,2,-1,1]\}$, algebraic multiplicity of $\lambda_2 = 1$

(h) $\lambda_1 = 0$, $E_0 = \{c[-1,1,1,0] + d[0,-1,0,1]\}$, algebraic multiplicity of $\lambda_1 = 2$;
$\lambda_2 = -3$, $E_{-3} = \{d[-1,0,2,2]\}$, algebraic multiplicity of $\lambda_2 = 2$

(4) (a) $\mathbf{P} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix}$

(b) $\mathbf{P} = \begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$

(c) Not diagonalizable

(d) $\mathbf{P} = \begin{bmatrix} 6 & 1 & 1 \\ 2 & 2 & 1 \\ 5 & 1 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

(e) Not diagonalizable

(f) Not diagonalizable

(g) $\mathbf{P} = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 0 & -1 \\ 0 & 3 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

(h) $\mathbf{P} = \begin{bmatrix} 6 & 3 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(i) $\mathbf{P} = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 2 & 0 & 2 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$,

$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(5) (a) $\begin{bmatrix} 32770 & -65538 \\ 32769 & -65537 \end{bmatrix}$ (b) $\begin{bmatrix} -17 & 6 & 24 \\ -15 & 6 & 20 \\ -9 & 3 & 13 \end{bmatrix}$ (c) $\mathbf{A}^{49} = \mathbf{A}$

(d) $\begin{bmatrix} 352246 & -4096 & 354294 \\ -175099 & 2048 & -175099 \\ -175099 & 4096 & -177147 \end{bmatrix}$ (e) $\begin{bmatrix} 4188163 & 6282243 & -9421830 \\ 4192254 & 6288382 & -9432060 \\ 4190208 & 6285312 & -9426944 \end{bmatrix}$

(6) $\mathbf{A}$ similar to $\mathbf{B}$ implies there is a nonsingular matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{P}^{-1}\mathbf{BP}$. Hence,

$$\begin{aligned} p_{\mathbf{A}}(x) &= |x\mathbf{I}_n - \mathbf{A}| = |x\mathbf{I}_n - \mathbf{P}^{-1}\mathbf{BP}| = |x\mathbf{P}^{-1}\mathbf{I}_n\mathbf{P} - \mathbf{P}^{-1}\mathbf{BP}| \\ &= |\mathbf{P}^{-1}(x\mathbf{I}_n - \mathbf{B})\mathbf{P}| = |\mathbf{P}^{-1}||(x\mathbf{I}_n - \mathbf{B})||\mathbf{P}| = \frac{1}{|\mathbf{P}|}|(x\mathbf{I}_n - \mathbf{B})||\mathbf{P}| \\ &= \frac{1}{|\mathbf{P}|}|\mathbf{P}||(x\mathbf{I}_n - \mathbf{B})| = |(x\mathbf{I}_n - \mathbf{B})| = p_{\mathbf{B}}(x). \end{aligned}$$

(7) (a) Suppose $\mathbf{A} = \mathbf{PDP}^{-1}$ for some diagonal matrix $\mathbf{D}$. Let $\mathbf{C}$ be the diagonal matrix with $c_{ii} = (d_{ii})^{\frac{1}{3}}$. Then $\mathbf{C}^3 = \mathbf{D}$. Clearly then, $(\mathbf{PCP}^{-1})^3 = \mathbf{A}$.

(b) If $\mathbf{A}$ has all eigenvalues nonnegative, then $\mathbf{A}$ has a square root. Proceed as in part (a), only taking square roots instead of cube roots.

(8) Let $\mathbf{B} = \begin{bmatrix} 15 & -14 & -14 \\ -13 & 16 & 17 \\ 20 & -22 & -23 \end{bmatrix}$. We find a matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{B} = \mathbf{PDP}^{-1}$.

If $\mathbf{C}$ is the diagonal matrix with $c_{ii} = \sqrt[3]{d_{ii}}$, then $\mathbf{C}^3 = \mathbf{D}$. So, if $\mathbf{A} = \mathbf{PCP}^{-1}$, then $\mathbf{A}^3 = (\mathbf{PCP}^{-1})^3 = \mathbf{PC}^3\mathbf{P}^{-1} = \mathbf{PDP}^{-1} = \mathbf{B}$. Hence, to solve this problem, we first need to find $\mathbf{P}$ and $\mathbf{D}$. To do this, we diagonalize $\mathbf{B}$.

Step 1: $p_{\mathbf{B}}(x) = |x\mathbf{I}_3 - \mathbf{B}| = x^3 - 8x^2 - x + 8 = (x-1)(x+1)(x-8)$.

Step 2: The three eigenvalues of $\mathbf{B}$ are $\lambda_1 = 1$, $\lambda_2 = -1$ and $\lambda_3 = 8$.

Step 3: Now we solve for fundamental eigenvectors.

Eigenvalue $\lambda_1 = 1$: $[(1\mathbf{I}_3 - \mathbf{B})|\mathbf{0}]$ reduces to $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$. This yields the fundamental eigenvector $[-1, -2, 1]$.

Eigenvalue $\lambda_2 = -1$: $[((-1)\mathbf{I}_3 - \mathbf{B})|\mathbf{0}]$ reduces to $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$. This produces the fundamental eigenvector $[0, -1, 1]$.

Eigenvalue $\lambda_3 = 8$: $[(8\mathbf{I}_3 - \mathbf{B})|\mathbf{0}]$ reduces to $\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$. This yields the fundamental eigenvector $[2, -1, 2]$.

Step 4: Since $n = 3$, and we have found 3 fundamental eigenvectors, $\mathbf{B}$ is diagonalizable.

Step 5: $\mathbf{P} = \begin{bmatrix} -1 & 0 & 2 \\ -2 & -1 & -1 \\ 1 & 1 & 2 \end{bmatrix}$.

Step 6: $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix}$. Also, $\mathbf{P}^{-1} = \begin{bmatrix} 1 & -2 & -2 \\ -3 & 4 & 5 \\ 1 & -1 & -1 \end{bmatrix}$.

The diagonal matrix $\mathbf{C}$ whose main diagonal entries are the cube roots of the eigenvalues of $\mathbf{B}$. That is,

$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Finally, $\mathbf{A} = \mathbf{PCP}^{-1} = \begin{bmatrix} 3 & -2 & -2 \\ -7 & 10 & 11 \\ 8 & -10 & -11 \end{bmatrix}$. Direct calculation of $\mathbf{A}^3$ verifies that $\mathbf{A}^3 = \mathbf{B}$.

(9) The characteristic polynomial of the matrix is $x^2 - (a+d)x + (ad - bc)$, whose discriminant simplifies to $(a-d)^2 + 4bc$.

(10) (a) Let $\mathbf{v}$ be an eigenvector for $\mathbf{A}$ corresponding to λ. Then $\mathbf{Av} = \lambda\mathbf{v}$,

$$\mathbf{A}^2\mathbf{v} = \mathbf{A}(\mathbf{Av}) = \mathbf{A}(\lambda\mathbf{v}) = \lambda(\mathbf{Av}) = \lambda(\lambda\mathbf{v}) = \lambda^2\mathbf{v},$$

and an analogous induction argument shows $\mathbf{A}^k\mathbf{v} = \lambda^k\mathbf{v}$, for any integer $k \geq 1$.

(b) Consider the matrix $\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Although $\mathbf{A}$ has no eigenvalues, $\mathbf{A}^4 = \mathbf{I}_2$ has 1 as an eigenvalue.

(11) $(-x)^n|\mathbf{A}^{-1}|p_{\mathbf{A}}(\frac{1}{x}) = (-x)^n|\mathbf{A}^{-1}|\,|\frac{1}{x}\mathbf{I}_n - \mathbf{A}| = (-x)^n|\mathbf{A}^{-1}(\frac{1}{x}\mathbf{I}_n - \mathbf{A})|$
$= (-x)^n|\frac{1}{x}\mathbf{A}^{-1} - \mathbf{I}_n| = |(-x)(\frac{1}{x}\mathbf{A}^{-1} - \mathbf{I}_n)| = |x\mathbf{I}_n - \mathbf{A}^{-1}| = p_{\mathbf{A}^{-1}}(x)$.

(12) Both parts are true, because $x\mathbf{I}_n - \mathbf{A}$ is also upper triangular, and so
$p_{\mathbf{A}}(x) = |x\mathbf{I}_n - \mathbf{A}| = (x - a_{11})(x - a_{22})\cdots(x - a_{nn})$, by Theorem 3.2.

(13) $p_{\mathbf{A}^T}(x) = |x\mathbf{I}_n - \mathbf{A}^T| = |x\mathbf{I}_n^T - \mathbf{A}^T| = |(x\mathbf{I}_n - \mathbf{A})^T| = |x\mathbf{I}_n - \mathbf{A}| = p_{\mathbf{A}}(x)$

(14) $\mathbf{A}^T\mathbf{X} = \mathbf{X}$ since the entries of each row of $\mathbf{A}^T$ sum to 1. Hence $\lambda = 1$ is an eigenvalue for $\mathbf{A}^T$, and hence for $\mathbf{A}$, by Exercise 13.

(15) Base Step: $k = 1$. Use the argument in the text directly after the definition of similarity.
Inductive Step: Assume that if $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$, then for some $k > 0$, $\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1}$. We must prove that if $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$, then $\mathbf{A}^{k+1} = \mathbf{PD}^{k+1}\mathbf{P}^{-1}$. But $\mathbf{A}^{k+1} = \mathbf{AA}^k = (\mathbf{PDP}^{-1})(\mathbf{PD}^k\mathbf{P}^{-1})$ (by the inductive hypothesis) $= \mathbf{PD}^{k+1}\mathbf{P}^{-1}$.

(16) Since $\mathbf{A}$ is upper triangular, so is $x\mathbf{I}_n - \mathbf{A}$. Thus, by Theorem 3.2, $|x\mathbf{I}_n - \mathbf{A}|$ equals the product of the main diagonal entries of $x\mathbf{I}_n - \mathbf{A}$, which is $(x - a_{11})(x - a_{22})\cdots(x - a_{nn})$. Hence the main diagonal entries of $\mathbf{A}$ are the eigenvalues of $\mathbf{A}$. Thus, $\mathbf{A}$ has n distinct eigenvalues. Then, by the Diagonalization Method given in Section 3.4, the necessary matrix $\mathbf{P}$ can be constructed (since only one fundamental eigenvector for each eigenvalue is needed in this case), and so $\mathbf{A}$ is diagonalizable.

(17) Assume $\mathbf{A}$ is an $n \times n$ matrix. Then $\mathbf{A}$ is singular $\iff |\mathbf{A}| = 0 \iff (-1)^n|\mathbf{A}| = 0 \iff |-\mathbf{A}| = 0 \iff |0\mathbf{I}_n - \mathbf{A}| = 0 \iff p_{\mathbf{A}}(0) = 0 \iff \lambda = 0$ is an eigenvalue for $\mathbf{A}$.

(18) Let $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$. Then $\mathbf{D} = \mathbf{D}^T = (\mathbf{P}^{-1}\mathbf{AP})^T = \mathbf{P}^T\mathbf{A}^T(\mathbf{P}^{-1})^T = ((\mathbf{P}^T)^{-1})^{-1}\mathbf{A}^T(\mathbf{P}^T)^{-1}$, so $\mathbf{A}^T$ is diagonalizable.

(19) $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP} \implies \mathbf{D}^{-1} = (\mathbf{P}^{-1}\mathbf{AP})^{-1} = \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{P}$. But $\mathbf{D}^{-1}$ is a diagonal matrix whose main diagonal entries are the reciprocals of the main diagonal entries of $\mathbf{D}$. (Since the main diagonal entries of $\mathbf{D}$ are the nonzero eigenvalues of $\mathbf{A}$, their reciprocals exist.) Thus, $\mathbf{A}^{-1}$ is diagonalizable.

(20) (a) The ith column of $\mathbf{AP} = \mathbf{A}$(ith column of $\mathbf{P}$) $= \mathbf{AP}_i$. Since $\mathbf{P}_i$ is an eigenvector for λ_i, we have $\mathbf{AP}_i = \lambda_i\mathbf{P}_i$, for each i.

(b) $\mathbf{P}^{-1}\lambda_i\mathbf{P}_i = \lambda_i(\mathbf{P}^{-1}$($i$th column of $\mathbf{P}$)$) = \lambda_i$(ith column of $\mathbf{I}_n$) $= \lambda_i\mathbf{e}_i$.

(c) The ith column of $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{-1}$($i$th column of $\mathbf{AP}$) $= \mathbf{P}^{-1}(\lambda_i \mathbf{P}_i)$ (by part (a)) $= (\lambda_i \mathbf{e}_i)$ (by part (b)). Hence, $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix with main diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$.

(21) Since $\mathbf{PD} = \mathbf{AP}$, the ith column of $\mathbf{PD}$ equals the ith column of $\mathbf{AP}$. But,
ith column of $\mathbf{PD}$
$= \mathbf{P}$(ith column of $\mathbf{D}$)
$= \mathbf{P}(d_{ii}\mathbf{e}_i) = d_{ii}\mathbf{P}\mathbf{e}_i = d_{ii}(\mathbf{P}$($i$th column of $\mathbf{I}_n$))
$= d_{ii}$(ith column of $\mathbf{PI}_n$)
$= d_{ii}$(ith column of $\mathbf{P}$) $= d_{ii}\mathbf{P}_i$.
Also, the ith column of $\mathbf{AP} = \mathbf{A}$(ith column of $\mathbf{P}$) $= \mathbf{AP}_i$. Thus, d_{ii} is an eigenvalue for the ith column of $\mathbf{P}$.

(22) Use induction on n.
Base Step ($n = 1$): $|\mathbf{C}| = a_{11}x + b_{11}$, which has degree 1 if $k = 1$ (implying $a_{11} \neq 0$), and degree 0 if $k = 0$.
Inductive Step: Assume true for $(n-1)\times(n-1)$ matrices and prove for $n \times n$ matrices. Using a cofactor expansion along the first row yields

$$|\mathbf{C}| = (a_{11}x + b_{11})(-1)^2|\mathbf{C}_{11}| + (a_{12}x + b_{12})(-1)^3|\mathbf{C}_{12}| + \cdots + (a_{1n}x + b_{1n})(-1)^{n+1}|\mathbf{C}_{1n}|.$$

Case 1: Suppose the first row of $\mathbf{A}$ is all zeroes (so each $a_{1i} = 0$). Then, each of the submatrices $\mathbf{A}_{1i}$ has at most k nonzero rows. Hence, by the inductive hypothesis, each $|\mathbf{C}_{1i}| = |\mathbf{A}_{1i}x + \mathbf{B}_{1i}|$ is a polynomial of degree at most k. Therefore,

$$|\mathbf{C}| = b_{11}(-1)^2|\mathbf{C}_{11}| + b_{12}(-1)^3|\mathbf{C}_{12}| + \cdots + b_{1n}(-1)^{n+1}|\mathbf{C}_{1n}|$$

is a sum of constants multiplied by such polynomials, and hence is a polynomial of degree at most k.
Case 2: Suppose some $a_{1i} \neq 0$. Then, at most $k-1$ of the rows from 2 through n of $\mathbf{A}$ have a nonzero entry. Hence, each submatrix $\mathbf{A}_{1i}$ has at most $k-1$ nonzero rows, and, by the inductive hypothesis, each of $|\mathbf{C}_{1i}| = |\mathbf{A}_{1i}x + \mathbf{B}_{1i}|$ is a polynomial of degree at most $k-1$. Therefore, each term $(a_{1i}x + b_{1i})(-1)^{i+1}|\mathbf{C}_{1i}|$ is a polynomial of degree at most k. The desired result easily follows.

(23) (a) $|x\mathbf{I}_2 - \mathbf{A}| = \begin{vmatrix} x - a_{11} & -a_{12} \\ -a_{21} & x - a_{22} \end{vmatrix} = (x - a_{11})(x - a_{22}) - a_{12}a_{21}$
$= x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{21}) = x^2 - (\text{trace}(\mathbf{A}))x + |\mathbf{A}|$.

(b) Use induction on n.
Base Step ($n = 1$): Now $p_{\mathbf{A}}(x) = x - a_{11}$, which is a first degree polynomial with the coefficient of its x term equal to 1.
Inductive Step: Assume that the statement is true for all $(n-1)\times(n-1)$ matrices, and prove it is true for an $n \times n$ matrix $\mathbf{A}$. Consider

$$\mathbf{C} = x\mathbf{I}_n - \mathbf{A} = \begin{bmatrix} (x - a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (x - a_{22}) & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & (x - a_{nn}) \end{bmatrix}.$$

Using a cofactor expansion along the first row, $p_{\mathbf{A}}(x) = |\mathbf{C}| =$
$(x - a_{11})(-1)^{1+1}|\mathbf{C}_{11}| + (-a_{12})(-1)^{1+2}|\mathbf{C}_{12}| + (-a_{13})(-1)^{1+3}|\mathbf{C}_{13}| + \cdots + (-a_{1n})(-1)^{1+n}|\mathbf{C}_{1n}|$.
Now $|\mathbf{C}_{11}| = p_{\mathbf{A}_{11}}(x)$, and so, by the inductive hypothesis, is an $n-1$ degree polynomial with

coefficient of x^{n-1} equal to 1. Hence, $(x - a_{11})|\mathbf{C}_{11}|$ is a degree n polynomial having coefficient 1 for its x^n term. But, for $j \geq 2$, $\mathbf{C}_{1j}$ is an $(n-1) \times (n-1)$ matrix having exactly $n-2$ rows containing entries which are linear polynomials, since the linear polynomials at the $(1,1)$ and (j,j) entries of $\mathbf{C}$ have been eliminated. Therefore, by Exercise 22, $|\mathbf{C}_{1j}|$ is a polynomial of degree at most $n-2$. Summing to get $|\mathbf{C}|$ then gives an n degree polynomial with coefficient of the x^n term equal to 1, since only the term $(x - a_{11})|\mathbf{C}_{11}|$ in $|\mathbf{C}|$ contributes to the x^n term in $p_{\mathbf{A}}(x)$.

(c) Note that the constant term of $p_{\mathbf{A}}(x)$ is $p_{\mathbf{A}}(0) = |0\mathbf{I}_n - \mathbf{A}| = |-\mathbf{A}| = (-1)^n|\mathbf{A}|$.

(d) Use induction on n.
Base Step ($n = 1$): Now $p_{\mathbf{A}}(x) = x - a_{11}$. The degree 0 term is $-a_{11} = -\operatorname{trace}(\mathbf{A})$.
Inductive Step: Assume that the statement is true for all $(n-1) \times (n-1)$ matrices, and prove it is true for an $n \times n$ matrix $\mathbf{A}$. We know that $p_{\mathbf{A}}(x)$ is a polynomial of degree n from part (b). Then, let $\mathbf{C} = x\mathbf{I}_n - \mathbf{A}$ as in part (b). Then, arguing as in part (b), only the term $(x - a_{11})|\mathbf{C}_{11}|$ in $|\mathbf{C}|$ contributes to the x^n and x^{n-1} terms in $p_{\mathbf{A}}(x)$, because all of the terms of the form $\pm a_{1j}|\mathbf{C}_{1j}|$ are polynomials of degree $\leq (n-2)$, for $j \geq 2$ (using Exercise 22). Also, since $|\mathbf{C}_{11}| = p_{\mathbf{A}_{11}}(x)$, the inductive hypothesis and part (b) imply that

$$|\mathbf{C}_{11}| = x^{n-1} - \operatorname{trace}(\mathbf{A}_{11})x^{n-2} + (\text{terms of degree} \leq (n-3)).$$

Hence,

$$\begin{aligned}(x - a_{11})|\mathbf{C}_{11}| &= (x - a_{11})(x^{n-1} - \operatorname{trace}(\mathbf{A}_{11})x^{n-2} + (\text{terms of degree} \leq (n-3)))\\ &= x^n - a_{11}x^{n-1} - \operatorname{trace}(\mathbf{A}_{11})x^{n-1} + (\text{terms of degree} \leq (n-2))\\ &= x^n - \operatorname{trace}(\mathbf{A})x^{n-1} + (\text{terms of degree} \leq (n-2)),\end{aligned}$$

since $a_{11} + \operatorname{trace}(\mathbf{A}_{11}) = \operatorname{trace}(\mathbf{A})$. This completes the proof.

(24) (a) T (b) F (c) T (d) T (e) F (f) T (g) T (h) F

Chapter 3 Review Exercises

(1) (a) $(3,4)$ minor $= |\mathbf{A}_{34}| = -30$
(b) $(3,4)$ cofactor $= \mathcal{A}_{34} = -|\mathbf{A}_{34}| = 30$
(c) $|\mathbf{A}| = -830$
(d) $|\mathbf{A}| = -830$

(2) $|\mathbf{A}| = -262$

(3) $|\mathbf{A}| = -42$

(4) Volume $= 45$

(5) (a) $|\mathbf{B}| = 60$ (b) $|\mathbf{B}| = -15$ (c) $|\mathbf{B}| = 15$

(6) (a) Yes (b) 4 (c) Yes (d) Yes

(7) 378

(8) (a) $|\mathbf{A}| = 0$ (b) No. A nontrivial solution is $\begin{bmatrix} 1 \\ -8 \end{bmatrix}$.

(9) From Exercise 17 in Section 3.3, $\mathbf{AB}^T = |\mathbf{A}|\,\mathbf{I}_n$. Hence, $\mathbf{A}\left(\frac{1}{|\mathbf{A}|}\mathbf{B}^T\right) = \mathbf{I}_n$. Therefore, by Theorem 2.10, $\left(\frac{1}{|\mathbf{A}|}\mathbf{B}^T\right)\mathbf{A} = \mathbf{I}_n$, and so $\mathbf{B}^T\mathbf{A} = |\mathbf{A}|\,\mathbf{I}_n = \mathbf{AB}^T$.

(10) $x_1 = -4$, $x_2 = -3$, $x_3 = 5$

(11) (a) The determinant of the given matrix is -289. Thus, we would need $|\mathbf{A}|^4 = -289$. But no real number raised to the fourth power is negative.

(b) The determinant of the given matrix is zero, making the given matrix singular. Hence it can not be the inverse of any matrix.

(12) $\mathbf{B}$ similar to $\mathbf{A}$ implies there is a matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$.

(a) We will prove that $\mathbf{B}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}$ by induction on k.
Base Step: $k = 1$. $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$ is given.
Inductive Step: Assume that $\mathbf{B}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}$ for some k. Then

$$\begin{aligned}\mathbf{B}^{k+1} &= \mathbf{BB}^k = (\mathbf{P}^{-1}\mathbf{AP})(\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}) = \mathbf{P}^{-1}\mathbf{A}(\mathbf{PP}^{-1})\mathbf{A}^k\mathbf{P} \\ &= \mathbf{P}^{-1}\mathbf{A}(\mathbf{I}_n)\mathbf{A}^k\mathbf{P} = \mathbf{P}^{-1}\mathbf{AA}^k\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}^{k+1}\mathbf{P},\end{aligned}$$

so $\mathbf{B}^{k+1}$ is similar to $\mathbf{A}^{k+1}$.

(b) $|\mathbf{B}^T| = |\mathbf{B}| = |\mathbf{P}^{-1}\mathbf{AP}| = |\mathbf{P}^{-1}||\mathbf{A}||\mathbf{P}| = \frac{1}{|\mathbf{P}|}|\mathbf{A}||\mathbf{P}| = |\mathbf{A}| = |\mathbf{A}^T|$.

(c) If $\mathbf{A}$ is nonsingular,

$$\mathbf{B}^{-1} = (\mathbf{P}^{-1}\mathbf{AP})^{-1} = \mathbf{P}^{-1}\mathbf{A}^{-1}(\mathbf{P}^{-1})^{-1} = \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{P},$$

so $\mathbf{B}$ is nonsingular.

Note that $\mathbf{A} = \mathbf{PBP}^{-1}$. So, if $\mathbf{B}$ is nonsingular,

$$\mathbf{A}^{-1} = (\mathbf{PBP}^{-1})^{-1} = (\mathbf{P}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{P}^{-1} = \mathbf{PB}^{-1}\mathbf{P}^{-1},$$

so $\mathbf{A}$ is nonsingular.

(d) By part (c), $\mathbf{A}^{-1} = (\mathbf{P}^{-1})^{-1}\mathbf{B}^{-1}(\mathbf{P}^{-1})$, proving that $\mathbf{A}^{-1}$ is similar to $\mathbf{B}^{-1}$.

(e) $\mathbf{B} + \mathbf{I}_n = \mathbf{P}^{-1}\mathbf{AP} + \mathbf{I}_n = \mathbf{P}^{-1}\mathbf{AP} + \mathbf{P}^{-1}\mathbf{I}_n\mathbf{P} = \mathbf{P}^{-1}(\mathbf{A} + \mathbf{I}_n)\mathbf{P}$.

(f) Exercise 28(c) in Section 1.5 states that trace($\mathbf{AB}$) = trace($\mathbf{BA}$) for any two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$. In this particular case,

$$\begin{aligned}\text{trace}(\mathbf{B}) &= \text{trace}(\mathbf{P}^{-1}(\mathbf{AP})) = \text{trace}((\mathbf{AP})\mathbf{P}^{-1}) = \text{trace}(\mathbf{A}(\mathbf{PP}^{-1})) \\ &= \text{trace}(\mathbf{AI}_n) = \text{trace}(\mathbf{A}).\end{aligned}$$

(g) If $\mathbf{B}$ is diagonalizable, then there is a matrix $\mathbf{Q}$ such that $\mathbf{D} = \mathbf{Q}^{-1}\mathbf{BQ}$ is diagonal. Thus, $\mathbf{D} = \mathbf{Q}^{-1}(\mathbf{P}^{-1}\mathbf{AP})\mathbf{Q} = (\mathbf{Q}^{-1}\mathbf{P}^{-1})\mathbf{A}(\mathbf{PQ}) = (\mathbf{PQ})^{-1}\mathbf{A}(\mathbf{PQ})$, and so $\mathbf{A}$ is diagonalizable. Since $\mathbf{A}$ is similar to $\mathbf{B}$ if and only if $\mathbf{B}$ is similar to $\mathbf{A}$ (see Exercise 13(d) in Section 3.3), an analogous argument shows that $\mathbf{A}$ is diagonalizable $\Longrightarrow$ $\mathbf{B}$ is diagonalizable.

(13) $\mathbf{A}(c\mathbf{X} + d\mathbf{Y}) = c\mathbf{AX} + d\mathbf{AY} = c\lambda\mathbf{X} + d\lambda\mathbf{Y} = \lambda(c\mathbf{X} + d\mathbf{Y})$. This, and the given fact that $c\mathbf{X} + d\mathbf{Y}$ is nonzero, prove that $c\mathbf{X} + d\mathbf{Y}$ is an eigenvector for $\mathbf{A}$ corresponding to the eigenvalue λ.

(14) (a) $p_{\mathbf{A}}(x) = x^3 - x^2 - 10x - 8 = (x+2)(x+1)(x-4)$;
Eigenvalues: $\lambda_1 = -2$, $\lambda_2 = -1$, $\lambda_3 = 4$;
Eigenspaces: $E_{-2} = \{a[-1, 3, 3] \mid a \in \mathbb{R}\}$, $E_{-1} = \{a[-2, 7, 7] \mid a \in \mathbb{R}\}$,
$E_4 = \{a[-3, 10, 11] \mid a \in \mathbb{R}\}$;

$$\mathbf{P} = \begin{bmatrix} -1 & -2 & -3 \\ 3 & 7 & 10 \\ 3 & 7 & 11 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

(b) $p_{\mathbf{A}}(x) = x^3 + x^2 - 21x - 45 = (x+3)^2(x-5)$;
Eigenvalues: $\lambda_1 = 5$, $\lambda_2 = -3$;
Eigenspaces: $E_5 = \{a[-1, 4, 4] \mid a \in \mathbb{R}\}$; $E_{-3} = \{a[-2, 1, 0] + b[2, 0, 1] \mid a, b \in \mathbb{R}\}$;

$$\mathbf{P} = \begin{bmatrix} -1 & -2 & 2 \\ 4 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

(15) (a) $p_{\mathbf{A}}(x) = x^3 - 1 = (x-1)(x^2+x+1)$. $\lambda = 1$ is the only eigenvalue, having algebraic multiplicity 1. Thus, at most 1 fundamental eigenvector will be produced, which is insufficient for diagonalization by Step 4 of the Diagonalization Method.

(b) $p_{\mathbf{A}}(x) = x^4+6x^3+9x^2 = x^2(x+3)^2$. Even though the eigenvalue $\lambda = -3$ has algebraic multiplicity 2, only 1 fundamental eigenvector is produced for λ because $(-3\mathbf{I}_4 - \mathbf{A})$ has rank 3. In fact, we get only 3 fundamental eigenvectors overall, which is insufficient for diagonalization by Step 4 of the Diagonalization Method.

(16) $$\mathbf{A}^{13} = \begin{bmatrix} -9565941 & 9565942 & 4782976 \\ -12754588 & 12754589 & 6377300 \\ 3188648 & -3188648 & -1594325 \end{bmatrix}$$

(17) (a) $\lambda_1 = 2$, $\lambda_2 = -1$, $\lambda_3 = 3$

(b) $E_2 = \{a[1, -2, 1, 1] \mid a \in \mathbb{R}\}$, $E_{-1} = \{a[1, 0, 0, 1] + b[3, 7, -3, 2] \mid a, b \in \mathbb{R}\}$,
$E_3 = \{a[2, 8, -4, 3] \mid a \in \mathbb{R}\}$

(c) $|\mathbf{A}| = 6$

(18)

(a) F	(d) F	(g) T	(j) T	(m) T	(p) F	(s) T	(v) T	(y) F
(b) F	(e) T	(h) T	(k) F	(n) F	(q) F	(t) F	(w) T	
(c) F	(f) T	(i) F	(l) F	(o) F	(r) T	(u) F	(x) T	(z) F

Chapter 4

Section 4.1

(1) Property (2): $u \oplus (v \oplus w) = (u \oplus v) \oplus w$
Property (5): $a \odot (u \oplus v) = (a \odot u) \oplus (a \odot v)$
Property (6): $(a + b) \odot u = (a \odot u) \oplus (b \odot u)$
Property (7): $(ab) \odot u = a \odot (b \odot u)$

(2) Let $\mathcal{W} = \{a[1, 3, 2] \,|\, a \in \mathbb{R}\}$. First, $\mathcal{W}$ is closed under addition because

$$a[1, 3, 2] + b[1, 3, 2] = (a + b)[1, 3, 2],$$

which has the correct form. Similarly, $c(a[1, 3, 2]) = (ca)[1, 3, 2]$, proving closure under scalar multiplication. Now, Properties (1), (2), (5), (6), (7), and (8) are true for all vectors in $\mathbb{R}^3$ by Theorem 1.3, and hence are true in $\mathcal{W}$. Property (3) is satisfied because $[0, 0, 0] = 0[1, 3, 2] \in \mathcal{W}$, and

$$[0, 0, 0] + a[1, 3, 2] = a[1, 3, 2] + [0, 0, 0] = [0, 0, 0].$$

Finally, Property (4) follows from the equation $(a[1, 3, 2]) + ((-a)[1, 3, 2]) = [0, 0, 0]$.

(3) Let $\mathcal{W} = \{\mathbf{f} \in \mathcal{P}_3 \,|\, \mathbf{f}(2) = 0\}$. First, $\mathcal{W}$ is closed under addition since the sum of polynomials of degree ≤ 3 has degree ≤ 3, and because if $\mathbf{f}, \mathbf{g} \in \mathcal{W}$, then $(\mathbf{f} + \mathbf{g})(2) = \mathbf{f}(2) + \mathbf{g}(2) = 0 + 0 = 0$, and so $(\mathbf{f} + \mathbf{g}) \in \mathcal{W}$. Similarly, if $\mathbf{f} \in \mathcal{W}$ and $c \in \mathbb{R}$, then $c\mathbf{f}$ has the proper degree, and $(c\mathbf{f})(2) = c\mathbf{f}(2) = c0 = 0$, thus establishing closure under scalar multiplication. Now Properties (2), (5), (6), (7), and (8) are true for all real-valued functions on $\mathbb{R}$, as shown in Example 6 in Section 4.1 of the text. Property (1) holds because for every $x \in \mathbb{R}$, $\mathbf{f}(x)$ and $\mathbf{g}(x)$ are both real numbers, and so $\mathbf{f}(x) + \mathbf{g}(x) = \mathbf{g}(x) + \mathbf{f}(x)$ by the commutative law of addition for real numbers, which means $\mathbf{f} + \mathbf{g} = \mathbf{g} + \mathbf{f}$. Also, Property (3) holds because the zero function $\mathbf{z}$ is a polynomial of degree 0, and $\mathbf{z}(2) = 0$, so $\mathbf{z} \in \mathcal{W}$. Finally, Property (4) is true since $-\mathbf{f}$ is the additive inverse of $\mathbf{f}$, $-\mathbf{f}$ has the correct degree, and $(-\mathbf{f})(2) = -\mathbf{f}(2) = -0 = 0$, so $-\mathbf{f} \in \mathcal{W}$.

(4) Clearly the operation $\oplus$ is commutative. Also, $\oplus$ is associative because

$$\begin{aligned}(\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z} &= (x^3 + y^3)^{1/3} \oplus \mathbf{z} = (((x^3 + y^3)^{1/3})^3 + z^3)^{1/3} = ((x^3 + y^3) + z^3)^{1/3} \\ &= (x^3 + (y^3 + z^3))^{1/3} = (x^3 + ((y^3 + z^3)^{1/3})^3)^{1/3} = \mathbf{x} \oplus (y^3 + z^3)^{1/3} \\ &= \mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z}).\end{aligned}$$

Clearly, the real number 0 acts as the additive identity, and, for any real number x, the additive inverse of x is $-x$, because $(x^3 + (-x)^3)^{1/3} = 0^{1/3} = 0$, the additive identity.

The first distributive law holds because

$$\begin{aligned}a \odot (\mathbf{x} \oplus \mathbf{y}) &= a \odot (x^3 + y^3)^{1/3} = a^{1/3}(x^3 + y^3)^{1/3} = (a(x^3 + y^3))^{1/3} \\ &= (ax^3 + ay^3)^{1/3} = ((a^{1/3}x)^3 + (a^{1/3}y)^3)^{1/3} = (a^{1/3}x) \oplus (a^{1/3}y) \\ &= (a \odot \mathbf{x}) \oplus (a \odot \mathbf{y}).\end{aligned}$$

Similarly, the other distributive law holds because

$$\begin{aligned}(a + b) \odot \mathbf{x} &= (a + b)^{1/3}x = (a + b)^{1/3}(x^3)^{1/3} = (ax^3 + bx^3)^{1/3} \\ &= ((a^{1/3}x)^3 + (b^{1/3}x)^3)^{1/3} = (a^{1/3}x) \oplus (b^{1/3}x) \\ &= (a \odot \mathbf{x}) \oplus (b \odot \mathbf{x}).\end{aligned}$$

Associativity of scalar multiplication holds since

$$(ab) \odot \mathbf{x} = (ab)^{1/3}x = a^{1/3}b^{1/3}x = a^{1/3}(b^{1/3}x) = a \odot (b^{1/3}x) = a \odot (b \odot \mathbf{x}).$$

Finally, $1 \odot \mathbf{x} = 1^{1/3}x = 1x = \mathbf{x}$.

(5) The set is not closed under addition. For example, $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$.

(6) The set does not contain the zero vector.

(7) Property (8) is not satisfied. For example, $1 \odot 5 = 0 \neq 5$.

(8) Properties (2), (3), and (6) are not satisfied. Property (4) makes no sense without Property (3). The following is a counterexample for Property (2):

$$3 \oplus (4 \oplus 5) = 3 \oplus 18 = 42, \text{ but } (3 \oplus 4) \oplus 5 = 14 \oplus 5 = 38.$$

(9) Properties (3) and (6) are not satisfied. Property (4) makes no sense without Property (3). The following is a counterexample for Property (6):

$$(1+2) \odot [3,4] = 3 \odot [3,4] = [9,12], \text{ but } (1 \odot [3,4]) \oplus (2 \odot [3,4]) = [3,4] \oplus [6,8] = [9,0].$$

(10) Not a vector space. Property (6) is not satisfied. For example, $(1+2) \odot 3 = 3 \odot 3 = 9$, but

$$(1 \odot 3) \oplus (2 \odot 4) = 3 \oplus 8 = (3^5 + 8^5)^{1/5} = 33011^{1/5} \approx 8.01183.$$

(11) Suppose $\mathbf{0}_1$ and $\mathbf{0}_2$ are both zero vectors. Then, $\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2$.

(12) Zero vector $= [2, -3]$; additive inverse of $[x, y] = [4 - x, -6 - y]$

(13) (a) Add $-\mathbf{v}$ to both sides.

(b) $a\mathbf{v} = b\mathbf{v} \implies a\mathbf{v} + (-(b\mathbf{v})) = b\mathbf{v} + (-(b\mathbf{v})) \implies a\mathbf{v} + (-1)(b\mathbf{v}) = \mathbf{0}$ (by Theorem 4.1, part (3), and Property (4)) $\implies a\mathbf{v} + (-b)\mathbf{v} = \mathbf{0}$ (by Property (7)) $\implies (a-b)\mathbf{v} = \mathbf{0}$ (by Property (6)) $\implies a - b = 0$ (by Theorem 4.1, part (4)) $\implies a = b$.

(c) Multiply both sides by the scalar $\frac{1}{a}$, and use Property (7).

(14) For the Closure Properties, and Properties (2), (5), (6), (7), (8), mimic the steps in Example 6 in Section 4.1 in the text, restricting the proof to polynomial functions only. For Property (1), mimic the step for Property (1) in the answer to Exercise 3 above. Also, Property (3) holds because the zero polynomial $\mathbf{z}$ (of degree zero) in $\mathcal{P}$ serves as an additive identity since $\mathbf{z} + \mathbf{p} = \mathbf{p} + \mathbf{z} = \mathbf{p}$, for any polynomial $\mathbf{p}$. Finally, Property (4) is true since, for any polynomial $\mathbf{p}$, $-\mathbf{p}$ is a polynomial that has the property $\mathbf{p} + (-\mathbf{p}) = (-\mathbf{p}) + \mathbf{p} = \mathbf{z}$ (the additive identity), and so $-\mathbf{p} \in \mathcal{P}$ serves as the additive inverse of $\mathbf{p}$.

(15) For the Closure Properties, and Properties (2), (5), (6), (7), (8), mimic the steps in Example 6 in Section 4.1 in the text, generalizing the proofs to allow the domain of the functions to be the set X rather than $\mathbb{R}$. For Property (1), mimic the step for Property (1) in the answer to Exercise 3 above. Also, Property (3) holds because the zero function $\mathbf{z}$ with domain X serves as an additive identity since $\mathbf{z} + \mathbf{f} = \mathbf{f} + \mathbf{z} = \mathbf{f}$, for any function $\mathbf{f} \in \mathcal{V}$. Finally, Property (4) is true since, for any function $\mathbf{f} \in \mathcal{V}$, $-\mathbf{f}$ is a function with domain X that has the property $\mathbf{f} + (-\mathbf{f}) = (-\mathbf{f}) + \mathbf{f} = \mathbf{z}$ (the additive identity), and so $-\mathbf{f} \in \mathcal{V}$ serves as the additive inverse of $\mathbf{f}$.

(16) Base Step: $n = 1$. Then $a_1\mathbf{v}_1 \in \mathcal{V}$, since $\mathcal{V}$ is closed under scalar multiplication.
Inductive Step: Assume $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \in \mathcal{V}$ if $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{V}$, $a_1, \ldots, a_n \in \mathbb{R}$. Prove $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n + a_{n+1}\mathbf{v}_{n+1} \in \mathcal{V}$ if $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1} \in \mathcal{V}$, $a_1, \ldots, a_n, a_{n+1} \in \mathbb{R}$. But,

$$\begin{aligned} a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n + a_{n+1}\mathbf{v}_{n+1} &= (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) + a_{n+1}\mathbf{v}_{n+1} \\ &= \mathbf{w} + a_{n+1}\mathbf{v}_{n+1} \qquad \text{(by the inductive hypothesis).} \end{aligned}$$

Also, $a_{n+1}\mathbf{v}_{n+1} \in \mathcal{V}$ since $\mathcal{V}$ is closed under scalar multiplication. Finally, $\mathbf{w} + a_{n+1}\mathbf{v}_{n+1} \in \mathcal{V}$ since $\mathcal{V}$ is closed under addition.

(17) $0\mathbf{v} = 0\mathbf{v} + \mathbf{0} = 0\mathbf{v} + (0\mathbf{v} + (-(0\mathbf{v}))) = (0\mathbf{v} + 0\mathbf{v}) + (-(0\mathbf{v})) = (0+0)\mathbf{v} + (-(0\mathbf{v})) = 0\mathbf{v} + (-(0\mathbf{v})) = \mathbf{0}$.

(18) Let $\mathcal{V}$ be a nontrivial vector space, and let $\mathbf{v} \in \mathcal{V}$ with $\mathbf{v} \neq \mathbf{0}$. Then the set $S = \{a\mathbf{v} \mid a \in \mathbb{R}\}$ is a subset of $\mathcal{V}$, because $\mathcal{V}$ is closed under scalar multiplication. Also, if $a \neq b$ then $a\mathbf{v} \neq b\mathbf{v}$, since $a\mathbf{v} = b\mathbf{v} \Rightarrow a\mathbf{v} - b\mathbf{v} = \mathbf{0} \Rightarrow (a-b)\mathbf{v} = \mathbf{0} \Rightarrow a = b$ (by part (4) of Theorem 4.1). Hence, the elements of S are distinct, making S an infinite set, since $\mathbb{R}$ is infinite. Thus, $\mathcal{V}$ has an infinite number of elements because it has an infinite subset.

(19) (a) F (b) F (c) T (d) T (e) F (f) T (g) T

Section 4.2

(1) In what follows, $\mathcal{V}$ represents the given subset.

(a) Not a subspace; no zero vector.

(b) Not a subspace; not closed under either operation.
Counterexample for scalar multiplication: $2[1,1] = [2,2] \notin \mathcal{V}$.

(c) Subspace. Nonempty: $[0,0] \in \mathcal{V}$.
The vectors in $\mathcal{V}$ are precisely those of the form $a[1,2]$, where a is a scalar.
Addition: $\mathcal{V}$ is closed under addition because $a_1[1,2] + a_2[1,2] = (a_1 + a_2)[1,2]$, which has the required form since $a_1 + a_2$ is a scalar.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication because $k(a[1,2]) = (ka)[1,2]$, which has the required form since ka is a scalar.

(d) Not a subspace; not closed under addition.
Counterexample: $[1,0] + [0,1] = [1,1] \notin \mathcal{V}$.

(e) Not a subspace; no zero vector.

(f) Subspace. Nonempty: $[0,0] \in \mathcal{V}$.
The vectors in $\mathcal{V}$ are precisely those of the form $[a,0] = a[1,0]$.
Addition: $\mathcal{V}$ is closed under addition since $a_1[1,0]+a_2[1,0] = (a_1+a_2)[1,0]$, which has the required form since $a_1 + a_2$ is a scalar.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication since $c(a[1,0]) = (ca)[1,0]$, which has the required form since ca is a scalar.

(g) Not a subspace; not closed under addition.
Counterexample: $[1,1] + [1,-1] = [2,0] \notin \mathcal{V}$.

(h) Subspace. Nonempty: $[0,0] \in \mathcal{V}$ since $0 = -3(0)$.
The vectors in $\mathcal{V}$ are precisely those of the form $[a, -3a] = a[1, -3]$.
Addition: $\mathcal{V}$ is closed under addition since $a_1[1,-3] + a_2[1,-3] = (a_1 + a_2)[1,-3]$, which has the required form since $a_1 + a_2$ is a scalar.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication since $c(a[1,-3]) = (ca)[1,-3]$, which has the required form since ca is a scalar.

(i) Not a subspace; no zero vector since $0 \neq 7(0) - 5$.

(j) Not a subspace; not closed under either operation.
Counterexample for addition: $[1,1] + [2,4] = [3,5] \notin \mathcal{V}$, since $5 \neq 3^2$.

(k) Not a subspace; not closed under either operation.
Counterexample for scalar multiplication: $2[0,-4] = [0,-8] \notin \mathcal{V}$, since $-8 < 2(0) - 5$.

(l) Not a subspace; not closed under either operation.
Counterexample for scalar multiplication: $2[0.75, 0] = [1.5, 0] \notin \mathcal{V}$.

(2) In what follows, $\mathcal{V}$ represents the given subset.

(a) Subspace. Nonempty: $\mathbf{O}_{22} \in \mathcal{V}$.
The matrices in $\mathcal{V}$ are precisely those of the form $\begin{bmatrix} a & -a \\ b & 0 \end{bmatrix} = a\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.
Addition: $\mathcal{V}$ is closed under matrix addition since

$$\begin{aligned} &\left(a_1\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + b_1\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) + \left(a_2\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + b_2\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) \\ = \; &(a_1 + a_2)\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + (b_1 + b_2)\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

which is of the required form since $a_1 + a_2$ and $b_1 + b_2$ are scalars.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication because

$$c\left(a\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = (ca)\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} + (cb)\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

which is of the required form since ca and cb are scalars.

(b) Not a subspace; not closed under addition.
Counterexample: $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \notin \mathcal{V}$.

(c) Subspace. Nonempty: $\mathbf{O}_{22} \in \mathcal{V}$.
The matrices in $\mathcal{V}$ are precisely those of the form

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Addition: $\mathcal{V}$ is closed under matrix addition since

$$\begin{aligned} &\left(a_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b_1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c_1\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) + \left(a_2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b_2\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ = \; &(a_1 + a_2)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (b_1 + b_2)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + (c_1 + c_2)\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned}$$

which is of the required form, since $a_1 + a_2$, $b_1 + b_2$, and $c_1 + c_2$ are scalars.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication because

$$k\left(a\begin{bmatrix}1&0\\0&0\end{bmatrix}+b\begin{bmatrix}0&1\\1&0\end{bmatrix}+c\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)=(ka_1)\begin{bmatrix}1&0\\0&0\end{bmatrix}+(kb_1)\begin{bmatrix}0&1\\1&0\end{bmatrix}+(kc_1)\begin{bmatrix}0&0\\0&1\end{bmatrix},$$

which is of the required form, since ka_1, kb_1, and kc_1 are scalars.

(d) Not a subspace; no zero vector since $\mathbf{O}_{22}$ is singular.

(e) Subspace. Nonempty: $\mathbf{O}_{22} \in \mathcal{V}$.

A typical element of $\mathcal{V}$ has the form $\begin{bmatrix}a&b\\c&d\end{bmatrix}$, where $d = -(a+b+c)$. Therefore, the matrices in $\mathcal{V}$ are precisely those of the form

$$\begin{bmatrix}a&b\\c&-(a+b+c)\end{bmatrix}=a\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b\begin{bmatrix}0&1\\0&-1\end{bmatrix}+c\begin{bmatrix}0&0\\1&-1\end{bmatrix}.$$

Addition: $\mathcal{V}$ is closed under addition because

$$\begin{aligned}&\left(a_1\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b_1\begin{bmatrix}0&1\\0&-1\end{bmatrix}+c_1\begin{bmatrix}0&0\\1&-1\end{bmatrix}\right)\\&\qquad+\left(a_2\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b_2\begin{bmatrix}0&1\\0&-1\end{bmatrix}+c_2\begin{bmatrix}0&0\\1&-1\end{bmatrix}\right)\\=\ &(a_1+a_2)\begin{bmatrix}1&0\\0&-1\end{bmatrix}+(b_1+b_2)\begin{bmatrix}0&1\\0&-1\end{bmatrix}+(c_1+c_2)\begin{bmatrix}0&0\\1&-1\end{bmatrix},\end{aligned}$$

which is of the required form, since $a_1 + a_2$, $b_1 + b_2$, and $c_1 + c_2$ are scalars.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication since

$$\begin{aligned}&k\left(a\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b\begin{bmatrix}0&1\\0&-1\end{bmatrix}+c\begin{bmatrix}0&0\\1&-1\end{bmatrix}\right)\\=\ &(ka)\begin{bmatrix}1&0\\0&-1\end{bmatrix}+(kb)\begin{bmatrix}0&1\\0&-1\end{bmatrix}+(kc)\begin{bmatrix}0&0\\1&-1\end{bmatrix},\end{aligned}$$

which is of the required form, since ka, kb, and kc are scalars.

(f) Subspace. Nonempty: $\mathbf{O}_{22} \in \mathcal{V}$.

The matrices in $\mathcal{V}$ are precisely those of the form

$$\begin{bmatrix}a&b\\c&-a\end{bmatrix}=a\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b\begin{bmatrix}0&1\\0&0\end{bmatrix}+c\begin{bmatrix}0&0\\1&0\end{bmatrix}.$$

Addition: $\mathcal{V}$ is closed under addition since

$$\begin{aligned}&\left(a_1\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b_1\begin{bmatrix}0&1\\0&0\end{bmatrix}+c_1\begin{bmatrix}0&0\\1&0\end{bmatrix}\right)\\&\qquad+\left(a_2\begin{bmatrix}1&0\\0&-1\end{bmatrix}+b_2\begin{bmatrix}0&1\\0&0\end{bmatrix}+c_2\begin{bmatrix}0&0\\1&0\end{bmatrix}\right)\\=\ &(a_1+a_2)\begin{bmatrix}1&0\\0&-1\end{bmatrix}+(b_1+b_2)\begin{bmatrix}0&1\\0&0\end{bmatrix}+(c_1+c_2)\begin{bmatrix}0&0\\1&0\end{bmatrix},\end{aligned}$$

which is of the required form, since $a_1 + a_2$, $b_1 + b_2$, and $c_1 + c_2$ are scalars.
Scalar multiplication: $\mathcal{V}$ is closed under scalar multiplication since

$$\begin{aligned} & k\left(a\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) \\ = \; & (ka)\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + (kb)\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + (kc)\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

which is of the required form, since ka, kb, and kc are scalars.

(g) Subspace. Let $\mathbf{B} = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}$. Nonempty: $\mathbf{O}_{22} \in \mathcal{V}$ because $\mathbf{O}_{22}\mathbf{B} = \mathbf{O}_{22}$.

Addition: If $\mathbf{A}, \mathbf{C} \in \mathcal{V}$, then $(\mathbf{A} + \mathbf{C})\mathbf{B} = \mathbf{AB} + \mathbf{CB} = \mathbf{O}_{22} + \mathbf{O}_{22} = \mathbf{O}_{22}$. Hence, $(\mathbf{A} + \mathbf{C}) \in \mathcal{V}$.
Scalar multiplication: If $\mathbf{A} \in \mathcal{V}$, then $(c\mathbf{A})\mathbf{B} = c(\mathbf{AB}) = c\mathbf{O}_{22} = \mathbf{O}_{22}$, and so $(c\mathbf{A}) \in \mathcal{V}$.

(h) Not a subspace; not closed under addition.

Counterexample: $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \notin \mathcal{V}$.

(3) In what follows, $\mathcal{V}$ represents the given subset, and $\mathbf{z}$ represents the zero polynomial.

(a) Subspace. Nonempty: Clearly, $\mathbf{z} \in \mathcal{V}$.
Addition:

$$\begin{aligned} & (ax^5 + bx^4 + cx^3 + dx^2 + ax + e) + (rx^5 + sx^4 + tx^3 + ux^2 + rx + w) \\ = \; & (a + r)x^5 + (b + s)x^4 + (c + t)x^3 + (d + u)x^2 + (a + r)x + (e + w) \in \mathcal{V}; \end{aligned}$$

Scalar multiplication:

$$\begin{aligned} & t(ax^5 + bx^4 + cx^3 + dx^2 + ax + e) \\ = \; & (ta)x^5 + (tb)x^4 + (tc)x^3 + (td)x^2 + (ta)x + (te) \in \mathcal{V}. \end{aligned}$$

(b) Subspace. Nonempty: $\mathbf{z} \in \mathcal{V}$. Suppose $\mathbf{p}, \mathbf{q} \in \mathcal{V}$.
Addition: $(\mathbf{p} + \mathbf{q})(3) = \mathbf{p}(3) + \mathbf{q}(3) = 0 + 0 = 0$. Hence $(\mathbf{p} + \mathbf{q}) \in \mathcal{V}$.
Scalar multiplication: $(c\mathbf{p})(3) = c\mathbf{p}(3) = c(0) = 0$. Thus $(c\mathbf{p}) \in \mathcal{V}$.

(c) Subspace. Nonempty: $\mathbf{z} \in \mathcal{V}$.
Addition:

$$\begin{aligned} & (ax^5 + bx^4 + cx^3 + dx^2 + ex - (a + b + c + d + e)) \\ & \quad + (rx^5 + sx^4 + tx^3 + ux^2 + wx - (r + s + t + u + w)) \\ = \; & (a + r)x^5 + (b + s)x^4 + (c + t)x^3 + (d + u)x^2 + (e + w)x \\ & \quad + (-(a + b + c + d + e) - (r + s + t + u + w)) \\ = \; & (a + r)x^5 + (b + s)x^4 + (c + t)x^3 + (d + u)x^2 + (e + w)x \\ & \quad + (-((a + r) + (b + s) + (c + t) + (d + u) + (e + w))), \end{aligned}$$

which is in $\mathcal{V}$.
Scalar multiplication:

$$\begin{aligned} & t(ax^5 + bx^4 + cx^3 + dx^2 + ex - (a + b + c + d + e)) \\ = \; & (ta)x^5 + (tb)x^4 + (tc)x^3 + (td)x^2 + (te)x - ((ta) + (tb) + (tc) + (td) + (te)), \end{aligned}$$

which is in $\mathcal{V}$.

(d) Subspace. Nonempty: $\mathbf{z} \in \mathcal{V}$. Suppose $\mathbf{p}, \mathbf{q} \in \mathcal{V}$.
Addition: $(\mathbf{p}+\mathbf{q})(3) = \mathbf{p}(3) + \mathbf{q}(3) = \mathbf{p}(5) + \mathbf{q}(5) = (\mathbf{p}+\mathbf{q})(5)$. Hence $(\mathbf{p}+\mathbf{q}) \in \mathcal{V}$.
Scalar multiplication: $(c\mathbf{p})(3) = c\mathbf{p}(3) = c\mathbf{p}(5) = (c\mathbf{p})(5)$. Thus $(c\mathbf{p}) \in \mathcal{V}$.

(e) Not a subspace; $\mathbf{z} \notin \mathcal{V}$.

(f) Not a subspace; not closed under scalar multiplication.
Counterexample: $-x^2$ has a relative maximum at $x = 0$, but $(-1)(-x^2) = x^2$ does not (it has a relative minimum instead).

(g) Subspace. Nonempty: $\mathbf{z} \in \mathcal{V}$. Suppose $\mathbf{p}, \mathbf{q} \in \mathcal{V}$.
Addition: $(\mathbf{p}+\mathbf{q})'(4) = \mathbf{p}'(4) + \mathbf{q}'(4) = 0 + 0 = 0$. Hence $(\mathbf{p}+\mathbf{q}) \in \mathcal{V}$.
Scalar multiplication: $(c\mathbf{p})'(4) = c\mathbf{p}'(4) = c(0) = 0$. Thus $(c\mathbf{p}) \in \mathcal{V}$.

(h) Not a subspace; $\mathbf{z} \notin \mathcal{V}$.

(4) Let $\mathcal{V}$ be the given set of vectors. Setting $a = b = c = 0$ shows that $\mathbf{0} \in \mathcal{V}$. Hence $\mathcal{V}$ is nonempty. Note that the vectors in $\mathcal{V}$ have the form $[a, b, 0, c, a-2b+c] = a[1,1,0,0,1] + b[0,1,0,0,-2] + c[0,0,0,1,1]$. For closure under addition, note that

$$\begin{aligned} & (a_1[1,1,0,0,1] + b_1[0,1,0,0,-2] + c_1[0,0,0,1,1]) \\ & \qquad + (a_2[1,1,0,0,1] + b_2[0,1,0,0,-2] + c_2[0,0,0,1,1]) \\ = \; & (a_1+a_2)[1,1,0,0,1] + (b_1+b_2)[0,1,0,0,-2] + (c_1+c_2)[0,0,0,1,1], \end{aligned}$$

which is clearly another vector of this same form, since a_1+a_2, b_1+b_2, and c_1+c_2 are scalars. Similarly, for closure under scalar multiplication, note that

$$\begin{aligned} & k(a[1,1,0,0,1] + b[0,1,0,0,-2] + c[0,0,0,1,1]) \\ = \; & (ka)[1,1,0,0,1] + (kb)[0,1,0,0,-2] + (kc)[0,0,0,1,1], \end{aligned}$$

which also has the correct form since ka, kb, and kc are scalars. Thus $\mathcal{V}$ is a subspace by Theorem 4.2.

(5) Let $\mathcal{V}$ be the given set of vectors. Setting $a = b = c = 0$ shows that $\mathbf{0} \in \mathcal{V}$. Hence $\mathcal{V}$ is nonempty. Note that the vectors in $\mathcal{V}$ have the form $[2a-3b, a-5c, a, 4c-b, c] = a[2,1,1,0,1] + b[-3,0,0,-1,0] + c[0,-5,0,4,1]$. For closure under addition, note that

$$\begin{aligned} & (a_1[2,1,1,0,1] + b_1[-3,0,0,-1,0] + c_1[0,-5,0,4,1]) \\ & \qquad + (a_2[2,1,1,0,1] + b_2[-3,0,0,-1,0] + c_2[0,-5,0,4,1]) \\ = \; & (a_1+a_2)[2,1,1,0,1] + (b_1+b_2)[-3,0,0,-1,0] + (c_1+c_2)[0,-5,0,4,1], \end{aligned}$$

which is clearly another vector of this same form, since a_1+a_2, b_1+b_2, and c_1+c_2 are scalars. Similarly, for closure under scalar multiplication,

$$\begin{aligned} & k(a[2,1,1,0,1] + b[-3,0,0,-1,0] + c[0,-5,0,4,1]) \\ = \; & (ka)[2,1,1,0,1] + (kb)[-3,0,0,-1,0] + (kc)[0,-5,0,4,1], \end{aligned}$$

which also has the correct form since ka, kb, and kc are scalars. Thus $\mathcal{V}$ is a subspace by Theorem 4.2.

(6) (a) Let $\mathcal{V} = \{\mathbf{x} \in \mathbb{R}^3 \,|\, \mathbf{x} \cdot [1,-1,4] = 0\}$. Clearly $[0,0,0] \in \mathcal{V}$ since $[0,0,0] \cdot [1,-1,4] = 0$. Hence $\mathcal{V}$ is nonempty. Next, if $\mathbf{x}, \mathbf{y} \in \mathcal{V}$, then

$$(\mathbf{x}+\mathbf{y}) \cdot [1,-1,4] = \mathbf{x} \cdot [1,-1,4] + \mathbf{y} \cdot [1,-1,4] = 0 + 0 = 0.$$

Thus, $(\mathbf{x}+\mathbf{y}) \in \mathcal{V}$, and so $\mathcal{V}$ is closed under addition. Also, if $\mathbf{x} \in \mathcal{V}$ and $c \in \mathbb{R}$, then

$$(c\mathbf{x}) \cdot [1,-1,4] = c(\mathbf{x} \cdot [1,-1,4]) = c(0) = 0.$$

Therefore $(c\mathbf{x}) \in \mathcal{V}$, so $\mathcal{V}$ is closed under scalar multiplication. Hence, by Theorem 4.2, $\mathcal{V}$ is a subspace of $\mathbb{R}^3$.

(b) Plane, since it contains the nonparallel vectors $[0,4,1]$ and $[1,1,0]$.

(7) (a) The zero function $\mathbf{z}(x) = 0$ is continuous, and so the set is nonempty. Also, from calculus, we know that the sum of continuous functions is continuous, as is any scalar multiple of a continuous function. Hence, both closure properties hold. Apply Theorem 4.2.

(b) Replace the word "continuous" everywhere in part (a) with the word "differentiable."

(c) Let $\mathcal{V}$ be the given set. The zero function $\mathbf{z}(x) = 0$ satisfies $\mathbf{z}(\frac{1}{2}) = 0$, and so $\mathcal{V}$ is nonempty. Furthermore, if $\mathbf{f}, \mathbf{g} \in \mathcal{V}$ and $c \in \mathbb{R}$, then

$$(\mathbf{f}+\mathbf{g})(\frac{1}{2}) = \mathbf{f}(\frac{1}{2}) + \mathbf{g}(\frac{1}{2}) = 0 + 0 = 0,$$

and

$$(c\mathbf{f})(\frac{1}{2}) = c\mathbf{f}(\frac{1}{2}) = c0 = 0.$$

Hence, both closure properties hold. Now use Theorem 4.2.

(d) Let $\mathcal{V}$ be the given set. The zero function $\mathbf{z}(x) = 0$ is continuous and satisfies $\int_0^1 \mathbf{z}(x)\,dx = 0$, and so $\mathcal{V}$ is nonempty. Also, from calculus, we know that the sum of continuous functions is continuous, as is any scalar multiple of a continuous function. Furthermore, if $\mathbf{f}, \mathbf{g} \in \mathcal{V}$ and $c \in \mathbb{R}$, then

$$\int_0^1 (\mathbf{f}+\mathbf{g})(x)\,dx = \int_0^1 (\mathbf{f}(x)+\mathbf{g}(x))\,dx = \int_0^1 \mathbf{f}(x)\,dx + \int_0^1 \mathbf{g}(x)\,dx = 0+0 = 0,$$

and

$$\int_0^1 (c\mathbf{f})(x)\,dx = \int_0^1 c\mathbf{f}(x)\,dx = c\int_0^1 \mathbf{f}(x)\,dx = c0 = 0.$$

Hence, both closure properties hold. Finally, apply Theorem 4.2.

(8) The zero function $\mathbf{z}(x) = 0$ is differentiable and satisfies $3(d\mathbf{z}/dx) - 2\mathbf{z} = 0$, and so $\mathcal{V}$ is nonempty. Also, from calculus, we know that the sum of differentiable functions is differentiable, as is any scalar multiple of a differentiable function. Furthermore, if $\mathbf{f}, \mathbf{g} \in \mathcal{V}$ and $c \in \mathbb{R}$, then

$$3(\mathbf{f}+\mathbf{g})' - 2(\mathbf{f}+\mathbf{g}) = (3\mathbf{f}' - 2\mathbf{f}) + (3\mathbf{g}' - 2\mathbf{g}) = 0+0 = 0,$$

and

$$3(c\mathbf{f})' - 2(c\mathbf{f}) = 3c\mathbf{f}' - 2c\mathbf{f} = c(3\mathbf{f}' - 2\mathbf{f}) = c0 = 0.$$

Hence, both closure properties hold. Now use Theorem 4.2.

(9) Let $\mathcal{V}$ be the given set. The zero function $\mathbf{z}(x) = 0$ is twice-differentiable and satisfies $\mathbf{z}'' + 2\mathbf{z}' - 9\mathbf{z} = 0$, and so $\mathcal{V}$ is nonempty. From calculus, sums and scalar multiples of twice-differentiable functions are twice-differentiable. Furthermore, if $\mathbf{f}, \mathbf{g} \in \mathcal{V}$ and $c \in \mathbb{R}$, then

$$(\mathbf{f}+\mathbf{g})'' + 2(\mathbf{f}+\mathbf{g})' - 9(\mathbf{f}+\mathbf{g}) = (\mathbf{f}'' + 2\mathbf{f}' - 9\mathbf{f}) + (\mathbf{g}'' + 2\mathbf{g}' - 9\mathbf{g}) = 0+0 = 0,$$

and

$$(c\mathbf{f})'' + 2(c\mathbf{f})' - 9(c\mathbf{f}) = c\mathbf{f}'' + 2c\mathbf{f}' - 9c\mathbf{f} = c(\mathbf{f}'' + 2\mathbf{f}' - 9\mathbf{f}) = c0 = 0.$$

Hence, both closure properties hold. Theorem 4.2 finishes the proof.

(10) The given subset does not contain the zero function.

(11) First note that $\mathbf{AO} = \mathbf{OA}$, and so $\mathbf{O} \in \mathcal{W}$. Thus $\mathcal{W}$ is nonempty. Next, let $\mathbf{B}_1, \mathbf{B}_2 \in \mathcal{W}$. Then

$$\mathbf{A}(\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{AB}_1 + \mathbf{AB}_2 = \mathbf{B}_1\mathbf{A} + \mathbf{B}_2\mathbf{A} = (\mathbf{B}_1 + \mathbf{B}_2)\mathbf{A}$$

and

$$\mathbf{A}(c\mathbf{B}_1) = c(\mathbf{AB}_1) = c(\mathbf{B}_1\mathbf{A}) = (c\mathbf{B}_1)\mathbf{A}.$$

(12) (a) Closure under addition is a required property of every vector space.

(b) Let $\mathbf{A}$ be a singular $n \times n$ matrix. Then $|\mathbf{A}| = 0$. But then $|c\mathbf{A}| = c^n|\mathbf{A}| = 0$, so $c\mathbf{A}$ is also singular.

(c) All 8 properties hold.

(d) For $n = 2$, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(e) No; if $|\mathbf{A}| \neq 0$ and $c = 0$, then $|c\mathbf{A}| = |\mathbf{O}_n| = 0$.

(13) (a) Since the given line goes through the origin, it must have either the form $y = mx$, or $x = 0$. In the first case, all vectors on the line have the form $[x, mx] = x[1, m]$. Notice that $x_1[1, m] + x_2[1, m] = (x_1 + x_2)[1, m]$, which also lies on $y = mx$, and $k(x[1, m]) = (kx)[1, m]$, which also lies on $y = mx$. Thus, both closure properties hold in this case. On the other hand, if the line has the form $x = 0$, then all vectors on the line have the form $[0, y]$, and these vectors clearly form a subspace of $\mathbb{R}^2$.

(b) The given subset does not contain the zero vector.

(14) Let $\mathbf{A}$ be an $m \times n$ matrix, and let $\mathcal{V}$ be the set of solutions of the homogeneous system $\mathbf{AX} = \mathbf{0}$. Now, $\mathcal{V}$ is nonempty since the zero n-vector is a solution of $\mathbf{AX} = \mathbf{0}$. For closure under addition, notice that if $\mathbf{X}_1$ and $\mathbf{X}_2$ are solutions of $\mathbf{AX} = \mathbf{0}$, then $\mathbf{A}(\mathbf{X}_1 + \mathbf{X}_2) = \mathbf{AX}_1 + \mathbf{AX}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$, so $\mathbf{X}_1 + \mathbf{X}_2$ is also a solution of $\mathbf{AX} = \mathbf{0}$. For closure under scalar multiplication, notice that if $\mathbf{X}_1$ is a solution of $\mathbf{AX} = \mathbf{0}$, then $\mathbf{A}(k\mathbf{X}_1) = k(\mathbf{AX}_1) = k\mathbf{0} = \mathbf{0}$, so $k\mathbf{X}_1$ is also a solution of $\mathbf{AX} = \mathbf{0}$. Thus, by Theorem 4.2, $\mathcal{V}$ is a subspace of $\mathbb{R}^n$.

(15) $S = \{\mathbf{0}\}$, the trivial subspace of $\mathbb{R}^n$.

(16) Let $\mathbf{a} \in \mathcal{V}$ with $\mathbf{a} \neq 0$, and let $\mathbf{b} \in \mathbb{R}$. Then $(\frac{b}{a})\mathbf{a} = \mathbf{b} \in \mathcal{V}$. (We have used bold-faced variables when we are considering objects as vectors rather than scalars. However, $\mathbf{a}$ and a have the same value as real numbers; similarly for $\mathbf{b}$ and b.)

(17) The given subset does not contain the zero vector.

(18) Note that $\mathbf{0} \in \mathcal{W}_1$ and $\mathbf{0} \in \mathcal{W}_2$. Hence $\mathbf{0} \in \mathcal{W}_1 \cap \mathcal{W}_2$, and so $\mathcal{W}_1 \cap \mathcal{W}_2$ is nonempty. Let $\mathbf{x}, \mathbf{y} \in \mathcal{W}_1 \cap \mathcal{W}_2$. Then $\mathbf{x}+\mathbf{y} \in \mathcal{W}_1$ since $\mathcal{W}_1$ is a subspace, and $\mathbf{x}+\mathbf{y} \in \mathcal{W}_2$ since $\mathcal{W}_2$ is a subspace. Hence, $\mathbf{x}+\mathbf{y} \in \mathcal{W}_1 \cap \mathcal{W}_2$. Similarly, if $c \in \mathbb{R}$, then $c\mathbf{x} \in \mathcal{W}_1$ and $c\mathbf{x} \in \mathcal{W}_2$ since $\mathcal{W}_1$ and $\mathcal{W}_2$ are subspaces. Thus $c\mathbf{x} \in \mathcal{W}_1 \cap \mathcal{W}_2$. Now apply Theorem 4.2.

(19) If $\mathcal{W}$ is a subspace, then $a\mathbf{w}_1, b\mathbf{w}_2 \in \mathcal{W}$ by closure under scalar multiplication.
Hence $a\mathbf{w}_1 + b\mathbf{w}_2 \in \mathcal{W}$ by closure under addition.
Conversely, setting $b = 0$ shows that $a\mathbf{w}_1 \in \mathcal{W}$ for all $\mathbf{w}_1 \in \mathcal{W}$. This establishes closure under scalar multiplication. Use $a = 1$ and $b = 1$ to get closure under addition.

(20) (a) Suppose that $\mathcal{W}$ is a subspace of a vector space $\mathcal{V}$. We give a proof by induction on n.

Base Step: If $n = 1$, then we must show that if $\mathbf{v}_1 \in \mathcal{W}$ and a_1 is a scalar, then $a_1\mathbf{v}_1 \in \mathcal{W}$. But this is certainly true since the subspace $\mathcal{W}$ is closed under scalar multiplication.

Inductive Step: Assume that the theorem is true for any linear combination of n vectors in $\mathcal{W}$. We must prove the theorem holds for a linear combination of $n+1$ vectors. Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$ are vectors in $\mathcal{W}$, and $a_1, a_2, \ldots, a_n, a_{n+1}$ are scalars. We must show that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n + a_{n+1}\mathbf{v}_{n+1} \in \mathcal{W}$. However, by the inductive hypothesis, we know that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \in \mathcal{W}$. Also, $a_{n+1}\mathbf{v}_{n+1} \in \mathcal{W}$, since $\mathcal{W}$ is closed under scalar multiplication. But since $\mathcal{W}$ is also closed under addition, the sum of any two vectors in $\mathcal{W}$ is again in $\mathcal{W}$, so $(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) + (a_{n+1}\mathbf{v}_{n+1}) \in \mathcal{W}$.

(b) Suppose $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}$. Then $\mathbf{w}_1 + \mathbf{w}_2 = 1\mathbf{w}_1 + 1\mathbf{w}_2$ is a finite linear combination of vectors of $\mathcal{W}$, so $\mathbf{w}_1 + \mathbf{w}_2 \in \mathcal{W}$. Also, $c_1\mathbf{w}_1$ (for $c_1 \in \mathbb{R}$) is a finite linear combination in $\mathcal{W}$, so $c_1\mathbf{w}_1 \in \mathcal{W}$. Now apply Theorem 4.2.

(21) Apply Theorems 4.4 and 4.3.

(22) (a) F (b) T (c) F (d) T (e) T (f) F (g) T (h) T

Section 4.3

(1) (a) $\{[a, b, -a+b] \mid a, b \in \mathbb{R}\}$

(b) $\{a[1, \frac{1}{3}, -\frac{2}{3}] \mid a \in \mathbb{R}\}$

(c) $\{[a, b, -b] \mid a, b \in \mathbb{R}\}$

(d) $\text{Span}(S) = \mathbb{R}^3$

(e) $\{[a, b, c, -2a+b+c] \mid a, b, c \in \mathbb{R}\}$

(f) $\{[a, b, 2a+b, a+b] \mid a, b \in \mathbb{R}\}$

(2) (a) $\{ax^3 + bx^2 + cx - (a+b+c) \mid a, b, c \in \mathbb{R}\}$

(b) $\{ax^3 + bx^2 + d \mid a, b, d \in \mathbb{R}\}$

(c) $\{ax^3 - ax + b \mid a, b \in \mathbb{R}\}$

(3) (a) $\left\{ \left[\begin{array}{cc} a & b \\ c & -a-b-c \end{array} \right] \middle| \, a, b, c \in \mathbb{R} \right\}$

(b) $\left\{ \left[\begin{array}{cc} a & b \\ a-b & -8a+3b \end{array} \right] \middle| \, a, b \in \mathbb{R} \right\}$

(c) $\left\{ \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \middle| \, a, b, c, d \in \mathbb{R} \right\} = \mathcal{M}_{22}$

(4) (a) $\mathcal{W} = \text{row space of } \mathbf{A} = \text{row space of } \left[\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{array} \right]$

$= \{a[1, 1, 0, 0] + b[1, 0, 1, 0] + c[0, 1, 1, 1] \mid a, b, c \in \mathbb{R}\}$

(b) $\mathbf{B} = \left[\begin{array}{cccc} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \end{array} \right]$

(c) Row space of $\mathbf{B} = \{a[1, 0, 0, -\frac{1}{2}] + b[0, 1, 0, \frac{1}{2}] + c[0, 0, 1, \frac{1}{2}] \mid a, b, c \in \mathbb{R}\}$
$= \{[a, b, c, -\frac{1}{2}a + \frac{1}{2}b + \frac{1}{2}c] \mid a, b, c \in \mathbb{R}\}$

(5) (a) $\mathcal{W}$ = row space of $\mathbf{A}$ = row space of $\begin{bmatrix} 2 & 1 & 0 & 3 & 4 \\ 3 & 1 & -1 & 4 & 2 \\ -4 & -1 & 7 & 0 & 0 \end{bmatrix}$

$= \{a[2,1,0,3,4] + b[3,1,-1,4,2] + c[-4,-1,7,0,0]\}$

(b) $\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 2 & -2 \\ 0 & 1 & 0 & -1 & 8 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$

(c) Row space of $\mathbf{B} = \{[a, b, c, 2a - b + c, -2a + 8b] \mid a, b, c \in \mathbb{R}\}$

(6) S spans $\mathbb{R}^3$ by the Simplified Span Method since

$$\begin{bmatrix} 1 & 3 & -1 \\ 2 & 7 & -3 \\ 4 & 8 & -7 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(7) S does not span $\mathbb{R}^3$ by the Simplified Span Method since

$$\begin{bmatrix} 1 & -2 & 2 \\ 3 & -4 & -1 \\ 1 & -4 & 9 \\ 0 & 2 & -7 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & -\frac{7}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, span(S) $= \{[5c, \frac{7}{2}c, c]\}$, so, for example, $[0, 0, 1]$ is not in span(S).

(8) $ax^2 + bx + c = a(x^2 + x + 1) + (b - a)(x + 1) + (c - b)1$.

(9) The set does not span $\mathcal{P}_2$ by the Simplified Span Method since

$$\begin{bmatrix} 1 & 4 & -3 \\ 2 & 1 & 5 \\ 0 & 7 & -11 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & \frac{23}{7} \\ 0 & 1 & -\frac{11}{7} \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, span(S) $= \{-\frac{23}{7}cx^2 + \frac{11}{7}cx + c\}$, so, for example, $0x^2 + 0x + 1$ is not in the span of the set.

(10) (a) $[-4, 5, -13] = 5[1, -2, -2] - 3[3, -5, 1] + 0[-1, 1, -5]$.

(b) S does not span $\mathbb{R}^3$ by the Simplified Span Method since

$$\begin{bmatrix} 1 & -2 & -2 \\ 3 & -5 & 1 \\ -1 & 1 & -5 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & 12 \\ 0 & 1 & 7 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, span(S) $= \{[-12c, -7c, c]\}$, so, for example, $[0, 0, 1]$ is not in span(S).

(11) One answer is

$$-1(x^3 - 2x^2 + x - 3) + 2(2x^3 - 3x^2 + 2x + 5) - 1(4x^2 + x - 3) + 0(4x^3 - 7x^2 + 4x - 1).$$

(12) Let $S_1 = \{[1, 1, 0, 0], [1, 0, 1, 0], [1, 0, 0, 1], [0, 0, 1, 1]\}$. The Simplified Span Method shows that S_1 spans $\mathbb{R}^4$. Thus, since $S_1 \subseteq S$, S also spans $\mathbb{R}^4$.

(13) Let $\mathbf{b} \in \mathbb{R}$. Then $\mathbf{b} = (\frac{b}{a})\mathbf{a} \in \text{span}(\{\mathbf{a}\})$. (We have used bold-faced variables when we are considering objects as vectors rather than scalars. However, $\mathbf{a}$ and a have the same value as real numbers; similarly for $\mathbf{b}$ and b.)

(14) Apply the Simplified Span Method to S. The matrix $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ -12 & 3 & -2 & 3 \end{bmatrix}$ represents the vectors in S. Then $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is the corresponding reduced row echelon form matrix, which clearly indicates the desired result.

(15) (a) $\{[-3, 2, 0], [4, 0, 5]\}$

(b) $E_1 = \{[-\frac{3}{2}b + \frac{4}{5}c, b, c]\} = \{b[-\frac{3}{2}, 1, 0] + c[\frac{4}{5}, 0, 1]\} = \{\frac{b}{2}[-3, 2, 0] + \frac{c}{5}[4, 0, 5]\}$, so the set in part (a) spans E_1.

(16) $\text{Span}(S_1) \subseteq \text{span}(S_2)$, since $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = (-a_1)(-\mathbf{v}_1) + \cdots + (-a_n)(-\mathbf{v}_n)$.
$\text{Span}(S_2) \subseteq \text{span}(S_1)$, since $a_1(-\mathbf{v}_1) + \cdots + a_n(-\mathbf{v}_n) = (-a_1)\mathbf{v}_1 + \cdots + (-a_n)\mathbf{v}_n$.

(17) If $\mathbf{u} = a\mathbf{v}$, then all vectors in span(S) are scalar multiples of $\mathbf{v} \neq 0$. Hence, span(S) is a line through the origin. If $\mathbf{u} \neq a\mathbf{v}$, then span($S$) is the set of all linear combinations of the vectors $\mathbf{u}, \mathbf{v}$, and since these vectors point in different directions, span(S) is a plane through the origin.

(18) First, suppose that $\text{span}(S) = \mathbb{R}^3$. Let $\mathbf{x}$ be a solution to $\mathbf{Ax} = \mathbf{0}$. Then $\mathbf{u}\cdot\mathbf{x} = \mathbf{v}\cdot\mathbf{x} = \mathbf{w}\cdot\mathbf{x} = 0$, since these are the three entries of $\mathbf{Ax}$. Because $\text{span}(S) = \mathbb{R}^3$, there exist a, b, c such that $a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{x}$. Hence, $\mathbf{x}\cdot\mathbf{x} = (a\mathbf{u} + b\mathbf{v} + c\mathbf{w})\cdot\mathbf{x} = a\mathbf{u}\cdot\mathbf{x} + b\mathbf{v}\cdot\mathbf{x} + c\mathbf{w}\cdot\mathbf{x} = 0$. Therefore, by part (3) of Theorem 1.5, $\mathbf{x} = \mathbf{0}$. Thus, $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution. Theorem 2.7 and Corollary 3.6 then show that $|\mathbf{A}| \neq 0$.

Next, suppose that $|\mathbf{A}| \neq 0$. Then Corollary 3.6 implies that $\text{rank}(\mathbf{A}) = 3$. This means that the reduced row echelon form for $\mathbf{A}$ has three nonzero rows. This can only occur if $\mathbf{A}$ row reduces to $\mathbf{I}_3$. Thus, $\mathbf{A}$ is row equivalent to $\mathbf{I}_3$. By Theorem 2.9, $\mathbf{A}$ and $\mathbf{I}_3$ have the same row space. Hence, span(S) = row space of $\mathbf{A}$ = row space of $\mathbf{I}_3 = \mathbb{R}^3$.

(19) Choose $n = \max(\text{degree}(\mathbf{p}_1), \ldots, \text{degree}(\mathbf{p}_k))$.

(20) This follows immediately from Theorem 1.15.

(21) Consider the set $\{\Psi_{ij}\}$ of $n \times n$ matrices, where each Ψ_{ij} has 1 as its (i, j) entry, and every remaining entry equal to 0. Any $n \times n$ matrix $\mathbf{A}$ can be expressed as a linear combination of the matrices in $\{\Psi_{ij}\}$ by letting the coefficient of Ψ_{ij} be a_{ij}, the (i, j) entry of $\mathbf{A}$. Therefore, the set $\{\Psi_{ij}\}$ spans $\mathcal{M}_{nn}$. But since each Ψ_{ij} is upper triangular or lower triangular (or both), $\mathcal{M}_{nn}$ is spanned by $\mathcal{U}_n \cup \mathcal{L}_n$.

(22) $S_1 \subseteq S_2 \subseteq \text{span}(S_2)$, by Theorem 4.5, part (1). Then, since span(S_2) is a subspace of $\mathcal{V}$ containing S_1 (by Theorem 4.5, part (2)), $\text{span}(S_1) \subseteq \text{span}(S_2)$ (by Theorem 4.5, part (3)).

(23) (a) Suppose S is a subspace. Then $S \subseteq \text{span}(S)$ by Theorem 4.5, part (1). Also, S is a subspace containing S, so $\text{span}(S) \subseteq S$ by Theorem 4.5, part (3). Hence, $S = \text{span}(S)$.

Conversely, if $\text{span}(S) = S$, then S is a subspace by Theorem 4.5, part (2).

(b) span({skew-symmetric 3×3 matrices}) = {skew-symmetric 3×3 matrices}

(24) If $\text{span}(S_1) = \text{span}(S_2)$, then $S_1 \subseteq \text{span}(S_1) = \text{span}(S_2)$ and $S_2 \subseteq \text{span}(S_2) = \text{span}(S_1)$.

Conversely, since $\text{span}(S_1)$ and $\text{span}(S_2)$ are subspaces, part (3) of Theorem 4.5 shows that if $S_1 \subseteq \text{span}(S_2)$ and $S_2 \subseteq \text{span}(S_1)$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$ and $\text{span}(S_2) \subseteq \text{span}(S_1)$, and so $\text{span}(S_1) = \text{span}(S_2)$.

(25) (a) Clearly $S_1 \cap S_2 \subseteq S_1$ and $S_1 \cap S_2 \subseteq S_2$. Corollary 4.6 then implies $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_1)$ and $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_2)$. The desired result easily follows.

(b) $S_1 = \{[1,0,0],[0,1,0]\}$, $S_2 = \{[0,1,0],[0,0,1]\}$

(c) $S_1 = \{[1,0,0],[0,1,0]\}$, $S_2 = \{[1,0,0],[1,1,0]\}$

(26) (a) Clearly $S_1 \subseteq S_1 \cup S_2$ and $S_2 \subseteq S_1 \cup S_2$. Corollary 4.6 then implies $\text{span}(S_1) \subseteq \text{span}(S_1 \cup S_2)$ and $\text{span}(S_2) \subseteq \text{span}(S_1 \cup S_2)$. The desired result easily follows.

(b) If $S_1 \subseteq S_2$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$, so $\text{span}(S_1) \cup \text{span}(S_2) = \text{span}(S_2)$. Also $S_1 \cup S_2 = S_2$.

(c) $S_1 = \{x^5\}$, $S_2 = \{x^4\}$

(27) Step 3 of the Diagonalization Method creates a set $S = \{\mathbf{X}_1, \ldots, \mathbf{X}_m\}$ of fundamental eigenvectors where each $\mathbf{X}_j$ is the particular solution of $\lambda\mathbf{I}_n - \mathbf{A} = \mathbf{0}$ obtained by letting the jth independent variable equal 1 while letting all other independent variables equal 0. Now, let $\mathbf{X} \in E_\lambda$. Then $\mathbf{X}$ is a solution of $\lambda\mathbf{I}_n - \mathbf{A} = \mathbf{0}$. Suppose $\mathbf{X}$ is obtained by setting the ith independent variable equal to k_i, for each i, with $1 \le i \le m$. Then $\mathbf{X} = k_1\mathbf{X}_1 + \ldots + k_m\mathbf{X}_m$. Hence $\mathbf{X}$ is a linear combination of the $\mathbf{X}_i$'s, and thus S spans E_λ.

(28) (a) Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two vectors in $\text{span}(S)$. By the definition of $\text{span}(S)$, both $\mathbf{v}_1$ and $\mathbf{v}_2$ can be expressed as finite linear combinations of vectors from S. That is, there are finite subsets $S_1 = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ and $S_2 = \{\mathbf{x}_1, \ldots, \mathbf{x}_l\}$ of S such that $\mathbf{v}_1 = a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$ and $\mathbf{v}_2 = b_1\mathbf{x}_1 + \cdots + b_l\mathbf{x}_l$ for some real numbers $a_1, \ldots, a_k, b_1, \ldots, b_l$.

(b) The natural thing to do at this point would be to combine the expressions for $\mathbf{v}_1$ and $\mathbf{v}_2$ by adding corresponding coefficients. However, each of the subsets $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ or $\{\mathbf{x}_1, \ldots, \mathbf{x}_l\}$ may contain elements not found in the other. Therefore, we create a larger set $S_3 = S_1 \cup S_2$ containing all of the vectors in both subsets. We rename the elements of the finite subset S_3 as $\{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$. Then $\mathbf{v}_1 = a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$ can be expressed as $c_1\mathbf{z}_1 + \cdots + c_m\mathbf{z}_m$, where $c_i = a_j$ if $\mathbf{z}_i = \mathbf{w}_j$, and $c_i = 0$ if $\mathbf{z}_i \notin \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$. Similarly, $\mathbf{v}_2 = b_1\mathbf{x}_1 + \cdots + b_l\mathbf{x}_l$ can be expressed as $d_1\mathbf{z}_1 + \cdots + d_m\mathbf{z}_m$, where $d_i = b_j$ if $\mathbf{z}_i = \mathbf{x}_j$, and $c_i = 0$ if $\mathbf{z}_i \notin \{\mathbf{x}_1, \ldots, \mathbf{x}_l\}$. In this way, both $\mathbf{v}_1$ and $\mathbf{v}_2$ can be expressed as linear combinations of the vectors in S_3.

(c) From parts (a) and (b), we have

$$\mathbf{v}_1 + \mathbf{v}_2 = \sum_{i=1}^{m} c_i\mathbf{z}_i + \sum_{i=1}^{m} d_i\mathbf{z}_i = \sum_{i=1}^{m} (c_i + d_i)\,\mathbf{z}_i,$$

a linear combination of the vectors $\mathbf{z}_i$ in the subset S_3 of S. Thus, $\mathbf{v}_1 + \mathbf{v}_2 \in \text{span}(S)$.

(29) If S contains a nonzero vector $\mathbf{v}$, then $\text{span}(S)$ contains $c\mathbf{v}$, for every scalar c. Now, if c and d are different scalars, we have $c\mathbf{v} \neq d\mathbf{v}$ (or else $(c-d)\mathbf{v} = \mathbf{0}$, and then $\mathbf{v} = \mathbf{0}$ by part (4) of Theorem 4.1). That is, distinct scalar multiples of $\mathbf{v}$ form distinct vectors. Because the number of scalars is infinite, the set of scalar multiples of $\mathbf{v}$ forms an infinite set of vectors, and so $\text{span}(S)$ is infinite.

(30) (a) F (b) T (c) F (d) F (e) F (f) T

Section 4.4

(1) (a) Linearly independent ($[0, 1, 1]$ is not the zero vector)
(b) Linearly independent (neither vector is a multiple of the other)
(c) Linearly dependent (each vector is a multiple of the other)
(d) Linearly dependent (contains $[0, 0, 0]$)
(e) Linearly dependent (Theorem 4.7)

(2) Linearly independent: (b), (c), (f). The others are linearly dependent, and the appropriate reduced row echelon form matrix for each is given.

(a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(3) Linearly independent: (a), (b). The others are linearly dependent, and the appropriate reduced row echelon form matrix for each is given.

(c) $\begin{bmatrix} 4 & 60 & -24 \\ -2 & -25 & 9 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -\frac{3}{5} \end{bmatrix}$.

(d) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$.

(4) (a) Linearly independent: $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$ row reduces to $\mathbf{I}_3$.

(b) Linearly independent:
$\begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

(c) Consider the polynomials in $\mathcal{P}_2$ as corresponding vectors in $\mathbb{R}^3$. Then, the set is linearly dependent by Theorem 4.7.

(d) Linearly dependent:

$\begin{bmatrix} 3 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & 0 & -5 & -10 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 & \frac{5}{2} \\ 0 & 1 & 0 & -\frac{13}{2} \\ 0 & 0 & 1 & \frac{5}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}$, although row reduction really is not necessary since the original matrix clearly has rank ≤ 3 due to the row of zeroes.

(e) Linearly independent. See the remarks just before Example 13 in the textbook. No polynomial in the set can be expressed as a finite linear combination of the others since no single power of x can be expressed as a combination of other powers of x.

(f) Linearly independent. Use an argument similar to that in Example 13 in the textbook.

(5) Solving the appropriate linear system yields
$$21\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} + 15\begin{bmatrix} 3 & 2 \\ -6 & 1 \end{bmatrix} - 18\begin{bmatrix} 4 & -1 \\ -5 & 2 \end{bmatrix} + 2\begin{bmatrix} 3 & -3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

(6) $\begin{bmatrix} 1 & 4 & 0 & 0 \\ 2 & 2 & 1 & 7 \\ -1 & -6 & 1 & 5 \\ 1 & 1 & -1 & 2 \\ 3 & 0 & 2 & -1 \\ 0 & 1 & 2 & 6 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, so the set is linearly independent by the Independence Test Method.

(7) (a) $[-2, 0, 1]$ is not a scalar multiple of $[1, 1, 0]$.

(b) $[0, 1, 0]$, $[0, 0, 1]$

(c) Any nonzero linear combination of $[1, 1, 0]$ and $[-2, 0, 1]$, other than $[1, 1, 0]$ and $[-2, 0, 1]$ themselves, will work. One possibility is $\mathbf{u} = [1, 1, 0] + [-2, 0, 1] = [-1, 1, 1]$.

(8) (a) Applying the Independence Test Method, we obtain a pivot in every column.

(b) $[-2, 0, 3, -4] = -1[2, -1, 0, 5] + 1[1, -1, 2, 0] + 1[-1, 0, 1, 1]$.

(c) No, because S is linearly independent (see Theorem 4.9).

(9) (a) $x^3 - 4x + 8 = 2(2x^3 - x + 3) - (3x^3 + 2x - 2)$. (Coefficients were obtained using Independence Test Method.)

(b) See part (a). Also, solving appropriate systems yields:
$4x^3 + 5x - 7 = -1(2x^3 - x + 3) + 2(3x^3 + 2x - 2)$;
$2x^3 - x + 3 = \frac{2}{3}(x^3 - 4x + 8) + \frac{1}{3}(4x^3 + 5x - 7)$;
$3x^3 + 2x - 2 = \frac{1}{3}(x^3 - 4x + 8) + \frac{2}{3}(4x^3 + 5x - 7)$

(c) No polynomial in S is a scalar multiple of any other polynomial in S.

(10) Following the hint in the textbook, let $\mathbf{A}$ be the matrix whose rows are the vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. By the Independence Test Method, S is linearly independent iff $\mathbf{A}^T$ row reduces to a matrix with a pivot in every column iff $\mathbf{A}^T$ has rank 3 iff $|\mathbf{A}^T| \neq 0$ iff $|\mathbf{A}| \neq 0$.

(11) In each part, there are many different correct answers. Only one possibility is given here.

(a) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ (b) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ (c) $\{1, x, x^2, x^3\}$

(d) $\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \right\}$

(e) $\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\}$

(Notice that each matrix is symmetric.)

(12) (a) If S is linearly dependent, then $a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n=\mathbf{0}$ for some $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}\subseteq S$ and $a_1,\ldots,a_n\in\mathbb{R}$, with some $a_i\neq 0$. Then

$$\mathbf{v}_i=-\frac{a_1}{a_i}\mathbf{v}_1-\cdots-\frac{a_{i-1}}{a_i}\mathbf{v}_{i-1}-\frac{a_{i+1}}{a_i}\mathbf{v}_{i+1}-\cdots-\frac{a_n}{a_i}\mathbf{v}_n.$$

So, if $\mathbf{w}\in\mathrm{span}(S)$, then $\mathbf{w}=b\mathbf{v}_i+c_1\mathbf{w}_1+\cdots+c_k\mathbf{w}_k$, for some $\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}\subseteq S-\{\mathbf{v}_i\}$ and $b,c_1,\ldots,c_k\in\mathbb{R}$. Hence,

$$\mathbf{w}=-\frac{ba_1}{a_i}\mathbf{v}_1-\cdots-\frac{ba_{i-1}}{a_i}\mathbf{v}_{i-1}-\frac{ba_{i+1}}{a_i}\mathbf{v}_{i+1}-\cdots-\frac{ba_n}{a_i}\mathbf{v}_n+c_1\mathbf{w}_1+\cdots+c_k\mathbf{w}_k\in\mathrm{span}(S-\{\mathbf{v}_i\}).$$

So, $\mathrm{span}(S)\subseteq\mathrm{span}(S-\{\mathbf{v}_i\})$. Also, $\mathrm{span}(S-\{\mathbf{v}_i\})\subseteq\mathrm{span}(S)$ by Corollary 4.6. Thus, we have $\mathrm{span}(S-\{\mathbf{v}_i\})=\mathrm{span}(S)$.

Conversely, if $\mathrm{span}(S-\{\mathbf{v}\})=\mathrm{span}(S)$ for some $\mathbf{v}\in S$, then by Theorem 4.5, part (1), we have $\mathbf{v}\in\mathrm{span}(S)=\mathrm{span}(S-\{\mathbf{v}\})$, and so S is linearly dependent by the Alternate Characterization of linear dependence in Table 4.1.

(b) Suppose $\mathbf{v}\in S$ is redundant. Then, $\mathbf{v}\in\mathrm{span}(S)=\mathrm{span}(S-\{\mathbf{v}\})$, and so $\mathbf{v}$ is a linear combination of vectors in $S-\{\mathbf{v}\}$.

Conversely, suppose $\mathbf{v}$ is a linear combination of vectors in $S-\{\mathbf{v}\}$. Then $\mathbf{v}=a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k$, for some $\mathbf{v}_1,\ldots,\mathbf{v}_k\in S-\{\mathbf{v}\}$. We need to show that $\mathrm{span}(S)=\mathrm{span}(S-\{\mathbf{v}\})$. Now, $\mathrm{span}(S-\{\mathbf{v}\})\subseteq\mathrm{span}(S)$ by Corollary 4.6. So, we need to show that $\mathrm{span}(S)\subseteq\mathrm{span}(S-\{\mathbf{v}\})$. Suppose $\mathbf{w}\in\mathrm{span}(S)$. Then $\mathbf{w}=b\mathbf{v}+c_1\mathbf{w}_1+\cdots+c_k\mathbf{w}_k$, for some $\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}\subseteq S-\{\mathbf{v}\}$ and $b,c_1,\ldots,c_k\in\mathbb{R}$. Hence,

$$\mathbf{w}=ba_1\mathbf{v}_1+\cdots+ba_n\mathbf{v}_n+c_1\mathbf{w}_1+\cdots+c_k\mathbf{w}_k\in\mathrm{span}(S-\{\mathbf{v}\}),$$

completing the proof.

(13) In each part, you can prove that the indicated vector is redundant by creating a matrix using the given vectors as rows, and showing that the Simplified Span Method leads to the same reduced row echelon form (except, perhaps, with an extra row of zeroes) with or without the indicated vector as one of the rows.

(a) Let $\mathbf{v}=[0,0,0,0]$. Any linear combination of the other three vectors becomes a linear combination of all four vectors when $1[0,0,0,0]$ is added to it.

(b) Let $\mathbf{v}=[0,0,-6,0]$. Let S be the given set of three vectors. To prove that $\mathbf{v}=[0,0,-6,0]$ is redundant, we need to show that $\mathrm{span}(S-\{\mathbf{v}\})=\mathrm{span}(S)$. We apply the Simplified Span Method to both $S-\{\mathbf{v}\}$ and S.

For $\mathrm{span}(S-\{\mathbf{v}\})$:

$\begin{bmatrix}1&1&0&0\\1&1&1&0\end{bmatrix}$ row reduces to $\begin{bmatrix}1&1&0&0\\0&0&1&0\end{bmatrix}$.

For $\mathrm{span}(S)$:

$\begin{bmatrix}1&1&0&0\\1&1&1&0\\0&0&-6&0\end{bmatrix}$ row reduces to $\begin{bmatrix}1&1&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}$.

Since the reduced row echelon form matrices are the same, except for the extra row of zeroes, the two spans are equal, and $\mathbf{v}=[0,0,-6,0]$ is a redundant vector.

For an alternate approach to proving that span$(S - \{\mathbf{v}\})$ = span(S), first note that $S - \{\mathbf{v}\} \subseteq S$, and so span$(S - \{\mathbf{v}\}) \subseteq$ span(S) by Corollary 4.6. Next, $[1, 1, 0, 0]$ and $[1, 1, 1, 0]$ are clearly in span$(S - \{\mathbf{v}\})$ by part (1) of Theorem 4.5. But $\mathbf{v} \in$ span$(S - \{\mathbf{v}\})$ as well, because

$$[0, 0, -6, 0] = (6)[1, 1, 0, 0] + (-6)[1, 1, 1, 0].$$

Hence, S is a subset of the subspace span$(S - \{\mathbf{v}\})$. Therefore, by part (3) of Theorem 4.5,

$$\text{span}(S) \subseteq \text{span}(S - \{\mathbf{v}\}).$$

Thus, since we have proven subset inclusion in both directions, span$(S - \{\mathbf{v}\})$ = span(S).

(c) Let S be the given set of vectors. Any of the given 16 vectors in S can be considered as the redundant vector. One clever way to see this is to consider the matrix

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

whose rows are 4 of the vectors in S. Now, $\mathbf{A}$ row reduces to $\mathbf{I}_4$. Thus, the four rows of $\mathbf{A}$ span $\mathbb{R}^4$, by the Simplified Span Method. Hence, any of the remaining 12 vectors in S are redundant since they are in the span of these 4 rows. Now, repeat this argument using $-\mathbf{A}$ to show that the four rows in $\mathbf{A}$ are also individually redundant. (Note that $\mathbf{A}$ and $-\mathbf{A}$ are row equivalent, so we do not have to perform a second row reduction.)

(14) Assume S_1 is linearly independent. Suppose $a_1(c\mathbf{v}_1) + a_2(c\mathbf{v}_2) + \cdots + a_n(c\mathbf{v}_n) = \mathbf{0}$. Divide both sides by c and use the fact that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent to obtain $a_1 = a_2 = \cdots = a_n = 0$. Thus S_2 is linearly independent. Conversely, assume S_2 is linearly independent. Suppose $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$. Multiply both sides by c and use the fact that $\{c\mathbf{v}_1, c\mathbf{v}_2, \ldots, c\mathbf{v}_n\}$ is linearly independent to obtain $a_1 = a_2 = \cdots = a_n = 0$. Thus S_1 is linearly independent.

(15) Notice that $x\mathbf{f}'(x)$ is not a scalar multiple of $\mathbf{f}(x)$. For if the terms of $\mathbf{f}(x)$ are written in order of descending degree, and the first two nonzero terms are, respectively, a_nx^n and a_kx^k, then the corresponding terms of $x\mathbf{f}'(x)$ are na_nx^n and ka_kx^k, which are different multiples of a_nx^n and a_kx^k, respectively, since $n \neq k$.

(16) Theorem 4.5, part (3) shows that span$(S) \subseteq \mathcal{W}$. Thus $\mathbf{v} \notin$ span(S). Now if $S \cup \{\mathbf{v}\}$ is linearly dependent, then there exists $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\} \subseteq S$ so that $a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n + b\mathbf{v} = \mathbf{0}$, with not all of $a_1, \ldots, a_n, b$ equal to zero. Also, $b \neq 0$ since S is linearly independent. Hence,

$$\mathbf{v} = -\frac{a_1}{b}\mathbf{w}_1 - \cdots - \frac{a_n}{b}\mathbf{w}_n \in \text{span}(S),$$

a contradiction.

(17) (a) Suppose $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. Then $a_1\mathbf{A}\mathbf{v}_1 + \cdots + a_k\mathbf{A}\mathbf{v}_k = \mathbf{0} \Longrightarrow a_1 = \cdots = a_k = 0$, since T is linearly independent.

(b) The converse to the statement is: If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent subset of $\mathbb{R}^m$, then $T = \{\mathbf{A}\mathbf{v}_1, \ldots, \mathbf{A}\mathbf{v}_k\}$ is a linearly independent subset of $\mathbb{R}^n$ with $\mathbf{A}\mathbf{v}_1, \ldots, \mathbf{A}\mathbf{v}_k$ distinct. We can construct specific counterexamples that contradict either (or both) of the conclusions of the converse. For a specific counterexample in which $\mathbf{A}\mathbf{v}_1, \ldots, \mathbf{A}\mathbf{v}_k$ are distinct but T is linearly dependent, let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $S = \{[1,0],[1,2]\}$. Then $T = \{[1,1],[3,3]\}$. Note that S is linearly independent, but the vectors $\mathbf{Av}_1$ and $\mathbf{Av}_2$ are distinct and T is linearly dependent. For a specific counterexample in which $\mathbf{Av}_1, \ldots, \mathbf{Av}_k$ are not distinct but T is linearly independent, use $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $S = \{[1,2],[2,1]\}$. Then $T = \{[3,3]\}$. Note that $\mathbf{Av}_1$ and $\mathbf{Av}_2$ both equal $[3,3]$, and that both S and T are linearly independent. For a final counterexample in which both parts of the conclusion are false, let $\mathbf{A} = \mathbf{O}_{22}$ and $S = \{\mathbf{e}_1, \mathbf{e}_2\}$. Then $\mathbf{Av}_1$ and $\mathbf{Av}_2$ both equal $\mathbf{0}$, S is linearly independent, but $T = \{\mathbf{0}\}$ is linearly dependent.

(c) The converse to the statement is: If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent subset of $\mathbb{R}^m$, then $T = \{\mathbf{Av}_1, \ldots, \mathbf{Av}_k\}$ is a linearly independent subset of $\mathbb{R}^n$ with $\mathbf{Av}_1, \ldots, \mathbf{Av}_k$ distinct. To prove that $\mathbf{Av}_1, \ldots, \mathbf{Av}_k$ are distinct, we assume $\mathbf{Av}_i = \mathbf{Av}_j$, with $i \neq j$, and then multiply both sides by $\mathbf{A}^{-1}$ to obtain $\mathbf{v}_i = \mathbf{v}_j$, a contradiction. To prove $\{\mathbf{Av}_1, \ldots, \mathbf{Av}_k\}$ is linearly independent, suppose $a_1\mathbf{Av}_1 + \cdots + a_k\mathbf{Av}_k = \mathbf{0}$. Then $a_1\mathbf{A}^{-1}\mathbf{Av}_1 + \cdots + a_k\mathbf{A}^{-1}\mathbf{Av}_k = \mathbf{0} \Longrightarrow a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0} \Longrightarrow a_1 = \cdots = a_k = 0$, since S is linearly independent.

(18) Case 1: S is empty. This case is obvious, since the empty set is the only subset of S.
Case 2: Suppose $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a finite nonempty linearly independent set in $\mathcal{V}$, and let S_1 be a subset of S. If S_1 is empty, it is linearly independent by definition. Suppose S_1 is nonempty. After renumbering the subscripts if necessary, we can assume $S_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, with $k \leq n$. If $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, then $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + 0\mathbf{v}_{k+1} + \cdots 0\mathbf{v}_n = \mathbf{0}$, and so $a_1 = \cdots = a_k = 0$, since S is linearly independent. Hence, S_1 is linearly independent.
Case 3: Suppose S is an infinite linearly independent subset of $\mathcal{V}$. Then any finite subset of S is linearly independent by the definition of linear independence for infinite subsets. If S_1 is an infinite subset, then any finite subset S_2 of S_1 is also a finite subset of S. Hence S_2 is linearly independent because S is. Thus, S_1 is linearly independent, by definition.

(19) First, suppose that $\mathbf{v}_1 \neq \mathbf{0}$ and that for every k, with $2 \leq k \leq n$, we have $\mathbf{v}_k \notin \text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\})$. It is enough to show that S is linearly independent by assuming that S is linearly *dependent* and showing that this leads to a contradiction. If S is linearly dependent, then there real numbers $a_1, \ldots, a_n$, not all zero, such that

$$\mathbf{0} = a_1\mathbf{v}_1 + \cdots + a_i\mathbf{v}_i + \cdots + a_n\mathbf{v}_n.$$

Suppose k is the largest subscript such that $a_k \neq 0$. Then

$$\mathbf{0} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k.$$

If $k = 1$, then $a_1\mathbf{v}_1 = \mathbf{0}$ with $a_1 \neq 0$. This implies that $\mathbf{v}_1 = \mathbf{0}$, a contradiction. If $k \geq 2$, then solving for $\mathbf{v}_k$ yields

$$\mathbf{v}_k = \left(-\frac{a_1}{a_k}\right)\mathbf{v}_1 + \cdots + \left(-\frac{a_{k-1}}{a_k}\right)\mathbf{v}_{k-1},$$

where $\mathbf{v}_k$ is expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$, contradicting the fact that $\mathbf{v}_k \notin \text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\})$.

Conversely, suppose S is linearly independent. Notice that $\mathbf{v}_1 \neq \mathbf{0}$ since any finite subset of $\mathcal{V}$ containing $\mathbf{0}$ is linearly dependent. We must prove that for each k with $1 \leq k \leq n$, $\mathbf{v}_k \notin \text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\})$. Now, by Corollary 4.6, $\text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\}) \subseteq \text{span}(S - \{\mathbf{v}_k\})$ since $\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\} \subseteq S - \{\mathbf{v}_k\}$. But, by the first "boxed" alternate characterization of linear independence following Example 10 in the textbook, $\mathbf{v}_k \notin \text{span}(S - \{\mathbf{v}_k\})$. Therefore, $\mathbf{v}_k \notin \text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\})$.

(20) (a) Reversing the order of the elements gives $\{72, 72x, 36x^2-30x, 12x^3-15x^2+4, 3x^4-5x^3+4x-8\}$. Each polynomial in the set now has a higher degree than those preceding it, and hence is not a linear combination of those preceding polynomials. So, by Exercise 19 this set is linearly independent.

(b) Reversing the order of the elements gives $\{\mathbf{f}^{(n)}, \ldots, \mathbf{f}^{(2)}, \mathbf{f}^{(1)}, \mathbf{f}\}$. But each $\mathbf{f}^{(k)} \neq a_{k+1}\mathbf{f}^{(k+1)} + \cdots + a_n\mathbf{f}^{(n)}$, for any $a_{k+1}, \ldots, a_n$, because all polynomials on the right side of the equation have lower degrees than $\mathbf{f}^{(k)}$. Thus, by Exercise 19 this set is linearly independent.

(21) (a) If S is linearly independent, then the zero vector has a unique expression by Theorem 4.9.

Conversely, suppose $\mathbf{v} \in \text{span}(S)$ has a unique expression of the form $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, where $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are in S. To show that S is linearly independent, it is enough to show that any finite subset $S_1 = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ of S is linearly independent. Suppose $b_1\mathbf{w}_1 + \cdots + b_k\mathbf{w}_k = \mathbf{0}$. Now $b_1\mathbf{w}_1 + \cdots + b_k\mathbf{w}_k + a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0} + \mathbf{v} = \mathbf{v}$. If $\mathbf{w}_i = \mathbf{v}_j$ for some j, then the uniqueness of the expression for $\mathbf{v}$ implies that $b_i + a_j = a_j$, yielding $b_i = 0$. If, instead, $\mathbf{w}_i$ does not equal any $\mathbf{v}_j$, then $b_i = 0$, again by the uniqueness assumption. Thus each b_i must equal zero. Hence, S_1 is linearly independent by the definition of linear independence.

(b) S is linearly dependent iff every vector $\mathbf{v}$ in span(S) can be expressed in more than one way as a linear combination of vectors in S (ignoring zero coefficients).

(22) Follow the hint in the text. The fundamental eigenvectors are found by solving the homogeneous system $(\lambda\mathbf{I}_n - \mathbf{A})\mathbf{v} = \mathbf{0}$. Each $\mathbf{v}_i$ has a "1" in the position corresponding to the ith independent variable for the system and a "0" in the position corresponding to every other independent variable. Hence, any linear combination $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$ is an n-vector having the value a_i in the position corresponding to the ith independent variable for the system, for each i. So if this linear combination equals $\mathbf{0}$, then each a_i must equal 0, proving linear independence.

(23) (a) Since $T \cup \{\mathbf{v}\}$ is linearly dependent, by the definition of linear independence, there is some finite subset $\{\mathbf{t}_1, \ldots, \mathbf{t}_k\}$ of T such that $a_1\mathbf{t}_1 + \cdots + a_k\mathbf{t}_k + b\mathbf{v} = \mathbf{0}$, with not all coefficients equal to zero. But if $b = 0$, then $a_1\mathbf{t}_1 + \cdots + a_k\mathbf{t}_k = \mathbf{0}$ with not all $a_i = 0$, contradicting the fact that T is linearly independent. Hence, $b \neq 0$, and thus $\mathbf{v} = (-\frac{a_1}{b})\mathbf{t}_1 + \cdots + (-\frac{a_k}{b})\mathbf{t}_k$. Hence, $\mathbf{v} \in \text{span}(T)$.

(b) The statement in part (b) is merely the contrapositive of the statement in part (a).

(24) Suppose that S is linearly independent, and suppose $\mathbf{v} \in \text{span}(S)$ can be expressed both as $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k$ and $\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_l\mathbf{v}_l$, for distinct $\mathbf{u}_1, \ldots, \mathbf{u}_k \in S$ and distinct $\mathbf{v}_1, \ldots, \mathbf{v}_l \in S$, where these expressions differ in at least one nonzero term. Since the $\mathbf{u}_i$'s might not be distinct from the $\mathbf{v}_i$'s, we consider the set $X = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\} \cup \{\mathbf{v}_1, \ldots, \mathbf{v}_l\}$ and label the distinct vectors in X as $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. Then we can express $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k$ as $\mathbf{v} = c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m$, and also $\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_l\mathbf{v}_l$ as $\mathbf{v} = d_1\mathbf{w}_1 + \cdots + d_m\mathbf{w}_m$, choosing the scalars c_i, d_i, $1 \leq i \leq m$ with $c_i = a_j$ if $\mathbf{w}_i = \mathbf{u}_j$, $c_i = 0$ otherwise, and $d_i = b_j$ if $\mathbf{w}_i = \mathbf{v}_j$, $d_i = 0$ otherwise. Since the original linear combinations for $\mathbf{v}$ are distinct, we know that $c_i \neq d_i$ for some i. Now, $\mathbf{v} - \mathbf{v} = \mathbf{0} = (c_1 - d_1)\mathbf{w}_1 + \cdots + (c_m - d_m)\mathbf{w}_m$. Since $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\} \subseteq S$, a linearly independent set, each $c_j - d_j = 0$ for every j with $1 \leq j \leq m$. But this is a contradiction since $c_i \neq d_i$.

Conversely, assume every vector in span(S) can be uniquely expressed as a linear combination of elements of S. Since $\mathbf{0} \in \text{span}(S)$, there is exactly one linear combination of elements of S that equals $\mathbf{0}$. Now, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is any finite subset of S, we have $\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n$. Because this representation is unique, it means that in any linear combination of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ that equals $\mathbf{0}$, the only possible coefficients are zeroes. Thus, by definition, S is linearly independent.

(25) (a) F (b) T (c) T (d) F (e) T (f) T (g) F (h) T (i) T

Section 4.5

(1) For each part, row reduce the matrix whose rows are the given vectors. The result will be $\mathbf{I}_4$. Hence, by the Simplified Span Method, the set of vectors spans $\mathbb{R}^4$. Row reducing the matrix whose columns are the given vectors also results in $\mathbf{I}_4$. Therefore, by the Independence Test Method, the set of vectors is linearly independent. (After the first row reduction, you could just apply part (1) of Theorem 4.12, and then skip the second row reduction.)

(2) Follow the procedure in the answer to Exercise 1.

(3) Follow the procedure in the answer to Exercise 1, except that row reduction results in $\mathbf{I}_5$ instead of $\mathbf{I}_4$.

(4) (a) Not a basis: $|S| < \dim(\mathbb{R}^4)$. (linearly independent but does not span)

(b) Not a basis: $|S| < \dim(\mathbb{R}^4)$. (linearly dependent and does not span)

(c) Basis: Follow the procedure in the answer to Exercise 1.

(d) Not a basis: $\begin{bmatrix} 1 & 3 & 2 & -1 \\ -2 & 0 & 6 & -10 \\ 0 & 6 & 10 & -12 \\ 2 & 10 & -3 & 31 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, and so does not span.

$\begin{bmatrix} 1 & -2 & 0 & 2 \\ 3 & 0 & 6 & 10 \\ 2 & 6 & 10 & -3 \\ -1 & -10 & -12 & 31 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, and so is linearly dependent.

(e) Not a basis: $|S| > \dim(\mathbb{R}^4)$. (linearly dependent but spans)

(5) (a) $\mathcal{W}$ is nonempty, since $\mathbf{0} \in \mathcal{W}$. Let $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{W}$, $c \in \mathbb{R}$. Then $\mathbf{A}(\mathbf{X}_1 + \mathbf{X}_2) = \mathbf{A}\mathbf{X}_1 + \mathbf{A}\mathbf{X}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Also, $\mathbf{A}(c\mathbf{X}_1) = c(\mathbf{A}\mathbf{X}_1) = c\mathbf{0} = \mathbf{0}$. Hence $\mathcal{W}$ is closed under addition and scalar multiplication, and so is a subspace by Theorem 4.2.

(b) $\mathbf{A}$ row reduces to $\begin{bmatrix} 1 & 0 & \frac{1}{5} & \frac{2}{5} & 1 \\ 0 & 1 & \frac{2}{5} & -\frac{1}{5} & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Thus, a basis for $\mathcal{W}$ is $\{[-1, -2, 5, 0, 0], [-2, 1, 0, 5, 0], [-1, 1, 0, 0, 1]\}$.

(c) $\dim(\mathcal{W}) = 3$; $\text{rank}(\mathbf{A}) = 2$; $3 + 2 = 5$

(6) (a) $\mathcal{W}$ is nonempty, since $\mathbf{0} \in \mathcal{W}$. Let $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{W}$, $c \in \mathbb{R}$. Then $\mathbf{A}(\mathbf{X}_1 + \mathbf{X}_2) = \mathbf{A}\mathbf{X}_1 + \mathbf{A}\mathbf{X}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Also, $\mathbf{A}(c\mathbf{X}_1) = c(\mathbf{A}\mathbf{X}_1) = c\mathbf{0} = \mathbf{0}$. Hence $\mathcal{W}$ is closed under addition and scalar multiplication, and so is a subspace by Theorem 4.2.

(b) $\mathbf{A}$ row reduces to $\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Thus, a basis for $\mathcal{W}$ is $\{[2, -3, 1, 0]\}$.

(c) $\dim(\mathcal{W}) = 1$; $\text{rank}(\mathbf{A}) = 3$; $1 + 3 = 4$

(7) These vectors are linearly independent by part (b) of Exercise 20 in Section 4.4. The given set has $n+1$ distinct elements, and $\dim(\mathcal{P}_n) = n+1$. Hence, by part (2) of Theorem 4.12, the set is a basis for $\mathcal{P}_n$.

(8) (a) Let $S = \{\mathbf{I}_2, \mathbf{A}, \mathbf{A}^2, \mathbf{A}^3, \mathbf{A}^4\}$. If S is linearly independent, then S is a basis for $\mathcal{W} = \text{span}(S) \subseteq \mathcal{M}_{22}$, and since $\dim(\mathcal{W}) = 5$ and $\dim(\mathcal{M}_{22}) = 4$, this contradicts Theorem 4.13. Hence S is linearly dependent. Now use the definition of linear dependence.

(b) Generalize the proof in part (a).

(9) (a) B is easily seen to be linearly independent using Exercise 19, Section 4.4. To show B spans $\mathcal{V}$, note that every $\mathbf{f} \in \mathcal{V}$ can be expressed as $\mathbf{f} = (x-2)\mathbf{g}$, for some $\mathbf{g} \in \mathcal{P}_4$.

(b) 5

(c) $\{(x-2)(x-3), x(x-2)(x-3), x^2(x-2)(x-3), x^3(x-2)(x-3)\}$

(d) 4

(10) (a) $[-1,1,1,-1]$ is not a scalar multiple of $[2,3,0,-1]$. Also, $[1,4,1,-2] = 1[2,3,0,-1] + 1[-1,1,1,-1]$, and $[3,2,-1,0] = 1[2,3,0,-1] - 1[-1,1,1,-1]$.

(b) We illustrate two approaches:

First approach: We check that B and S both span the same subspace of $\mathbb{R}^4$ by using the Simplified Span Method and showing that both S and B produce the same bases for their spans.

For S:

$$\begin{bmatrix} 1 & 4 & 1 & -2 \\ -1 & 1 & 1 & -1 \\ 3 & 2 & -1 & 0 \\ 2 & 3 & 0 & -1 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & -\frac{3}{5} & \frac{2}{5} \\ 0 & 1 & \frac{2}{5} & -\frac{3}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For B:

$$\begin{bmatrix} 2 & 3 & 0 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & -\frac{3}{5} & \frac{2}{5} \\ 0 & 1 & \frac{2}{5} & -\frac{3}{5} \end{bmatrix}.$$

Since both row reduced matrices have identical nonzero rows, $\text{span}(S)$ and $\text{span}(B)$ have identical bases (the set of nonzero rows of the row reduced matrices). Hence, they must be the same subspaces. Also, since the basis produced for $\text{span}(S)$ contains two vectors, $\dim(\text{span}(S)) = 2$. Finally, because B is a linearly independent subset of $\text{span}(S)$ with $|B| = \dim(\text{span}(S))$, part (2) of Theorem 4.12 implies that B is a basis for $\text{span}(S)$.

Second approach: In part (a) we showed that no larger subset of S containing B is linearly independent; that is, each of the remaining vectors in S can be expressed as a linear combination of the vectors in B. Therefore, $S \subseteq \text{span}(B)$. Then, part (3) of Theorem 4.5 implies that $\text{span}(S) \subseteq \text{span}(B)$. We also know that $B \subseteq S$, and so Corollary 4.6 implies that $\text{span}(B) \subseteq \text{span}(S)$. Therefore, $\text{span}(B) = \text{span}(S)$. Since B was shown to be linearly independent in part (a), B is a basis for $\text{span}(S)$. Hence, $\dim(\text{span}(S)) = |B| = 2$.

(c) No; $\dim(\text{span}(S)) = 2 \neq 4 = \dim(\mathbb{R}^4)$

(11) (a) Solving an appropriate homogeneous system shows that B is linearly independent. Also,

$$x - 1 = \tfrac{1}{12}(x^3 - x^2 + 2x + 1) + \tfrac{1}{12}(2x^3 + 4x - 7) - \tfrac{1}{12}(3x^3 - x^2 - 6x + 6)$$

and

$$x^3 - 3x^2 - 22x + 34 = 1(x^3 - x^2 + 2x + 1) - 3(2x^3 + 4x - 7) + 2(3x^3 - x^2 - 6x + 6).$$

(b) We illustrate two approaches:
First approach: We check that B and S both span the same subspace of $\mathcal{P}_3$ by using the Simplified Span Method and showing that both S and B produce the same bases for their spans.
For S:

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 \\ 2 & 0 & 4 & -7 \\ 1 & -3 & -22 & 34 \\ 3 & -1 & -6 & 6 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{2} \\ 0 & 1 & 0 & -\frac{9}{2} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For B:

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & 0 & 4 & -7 \\ 3 & -1 & -6 & 6 \end{bmatrix} \text{ row reduces to } \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{2} \\ 0 & 1 & 0 & -\frac{9}{2} \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Since both row reduced matrices have identical nonzero rows, span(S) and span(B) have identical bases (the set of nonzero rows of the row reduced matrices). Hence, they must be the same subspaces. Also, since the basis produced for span(S) contains three vectors, dim(span(S)) = 3. Finally, because B is a linearly independent subset of span(S) with $|B| = \dim(\text{span}(S))$, part (2) of Theorem 4.12 implies that B is a basis for span(S).

Second approach: In part (a) we showed that no larger subset of S containing B is linearly independent; that is, each of the remaining vectors in S can be expressed as a linear combination of the vectors in B. Therefore, $S \subseteq \text{span}(B)$. Then, part (3) of Theorem 4.5 implies that $\text{span}(S) \subseteq \text{span}(B)$. We also know that $B \subseteq S$, and so Corollary 4.6 implies that $\text{span}(B) \subseteq \text{span}(S)$. Therefore, $\text{span}(B) = \text{span}(S)$. Since B was shown to be linearly independent in part (a), B is a basis for span(S). Hence, $\dim(\text{span}(S)) = |B| = 3$.

(c) No; $\dim(\text{span}(S)) = 3 \neq 4 = \dim(\mathcal{P}_3)$

(12) (a) Let $\mathcal{V} = \mathbb{R}^3$, and let $S = \{[1, 0, 0], [2, 0, 0], [3, 0, 0]\}$.

(b) Let $\mathcal{V} = \mathbb{R}^3$, and let $T = \{[1, 0, 0], [2, 0, 0], [3, 0, 0]\}$.

(13) If S spans $\mathcal{V}$, then S is a basis by part (1) of Theorem 4.12. If S is linearly independent, then S is a basis by part (2) of Theorem 4.12.

(14) By part (1) of Theorem 4.12, if S spans $\mathcal{V}$, then S is a basis for $\mathcal{V}$, so S is linearly independent. Similarly, by part (2) of Theorem 4.12, if S is linearly independent, then S is a basis for $\mathcal{V}$, so S spans $\mathcal{V}$.

(15) (a) Suppose $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.
Linear independence of B_1: $a_1\mathbf{A}\mathbf{v}_1 + \cdots + a_n\mathbf{A}\mathbf{v}_n = \mathbf{0}$
$\Longrightarrow \mathbf{A}^{-1}(a_1\mathbf{A}\mathbf{v}_1 + \cdots + a_n\mathbf{A}\mathbf{v}_n) = \mathbf{A}^{-1}\mathbf{0} \Longrightarrow a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$
$\Longrightarrow a_1 = \cdots = a_n = 0$, since B is linearly independent.
B_1 is a basis by part (2) of Theorem 4.12.

(b) Similar to part (a).

(c) Note that $\mathbf{A}\mathbf{e}_i = (i\text{th column of } \mathbf{A})$.

(d) Note that $\mathbf{e}_i\mathbf{A} = (i\text{th row of } \mathbf{A})$.

(16) The dimension of a proper nontrivial subspace must be 1 or 2, by Theorem 4.13. If the basis for the subspace contains one [two] element(s), then the subspace is a line [plane] through the origin as in Exercise 17 of Section 4.3.

(17) (a) If $S = \{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$, let $n = \max(\text{degree}(\mathbf{f}_1), \ldots, \text{degree}(\mathbf{f}_k))$. Then $S \subseteq \mathcal{P}_n$.

(b) Apply Theorem 4.5, part (3).

(c) $x^{n+1} \notin \text{span}(S)$.

(18) (a) Suppose $\mathcal{V}$ is finite dimensional. Since $\mathcal{V}$ has an infinite linearly independent subset, part (2) of Theorem 4.12 is contradicted.

(b) $\{1, x, x^2, \ldots\}$ is a linearly independent subset of $\mathcal{P}$, and hence, of any vector space containing $\mathcal{P}$. Apply part (a).

(19) Suppose B is a basis for $\mathcal{V}$ and that C is a subset of $\mathcal{V}$ containing B with $C \neq B$. Let $\mathbf{v} \in C$ with $\mathbf{v} \notin B$. Then $\mathbf{v} \in \text{span}(B)$ because B spans $\mathcal{V}$. Thus, $\mathbf{v}$ is a linear combination of vectors in B, which means that $\mathbf{v}$ is a linear combination of vectors in C other than $\mathbf{v}$ itself. Hence, C is linearly dependent by Theorem 4.8.

(20) Suppose B is a basis for $\mathcal{V}$. Suppose that C is a subset of B with $C \neq B$. Let $\mathbf{v} \in B$ with $\mathbf{v} \notin C$. Now B is linearly independent, so by the Alternate Characterization of linear independence in Table 4.1, $\mathbf{v} \notin \text{span}(B - \{\mathbf{v}\})$. But $C \subseteq B - \{\mathbf{v}\}$, implying $\text{span}(C) \subseteq \text{span}(B - \{\mathbf{v}\})$ (Corollary 4.6). Hence $\mathbf{v} \notin \text{span}(C)$. Therefore, C does not span $\mathcal{V}$.

(21) Let $\mathcal{V}$ be a finite dimensional vector space and let $\mathcal{W}$ be a subspace of $\mathcal{V}$. Consider the set A of nonnegative integers defined by

$$A = \{k \mid \text{a set } T \text{ exists with } T \subseteq \mathcal{W},\ |T| = k, \text{ and } T \text{ linearly independent}\}.$$

(a) The empty set is linearly independent and is a subset of $\mathcal{W}$. Hence $0 \in A$. Therefore, A is not the empty set.

(b) Since $\mathcal{V}$ and $\mathcal{W}$ share the same operations, every linearly independent subset T of $\mathcal{W}$ is also a linearly independent subset of $\mathcal{V}$ (by the definition of linear independence). Hence, using part (2) of Theorem 4.12 on T in $\mathcal{V}$ shows that $k = |T| \leq \dim(\mathcal{V})$. Therefore, A contains no numbers larger than $\dim(\mathcal{V})$, and so has a finite number of elements.

(c) Because A is a finite nonempty set of numbers, it has a largest number n. Therefore, because $n \in A$, the definition of the set A implies that there must exist some set T such that $T \subseteq \mathcal{W}$, $|T| = n$, and T is linearly independent. Let $T = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be such a set.

Now, we are given that $T \subseteq \mathcal{W}$, and so $\text{span}(T) \subseteq \mathcal{W}$ by part (3) of Theorem 4.5.

(d) Suppose $\mathbf{w} \in \mathcal{W}$ with $\mathbf{w} \notin \text{span}(T)$. Then $\mathbf{w} \notin T$, and so $T \cup \{\mathbf{w}\}$ is a subset of $\mathcal{W}$ containing $n+1$ vectors. But since n is the largest element of A, $n+1$ is not in A. Hence, the set $T \cup \{\mathbf{w}\}$ must fail to satisfy at least one of the conditions in the description of A. Since $T \cup \{\mathbf{w}\} \subseteq \mathcal{W}$, we must have that $T \cup \{\mathbf{w}\}$ is not linearly independent. Therefore, $T \cup \{\mathbf{w}\}$ is linearly dependent.

(e) Suppose $\mathbf{w} \in \mathcal{W}$, but $\mathbf{w} \notin \text{span}(T)$ as in part (d). Because $T \cup \{\mathbf{w}\} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}\}$ is linearly dependent (by part (d)), there exist scalars $a_1, \ldots, a_n, b$ such that $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n + b\mathbf{w} = \mathbf{0}$, with not all coefficients equal to zero. But if $b = 0$, then $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$ with not all $a_i = 0$, contradicting the fact that T is linearly independent. Hence, $b \neq 0$, and thus

$$\mathbf{w} = (-\frac{a_1}{b})\mathbf{v}_1 + \cdots + (-\frac{a_n}{b})\mathbf{v}_n.$$

(f) We assumed that $\mathbf{w} \notin \text{span}(T)$. However, the conclusion of part (e) shows that $\mathbf{w} \in \text{span}(T)$. Since both clearly can not be true, we have a contradiction. Hence, our assumption that there is a $\mathbf{w} \in \mathcal{W}$ with $\mathbf{w} \notin \text{span}(T)$ is false. That is, there are no vectors in $\mathcal{W}$ outside of span(T). Thus, $\mathcal{W} \subseteq \text{span}(T)$.

(g) From the conclusions of parts (c) and (f), $\mathcal{W} = \text{span}(T)$. Now, T is a basis for $\mathcal{W}$, since it is linearly independent by assumption (part (c)), and spans $\mathcal{W}$. Hence, $\mathcal{W}$ is finite dimensional because T is a finite basis (containing n elements). Also, using part (b),

$$\dim(\mathcal{W}) = |T| = n \leq \dim(\mathcal{V}).$$

(h) Suppose $\dim(\mathcal{W}) = \dim(\mathcal{V})$, and let B be a basis for $\mathcal{W}$. Then B is a linearly independent subset of $\mathcal{V}$ (see the discussion in part (b)) with $|B| = \dim(\mathcal{W}) = \dim(\mathcal{V})$. Part (2) of Theorem 4.12 then shows that B is a basis for $\mathcal{V}$. Hence, $\mathcal{W} = \text{span}(B) = \mathcal{V}$.

(i) The converse of part (h) is "If $\mathcal{W} = \mathcal{V}$, then $\dim(\mathcal{W}) = \dim(\mathcal{V})$." This is obviously true, since the dimension of a finite dimensional vector space is unique.

(22) We need to find a finite basis for $\mathcal{V}$. If S itself is linearly independent, then S is a finite basis, and we are finished. If S is linearly dependent, then part (a) of Exercise 12 in Section 4.4 shows that there is a redundant vector $\mathbf{v} \in S$ such that $\text{span}(S - \{\mathbf{v}\}) = \text{span}(S) = \mathcal{V}$. Let $S_1 = S - \{\mathbf{v}\}$.
If S_1 is linearly independent, we are finished, since it is a finite basis for $\mathcal{V}$. Otherwise, repeat the above process, producing a subset S_2 of S_1 with $\text{span}(S_2) = \mathcal{V}$. Continue in a similar manner until a linearly independent subset S_i is produced with $\text{span}(S_i) = \mathcal{V}$. Then S_i is a finite basis for $\mathcal{V}$. This will take a finite number of steps because there are at most $|S|$ vectors overall that can be removed. (This solution can be expressed more formally using induction.)

(23) Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\} \subseteq \mathbb{R}^n$ be a basis for $\mathcal{V}$. Let $\mathbf{A}$ be the $(n-1) \times n$ matrix whose rows are the vectors in B, and consider the homogeneous system $\mathbf{AX} = \mathbf{0}$. Now, by Corollary 2.3, the system has a nontrivial solution $\mathbf{x}$. Therefore, $\mathbf{Ax} = \mathbf{0}$. That is, $\mathbf{x} \cdot \mathbf{v}_i = 0$ for each $1 \leq i \leq n-1$. Now, suppose $\mathbf{v} \in \mathcal{V}$. Then there are $a_1, \ldots, a_{n-1} \in \mathbb{R}$ such that $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_{n-1}\mathbf{v}_{n-1}$. Hence,

$$\mathbf{x} \cdot \mathbf{v} = \mathbf{x} \cdot a_1\mathbf{v}_1 + \cdots + \mathbf{x} \cdot a_{n-1}\mathbf{v}_{n-1} = 0 + \cdots + 0 = 0.$$

Therefore $\mathcal{V} \subseteq \{\mathbf{v} \in \mathbb{R}^n \,|\, \mathbf{x} \cdot \mathbf{v} = 0\}$. Let $\mathcal{W} = \{\mathbf{v} \in \mathbb{R}^n \,|\, \mathbf{x} \cdot \mathbf{v} = 0\}$. So, $\mathcal{V} \subseteq \mathcal{W}$. Now, it is easy to see that $\mathcal{W}$ is a subspace of $\mathbb{R}^n$. (Clearly $\mathbf{0} \in \mathcal{W}$ since $\mathbf{x} \cdot \mathbf{0} = 0$. Hence $\mathcal{W}$ is nonempty. Next, if $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}$, then

$$\mathbf{x} \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{x} \cdot \mathbf{w}_1 + \mathbf{x} \cdot \mathbf{w}_2 = 0 + 0 = 0.$$

Thus, $(\mathbf{w}_1 + \mathbf{w}_2) \in \mathcal{W}$, and so $\mathcal{W}$ is closed under addition. Also, if $\mathbf{w} \in \mathcal{W}$ and $c \in \mathbb{R}$, then

$$\mathbf{x} \cdot (c\mathbf{w}) = c(\mathbf{x} \cdot \mathbf{w}) = c(0) = 0.$$

Therefore $(c\mathbf{w}) \in \mathcal{W}$, and $\mathcal{W}$ is closed under scalar multiplication. Hence, by Theorem 4.2, $\mathcal{W}$ is a subspace of $\mathbb{R}^n$.) If $\mathcal{V} = \mathcal{W}$, we are done. Suppose, then, that $\mathcal{V} \neq \mathcal{W}$. Then, by Theorem 4.13,

$$n - 1 = \dim(\mathcal{V}) < \dim(\mathcal{W}) \leq \dim(\mathbb{R}^n) = n.$$

Hence, $\dim(\mathcal{W}) = n$, and so $\mathcal{W} = \mathbb{R}^n$, by Theorem 4.13. But $\mathbf{x} \notin \mathcal{W}$, since $\mathbf{x} \cdot \mathbf{x} \neq 0$, because $\mathbf{x}$ is nontrivial. This contradiction completes the proof.

(24) (a) T (b) F (c) F (d) F (e) F (f) T (g) F (h) F (i) T

Section 4.6

Problems 1 through 12 ask you to construct a basis for some vector space. The answers here are those you will most likely get using the techniques in the textbook. However, other answers are possible.

(1) (a) $\{[1,0,0,2,-2],[0,1,0,0,1],[0,0,1,-1,0]\}$
(b) $\{[1,0,-\frac{1}{5},\frac{6}{5},\frac{3}{5}],[0,1,-\frac{1}{5},-\frac{9}{5},-\frac{2}{5}]\}$
(c) $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_4,\mathbf{e}_5\}$
(d) $\{[1,0,0,-2,-\frac{13}{4}],[0,1,0,3,\frac{9}{2}],[0,0,1,0,-\frac{1}{4}]\}$

(2) $\{x^3-3x, x^2-x, 1\}$

(3) $\left\{\begin{bmatrix} 1 & 0 \\ \frac{4}{3} & \frac{1}{3} \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

(4) (a) $\{[1,3,-2],\ [2,1,4],\ [0,1,-1]\}$
(b) $\{[1,4,-2],\ [2,-8,5],\ [0,-7,2]\}$
(c) $\{[3,-2,2],\ [1,2,-1],\ [3,-2,7]\}$
(d) $\{[3,1,0],\ [2,-1,7],\ [1,5,7]\}$

(5) (a) $\{x^3-8x^2+1, 3x^3-2x^2+x, 4x^3+2x-10, x^3-20x^2-x+12\}$
(b) $\{-2x^3+x+2, 3x^3-x^2+4x+6, 8x^3+x^2+6x+10\}$

(6) (a) $\{[3,1,-2],\ [6,2,-3]\}$
(b) $\{[4,7,1],\ [1,0,0]\}$
(c) $\{\mathbf{e}_1,\mathbf{e}_3\}$
(d) $\{[1,-3,0],[0,1,1]\}$

(7) (a) $\{x^3,x^2,x\}$
(b) $\{1,x,x^2\}$
(c) $\{x^3+x^2,x,1\}$
(d) $\{8x^3, 8x^3+x^2, 8x^3+x, 8x^3+1\}$

(8) (a) One answer is $\{\mathbf{\Psi}_{ij}\,|\,1\le i,j\le 3\}$, where $\mathbf{\Psi}_{ij}$ is the 3×3 matrix with (i,j) entry $=1$ and all other entries 0.

(b) $\left\{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},\right.$

$\left.\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\}$

(d) $\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},\right.$

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

(9) Let $\mathcal{V}$ be the subspace of $\mathcal{M}_{22}$ consisting of all 2×2 symmetric matrices. Let S be the set of nonsingular matrices in $\mathcal{V}$, and let $\mathcal{W} = \text{span}(S) = \text{span}(\{\text{nonsingular, symmetric } 2 \times 2 \text{ matrices}\})$. Using the strategy from Exercise 7, we reduce S to a basis for $\mathcal{W}$ using the Independence Test Method, even though S is infinite.

The strategy is to guess a *finite* subset Y of S that spans $\mathcal{W}$. We then use the Independence Test Method on Y to find the desired basis. We try to pick vectors for Y whose forms are as simple as possible to make computation easier. In this case, we choose the set of all nonsingular symmetric 2×2 matrices having only zeroes and ones as entries. That is,

$$Y = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\}.$$

Now, before continuing, we must ensure that $\text{span}(Y) = \mathcal{W}$. That is, we must show every nonsingular symmetric 2×2 matrix is in $\text{span}(Y)$. In fact, we will show every symmetric 2×2 matrix is in $\text{span}(Y)$ by finding real numbers w, x, y, and z so that

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = w\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + y\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + z\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Thus, we must prove that the system

$$\begin{cases} w + x = a \\ x + y + z = b \\ x + y + z = b \\ w + z = c \end{cases}$$

has solutions for w, x, y, and z in terms of a, b, and c. But $w = 0$, $x = a$, $y = b - a - c$, $z = c$ certainly satisfies the system. Hence, $\mathcal{V} \subseteq \text{span}(Y)$. Since $\text{span}(Y) \subseteq \mathcal{V}$, we have $\text{span}(Y) = \mathcal{V} = \mathcal{W}$.

We can now use the Independence Test Method on Y. We express the matrices in Y as corresponding vectors in $\mathbb{R}^4$ and create the matrix with these vectors as columns, as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \text{which reduces to} \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then, the desired basis is

$$B = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\},$$

the elements of Y corresponding to the pivot columns of C.

(10) (a) $\{[1, -3, 0, 1, 4], [2, 2, 1, -3, 1], [1, 0, 0, 0, 0], [0, 1, 0, 0, 0], [0, 0, 1, 0, 0]\}$

(b) $\{[1,1,1,1,1],[0,1,1,1,1],[0,0,1,1,1],[0,0,1,0,0],[0,0,0,1,0]\}$

(c) $\{[1,0,-1,0,0],[0,1,-1,1,0],[2,3,-8,-1,0],[1,0,0,0,0],[0,0,0,0,1]\}$

(11) (a) $\{x^3-x^2, x^4-3x^3+5x^2-x, x^4, x^3, 1\}$

(b) $\{3x-2, x^3-6x+4, x^4, x^2, x\}$

(c) $\{x^4-x^3+x^2-x+1, x^3-x^2+x-1, x^2-x+1, x^2, x\}$

(12) (a) $\left\{\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$

(b) $\left\{\begin{bmatrix} 0 & -2 \\ 1 & 0 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & -3 \\ 0 & 1 \\ 3 & -6 \end{bmatrix}, \begin{bmatrix} 4 & -13 \\ 2 & 3 \\ 7 & -14 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix} 3 & -1 \\ 2 & -6 \\ 5 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ -4 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 6 & 2 \\ -2 & -9 \\ 10 & 2 \end{bmatrix}, \begin{bmatrix} 3 & -4 \\ 8 & -9 \\ 5 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$

(13) (a) 2 (b) 8 (c) 4 (d) 3

(14) (a) A basis for $\mathcal{U}_n$ is $\{\boldsymbol{\Psi}_{ij} \,|\, 1 \le i \le j \le n\}$, where $\boldsymbol{\Psi}_{ij}$ is the $n \times n$ matrix having a zero in every entry, except for the (i,j)th entry, which equals 1. Since there are $n-i+1$ entries in the ith row of an $n \times n$ matrix on or above the main diagonal, this basis contains

$$\sum_{i=1}^{n}(n-i+1) = \sum_{i=1}^{n} n - \sum_{i=1}^{n} i + \sum_{i=1}^{n} 1 = n^2 - \frac{n(n+1)}{2} + n = \frac{n^2+n}{2}$$

elements.

Similarly, a basis for $\mathcal{L}_n$ is $\{\boldsymbol{\Psi}_{ij} \,|\, 1 \le j \le i \le n\}$, and a basis for the symmetric $n \times n$ matrices is $\{\boldsymbol{\Psi}_{ij} + \boldsymbol{\Psi}_{ij}^T \,|\, 1 \le i \le j \le n\}$.

(b) $(n^2-n)/2$

(15) Let $\mathbf{B}$ be the reduced row echelon form of $\mathbf{A}$. Each basis element of $S_{\mathbf{A}}$ is found by setting one independent variable of the system $\mathbf{AX} = \mathbf{O}$ equal to 1 and the remaining independent variables equal to zero. (The values of the dependent variables are then determined from these values.) Since there is one independent variable for each nonpivot column in $\mathbf{B}$, $\dim(S_{\mathbf{A}})$ = number of nonpivot columns of $\mathbf{B}$. But, clearly, $\text{rank}(\mathbf{A})$ = number of pivot columns in $\mathbf{B}$.

So, $\dim(S_{\mathbf{A}}) + \text{rank}(\mathbf{A})$ = total number of columns in $\mathbf{B} = n$.

(16) Let $\mathcal{V}$ and S be as given in the statement of the theorem. Let $A = \{k \,|\, \text{a set } T \text{ exists with } T \subseteq S,\ |T| = k, \text{ and } T \text{ linearly independent}\}$.

The empty set is linearly independent and is a subset of S. Hence $0 \in A$, and A is nonempty.

Suppose $k \in A$. Then, every linearly independent subset T of S is also a linearly independent subset of $\mathcal{V}$. Hence, using part (2) of Theorem 4.12 on T shows that $k = |T| \le \dim(\mathcal{V})$.

Suppose n is the largest number in A, which must exist, since A is nonempty and all its elements are $\le \dim(\mathcal{V})$. Let B be a linearly independent subset of S with $|B| = n$, which exists because $n \in A$. We want to show that B is a basis for $\mathcal{V} = \text{span}(S)$. Now, B is given as linearly independent. We must show that B spans $\mathcal{V}$.

Now, $B \subseteq S$, and so $\text{span}(B) \subseteq \text{span}(S)$ (by Corollary 4.6). We need to show that $\text{span}(S) \subseteq \text{span}(B)$. To do this, we will prove that $S \subseteq \text{span}(B)$, for then, parts (2) and (3) of Theorem 4.5 will give us that $\text{span}(S) \subseteq \text{span}(B)$.

We prove that $S \subseteq \text{span}(B)$: Suppose $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathbf{v} \in S$. If $\mathbf{v} \in B$, then, clearly $\mathbf{v} \in \text{span}(B)$, and we are finished. So, suppose that $\mathbf{v} \notin B$. Then $B \cup \{\mathbf{v}\}$ is a subset of S containing $n+1$ elements. Therefore, since n is the largest number in A, $B \cup \{\mathbf{v}\}$ must be linearly dependent. Hence, there exist scalars $a_1, \ldots, a_n, b$ such that $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n + b\mathbf{v} = \mathbf{0}$, with not all coefficients equal to zero. But if $b = 0$, then $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$ with not all $a_i = 0$, contradicting the fact that B is linearly independent. Hence, $b \neq 0$, and thus

$$\mathbf{v} = (-\frac{a_1}{b})\mathbf{v}_1 + \cdots + (-\frac{a_n}{b})\mathbf{v}_n.$$

This shows that $\mathbf{v} \in \text{span}(B)$, completing the proof that B spans $\mathcal{V}$. Hence, the subset B of S is the desired basis for $\mathcal{V}$, finishing the proof of Theorem 4.14.

(17) (a) Let B' be a basis for $\mathcal{W}$. Expand B' to a basis B for $\mathcal{V}$ (using Theorem 4.15).

(b) No; consider the subspace $\mathcal{W}$ of $\mathbb{R}^3$ given by $\mathcal{W} = \{[a,0,0] \,|\, a \in \mathbb{R}\}$. No subset of $B = \{[1,1,0],[1,-1,0],[0,0,1]\}$ (a basis for $\mathbb{R}^3$) is a basis for $\mathcal{W}$.

(c) Yes; consider $\mathcal{Y} = \text{span}(B')$.

(18) (a) Let B_1 be a basis for $\mathcal{W}$. Expand B_1 to a basis for $\mathcal{V}$ (using Theorem 4.15). Let $B_2 = B - B_1$, and let $\mathcal{W}' = \text{span}(B_2)$. Suppose $B_1 = \{\mathbf{v}_1, \ldots \mathbf{v}_k\}$ and $B_2 = \{\mathbf{u}_1, \ldots, \mathbf{u}_l\}$. Since B is a basis for $\mathcal{V}$, all $\mathbf{v} \in \mathcal{V}$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{w}'$ with $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k \in \mathcal{W}$ and $\mathbf{w}' = b_1\mathbf{u}_1 + \cdots + b_l\mathbf{u}_l \in \mathcal{W}'$.

For uniqueness, suppose $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_1' = \mathbf{w}_2 + \mathbf{w}_2'$, with $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}$ and $\mathbf{w}_1', \mathbf{w}_2' \in \mathcal{W}'$. Then $\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{w}_2' - \mathbf{w}_1' \in \mathcal{W} \cap \mathcal{W}' = \{\mathbf{0}\}$. Hence, $\mathbf{w}_1 = \mathbf{w}_2$ and $\mathbf{w}_1' = \mathbf{w}_2'$.

(b) In $\mathbb{R}^3$, consider $\mathcal{W} = \{[a,b,0] \,|\, a,b \in \mathbb{R}\}$. We could let $\mathcal{W}' = \{[0,0,c] \,|\, c \in \mathbb{R}\}$ or $\mathcal{W}' = \{[0,c,c] \,|\, c \in \mathbb{R}\}$.

(19) (a) Let $\mathbf{A}$ be the matrix whose rows are the n-vectors in S. Let $\mathbf{C}$ be the reduced row echelon form matrix for $\mathbf{A}$. If the nonzero rows of $\mathbf{C}$ form the standard basis for $\mathbb{R}^n$, then $\text{span}(S) = \text{span}(\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}) = \mathbb{R}^n$.

Conversely, if $\text{span}(S) = \mathbb{R}^n$, then there must be at least n nonzero rows of $\mathbf{C}$ by part (1) of Theorem 4.12. But this means there are at least n rows with pivots, and hence all n columns have pivots. Thus the nonzero rows of $\mathbf{C}$ are $\mathbf{e}_1, \ldots, \mathbf{e}_n$, which form the standard basis for $\mathbb{R}^n$.

(b) Let $\mathbf{A}$ and $\mathbf{C}$ be as in the answer to part (a). Note that $\mathbf{A}$ and $\mathbf{C}$ are $n \times n$ matrices. Now, since $|B| = n$, B is a basis for $\mathbb{R}^n \iff \text{span}(B) = \mathbb{R}^n$ (by Theorem 4.12) $\iff$ the rows of $\mathbf{C}$ form the standard basis for $\mathbb{R}^n$ (from part (a)) $\iff \mathbf{C} = \mathbf{I}_n \iff |\mathbf{A}| \neq 0$.

(20) Let $\mathbf{A}$ be an $m \times n$ matrix and let $S \subseteq \mathbb{R}^n$ be the set of the m rows of $\mathbf{A}$.

(a) Suppose $\mathbf{C}$ is the reduced row echelon form of $\mathbf{A}$. Then the Simplified Span Method shows that the nonzero rows of $\mathbf{C}$ form a basis for $\text{span}(S)$. But the number of such rows is $\text{rank}(\mathbf{A})$. Hence, $\dim(\text{span}(S)) = \text{rank}(\mathbf{A})$.

(b) Let $\mathbf{D}$ be the reduced row echelon form of $\mathbf{A}^T$. By the Independence Test Method, the pivot columns of $\mathbf{D}$ correspond to the columns of $\mathbf{A}^T$ which make up a basis for $\text{span}(S)$. Hence, the number of pivot columns of $\mathbf{D}$ equals $\dim(\text{span}(S))$. However, the number of pivot columns of $\mathbf{D}$ equals the number of nonzero rows of $\mathbf{D}$ (each pivot column shares the single pivot element with a unique pivot row, and vice-versa), which equals $\text{rank}(\mathbf{A}^T)$.

(c) Both ranks equal dim(span(S)), and so they must be equal.

(21) Use the fact that $\sin(\alpha_i + \beta_j) = (\cos\alpha_i)\sin\beta_j + (\sin\alpha_i)\cos\beta_j$. Note that the ith row of $\mathbf{A}$ is then $(\cos\alpha_i)\mathbf{x}_1 + (\sin\alpha_i)\mathbf{x}_2$. Hence each row of $\mathbf{A}$ is a linear combination of $\{\mathbf{x}_1, \mathbf{x}_2\}$, and so $\dim(\text{span}(\{\text{rows of } \mathbf{A}\})) \le 2 < n$. Hence, $\mathbf{A}$ has less than n pivots when row reduced, and so $|\mathbf{A}| = 0$.

(22) (a) T (b) T (c) F (d) T (e) F (f) F (g) F

Section 4.7

(1) (a) $[\mathbf{v}]_B = [7, -1, -5]$
(b) $[\mathbf{v}]_B = [3, -2, -1, -2]$
(c) $[\mathbf{v}]_B = [-2, 4, -5]$
(d) $[\mathbf{v}]_B = [10, 3, 2, -1]$
(e) $[\mathbf{v}]_B = [4, -5, 3]$
(f) $[\mathbf{v}]_B = [-3, 2, 3]$
(g) $[\mathbf{v}]_B = [-1, 4, -2]$
(h) $[\mathbf{v}]_B = [2, -3, 1]$
(i) $[\mathbf{v}]_B = [5, 2, -6]$
(j) $[\mathbf{v}]_B = [5, -2]$

(2) (a) $\begin{bmatrix} -102 & 20 & 3 \\ 67 & -13 & -2 \\ 36 & -7 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} -2 & -6 & -1 \\ 3 & 6 & 1 \\ -6 & -7 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 20 & -30 & -69 \\ 24 & -24 & -80 \\ -9 & 11 & 31 \end{bmatrix}$

(d) $\begin{bmatrix} -1 & -4 & 2 & -9 \\ 4 & 5 & 1 & 3 \\ 0 & 2 & -3 & 1 \\ -4 & -13 & 13 & -15 \end{bmatrix}$

(e) $\begin{bmatrix} 3 & 2 \\ -2 & -3 \end{bmatrix}$

(f) $\begin{bmatrix} 6 & 1 & 2 \\ 1 & 1 & 2 \\ -1 & -1 & -3 \end{bmatrix}$

(g) $\begin{bmatrix} -3 & -2 & 4 \\ 2 & 3 & -5 \\ 2 & 1 & 3 \end{bmatrix}$

(3) $[(2,6)]_B = (2,2)$ (see Figure 11)

(4) (a) $\mathbf{P} = \begin{bmatrix} 13 & 31 \\ -18 & -43 \end{bmatrix}$, $\mathbf{Q} = \begin{bmatrix} -11 & -8 \\ 29 & 21 \end{bmatrix}$, $\mathbf{T} = \begin{bmatrix} 1 & 3 \\ -1 & -4 \end{bmatrix}$

(b) $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, $\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{T} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

(c) $\mathbf{P} = \begin{bmatrix} 2 & 8 & 13 \\ -6 & -25 & -43 \\ 11 & 45 & 76 \end{bmatrix}$, $\mathbf{Q} = \begin{bmatrix} -24 & -2 & 1 \\ 30 & 3 & -1 \\ 139 & 13 & -5 \end{bmatrix}$, $\mathbf{T} = \begin{bmatrix} -25 & -97 & -150 \\ 31 & 120 & 185 \\ 145 & 562 & 868 \end{bmatrix}$

(d) $\mathbf{P} = \begin{bmatrix} -3 & -45 & -77 & -38 \\ -41 & -250 & -420 & -205 \\ 19 & 113 & 191 & 93 \\ -4 & -22 & -37 & -18 \end{bmatrix}$, $\mathbf{Q} = \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 2 & -1 \\ 1 & 2 & 6 & 6 \\ 3 & -1 & 6 & 36 \end{bmatrix}$, $\mathbf{T} = \begin{bmatrix} 4 & 2 & 3 & 1 \\ 1 & -2 & -1 & -1 \\ 5 & 1 & 7 & 2 \\ 2 & 1 & 3 & 1 \end{bmatrix}$

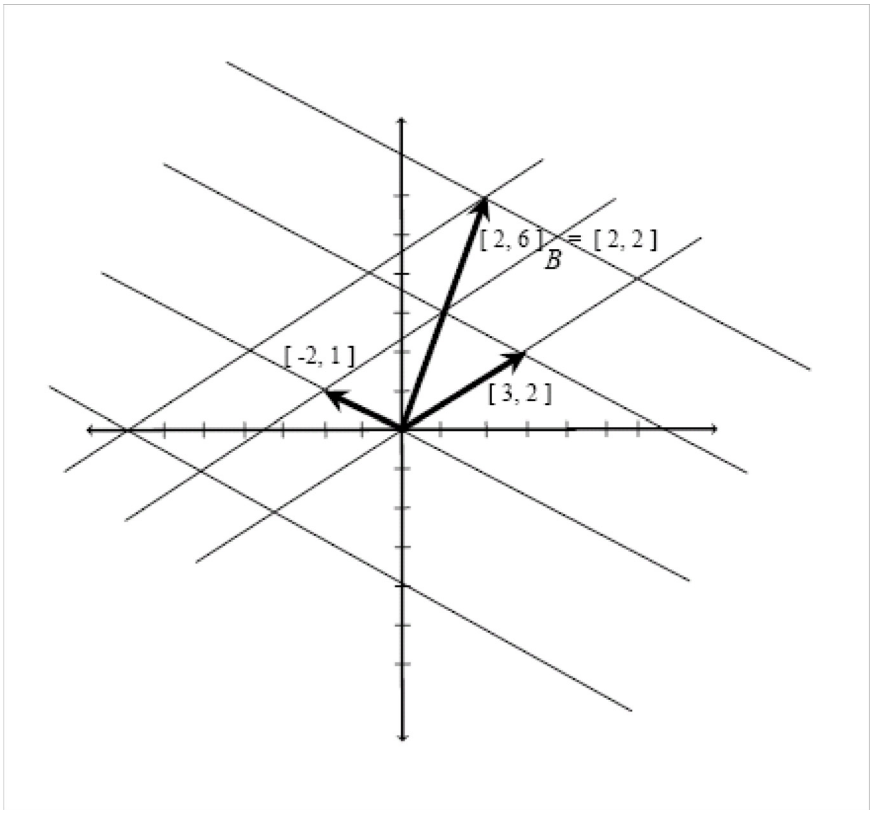


Figure 11: $(2, 6)$ in B-coordinates

(5) (a) $C = ([1, -4, 0, -2, 0], [0, 0, 1, 4, 0], [0, 0, 0, 0, 1])$; $\mathbf{P} = \begin{bmatrix} 1 & 6 & 3 \\ 1 & 5 & 3 \\ 1 & 3 & 2 \end{bmatrix}$;

$\mathbf{Q} = \mathbf{P}^{-1} = \begin{bmatrix} 1 & -3 & 3 \\ 1 & -1 & 0 \\ -2 & 3 & -1 \end{bmatrix}$; $[\mathbf{v}]_B = [17, 4, -13]$; $[\mathbf{v}]_C = [2, -2, 3]$

(b) $C = ([1, 0, 17, 0, 21], [0, 1, 3, 0, 5], [0, 0, 0, 1, 2])$; $\mathbf{P} = \begin{bmatrix} 1 & 3 & 1 \\ -5 & -14 & -4 \\ 0 & 2 & 3 \end{bmatrix}$;

$\mathbf{Q} = \mathbf{P}^{-1} = \begin{bmatrix} -34 & -7 & 2 \\ 15 & 3 & -1 \\ -10 & -2 & 1 \end{bmatrix}$; $[\mathbf{v}]_B = [5, -2, 3]$; $[\mathbf{v}]_C = [2, -9, 5]$

(c) $C = ([1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1])$;

$$\mathbf{P} = \begin{bmatrix} 3 & 6 & -4 & -2 \\ -1 & 7 & -3 & 0 \\ 4 & -3 & 3 & 1 \\ 6 & -2 & 4 & 2 \end{bmatrix}; \mathbf{Q} = \mathbf{P}^{-1} = \begin{bmatrix} 1 & -4 & -12 & 7 \\ -2 & 9 & 27 & -\frac{31}{2} \\ -5 & 22 & 67 & -\frac{77}{2} \\ 5 & -23 & -71 & 41 \end{bmatrix};$$

$[\mathbf{v}]_B = [2, 1, -3, 7]$; $[\mathbf{v}]_C = [10, 14, 3, 12]$

(6) For each i, $\mathbf{v}_i = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 1\mathbf{v}_i + \cdots + 0\mathbf{v}_n$, and so $[\mathbf{v}_i]_B = [0, 0, \ldots, 1, \ldots 0] = \mathbf{e}_i$.

(7) (a) Transition matrices to C_1, C_2, C_3, C_4, and C_5, respectively:

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

(b) Let $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ and $C = (\mathbf{v}_{\alpha_1}, \ldots, \mathbf{v}_{\alpha_n})$, where $\alpha_1, \alpha_2, \cdots, \alpha_n$ is a rearrangement of the numbers $1, 2, \cdots, n$. Let $\mathbf{P}$ be the transition matrix from B to C. The goal is to show, for $1 \le i \le n$, that the ith row of $\mathbf{P}$ equals $\mathbf{e}_{\alpha_i}$. This is true iff (for $1 \le i \le n$) the α_ith column of $\mathbf{P}$ equals $\mathbf{e}_i$. Now the α_ith column of $\mathbf{P}$ is, by definition, $[\mathbf{v}_{\alpha_i}]_C$. But since $\mathbf{v}_{\alpha_i}$ is the ith vector in C, $[\mathbf{v}_{\alpha_i}]_C = \mathbf{e}_i$, completing the proof.

(8) (a) Following the Transition Matrix Method, it is necessary to row reduce

$$\left[\begin{array}{cccc|c} \text{First} & \text{Second} & & n\text{th} & \\ \text{vector} & \text{vector} & \cdots & \text{vector} & \mathbf{I}_n \\ \text{in} & \text{in} & & \text{in} & \\ C & C & & C & \end{array}\right].$$

The algorithm from Section 2.4 for calculating inverses shows that the result obtained in the last n columns is the desired inverse.

(b) Following the Transition Matrix Method, it is necessary to row reduce

$$\left[\begin{array}{c|cccc} & \text{First} & \text{Second} & & n\text{th} \\ \mathbf{I}_n & \text{vector} & \text{vector} & \cdots & \text{vector} \\ & \text{in} & \text{in} & & \text{in} \\ & B & B & & B \end{array}\right].$$

Clearly this is already in row reduced form, and so the result is as desired.

(9) Let S be the standard basis for $\mathbb{R}^n$. Exercise 8(b) shows that $\mathbf{P}$ is the transition matrix from B to S, and Exercise 8(a) shows that $\mathbf{Q}^{-1}$ is the transition matrix from S to C. Theorem 4.18 finishes the proof.

(10) $C = ([-142, 64, 167], [-53, 24, 63], [-246, 111, 290])$

(11) (a) Easily verified with straightforward computations.

(b) $[\mathbf{v}]_B = [1, 2, -3]$; $\mathbf{Av} = [14, 2, 5]$; $\mathbf{D}[\mathbf{v}]_B = [\mathbf{Av}]_B = [2, -2, 3]$

(12) (a) $\lambda_1 = 2$, $\mathbf{v}_1 = [2, 1, 5]$; $\lambda_2 = -1$, $\mathbf{v}_2 = [19, 2, 31]$; $\lambda_3 = -3$, $\mathbf{v}_3 = [6, -2, 5]$

(b) $\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 19 & 6 \\ 1 & 2 & -2 \\ 5 & 31 & 5 \end{bmatrix}$

(13) Let $\mathcal{V}$, B, and the a_i's and $\mathbf{w}$'s be as given in the statement of the theorem.
Proof of Part (1): Suppose that $[\mathbf{w}_1]_B = [b_1, \ldots, b_n]$ and $[\mathbf{w}_2]_B = [c_1, \ldots, c_n]$. Then,

$$\mathbf{w}_1 = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n \text{ and } \mathbf{w}_2 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Hence,

$$\mathbf{w}_1 + \mathbf{w}_2 = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n + c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = (b_1 + c_1)\mathbf{v}_1 + \cdots + (b_n + c_n)\mathbf{v}_n,$$

implying

$$[\mathbf{w}_1 + \mathbf{w}_2]_B = [(b_1 + c_1), \ldots, (b_n + c_n)],$$

which equals $[\mathbf{w}_1]_B + [\mathbf{w}_2]_B$.
Proof of Part (2): Suppose that $[\mathbf{w}_1]_B = [b_1, \ldots, b_n]$. Then, $\mathbf{w}_1 = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$. Hence,

$$a_1\mathbf{w}_1 = a_1(b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n) = a_1b_1\mathbf{v}_1 + \cdots + a_1b_n\mathbf{v}_n,$$

implying $[a_1\mathbf{w}_1]_B = [a_1b_1, \ldots, a_1b_n]$, which equals $a_1[b_1, \ldots, b_n] = a_1[\mathbf{w}]_B$.
Proof of Part (3): Use induction on k.
Base Step ($k = 1$): This is just part (2).
Inductive Step: Assume true for any linear combination of k vectors, and prove true for a linear combination of $k + 1$ vectors. Now,

$$\begin{aligned} [a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k + a_{k+1}\mathbf{w}_{k+1}]_B &= [a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k]_B + [a_{k+1}\mathbf{w}_{k+1}]_B \quad \text{(by part (1))} \\ &= [a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k]_B + a_{k+1}[\mathbf{w}_{k+1}]_B \quad \text{(by part (2))} \\ &= a_1[\mathbf{w}_1]_B + \cdots + a_k[\mathbf{w}_k]_B + a_{k+1}[\mathbf{w}_{k+1}]_B \\ &\qquad \text{(by the inductive hypothesis).} \end{aligned}$$

(14) Let $\mathbf{v} \in \mathcal{V}$. Then, using Theorem 4.17, $(\mathbf{QP})[\mathbf{v}]_B = \mathbf{Q}[\mathbf{v}]_C = [\mathbf{v}]_D$. Hence, by Theorem 4.17, $\mathbf{QP}$ is the (unique) transition matrix from B to D.

(15) Let B, C, $\mathcal{V}$, and $\mathbf{P}$ be as given in the statement of the theorem. Let $\mathbf{Q}$ be the transition matrix from C to B. Then, by Theorem 4.18, $\mathbf{QP}$ is the transition matrix from B to B. But, clearly, since for all $\mathbf{v} \in \mathcal{V}$, $\mathbf{I}[\mathbf{v}]_B = [\mathbf{v}]_B$, Theorem 4.17 implies that $\mathbf{I}$ must be the transition matrix from B to B. Hence, $\mathbf{QP} = \mathbf{I}$, finishing the proof of the theorem.

(16) Let B, $\mathcal{V}$, and $\mathbf{P}$ be as given in the statement of the problem. Let $C = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, where $\mathbf{v}_i$ is the vector in $\mathcal{V}$ such that $[\mathbf{v}_i]_B = i$th column of $\mathbf{P}^{-1}$. Then C is a basis for $\mathcal{V}$ because it has n linearly independent vectors, since the columns of $\mathbf{P}^{-1}$ are linearly independent. By definition, $\mathbf{P}^{-1}$ is the transition matrix from C to B. Hence, by Theorem 4.19, $\mathbf{P}$ is the transition matrix from B to C.

(17) (a) F (b) T (c) T (d) F (e) F (f) T (g) F

Chapter 4 Review Exercises

(1) This is a vector space. To prove this, modify the proof in Example 7 in Section 4.1.

(2) Zero vector $= [-4, 5]$; additive inverse of $[x, y]$ is $[-x-8, -y+10]$

(3) (a), (b), (d), (f) and (g) are not subspaces; (c) and (e) are subspaces

(4) (a) $\text{span}(S) = \{[a, b, c, 5a-3b+c] \mid a, b, c \in \mathbb{R}\} \neq \mathbb{R}^4$

(b) Basis $= \{[1, 0, 0, 5], [0, 1, 0, -3], [0, 0, 1, 1]\}$; $\dim(\text{span}(S)) = 3$

(5) (a) $\text{span}(S) = \{ax^3 + bx^2 + (-3a+2b)x + c \mid a, b, c \in \mathbb{R}\} \neq \mathcal{P}_3$

(b) Basis $= \{x^3 - 3x, x^2 + 2x, 1\}$; $\dim(\text{span}(S)) = 3$

(6) (a) $\text{span}(S) = \left\{ \left[\begin{array}{ccc} a & b & 4a-3b \\ c & -2a+b-c & d \end{array} \right] \middle| \; a, b, c, d \in \mathbb{R} \right\} \neq \mathcal{M}_{23}$

(b) Basis $= \left\{ \left[\begin{array}{ccc} 1 & 0 & 4 \\ 0 & -2 & 0 \end{array} \right], \left[\begin{array}{ccc} 0 & 1 & -3 \\ 0 & 1 & 0 \end{array} \right], \left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & -1 & 0 \end{array} \right], \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right] \right\}$; $\dim(\text{span}(S)) = 4$

(7) (a) S is linearly independent

(b) S itself is a basis for $\text{span}(S)$. S spans $\mathbb{R}^3$.

(c) No, by Theorem 4.9

(8) (a) S is linearly dependent; $x^3 - 2x^2 - x + 2 = 3(-5x^3 + 2x^2 + 5x - 2) + 8(2x^3 - x^2 - 2x + 1)$

(b) The subset

$$\{-5x^3 + 2x^2 + 5x - 2,\ 2x^3 - x^2 - 2x + 1,\ -2x^3 + 2x^2 + 3x - 5\}$$

of S is a basis for $\text{span}(S)$. S does not span $\mathcal{P}_3$.

(c) Yes, by part (b) of Exercise 21 in Section 4.4. One such alternate linear combination is

$$-5\left(-5x^3 + 2x^2 + 5x - 2\right) - 5\left(2x^3 - x^2 - 2x + 1\right) + 1\left(x^3 - 2x^2 - x + 2\right) - 1\left(-2x^3 + 2x^2 + 3x - 5\right).$$

(9) (a) Linearly dependent;

$$\left[\begin{array}{cc} 4 & 6 \\ 3 & 4 \\ 2 & 5 \end{array} \right] = -\left[\begin{array}{cc} -10 & -14 \\ -10 & -8 \\ -6 & -12 \end{array} \right] - 2\left[\begin{array}{cc} 7 & 12 \\ 5 & 7 \\ 3 & 10 \end{array} \right] + \left[\begin{array}{cc} 8 & 16 \\ 3 & 10 \\ 2 & 13 \end{array} \right]$$

(b) $\left\{ \left[\begin{array}{cc} 4 & 0 \\ 11 & -2 \\ 6 & -1 \end{array} \right], \left[\begin{array}{cc} -10 & -14 \\ -10 & -8 \\ -6 & -12 \end{array} \right], \left[\begin{array}{cc} 7 & 12 \\ 5 & 7 \\ 3 & 10 \end{array} \right], \left[\begin{array}{cc} 8 & 16 \\ 3 & 10 \\ 2 & 13 \end{array} \right], \left[\begin{array}{cc} 6 & 11 \\ 4 & 7 \\ 3 & 9 \end{array} \right] \right\}$; S does not span $\mathcal{M}_{32}$

(c) Yes, by part (b) of Exercise 21 in Section 4.4. One such alternate linear combination is

$$2\left[\begin{array}{cc} 4 & 0 \\ 11 & -2 \\ 6 & -1 \end{array} \right] + \left[\begin{array}{cc} -10 & -14 \\ -10 & -8 \\ -6 & -12 \end{array} \right] - \left[\begin{array}{cc} 7 & 12 \\ 5 & 7 \\ 3 & 10 \end{array} \right] - \left[\begin{array}{cc} 8 & 16 \\ 3 & 10 \\ 2 & 13 \end{array} \right] + \left[\begin{array}{cc} 4 & 6 \\ 3 & 4 \\ 2 & 5 \end{array} \right] + \left[\begin{array}{cc} 6 & 11 \\ 4 & 7 \\ 3 & 9 \end{array} \right].$$

(10) $\mathbf{v} = 1\mathbf{v}$. Also, since $\mathbf{v} \in \text{span}(S)$, $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, for some $a_1, \ldots, a_n$. These are 2 different ways to express $\mathbf{v}$ as a linear combination of vectors in T.

(11) Use the same technique as in Example 13 in Section 4.4.

(12) (a) The matrix whose rows are the given vectors row reduces to $\mathbf{I}_4$, so the Simplified Span Method shows that the given set spans $\mathbb{R}^4$. Since the set has 4 vectors and $\dim(\mathbb{R}^4) = 4$, part (1) of Theorem 4.12 shows that the given set is a basis for $\mathbb{R}^4$.

(b) Similar to part (a). The matrix $\begin{bmatrix} 2 & 2 & 13 \\ 1 & 0 & 3 \\ 4 & 1 & 16 \end{bmatrix}$ row reduces to $\mathbf{I}_3$.

(c) Similar to part (a). The matrix $\begin{bmatrix} 1 & 5 & 0 & 3 \\ 6 & -1 & 4 & 3 \\ 7 & -4 & 7 & -1 \\ -3 & 7 & -2 & 4 \end{bmatrix}$ row reduces to $\mathbf{I}_4$.

(13) (a) $\mathcal{W}$ nonempty: $\mathbf{0} \in \mathcal{W}$ because $\mathbf{A0} = \mathbf{O}$.
Closure under addition: If $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{W}$, then $\mathbf{A}(\mathbf{X}_1 + \mathbf{X}_2) = \mathbf{AX}_1 + \mathbf{AX}_2 = \mathbf{O} + \mathbf{O} = \mathbf{O}$.
Closure under scalar multiplication: If $\mathbf{X} \in \mathcal{W}$, then $\mathbf{A}(c\mathbf{X}) = c\mathbf{AX} = c\mathbf{O} = \mathbf{O}$.

(b) Basis $= \{[3, 1, 0, 0], [-2, 0, 1, 1]\}$

(c) $\dim(\mathcal{W}) = 2$, $\text{rank}(\mathbf{A}) = 2$; $2 + 2 = 4$

(14) (a) First, we use direct computation to check that every polynomial in B is in $\mathcal{V}$:
If $\mathbf{p}(x) = x^3 - 3x$, then $\mathbf{p}'(x) = 3x^2 - 3$, and so $\mathbf{p}'(1) = 3 - 3 = 0$.
If $\mathbf{p}(x) = x^2 - 2x$, then $\mathbf{p}'(x) = 2x - 2$, and so $\mathbf{p}'(1) = 2 - 2 = 0$.
If $\mathbf{p}(x) = 1$, then $\mathbf{p}'(x) = 0$, and so $\mathbf{p}'(1) = 0$.
Thus, every polynomial in B is actually in $\mathcal{V}$.

Next, we convert the polynomials in B to 4-vectors, and use the Independence Test Method. Now,

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ clearly row reduces to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, so B is linearly independent. Finally, since the polynomial $x \notin \mathcal{V}$, $\dim(\mathcal{V}) < \dim(\mathcal{P}_3) = 4$ by Theorem 4.13. But, part (2) of Theorem 4.12 shows that $|B| \le \dim(\mathcal{V})$. Hence, $3 = |B| \le \dim(\mathcal{V}) < \dim(\mathcal{P}_3) = 4$, and so $\dim(\mathcal{V}) = 3$. Part (2) of Theorem 4.12 then implies that B is a basis for $\mathcal{V}$.

(b) $C = \{1, x^3 - 3x^2 + 3x\}$ is a basis for $\mathcal{W}$ and $\dim(\mathcal{W}) = 2$.

(15) $T = \{[2, -3, 0, 1], [4, 3, 0, 4], [1, 0, 2, 1]\}$

(16) $T = \{x^2 - 2x, x^3 - x, 2x^3 - 2x^2 + 1\}$

(17) $\{[2, 1, -1, 2], [1, -2, 2, -4], [0, 1, 0, 0], [0, 0, 1, 0]\}$

(18) $\left\{ \begin{bmatrix} 3 & -1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & -1 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 0 & 1 \\ 4 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

(19) $\{x^4 + 3x^2 + 1, x^3 - 2x^2 - 1, x + 3\}$

(20) (a) $[\mathbf{v}]_B = [-3, -1, -2]$ (b) $[\mathbf{v}]_B = [4, 2, -3]$ (c) $[\mathbf{v}]_B = [-3, 5, -1]$

(21) (a) $[\mathbf{v}]_B = [27, -62, 6]$; $\mathbf{P} = \begin{bmatrix} 4 & 2 & 5 \\ -1 & 0 & 1 \\ 3 & 1 & -2 \end{bmatrix}$; $[\mathbf{v}]_C = [14, -21, 7]$

(b) $[\mathbf{v}]_B = [-4, -1, 3]$; $\mathbf{P} = \begin{bmatrix} 4 & 1 & -2 \\ 1 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix}$; $[\mathbf{v}]_C = [-23, -6, 6]$

(c) $[\mathbf{v}]_B = [4, -2, 1, -3]$; $\mathbf{P} = \begin{bmatrix} 6 & 2 & 2 & -1 \\ 5 & 0 & 1 & -1 \\ -3 & 1 & -1 & 0 \\ -5 & 2 & 0 & 1 \end{bmatrix}$; $[\mathbf{v}]_C = [25, 24, -15, -27]$

(22) (a) $\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

(b) $\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

(c) $\mathbf{R} = \mathbf{QP} = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & -1 & 1 & 1 \end{bmatrix}$

(d) $\mathbf{R}^{-1} = \begin{bmatrix} 0 & -3 & 2 & 2 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 2 & -2 & -1 \end{bmatrix}$

(23) (a) $\lambda_1 = 0$, $\mathbf{v}_1 = [4, 4, 13]$; $\lambda_2 = 2$, $\mathbf{v}_2 = [-3, 2, 0]$, $\mathbf{v}_3 = [3, 0, 4]$

(b) $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 4 & -3 & 3 \\ 4 & 2 & 0 \\ 13 & 0 & 4 \end{bmatrix}$

(24) (a) $C = \{[1, 0, 0, -7], [0, 1, 0, 6], [0, 0, 1, 4]\}$

(b) $\mathbf{P} = \begin{bmatrix} 1 & -1 & 3 \\ -2 & -1 & -5 \\ 1 & 1 & 2 \end{bmatrix}$

(c) $[\mathbf{v}]_B = [25, -2, -10]$; (Note: $\mathbf{P}^{-1} = \begin{bmatrix} 3 & 5 & 8 \\ -1 & -1 & -1 \\ -1 & -2 & -3 \end{bmatrix}$.)

(d) $\mathbf{v} = [-3, 2, 3, 45]$

(25) By part (b) of Exercise 8 in Section 4.7, the matrix $\mathbf{C}$ is the transition matrix from C-coordinates to standard coordinates. Hence, by Theorem 4.18, $\mathbf{CP}$ is the transition matrix from B-coordinates to standard coordinates. However, again by part (b) of Exercise 8 in Section 4.7, this transition matrix is the matrix whose columns are the vectors in B.

(26) (a) T (b) T (c) F (d) T (e) F (f) T (g) T (h) F (i) F (j) T (k) F (l) F (m) T (n) T (o) F (p) F (q) T (r) F (s) T (t) T (u) T (v) T (w) F (x) T (y) T (z) F

Chapter 5

Section 5.1

(1) (a) Linear operator:

$$\begin{aligned} f([x,y]+[w,z]) &= f([x+w,y+z]) \\ &= [3(x+w)-4(y+z),-(x+w)+2(y+z)] \\ &= [(3x-4y)+(3w-4z),(-x+2y)+(-w+2z)] \\ &= [3x-4y,-x+2y]+[3w-4z,-w+2z]=f([x,y])+f([w,z]). \end{aligned}$$

Also,

$$\begin{aligned} f(c[x,y]) &= f([cx,cy])=[3(cx)-4(cy),-(cx)+2(cy)] \\ &= [c(3x-4y),c(-x+2y)] \\ &= c[3x-4y,-x+2y]=cf([x,y]). \end{aligned}$$

(b) Not a linear transformation: $h([0,0,0,0])=[2,-1,0,-3]\neq\mathbf{0}$.

(c) Linear operator:

$$\begin{aligned} k([a,b,c]+[d,e,f]) &= k([a+d,b+e,c+f]) \\ &= [b+e,c+f,a+d] \\ &= [b,c,a]+[e,f,d]=k([a,b,c])+k([d,e,f]). \end{aligned}$$

Also,

$$\begin{aligned} k(d[a,b,c]) &= k([da,db,dc])=[db,dc,da] \\ &= d[b,c,a]=dk([a,b,c]). \end{aligned}$$

(d) Linear operator:

$$\begin{aligned} l\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}+\begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) &= l\left(\begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}\right) \\ &= \begin{bmatrix} (a+e)-2(c+g)+(d+h) & 3(b+f) \\ -4(a+e) & (b+f)+(c+g)-3(d+h) \end{bmatrix} \\ &= \begin{bmatrix} (a-2c+d)+(e-2g+h) & 3b+3f \\ -4a-4e & (b+c-3d)+(f+g-3h) \end{bmatrix} \\ &= \begin{bmatrix} a-2c+d & 3b \\ -4a & b+c-3d \end{bmatrix}+\begin{bmatrix} e-2g+h & 3f \\ -4e & f+g-3h \end{bmatrix} \\ &= l\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)+l\left(\begin{bmatrix} e & f \\ g & h \end{bmatrix}\right). \end{aligned}$$

Also,

$$\begin{aligned} l\left(s\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) &= l\left(\begin{bmatrix} sa & sb \\ sc & sd \end{bmatrix}\right) \\ &= \begin{bmatrix} (sa)-2(sc)+(sd) & 3(sb) \\ -4(sa) & (sb)+(sc)-3(sd) \end{bmatrix} \\ &= \begin{bmatrix} s(a-2c+d) & s(3b) \\ s(-4a) & s(b+c-3d) \end{bmatrix} \\ &= s\begin{bmatrix} a-2c+d & 3b \\ -4a & b+c-3d \end{bmatrix} \\ &= sl\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right). \end{aligned}$$

(e) Not a linear transformation: $n\left(2\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right) = n\left(\begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}\right) = 4$, but $2n\left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right) = 2(1) = 2$.

(f) Not a linear transformation: $r(8x^3) = 2x^2$, but $8\ r(x^3) = 8(x^2)$.

(g) Not a linear transformation: $s(\mathbf{0}) = [1,0,1] \neq \mathbf{0}$.

(h) Linear transformation:

$$\begin{aligned} t\left(\begin{array}{l} (ax^3+bx^2+cx+d) \\ \quad +(ex^3+fx^2+gx+h) \end{array}\right) &= t((a+e)x^3+(b+f)x^2+(c+g)x+(d+h)) \\ &= (a+e)+(b+f)+(c+g)+(d+h) \\ &= (a+b+c+d)+(e+f+g+h) \\ &= t(ax^3+bx^2+cx+d)+t(ex^3+fx^2+gx+h). \end{aligned}$$

Also,

$$\begin{aligned} t(s(ax^3+bx^2+cx+d)) &= t((sa)x^3+(sb)x^2+(sc)x+(sd)) \\ &= (sa)+(sb)+(sc)+(sd) \\ &= s(a+b+c+d) \\ &= st(ax^3+bx^2+cx+d). \end{aligned}$$

(i) Not a linear transformation: $u((-1)[0,1,0,0]) = u([0,-1,0,0]) = 1$, but $(-1)\ u([0,1,0,0]) = (-1)(1) = -1$.

(j) Not a linear transformation: $v(x^2+(x+1)) = v(x^2+x+1) = 1$, but $v(x^2)+v(x+1) = 0+0 = 0$.

(k) Linear transformation:

$$\begin{aligned} g\left(\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} + \begin{bmatrix} u & v \\ w & x \\ y & z \end{bmatrix}\right) &= g\left(\begin{bmatrix} a+u & b+v \\ c+w & d+x \\ e+y & f+z \end{bmatrix}\right) \\ &= (a+u)x^4-(c+w)x^2+(e+y) \\ &= (ax^4-cx^2+e)+(ux^4-wx^2+y) \\ &= g\left(\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\right) + g\left(\begin{bmatrix} u & v \\ w & x \\ y & z \end{bmatrix}\right). \end{aligned}$$

Also,

$$\begin{aligned} g\left(s\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\right) &= g\left(\begin{bmatrix} sa & sb \\ sc & sd \\ se & sf \end{bmatrix}\right) \\ &= (sa)x^4 - (sc)x^2 + (se) \\ &= s(ax^4 - cx^2 + e) \\ &= sg\left(\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\right). \end{aligned}$$

(l) Not a linear transformation: $e([3,0]+[0,4]) = e([3,4]) = \sqrt{3^2+4^2} = 5$, but $e([3,0]) + e([0,4]) = \sqrt{3^2+0^2} + \sqrt{0^2+4^2} = 3+4 = 7$.

(2) (a) $i(\mathbf{v}+\mathbf{w}) = \mathbf{v}+\mathbf{w} = i(\mathbf{v}) + i(\mathbf{w})$; also, $i(c\mathbf{v}) = c\mathbf{v} = c\, i(\mathbf{v})$.

(b) $z(\mathbf{v}+\mathbf{w}) = \mathbf{0} = \mathbf{0}+\mathbf{0} = z(\mathbf{v}) + z(\mathbf{w})$; also, $z(c\mathbf{v}) = \mathbf{0} = c\mathbf{0} = c\, z(\mathbf{v})$.

(3) $f(\mathbf{v}+\mathbf{w}) = k(\mathbf{v}+\mathbf{w}) = k\mathbf{v} + k\mathbf{w} = f(\mathbf{v}) + f(\mathbf{w})$; $f(a\mathbf{v}) = ka\mathbf{v} = a(k\mathbf{v}) = af(\mathbf{v})$.

(4) (a) For addition:

$$\begin{aligned} f([u,v,w]+[x,y,z]) &= f([u+x, v+y, w+z]) \\ &= [-(u+x), v+y, w+z] \\ &= [(-u)+(-x), v+y, w+z] \\ &= [-u,v,w] + [-x,y,z] = f([u,v,w]) + f([x,y,z]). \end{aligned}$$

For scalar multiplication:

$$\begin{aligned} f(c[x,y,z]) &= f([cx,cy,cz]) = [-(cx), cy, cz] \\ &= [c(-x), cy, cz] = c[-x,y,z] = cf([x,y,z]). \end{aligned}$$

(b) $f([x,y,z]) = [x,-y,z]$. It is a linear operator.

(c) Through the y-axis: $g([x,y]) = [-x,y]$; Through the x-axis: $f([x,y]) = [x,-y]$; both are linear operators.

(5) First, h is a linear operator because

$$\begin{aligned} h([a_1,a_2,a_3]+[b_1,b_2,b_3]) &= h([a_1+b_1, a_2+b_2, a_3+b_3]) \\ &= [a_1+b_1, a_2+b_2, 0] \\ &= [a_1,a_2,0] + [b_1,b_2,0] = h([a_1,a_2,a_3]) + h([b_1,b_2,b_3]), \quad \text{and} \end{aligned}$$

$$h(c[a_1,a_2,a_3]) = h([ca_1,ca_2,ca_3]) = [ca_1,ca_2,0] = c[a_1,a_2,0] = ch([a_1,a_2,a_3]).$$

Also, j is a linear operator because

$$\begin{aligned} j([a_1,a_2,a_3,a_4]+[b_1,b_2,b_3,b_4]) &= j([a_1+b_1, a_2+b_2, a_3+b_3, a_4+b_4]) \\ &= [0, a_2+b_2, 0, a_4+b_4] \\ &= [0,a_2,0,a_4] + [0,b_2,0,b_4] \\ &= j([a_1,a_2,a_3,a_4]) + j([b_1,b_2,b_3,b_4]), \quad \text{and} \end{aligned}$$

$$\begin{aligned} j(c[a_1,a_2,a_3,a_4]) &= j([ca_1,ca_2,ca_3,ca_4]) = [0,ca_2,0,ca_4] \\ &= c[0,a_2,0,a_4] = cj([a_1,a_2,a_3,a_4]). \end{aligned}$$

(6) For addition:

$$\begin{aligned} f\left(\begin{array}{l} [a_1,a_2,\ldots,a_i,\ldots,a_n] \\ \quad +[b_1,b_2,\ldots,b_i,\ldots,b_n] \end{array}\right) &= f([a_1+b_1,a_2+b_2,\ldots,a_i+b_i,\ldots,a_n+b_n]) \\ &= a_i+b_i \\ &= f([a_1,a_2,\ldots,a_i,\ldots,a_n]) + f([b_1,b_2,\ldots,b_i,\ldots,b_n]). \end{aligned}$$

For scalar multiplication:

$$\begin{aligned} f(c[a_1,a_2,\ldots,a_i,\ldots,a_n]) &= f([ca_1,ca_2,\ldots,ca_i,\ldots,ca_n]) \\ &= ca_i = cf([a_1,a_2,\ldots,a_i,\ldots,a_n]). \end{aligned}$$

(7) For addition:

$$\begin{aligned} g(\mathbf{y}+\mathbf{z}) &= \mathbf{proj}_{\mathbf{x}}(\mathbf{y}+\mathbf{z}) = \frac{\mathbf{x}\cdot(\mathbf{y}+\mathbf{z})}{\|\mathbf{x}\|^2}\mathbf{x} = \frac{(\mathbf{x}\cdot\mathbf{y})+(\mathbf{x}\cdot\mathbf{z})}{\|\mathbf{x}\|^2}\mathbf{x} \\ &= \frac{(\mathbf{x}\cdot\mathbf{y})}{\|\mathbf{x}\|^2}\mathbf{x} + \frac{(\mathbf{x}\cdot\mathbf{z})}{\|\mathbf{x}\|^2}\mathbf{x} = g(\mathbf{y})+g(\mathbf{z}). \end{aligned}$$

For scalar multiplication:

$$g(c\mathbf{y}) = \frac{\mathbf{x}\cdot(c\mathbf{y})}{\|\mathbf{x}\|^2}\mathbf{x} = c\left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x} = cg(\mathbf{y}).$$

(8) $L(\mathbf{y}_1+\mathbf{y}_2) = \mathbf{x}\cdot(\mathbf{y}_1+\mathbf{y}_2) = (\mathbf{x}\cdot\mathbf{y}_1)+(\mathbf{x}\cdot\mathbf{y}_2) = L(\mathbf{y}_1)+L(\mathbf{y}_2)$;
$L(c\mathbf{y}_1) = \mathbf{x}\cdot(c\mathbf{y}_1) = c(\mathbf{x}\cdot\mathbf{y}_1) = cL(\mathbf{y}_1)$

(9) Follow the hint in the text, and use the sum of angle identities for sine and cosine.

(10) (a) Use Example 10.

(b) Use the hint in Exercise 9.

(c) $\begin{bmatrix} \sin\theta & 0 & \cos\theta \\ 0 & 1 & 0 \\ \cos\theta & 0 & -\sin\theta \end{bmatrix}$

(11) The proofs follow the proof in Example 10, but with the matrix $\mathbf{A}$ replaced by the specific matrices of this exercise.

(12) $f(\mathbf{A}+\mathbf{B}) = \mathrm{trace}(\mathbf{A}+\mathbf{B}) = \sum_{i=1}^n(a_{ii}+b_{ii}) = \sum_{i=1}^n a_{ii}+\sum_{i=1}^n b_{ii} = \mathrm{trace}(\mathbf{A})+\mathrm{trace}(\mathbf{B}) = f(\mathbf{A})+f(\mathbf{B})$;
$f(c\mathbf{A}) = \mathrm{trace}(c\mathbf{A}) = \sum_{i=1}^n ca_{ii} = c\sum_{i=1}^n a_{ii} = c\,\mathrm{trace}(\mathbf{A}) = cf(\mathbf{A})$.

(13) $g(\mathbf{A}_1+\mathbf{A}_2) = (\mathbf{A}_1+\mathbf{A}_2)+(\mathbf{A}_1+\mathbf{A}_2)^T = (\mathbf{A}_1+\mathbf{A}_1^T)+(\mathbf{A}_2+\mathbf{A}_2^T) = g(\mathbf{A}_1)+g(\mathbf{A}_2)$;
$g(c\mathbf{A}) = c\mathbf{A}+(c\mathbf{A})^T = c(\mathbf{A}+\mathbf{A}^T) = cg(\mathbf{A})$. The proof for h is done similarly.

(14) (a) $f(\mathbf{p}_1+\mathbf{p}_2) = \int(\mathbf{p}_1+\mathbf{p}_2)\,dx = \int\mathbf{p}_1\,dx+\int\mathbf{p}_2\,dx = f(\mathbf{p}_1)+f(\mathbf{p}_2)$ (since the constant of integration is assumed to be zero for all these indefinite integrals);
$f(c\mathbf{p}) = \int(c\mathbf{p})\,dx = c\int\mathbf{p}\,dx = cf(\mathbf{p})$ (since the constant of integration is assumed to be zero for these indefinite integrals).

(b) $g(\mathbf{p}_1 + \mathbf{p}_2) = \int_a^b (\mathbf{p}_1 + \mathbf{p}_2)\ dx = \int_a^b \mathbf{p}_1\ dx + \int_a^b \mathbf{p}_2\ dx = g(\mathbf{p}_1) + g(\mathbf{p}_2)$;
$g(c\mathbf{p}) = \int_a^b (c\mathbf{p})\ dx = c\int_a^b \mathbf{p}\ dx = cg(\mathbf{p})$.

(15) Base Step: $k = 1$: An argument similar to that in Example 3 in Section 5.1 shows that $M : \mathcal{V} \longrightarrow \mathcal{V}$ given by $M(f) = f'$ is a linear operator.
Inductive Step: Assume that $L_n : \mathcal{V} \longrightarrow \mathcal{V}$ given by $L_n(f) = f^{(n)}$ is a linear operator. We need to show that $L_{n+1} : \mathcal{V} \longrightarrow \mathcal{V}$ given by $L_{n+1}(f) = f^{(n+1)}$ is a linear operator. But since $L_{n+1} = L_n \circ M$, Theorem 5.2 assures us that L_{n+1} is a linear operator.

(16) $f(\mathbf{A}_1 + \mathbf{A}_2) = \mathbf{B}(\mathbf{A}_1 + \mathbf{A}_2) = \mathbf{B}\mathbf{A}_1 + \mathbf{B}\mathbf{A}_2 = f(\mathbf{A}_1) + f(\mathbf{A}_2)$;
$f(c\mathbf{A}) = \mathbf{B}(c\mathbf{A}) = c(\mathbf{B}\mathbf{A}) = cf(\mathbf{A})$.

(17) $f(\mathbf{A}_1 + \mathbf{A}_2) = \mathbf{B}^{-1}(\mathbf{A}_1 + \mathbf{A}_2)\mathbf{B} = \mathbf{B}^{-1}\mathbf{A}_1\mathbf{B} + \mathbf{B}^{-1}\mathbf{A}_2\mathbf{B} = f(\mathbf{A}_1) + f(\mathbf{A}_2)$;
$f(c\mathbf{A}) = \mathbf{B}^{-1}(c\mathbf{A})\mathbf{B} = c(\mathbf{B}^{-1}\mathbf{A}\mathbf{B}) = cf(\mathbf{A})$

(18) (a) $L(\mathbf{p} + \mathbf{q}) = (\mathbf{p} + \mathbf{q})(a) = \mathbf{p}(a) + \mathbf{q}(a) = L(\mathbf{p}) + L(\mathbf{q})$.
Similarly, $L(c\mathbf{p}) = (c\mathbf{p})(a) = c(\mathbf{p}(a)) = cL(\mathbf{p})$.

(b) For all $x \in \mathbb{R}$, $L((\mathbf{p} + \mathbf{q})(x)) = (\mathbf{p} + \mathbf{q})(x + a) = \mathbf{p}(x + a) + \mathbf{q}(x + a) = L(\mathbf{p}(x)) + L(\mathbf{q}(x))$.
Also, $L((c\mathbf{p})(x)) = (c\mathbf{p})(x + a) = c(\mathbf{p}(x + a)) = cL(\mathbf{p}(x))$.

(19) Let $\mathbf{p} = a_nx^n + \cdots + a_1x + a_0$ and let $\mathbf{q} = b_nx^n + \cdots + b_1x + b_0$. Then

$$\begin{aligned} L(\mathbf{p} + \mathbf{q}) &= L(a_nx^n + \cdots + a_1x + a_0 + b_nx^n + \cdots + b_1x + b_0) \\ &= L((a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0)) \\ &= (a_n + b_n)\mathbf{A}^n + \cdots + (a_1 + b_1)\mathbf{A} + (a_0 + b_0)\mathbf{I}_n \\ &= (a_n\mathbf{A}^n + \cdots + a_1\mathbf{A} + a_0\mathbf{I}_n) + (b_n\mathbf{A}^n + \cdots + b_1\mathbf{A} + b_0\mathbf{I}_n) \\ &= L(\mathbf{p}) + L(\mathbf{q}). \end{aligned}$$

Similarly,

$$\begin{aligned} L(c\mathbf{p}) &= L(c(a_nx^n + \cdots + a_1x + a_0)) \\ &= L(ca_nx^n + \cdots + ca_1x + ca_0) \\ &= ca_n\mathbf{A}^n + \cdots + ca_1\mathbf{A} + ca_0\mathbf{I}_n \\ &= c(a_n\mathbf{A}^n + \cdots + a_1\mathbf{A} + a_0\mathbf{I}_n) \\ &= cL(\mathbf{p}). \end{aligned}$$

(20) $L(x \oplus y) = L(xy) = \ln(xy) = \ln(x) + \ln(y) = L(x) + L(y)$;
$L(a \odot x) = L(x^a) = \ln(x^a) = a\ln(x) = aL(x)$

(21) $f(\mathbf{0}) = \mathbf{0} + \mathbf{x} = \mathbf{x} \neq \mathbf{0}$, contradicting Theorem 5.1, part (1).

(22) $f(\mathbf{0}) = \mathbf{A}\mathbf{0} + \mathbf{y} = \mathbf{y} \neq \mathbf{0}$, contradicting Theorem 5.1, part (1).

(23) $f\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 1$, but $f\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) + f\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 0 + 0 = 0$. (Note: For any $n > 1$, $f(2\mathbf{I}_n) = 2^n$, but $2f(\mathbf{I}_n) = 2(1) = 2$.)

(24) $L_2(\mathbf{v}+\mathbf{w}) = L_1(2(\mathbf{v}+\mathbf{w})) = 2L_1(\mathbf{v}+\mathbf{w}) = 2(L_1(\mathbf{v})+L_1(\mathbf{w})) = 2L_1(\mathbf{v})+2L_1(\mathbf{w}) = L_2(\mathbf{v})+L_2(\mathbf{w})$; $L_2(c\mathbf{v}) = L_1(2c\mathbf{v}) = 2cL_1(\mathbf{v}) = c(2L_1(\mathbf{v})) = cL_2(\mathbf{v})$.

(25) $L([-3,2,4]) = [12,-11,14]$;
$$L\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -2 & 3 & 0 \\ 1 & -2 & -1 \\ 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2x+3y \\ x-2y-z \\ y+3z \end{bmatrix}$$

(26) $L(\mathbf{i}) = \frac{7}{5}\mathbf{i} - \frac{11}{5}\mathbf{j}$; $L(\mathbf{j}) = -\frac{2}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}$

(27) $L(\mathbf{x}-\mathbf{y}) = L(\mathbf{x}+(-\mathbf{y})) = L(\mathbf{x}) + L((-1)\mathbf{y}) = L(\mathbf{x}) + (-1)L(\mathbf{y}) = L(\mathbf{x}) - L(\mathbf{y})$.

(28) Follow the hint in the text. Letting $a = b = 0$ proves property (1) of a linear transformation. Letting $b = 0$ proves property (2).

(29) Suppose part (3) of Theorem 5.1 is true for some n. We prove it true for $n+1$.

$$\begin{aligned} L(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n + a_{n+1}\mathbf{v}_{n+1}) &= L(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) + L(a_{n+1}\mathbf{v}_{n+1}) \\ &\qquad \text{(by property (1) of a linear transformation)} \\ &= (a_1L(\mathbf{v}_1) + \cdots + a_nL(\mathbf{v}_n)) + L(a_{n+1}\mathbf{v}_{n+1}) \\ &\qquad \text{(by the inductive hypothesis)} \\ &= a_1L(\mathbf{v}_1) + \cdots + a_nL(\mathbf{v}_n) + a_{n+1}L(\mathbf{v}_{n+1}) \\ &\qquad \text{(by property (2) of a linear transformation),} \end{aligned}$$

and we are done.

(30) (a) $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}_\mathcal{V} \implies L(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = L(\mathbf{0}_\mathcal{V})$
$\implies a_1L(\mathbf{v}_1) + \cdots + a_nL(\mathbf{v}_n) = \mathbf{0}_\mathcal{W} \implies a_1 = a_2 = \cdots = a_n = 0$.

(b) Consider the zero linear transformation.

(31) $(L_2 \circ L_1)(c\mathbf{v}) = L_2(L_1(c\mathbf{v})) = L_2(cL_1(\mathbf{v})) = cL_2(L_1(\mathbf{v})) = c(L_2 \circ L_1)(\mathbf{v})$.

(32) Let $c = L(1)$. Then $L(x) = L(x1) = xL(1) = xc = cx$.

(33) $\mathbf{0}_\mathcal{W} \in \mathcal{W}'$, so $\mathbf{0}_\mathcal{V} \in L^{-1}(\{\mathbf{0}_\mathcal{W}\}) \subseteq L^{-1}(\mathcal{W}')$. Hence, $L^{-1}(\mathcal{W}')$ is nonempty.
Also, $\mathbf{x},\mathbf{y} \in L^{-1}(\mathcal{W}') \implies L(\mathbf{x}), L(\mathbf{y}) \in \mathcal{W}' \implies L(\mathbf{x}) + L(\mathbf{y}) \in \mathcal{W}'$
$\implies L(\mathbf{x}+\mathbf{y}) \in \mathcal{W}' \implies \mathbf{x}+\mathbf{y} \in L^{-1}(\mathcal{W}')$.
Finally, $\mathbf{x} \in L^{-1}(\mathcal{W}') \implies L(\mathbf{x}) \in \mathcal{W}' \implies aL(\mathbf{x}) \in \mathcal{W}'$ (for any $a \in \mathbb{R}$)
$\implies L(a\mathbf{x}) \in \mathcal{W}' \implies a\mathbf{x} \in L^{-1}(\mathcal{W}')$.
Hence $L^{-1}(\mathcal{W}')$ is a subspace by Theorem 4.2.

(34) (a) For $L_1 \oplus L_2$:

$$\begin{aligned} L_1 \oplus L_2(\mathbf{x}+\mathbf{y}) &= L_1(\mathbf{x}+\mathbf{y}) + L_2(\mathbf{x}+\mathbf{y}) \\ &= (L_1(\mathbf{x}) + L_1(\mathbf{y})) + (L_2(\mathbf{x}) + L_2(\mathbf{y})) \\ &= (L_1(\mathbf{x}) + L_2(\mathbf{x})) + (L_1(\mathbf{y}) + L_2(\mathbf{y})) \\ &= L_1 \oplus L_2(\mathbf{x}) + L_1 \oplus L_2(\mathbf{y}). \end{aligned}$$

Also,

$$\begin{aligned} L_1 \oplus L_2(c\mathbf{x}) &= L_1(c\mathbf{x}) + L_2(c\mathbf{x}) \\ &= cL_1(\mathbf{x}) + cL_2(\mathbf{x}) \\ &= c(L_1(\mathbf{x}) + L_2(\mathbf{x})) \\ &= cL_1 \oplus L_2(\mathbf{x}). \end{aligned}$$

For $c \odot L_1$:

$$\begin{aligned} c \odot L_1(\mathbf{x}+\mathbf{y}) &= c(L_1(\mathbf{x}+\mathbf{y})) \\ &= c(L_1(\mathbf{x}) + L_1(\mathbf{y})) \\ &= cL_1(\mathbf{x}) + cL_2(\mathbf{y}) \\ &= c \odot L_1(\mathbf{x}) + c \odot L_1(x). \end{aligned}$$

Also,

$$\begin{aligned} c \odot L_1(a\mathbf{x}) &= cL_1(a\mathbf{x}) \\ &= c(aL_1(\mathbf{x})) \\ &= a(cL_1(\mathbf{x})) \\ &= a(c \odot L_1(\mathbf{x})). \end{aligned}$$

(b) Use Theorem 4.2: the existence of the zero linear transformation shows that the set is nonempty, while part (a) proves closure under addition and scalar multiplication.

(35) Define a line l parametrically by $\{t\mathbf{x} + \mathbf{y} \,|\, t \in \mathbb{R}\}$, for some fixed $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Since $L(t\mathbf{x} + \mathbf{y}) = tL(\mathbf{x}) + L(\mathbf{y})$, L maps l to the set $\{tL(\mathbf{x}) + L(\mathbf{y}) \,|\, t \in \mathbb{R}\}$. If $L(\mathbf{x}) \neq 0$, this set represents a line; otherwise, it represents a point.

(36) (a) F (b) T (c) F (d) F (e) T (f) F (g) T (h) T

Section 5.2

(1) For each operator, check that the ith column of the given matrix (for $1 \leq i \leq 3$) is the image of $\mathbf{e}_i$. This is easily done by inspection.

(2) (a) $\begin{bmatrix} -6 & 4 & -1 \\ -2 & 3 & -5 \\ 3 & -1 & 7 \end{bmatrix}$ (c) $\begin{bmatrix} 4 & -1 & 3 & 3 \\ 1 & 3 & -1 & 5 \\ -2 & -7 & 5 & -1 \end{bmatrix}$ (d) $\begin{bmatrix} -3 & 0 & -2 & 0 \\ 0 & -1 & 0 & 4 \\ 0 & 4 & -1 & 3 \\ -6 & -1 & 0 & 2 \end{bmatrix}$

(b) $\begin{bmatrix} 3 & -5 & 1 & -2 \\ 5 & 1 & -2 & 8 \end{bmatrix}$

(3) (a) $\begin{bmatrix} -47 & 128 & -288 \\ -18 & 51 & -104 \end{bmatrix}$ (c) $\begin{bmatrix} 22 & 14 \\ 62 & 39 \\ 68 & 43 \end{bmatrix}$ (e) $\begin{bmatrix} 5 & 6 & 0 \\ -11 & -26 & -6 \\ -14 & -19 & -1 \\ 6 & 3 & -2 \\ -1 & 1 & 1 \\ 11 & 13 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} -3 & 5 \\ 2 & -1 \\ 4 & -2 \end{bmatrix}$ (d) $\begin{bmatrix} -32 & 16 & 24 & 15 \\ -10 & 7 & 12 & 6 \\ -121 & 63 & 96 & 58 \end{bmatrix}$

(4) (a) $\begin{bmatrix} -202 & -32 & -43 \\ -146 & -23 & -31 \\ 83 & 14 & 18 \end{bmatrix}$ (b) $\begin{bmatrix} 21 & 7 & 21 & 16 \\ -51 & -13 & -51 & -38 \end{bmatrix}$ (c) $\begin{bmatrix} 98 & -91 & 60 & -76 \\ 28 & -25 & 17 & -21 \\ 135 & -125 & 82 & -104 \end{bmatrix}$

(5) See the answer for Exercise 3(d).

(6) (a) $\begin{bmatrix} 67 & -123 \\ 37 & -68 \end{bmatrix}$ (b) $\begin{bmatrix} -7 & 2 & 10 \\ 5 & -2 & -9 \\ -6 & 1 & 8 \end{bmatrix}$ (c) $\begin{bmatrix} 8 & -9 & 9 & -14 \\ 8 & -8 & 5 & -7 \\ 6 & -6 & -1 & 6 \\ 1 & -1 & -2 & 5 \end{bmatrix}$

(7) (a) $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$; $12x^2 - 10x + 6$ (b) $\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; $\frac{2}{3}x^3 - \frac{1}{2}x^2 + 5x$

(8) (a) $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{2}\sqrt{3} - 9 & -\frac{13}{2} \\ \frac{25}{2} & \frac{1}{2}\sqrt{3} + 9 \end{bmatrix}$

(9) (a) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ (b) $\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 0 & 1 & -1 \end{bmatrix}$

(10) $\begin{bmatrix} -12 & 12 & -2 \\ -4 & 6 & -2 \\ -10 & -3 & 7 \end{bmatrix}$

(11) (a) Matrix for L_1: $\begin{bmatrix} 1 & -1 & -1 \\ 0 & 2 & 3 \\ 1 & 3 & 0 \\ -2 & 0 & 1 \end{bmatrix}$; matrix for L_2: $\begin{bmatrix} 0 & 2 & -2 & 3 \\ 1 & 0 & -1 & 1 \end{bmatrix}$

(b) Matrix for $L_2 \circ L_1$: $\begin{bmatrix} -8 & -2 & 9 \\ -2 & -4 & 0 \end{bmatrix}$

(c) Easily verified with straightforward computations.

(12) (a) $\mathbf{A}^2$ represents a rotation through the angle θ twice, which corresponds to a rotation through the angle 2θ.

(b) Generalize the argument in part (a).

(13) (a) $\mathbf{I}_n$ (b) $\mathbf{O}_n$ (c) $c\mathbf{I}_n$

(d) The $n \times n$ matrix whose columns are $\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n, \mathbf{e}_1$, respectively

(e) The $n \times n$ matrix whose columns are $\mathbf{e}_n$, $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_{n-1}$, respectively

(14) Let $\mathbf{A}$ be the matrix for L (with respect to the standard bases). Then $\mathbf{A}$ is a $1 \times n$ matrix (an n-vector). If we let $\mathbf{x} = \mathbf{A}$, then $\mathbf{Ay} = \mathbf{x} \cdot \mathbf{y}$, for all $\mathbf{y} \in \mathbb{R}^n$.

(15) $[(L_2 \circ L_1)(\mathbf{v})]_D = [L_2(L_1(\mathbf{v}))]_D = \mathbf{A}_{CD}[L_1(\mathbf{v})]_C = \mathbf{A}_{CD}(\mathbf{A}_{BC}[\mathbf{v}]_B) = (\mathbf{A}_{CD}\mathbf{A}_{BC})[\mathbf{v}]_B$. Now apply the uniqueness condition in Theorem 5.5.

(16) (a) $\begin{bmatrix} 0 & -4 & -13 \\ -6 & 5 & 6 \\ 2 & -2 & -3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) The vectors in B are eigenvectors for the matrix from part (a) corresponding to the eigenvalues, 2, -1, and 1, respectively.

(17) Easily verified with straightforward computations.

(18) (a) $p_{\mathbf{A}_{BB}}(x) = x(x-1)^2$

(b) $E_1 = \{b[2,1,0] + c[2,0,1]\}$; basis for $E_1 = \{[2,1,0], [2,0,1]\}$;
$E_0 = \{c[-1,2,2]\}$; basis for $E_0 = \{[-1,2,2]\}$;
$C = ([2,1,0], [2,0,1], [-1,2,2])$
(Note: The remaining answers will vary if the vectors in C are ordered differently.)

(c) Using the basis C ordered as $([2,1,0], [2,0,1], [-1,2,2])$, the answer is $\mathbf{P} = \begin{bmatrix} 2 & 2 & -1 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}$, with $\mathbf{P}^{-1} = \frac{1}{9}\begin{bmatrix} 2 & 5 & -4 \\ 2 & -4 & 5 \\ -1 & 2 & 2 \end{bmatrix}$. Another possible answer, using $C = ([-1,2,2], [2,1,0], [2,0,1])$, is $\mathbf{P} = \begin{bmatrix} -1 & 2 & 2 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, with $\mathbf{P}^{-1} = \frac{1}{9}\begin{bmatrix} -1 & 2 & 2 \\ 2 & 5 & 4 \\ 2 & -4 & 5 \end{bmatrix}$.

(d) Using $C = ([2,1,0], [2,0,1], [-1,2,2])$ produces $\mathbf{A}_{CC} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Using $C = ([-1,2,2], [2,1,0], [2,0,1])$ yields $\mathbf{A}_{CC} = \mathbf{P}^{-1}\mathbf{A}_{BB}\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ instead.

(e) L is a projection onto the plane formed by $[2,1,0]$ and $[2,0,1]$, while $[-1,2,2]$ is orthogonal to this plane.

(19) $[L(\mathbf{v})]_C = \frac{1}{k}[L(\mathbf{v})]_B = \frac{1}{k}\mathbf{A}_{BB}[\mathbf{v}]_B = \frac{1}{k}\mathbf{A}_{BB}(k[\mathbf{v}]_C) = \mathbf{A}_{BB}[\mathbf{v}]_C$. By Theorem 5.5, $\mathbf{A}_{CC} = \mathbf{A}_{BB}$.

(20) Find the matrix for L with respect to B, and then apply Theorem 5.6.

(21) Let $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $\mathbf{P} = \frac{1}{\sqrt{1+m^2}}\begin{bmatrix} 1 & -m \\ m & 1 \end{bmatrix}$. Then $\mathbf{A}$ represents a reflection through the x-axis, and $\mathbf{P}$ represents a counterclockwise rotation putting the x-axis on the line $y = mx$ (using Example 9, Section 5.1). Thus, the desired matrix is $\mathbf{PAP}^{-1}$.

An alternate solution is obtained by finding the images of the standard basis vectors using Exercise 21 in Section 1.2. A vector in the direction of the line $y = mx$ is $[1, m]$. Then for any vector $\mathbf{x}$,

$$\mathbf{p} = \mathbf{proj}_{[1,m]}\mathbf{x} = \frac{[1,m] \cdot \mathbf{x}}{1+m^2},$$

and so by Exercise 21 in Section 1.2, the reflection of $\mathbf{x}$ through $y = mx$ is

$$2\mathbf{p} - \mathbf{x} = 2\left(\frac{[1, m] \cdot \mathbf{x}}{1 + m^2}\right)[1, m] - \mathbf{x}.$$

Therefore,

$$L([1, 0]) = 2\left(\frac{[1, m] \cdot [1, 0]}{1 + m^2}\right)[1, m] - [1, 0], \text{ which simplifies to } \left[\frac{1 - m^2}{1 + m^2}, \frac{2m}{1 + m^2}\right],$$

the first column of the desired matrix. Similarly,

$$L([0, 1]) = 2\left(\frac{[1, m] \cdot [0, 1]}{1 + m^2}\right)[1, m] - [0, 1], \text{ which simplifies to } \left[\frac{2m}{1 + m^2}, \frac{m^2 - 1}{1 + m^2}\right],$$

the second column of the desired matrix.

(22) Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for $\mathcal{Y}$. Extend this to a basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for $\mathcal{V}$. By Theorem 5.4, there is a unique linear transformation

$$L' : \mathcal{V} \longrightarrow \mathcal{W} \text{ such that } L'(\mathbf{v}_i) = \begin{cases} L(\mathbf{v}_i) & \text{if } i \leq k \\ \mathbf{0} & \text{if } i > k \end{cases}.$$

Clearly, L' agrees with L on $\mathcal{Y}$.

(23) Let $L : \mathcal{V} \longrightarrow \mathcal{W}$ be a linear transformation such that $L(\mathbf{v}_1) = \mathbf{w}_1$, $L(\mathbf{v}_2) = \mathbf{w}_2, \ldots, L(\mathbf{v}_n) = \mathbf{w}_n$, with $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ a basis for $\mathcal{V}$. If $\mathbf{v} \in \mathcal{V}$, then $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, for unique $c_1, c_2, \ldots, c_n \in \mathbb{R}$ (by Theorem 4.9). But then

$$\begin{aligned} L(\mathbf{v}) &= L(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) \\ &= c_1L(\mathbf{v}_1) + c_2L(\mathbf{v}_2) + \cdots + c_nL(\mathbf{v}_n) \\ &= c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n. \end{aligned}$$

Hence $L(\mathbf{v})$ is determined uniquely, for each $\mathbf{v} \in \mathcal{V}$, and so L is uniquely determined.

(24) (a) T (b) T (c) F (d) F (e) T (f) F (g) T (h) T (i) T (j) F

Section 5.3

(1) (a) Yes, because $L([1, -2, 3]) = [0, 0, 0]$

(b) No, because $L([2, -1, 4]) = [5, -2, -1]$

(c) No; the system $\begin{cases} 5x_1 + x_2 - x_3 = 2 \\ -3x_1 + x_3 = -1 \\ x_1 - x_2 - x_3 = 4 \end{cases}$ has no solutions.

Note: $\left[\begin{array}{rrr|r} 5 & 1 & -1 & 2 \\ -3 & 0 & 1 & -1 \\ 1 & -1 & -1 & 4 \end{array}\right]$ row reduces to $\left[\begin{array}{rrr|r} 1 & 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 4 \end{array}\right]$,

where we have *not* row reduced beyond the augmentation bar.

(d) Yes, because $L([-4,4,0]) = [-16,12,-8]$

(2) (a) No, since $L(x^3 - 5x^2 + 3x - 6) = 6x^3 + 4x - 9 \neq 0$

(b) Yes, because $L(4x^3 - 4x^2) = 0$

(c) Yes, because, for example, $L(x^3 + 4x + 3) = 8x^3 - x - 1$

(d) No, since every element of range(L) has a zero coefficient for its x^2 term.

(3) In each part, we give the reduced row echelon form for the matrix for the linear transformation as well.

(a) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$; $\dim(\ker(L)) = 1$, basis for $\ker(L) = \{[-2,3,1]\}$,

$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[1,-2,3],[-1,3,-3]\}$

(b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $\dim(\ker(L)) = 1$, basis for $\ker(L) = \{[-1,2,1]\}$,

$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[4,7,-2,3],[-2,1,-1,-2]\}$

(c) $\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -2 \end{bmatrix}$; $\dim(\ker(L)) = 1$, basis for $\ker(L) = \{[-5,2,1]\}$,

$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[3,2],[2,1]\}$

(d) $\begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$; $\dim(\ker(L)) = 2$, basis for $\ker(L) = \{[1,-3,1,0],[-1,2,0,1]\}$,

$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[-14,-4,-6,3,4],\ [-8,-1,2,-7,2]\}$

(4) Students can generally solve these problems by inspection.

(a) $\dim(\ker(L)) = 2$, basis for $\ker(L) = \{[1,0,0],\ [0,0,1]\}$,
$\dim(\text{range}(L)) = 1$, basis for $\text{range}(L) = \{[0,1]\}$

(b) $\dim(\ker(L)) = 0$, basis for $\ker(L) = \{\}$,
$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[1,1,0],[0,1,1]\}$

(c) $\dim(\ker(L)) = 2$, basis for $\ker(L) = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$,

$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$

(d) $\dim(\ker(L)) = 2$, basis for $\ker(L) = \{x^4, x^3\}$,
$\dim(\text{range}(L)) = 3$, basis for $\text{range}(L) = \{x^2, x, 1\}$

(e) $\dim(\ker(L)) = 0$, basis for $\ker(L) = \{\}$,
$\dim(\text{range}(L)) = 3$, basis for $\text{range}(L) = \{x^3, x^2, x\}$

(f) $\dim(\ker(L)) = 1$, basis for $\ker(L) = \{[0,1,1]\}$,
$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[1,0,1],[0,0,-1]\}$
(A simpler basis for $\text{range}(L) = \{[1,0,0],[0,0,1]\}$.)

(g) $\dim(\ker(L)) = 0$, basis for $\ker(L) = \{\}$ (empty set),
$\dim(\text{range}(L)) = 4$, basis for $\text{range}(L) =$ standard basis for $\mathcal{M}_{22}$

(h) $\dim(\ker(L)) = 6$, basis for $\ker(L)$

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

$$\dim(\text{range}(L)) = 3, \text{ basis for range}(L) = \left\{ \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \right\}$$

(i) $\dim(\ker(L)) = 1$, basis for $\ker(L) = \{x^2 - 2x + 1\}$,
$\dim(\text{range}(L)) = 2$, basis for $\text{range}(L) = \{[1,2],[1,1]\}$
(A simpler basis for $\text{range}(L) =$ standard basis for $\mathbb{R}^2$.)

(j) $\dim(\ker(L)) = 2$, basis for $\ker(L) = \{(x+1)x(x-1),\ x^2(x+1)(x-1)\}$,
$\dim(\text{range}(L)) = 3$, basis for $\text{range}(L) = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$

(5) (a) $\ker(L) = \mathcal{V}$, $\text{range}(L) = \{\mathbf{0}_\mathcal{W}\}$ (b) $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$, $\text{range}(L) = \mathcal{V}$

(6) $L(\mathbf{A}+\mathbf{B}) = \text{trace}(\mathbf{A}+\mathbf{B}) = \sum_{i=1}^{3}(a_{ii}+b_{ii}) = \sum_{i=1}^{3} a_{ii} + \sum_{i=1}^{3} b_{ii} = \text{trace}(\mathbf{A}) + \text{trace}(\mathbf{B}) = L(\mathbf{A}) + L(\mathbf{B})$;
$L(c\mathbf{A}) = \text{trace}(c\mathbf{A}) = \sum_{i=1}^{3} ca_{ii} = c\sum_{i=1}^{3} a_{ii} = c(\text{trace}(\mathbf{A})) = cL(\mathbf{A})$;

$$\ker(L) = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & -a-e \end{bmatrix} \middle| \, a,b,c,d,e,f,g,h \in \mathbb{R} \right\},$$

$\dim(\ker(L)) = 8$, $\text{range}(L) = \mathbb{R}$, $\dim(\text{range}(L)) = 1$

(7) $\text{range}(L) = \mathcal{V}$, $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$

(8) $\ker(L) = \{\mathbf{0}\}$, $\text{range}(L) = \{ax^4 + bx^3 + cx^2\}$, $\dim(\ker(L)) = 0$, $\dim(\text{range}(L)) = 3$

(9) $\ker(L) = \{dx + e \,|\, d, e \in \mathbb{R}\}$, $\text{range}(L) = \mathcal{P}_2$, $\dim(\ker(L)) = 2$, $\dim(\text{range}(L)) = 3$

(10) When $k \le n$, $\ker(L) =$ all polynomials of degree less than k, $\dim(\ker(L)) = k$, $\text{range}(L) = \mathcal{P}_{n-k}$, and $\dim(\text{range}(L)) = n - k + 1$. When $k > n$, $\ker(L) = \mathcal{P}_n$, $\dim(\ker(L)) = n + 1$, $\text{range}(L) = \{\mathbf{0}\}$, and $\dim(\text{range}(L)) = 0$.

(11) Since $\text{range}(L) = \mathbb{R}$, $\dim(\text{range}(L)) = 1$. By the Dimension Theorem, $\dim(\ker(L)) = n$. Since the given set is clearly a linearly independent subset of $\ker(L)$ containing n distinct elements, it must be a basis for $\ker(L)$ by part (2) of Theorem 4.12.

(12) $\ker(L) = \{[0,0,\ldots,0]\}$, $\text{range}(L) = \mathbb{R}^n$
(Note: Every vector $\mathbf{X}$ is in the range since $L(\mathbf{A}^{-1}\mathbf{X}) = \mathbf{A}(\mathbf{A}^{-1}\mathbf{X}) = \mathbf{X}$.)

(13) $\ker(L) = \{\mathbf{0}_\mathcal{V}\}$ iff $\dim(\ker(L)) = 0$ iff $\dim(\text{range}(L)) = \dim(\mathcal{V}) - 0 = \dim(\mathcal{V})$ iff $\text{range}(L) = \mathcal{V}$.

(14) For $\ker(L)$: $\mathbf{0}_{\mathcal{V}} \in \ker(L)$, so $\ker(L) \neq \{\}$. If $\mathbf{v}_1, \mathbf{v}_2 \in \ker(L)$, then

$$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) = \mathbf{0}_{\mathcal{V}} + \mathbf{0}_{\mathcal{V}} = \mathbf{0}_{\mathcal{V}},$$

and so $\mathbf{v}_1 + \mathbf{v}_2 \in \ker(L)$. Similarly, $L(a\mathbf{v}_1) = \mathbf{0}_{\mathcal{V}}$, and so $a\mathbf{v}_1 \in \ker(L)$.

For range(L): $\mathbf{0}_{\mathcal{W}} \in \text{range}(L)$ since $\mathbf{0}_{\mathcal{W}} = L(\mathbf{0}_{\mathcal{V}})$. If $\mathbf{w}_1, \mathbf{w}_2 \in \text{range}(L)$, then there exist $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ such that $L(\mathbf{v}_1) = \mathbf{w}_1$ and $L(\mathbf{v}_2) = \mathbf{w}_2$. Hence,

$$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2.$$

Thus $\mathbf{w}_1 + \mathbf{w}_2 \in \text{range}(L)$. Similarly, $L(a\mathbf{v}_1) = aL(\mathbf{v}_1) = a\mathbf{w}_1$, so $a\mathbf{w}_1 \in \text{range}(L)$.

(15) (a) If $\mathbf{v} \in \ker(L_1)$, then $(L_2 \circ L_1)(\mathbf{v}) = L_2(L_1(\mathbf{v})) = L_2(\mathbf{0}_{\mathcal{W}}) = \mathbf{0}_{\mathcal{X}}$.

(b) If $\mathbf{x} \in \text{range}(L_2 \circ L_1)$, then there is a $\mathbf{v}$ such that $(L_2 \circ L_1)(\mathbf{v}) = \mathbf{x}$. Hence $L_2(L_1(\mathbf{v})) = \mathbf{x}$, so $\mathbf{x} \in \text{range}(L_2)$.

(c) $\dim(\text{range}(L_2 \circ L_1)) = \dim(\mathcal{V}) - \dim(\ker(L_2 \circ L_1)) \leq \dim(\mathcal{V}) - \dim(\ker(L_1))$ (by part (a)) $= \dim(\text{range}(L_1))$

(16) Consider $L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. Then, $\ker(L) = \text{range}(L) = \{[a, a] \mid a \in \mathbb{R}\}$.

(17) (a) Let $\mathbf{P}$ be the transition matrix from the standard basis to B, and let $\mathbf{Q}$ be the transition matrix from the standard basis to C. Note that both $\mathbf{P}$ and $\mathbf{Q}$ are nonsingular by Theorem 4.19. Then, by Theorem 5.6, $\mathbf{B} = \mathbf{QAP}^{-1}$. Hence, rank$(\mathbf{B})$ = rank$(\mathbf{QAP}^{-1})$ = rank$(\mathbf{AP}^{-1})$ (by part (a) of Review Exercise 16 in Chapter 2) = rank$(\mathbf{A})$ (by part (b) of Review Exercise 16 in Chapter 2).

(b) The comments prior to Theorem 5.9 in the text show that parts (a) and (b) of the theorem are true if $\mathbf{A}$ is the matrix for L with respect to the standard bases. However, part (a) of this exercise shows that if $\mathbf{B}$ is the matrix for L with respect to any bases, rank$(\mathbf{B})$ = rank$(\mathbf{A})$. Hence, rank$(\mathbf{B})$ can be substituted for any occurrence of rank$(\mathbf{A})$ in the statement of Theorem 5.9. Thus, parts (a) and (b) of Theorem 5.9 hold when the matrix for L with respect to any bases is used. Part (c) of Theorem 5.9 follows immediately from parts (a) and (b).

(18) (a) By Theorem 4.15, there are vectors $\mathbf{q}_1, \ldots, \mathbf{q}_t$ in $\mathcal{V}$ such that $B = \{\mathbf{k}_1, \ldots, \mathbf{k}_s, \mathbf{q}_1, \ldots, \mathbf{q}_t\}$ is a basis for $\mathcal{V}$. Hence, $\dim(\mathcal{V}) = s + t$.

(b) If $\mathbf{v} \in \mathcal{V}$, then $\mathbf{v} = a_1\mathbf{k}_1 + \cdots + a_s\mathbf{k}_s + b_1\mathbf{q}_1 + \cdots + b_t\mathbf{q}_t$ for some $a_1, \ldots, a_s, b_1, \ldots, b_t \in \mathbb{R}$. Hence,

$$\begin{aligned} L(\mathbf{v}) &= L(a_1\mathbf{k}_1 + \cdots + a_s\mathbf{k}_s + b_1\mathbf{q}_1 + \cdots + b_t\mathbf{q}_t) \\ &= a_1L(\mathbf{k}_1) + \cdots + a_sL(\mathbf{k}_s) + b_1L(\mathbf{q}_1) + \cdots + b_tL(\mathbf{q}_t) \\ &= a_1\mathbf{0} + \cdots + a_s\mathbf{0} + b_1L(\mathbf{q}_1) + \cdots + b_tL(\mathbf{q}_t) \\ &= b_1L(\mathbf{q}_1) + \cdots + b_tL(\mathbf{q}_t). \end{aligned}$$

(c) By part (b), any element of range(L) is a linear combination of $L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)$, so range$(L)$ is spanned by $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$. Hence, by part (1) of Theorem 4.12, range(L) is finite dimensional and $\dim(\text{range}(L)) \leq t$.

(d) Suppose $c_1L(\mathbf{q}_1) + \cdots + c_tL(\mathbf{q}_t) = \mathbf{0}_{\mathcal{W}}$. Then $L(c_1\mathbf{q}_1 + \cdots + c_t\mathbf{q}_t) = \mathbf{0}_{\mathcal{W}}$, and so $c_1\mathbf{q}_1 + \cdots + c_t\mathbf{q}_t \in \ker(L)$.

(e) Every element of $\ker(L)$ can be expressed as a linear combination of the basis vectors $\mathbf{k}_1, \ldots, \mathbf{k}_s$ for $\ker(L)$.

(f) The fact that $c_1\mathbf{q}_1 + \cdots + c_t\mathbf{q}_t = d_1\mathbf{k}_1 + \cdots + d_s\mathbf{k}_s$ implies that

$$d_1\mathbf{k}_1 + \cdots + d_s\mathbf{k}_s - c_1\mathbf{q}_1 - \cdots - c_t\mathbf{q}_t = \mathbf{0}_{\mathcal{W}}.$$

But since $B = \{\mathbf{k}_1, \ldots, \mathbf{k}_s, \mathbf{q}_1, \ldots, \mathbf{q}_t\}$ is linearly independent,

$$d_1 = \cdots = d_s = c_1 = \cdots = c_t = 0,$$

by the definition of linear independence.

(g) In part (d) we assumed that $c_1L(\mathbf{q}_1) + \cdots + c_tL(\mathbf{q}_t) = \mathbf{0}_{\mathcal{W}}$. This led to the conclusion in part (f) that $c_1 = \cdots = c_t = 0$. Hence $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ is linearly independent by definition.

(h) Part (c) proved that $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ spans range(L) and part (g) shows that $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ is linearly independent. Hence, $\{L(\mathbf{q}_1), \ldots, L(\mathbf{q}_t)\}$ is a basis for range(L).

(i) We assumed that $\dim(\ker(L)) = s$. Part (h) shows that $\dim(\text{range}(L)) = t$.
Therefore, $\dim(\ker(L)) + \dim(\text{range}(L)) = s + t = \dim(\mathcal{V})$ (by part (a)).

(19) From Theorem 4.13, we have $\dim(\ker(L)) \leq \dim(\mathcal{V})$. The Dimension Theorem states that $\dim(\text{range}(L))$ is finite and $\dim(\ker(L)) + \dim(\text{range}(L)) = \dim(\mathcal{V})$. Since $\dim(\ker(L))$ is nonnegative, it follows that $\dim(\text{range}(L)) \leq \dim(\mathcal{V})$.

(20) (a) F (b) F (c) T (d) F (e) T (f) F (g) F (h) F

Section 5.4

(1) (a) Not one-to-one, because $L([1,0,0]) = L([0,0,0]) = [0,0,0,0]$;
not onto, because $[0,0,0,1]$ is not in range(L)

(b) Not one-to-one, because $L([1,-1,1]) = [0,0]$; onto, because $L([a,0,b]) = [a,b]$

(c) One-to-one, because $L([x,y,z]) = [0,0,0]$ implies that $[2x, x+y+z, -y] = [0,0,0]$, which gives $x = y = z = 0$;
onto, because every vector $[a,b,c]$ can be expressed as $[2x, x+y+z, -y]$,
where $x = \frac{a}{2}$, $y = -c$, and $z = b - \frac{a}{2} + c$

(d) Not one-to-one, because $L(1) = 0$; onto, because $L(ax^3 + bx^2 + cx) = ax^2 + bx + c$

(e) One-to-one, because $L(ax^2 + bx + c) = 0$ implies that $a + b = b + c = a + c = 0$, which gives $b = c$ and hence $a = b = c = 0$;
onto, because every polynomial $Ax^2 + Bx + C$ can be expressed as $(a+b)x^2 + (b+c)x + (a+c)$, where $a = (A - B + C)/2$, $b = (A + B - C)/2$, and $c = (-A + B + C)/2$

(f) One-to-one, because $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} d & b+c \\ b-c & a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\Longrightarrow d = a = b + c = b - c = 0 \Longrightarrow a = b = c = d = 0$;
onto, because $L\left(\begin{bmatrix} z & \frac{x+y}{2} \\ \frac{x-y}{2} & w \end{bmatrix}\right) = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$

(g) Not one-to-one, because $L\left(\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}\right) = L\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;

onto, because every 2×2 matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ can be expressed as $\begin{bmatrix} a & -c \\ 2e & d+f \end{bmatrix}$,
where $a = A$, $c = -B$, $e = C/2$, $d = D$, and $f = 0$

(h) One-to-one, because $L(ax^2 + bx + c) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ implies that $a + c = b - c = -3a = 0$, which gives $a = b = c = 0$;
not onto, because $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not in range(L)

(2) (a) One-to-one; onto;
the matrix row reduces to $\mathbf{I}_2$, which means that $\dim(\ker(L)) = 0$ and $\dim(\text{range}(L)) = 2$.

(b) One-to-one; not onto;
the matrix row reduces to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, which means that $\dim(\ker(L)) = 0$ and $\dim(\text{range}(L)) = 2$.

(c) Not one-to-one; not onto;
the matrix row reduces to $\begin{bmatrix} 1 & 0 & -\frac{2}{5} \\ 0 & 1 & -\frac{6}{5} \\ 0 & 0 & 0 \end{bmatrix}$, which means that $\dim(\ker(L)) = 1$ and $\dim(\text{range}(L)) = 2$.

(d) Not one-to-one; onto;
the matrix row reduces to $\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$, which means that $\dim(\ker(L)) = 1$ and $\dim(\text{range}(L)) = 3$.

(3) (a) One-to-one; onto;
the matrix row reduces to $\mathbf{I}_3$ which means that $\dim(\ker(L)) = 0$ and $\dim(\text{range}(L)) = 3$.

(b) Not one-to-one; not onto;
the matrix row reduces to $\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, which means that $\dim(\ker(L)) = 1$ and $\dim(\text{range}(L)) = 3$.

(c) Not one-to-one; not onto;
the matrix row reduces to $\begin{bmatrix} 1 & 0 & -\frac{10}{11} & \frac{19}{11} \\ 0 & 1 & \frac{3}{11} & -\frac{9}{11} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, which means that $\dim(\ker(L)) = 2$ and $\dim(\text{range}(L)) = 2$.

(4) (a) Let $L: \mathbb{R}^n \to \mathbb{R}^m$. Then $\dim(\text{range}(L)) = n - \dim(\ker(L)) \leq n < m$, so L is not onto.

(b) Let $L: \mathbb{R}^m \to \mathbb{R}^n$. Then $\dim(\ker(L)) = m - \dim(\text{range}(L)) \geq m - n > 0$, so L is not one-to-one.

(5) (a) $L(\mathbf{I}_n) = \mathbf{A}\mathbf{I}_n - \mathbf{I}_n\mathbf{A} = \mathbf{A} - \mathbf{A} = \mathbf{O}_n$. Hence $\mathbf{I}_n \in \ker(L)$, and so L is not one-to-one by part (1) of Theorem 5.12.

(b) Use part (a) of this exercise, part (2) of Theorem 5.12, and the Dimension Theorem.

(6) If $L(\mathbf{A}) = \mathbf{O}_3$, then $\mathbf{A}^T = -\mathbf{A}$, so $\mathbf{A}$ is skew-symmetric. But since $\mathbf{A}$ is also upper triangular, $\mathbf{A} = \mathbf{O}_3$. Thus, $\ker(L) = \{\mathbf{0}\}$ and L is one-to-one. By the Dimension Theorem, $\dim(\text{range}(L)) = \dim(\mathcal{U}_3) = 6$ (since $\dim(\ker(L)) = 0$). Hence, $\text{range}(L) \neq \mathcal{M}_{33}$, and L is not onto.

(7) (a) No, by Corollary 5.13, because $\dim(\mathbb{R}^6) = \dim(\mathcal{P}_5)$.

(b) No, by Corollary 5.13, because $\dim(\mathcal{M}_{22}) = \dim(\mathcal{P}_3)$.

(8) (a) Clearly, multiplying a nonzero polynomial by x produces another nonzero polynomial. Hence, $\ker(L) = \{\mathbf{0}_{\mathcal{P}}\}$, and L is one-to-one. But, every nonzero element of $\text{range}(L)$ has degree at least 1, since L multiplies by x. Hence, the constant polynomial 1 is not in $\text{range}(L)$, and L is not onto.

(b) Corollary 5.13 requires that the domain and codomain of the linear transformation be finite dimensional. However, $\mathcal{P}$ is infinite dimensional.

(9) (a) Consider $L\colon \mathcal{P}_2 \to \mathbb{R}^3$ given by $L(\mathbf{p}) = [\mathbf{p}(x_1), \mathbf{p}(x_2), \mathbf{p}(x_3)]$. Now, $\dim(\mathcal{P}_2) = \dim(\mathbb{R}^3) = 3$. Hence, by Corollary 5.13, if L is either one-to-one or onto, it has the other property as well.

We will show that L is one-to-one using part (1) of Theorem 5.12. If $\mathbf{p} \in \ker(L)$, then $L(\mathbf{p}) = \mathbf{0}$, and so $\mathbf{p}(x_1) = \mathbf{p}(x_2) = \mathbf{p}(x_3) = 0$. Hence $\mathbf{p}$ is a polynomial of degree ≤ 2 touching the x-axis at $x = x_1$, $x = x_2$, and $x = x_3$. Since the graph of $\mathbf{p}$ must be either a parabola or a line, it cannot touch the x-axis at three distinct points unless its graph is the line $y = 0$. That is, $\mathbf{p} = \mathbf{0}$ in $\mathcal{P}_2$. Therefore, $\ker(L) = \{\mathbf{0}\}$, and L is one-to-one.

Now, by Corollary 5.13, L is onto. Thus, given any 3-vector $[a, b, c]$, there is some $\mathbf{p} \in \mathcal{P}_2$ such that $\mathbf{p}(x_1) = a$, $\mathbf{p}(x_2) = b$, and $\mathbf{p}(x_3) = c$.

(b) From part (a), L is an isomorphism. Hence, any $[a, b, c] \in \mathbb{R}^3$ has a unique pre-image under L in $\mathcal{P}_2$.

(c) Generalize parts (a) and (b) using $L\colon \mathcal{P}_n \to \mathbb{R}^{n+1}$ given by $L(\mathbf{p}) = [\mathbf{p}(x_1), \ldots, \mathbf{p}(x_n), \mathbf{p}(x_{n+1})]$. Note that any polynomial in $\ker(L)$ has $n+1$ distinct roots, and so must be trivial. Hence L is one-to-one, and therefore by Corollary 5.13, L is onto. Thus, given any $(n+1)$-vector $[a_1, \ldots, a_n, a_{n+1}]$, there is a unique $\mathbf{p} \in \mathcal{P}_n$ such that $\mathbf{p}(x_1) = a_1, \ldots, \mathbf{p}(x_n) = a_n$, and $\mathbf{p}(x_{n+1}) = a_{n+1}$.

(10) Suppose L is not one-to-one. Then $\ker(L)$ is nontrivial. Let $\mathbf{v} \in \ker(L)$ with $\mathbf{v} \neq \mathbf{0}$. Then $T = \{\mathbf{v}\}$ is linearly independent. But $L(T) = \{\mathbf{0}_{\mathcal{W}}\}$, which is linearly dependent, a contradiction.

(11) (a) Suppose $\mathbf{w} \in L(\text{span}(S))$. Then there is a vector $\mathbf{v} \in \text{span}(S)$ such that $L(\mathbf{v}) = \mathbf{w}$. There are also vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in S$ and scalars $a_1, \ldots, a_k$ such that $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{v}$. Hence,

$$\begin{aligned} \mathbf{w} &= L(\mathbf{v}) = L(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) \\ &= L(a_1\mathbf{v}_1) + \cdots + L(a_k\mathbf{v}_k) \\ &= a_1L(\mathbf{v}_1) + \cdots + a_kL(\mathbf{v}_k), \end{aligned}$$

which shows that $\mathbf{w} \in \text{span}(L(S))$. Hence, $L(\text{span}(S)) \subseteq \text{span}(L(S))$.

Next, $S \subseteq \text{span}(S)$ (by Theorem 4.5)
$\implies L(S) \subseteq L(\text{span}(S))$
$\implies \text{span}(L(S)) \subseteq L(\text{span}(S))$ (by part (3) of Theorem 4.5, since $L(\text{span}(S))$ is a subspace of $\mathcal{W}$ by part (1) of Theorem 5.3). Therefore, $L(\text{span}(S)) = \text{span}(L(S))$.

(b) By part (a), $L(S)$ spans $L(\text{span}(S)) = L(\mathcal{V}) = \text{range}(L)$.

(c) $\mathcal{W} = \text{span}(L(S)) = L(\text{span}(S))$ (by part (a)) $\subseteq L(\mathcal{V})$ (because $\text{span}(S) \subseteq \mathcal{V}) = \text{range}(L)$. Thus, L is onto.

(12) (a) F (b) F (c) T (d) T (e) T (f) T (g) T (h) F (i) F

Section 5.5

(1) In each part, let **A** represent the given matrix for L_1 and let **B** represent the given matrix for L_2. By Theorem 5.16, L_1 is an isomorphism if and only if **A** is nonsingular, and L_2 is an isomorphism if and only if **B** is nonsingular. In each part, we will give $|\mathbf{A}|$ and $|\mathbf{B}|$ to show that **A** and **B** are nonsingular.

(a) $|\mathbf{A}| = 1$, $|\mathbf{B}| = 3$, $L_1^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & -2 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$L_2^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{3} \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $(L_2 \circ L_1)\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & -2 & 1 \\ 1 & 4 & -2 \\ 0 & 3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$(L_2 \circ L_1)^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = (L_1^{-1} \circ L_2^{-1})\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \\ 1 & 0 & \frac{2}{3} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(b) $|\mathbf{A}| = -1$, $|\mathbf{B}| = -1$, $L_1^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & -2 & 1 \\ 0 & 1 & 0 \\ 1 & -8 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$L_2^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $(L_2 \circ L_1)\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$(L_2 \circ L_1)^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = (L_1^{-1} \circ L_2^{-1})\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(c) $|\mathbf{A}| = -1$, $|\mathbf{B}| = 1$, $L_1^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2 & -4 & -1 \\ 7 & -13 & -3 \\ 5 & -10 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$L_2^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & -3 \\ -1 & -2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$(L_2 \circ L_1)\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 29 & -6 & -4 \\ 21 & -5 & -2 \\ 38 & -8 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$(L_2 \circ L_1)^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = (L_1^{-1} \circ L_2^{-1})\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -9 & -2 & 8 \\ -29 & -7 & 26 \\ -22 & -4 & 19 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(2) L is clearly a linear operator, and L is invertible ($L^{-1} = L$). By Theorem 5.15, L is an isomorphism.

(3) (a) L_1 is a linear operator:

$$\begin{aligned} L_1(\mathbf{B}+\mathbf{C}) &= \mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC} = L_1(\mathbf{B})+L_1(\mathbf{C}); \\ L_1(c\mathbf{B}) &= \mathbf{A}(c\mathbf{B}) = c(\mathbf{AB}) = cL_1(\mathbf{B}). \end{aligned}$$

Note that L_1 is invertible ($L_1^{-1}(\mathbf{B}) = \mathbf{A}^{-1}\mathbf{B}$). Use Theorem 5.15.

(b) L_2 is a linear operator:

$$\begin{aligned} L_2(\mathbf{B}+\mathbf{C}) &= \mathbf{A}(\mathbf{B}+\mathbf{C})\mathbf{A}^{-1} = (\mathbf{AB}+\mathbf{AC})\mathbf{A}^{-1} \\ &= \mathbf{ABA}^{-1}+\mathbf{ACA}^{-1} = L_2(\mathbf{B})+L_2(\mathbf{C}); \\ L_2(c\mathbf{B}) &= \mathbf{A}(c\mathbf{B})\mathbf{A}^{-1} = c(\mathbf{ABA}^{-1}) = cL_2(\mathbf{B}). \end{aligned}$$

Note that L_2 is invertible ($L_2^{-1}(\mathbf{B}) = \mathbf{A}^{-1}\mathbf{BA}$). Use Theorem 5.15.

(4) L is a linear operator:

$$\begin{aligned} L(\mathbf{p}+\mathbf{q}) &= (\mathbf{p}+\mathbf{q})+(\mathbf{p}+\mathbf{q})' = \mathbf{p}+\mathbf{q}+\mathbf{p}'+\mathbf{q}' \\ &= (\mathbf{p}+\mathbf{p}')+(\mathbf{q}+\mathbf{q}') = L(\mathbf{p})+L(q); \\ L(c\mathbf{p}) &= (c\mathbf{p})+(c\mathbf{p})' = c\mathbf{p}+c\mathbf{p}' = c(\mathbf{p}+\mathbf{p}') = cL(\mathbf{p}). \end{aligned}$$

Now, if $L(\mathbf{p}) = \mathbf{0}$, then $\mathbf{p}+\mathbf{p}' = \mathbf{0} \Longrightarrow \mathbf{p} = -\mathbf{p}'$. But if $\mathbf{p}$ is not a constant polynomial, then $\mathbf{p}$ and $\mathbf{p}'$ have different degrees, a contradiction. Hence, $\mathbf{p}' = \mathbf{0}$, and so $\mathbf{p} = \mathbf{0}$. Thus, $\ker(L) = \{\mathbf{0}\}$ and L is one-to-one. Then, by Corollary 5.13, L is an isomorphism.

(5) (a) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(b) Use Theorem 5.16 (since the matrix in part (a) is nonsingular).

(c) If $\mathbf{A}$ is the matrix for R in part (a), then $\mathbf{A}^2 = \mathbf{I}_2$, so $R = R^{-1}$.

(d) Performing the reflection R twice in succession gives the identity mapping.

(6) The change of basis is performed by multiplying by a nonsingular transition matrix. Use Theorem 5.16. (Note that multiplication by a matrix is a linear transformation.)

(7) $(M \circ L_1)(\mathbf{v}) = (M \circ L_2)(\mathbf{v}) \implies (M^{-1} \circ M \circ L_1)(\mathbf{v}) = (M^{-1} \circ M \circ L_2)(\mathbf{v}) \implies L_1(\mathbf{v}) = L_2(\mathbf{v})$

(8) Suppose $\dim(\mathcal{V}) = n > 0$ and $\dim(\mathcal{W}) = k > 0$.

Suppose $\mathbf{A}_{BC}$ is nonsingular, then $\mathbf{A}_{BC}$ is square, $n = k$, $\mathbf{A}_{BC}^{-1}$ exists, and $\mathbf{A}_{BC}^{-1}\mathbf{A}_{BC} = \mathbf{A}_{BC}\mathbf{A}_{BC}^{-1} = \mathbf{I}_n$. Let K be the linear operator from $\mathcal{W}$ to $\mathcal{V}$ whose matrix with respect to C and B is $\mathbf{A}_{BC}^{-1}$. Then, for all $\mathbf{v} \in \mathcal{V}$, $[(K \circ L)(\mathbf{v})]_B = \mathbf{A}_{BC}^{-1}\mathbf{A}_{BC}[\mathbf{v}]_B = \mathbf{I}_n[\mathbf{v}]_B = [\mathbf{v}]_B$. Therefore, $(K \circ L)(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in \mathcal{V}$, since coordinatizations are unique.

Similarly, for all $\mathbf{w} \in \mathcal{W}$, $[(L \circ K)(\mathbf{w})]_C = \mathbf{A}_{BC}\mathbf{A}_{BC}^{-1}[\mathbf{w}]_C = \mathbf{I}_n[\mathbf{w}]_C = [\mathbf{w}]_C$. Hence $(L \circ K)(\mathbf{w}) = \mathbf{w}$ for all $\mathbf{w} \in \mathcal{W}$. Therefore, K acts as an inverse for L. Hence, L is an isomorphism by Theorem 5.15.

(9) In all parts, use Corollary 5.20. In part (c), the answer to the question is yes.

(10) If L is an isomorphism, $(L \circ L)^{-1} = L^{-1} \circ L^{-1}$, so $L \circ L$ is an isomorphism by Theorem 5.15.

Conversely, if $L \circ L$ is an isomorphism, it is easy to prove L is one-to-one and onto directly. Or, let $f = L \circ (L \circ L)^{-1}$ and $g = (L \circ L)^{-1} \circ L$. Then clearly, $L \circ f = I$ (identity operator on $\mathcal{V}$) $= g \circ L$. But $f = I \circ f = g \circ L \circ f = g \circ I = g$, and so $f = g = L^{-1}$. Hence L is an isomorphism by Theorem 5.15.

(11) (a) Let $\mathbf{v} \in \mathcal{V}$ with $\mathbf{v} \neq \mathbf{0}_{\mathcal{V}}$. Since $(L \circ L)(\mathbf{v}) = \mathbf{0}_{\mathcal{V}}$, either $L(\mathbf{v}) = \mathbf{0}_{\mathcal{V}}$ or $L(L(\mathbf{v})) = \mathbf{0}_{\mathcal{V}}$, with $L(\mathbf{v}) \neq \mathbf{0}_{\mathcal{V}}$. Hence, $\ker(L) \neq \{\mathbf{0}_{\mathcal{V}}\}$, since one of the two vectors $\mathbf{v}$ or $L(\mathbf{v})$ is nonzero and is in $\ker(L)$.

(b) Proof by contrapositive: The goal is to prove that if L is an isomorphism, then $L \circ L \neq L$ or L is the identity transformation. Assume L is an isomorphism and $L \circ L = L$ (see "If A, Then B or C Proofs" in Section 1.3). Then L is the identity transformation because

$$(L \circ L)(\mathbf{v}) = L(\mathbf{v}) \implies (L^{-1} \circ L \circ L)(\mathbf{v}) = (L^{-1} \circ L)(\mathbf{v}) \implies L(\mathbf{v}) = \mathbf{v}.$$

(12) Range(L) = column space of $\mathbf{A}$ (see comments just before the Range Method in Section 5.3 of the textbook).
Now, L is an isomorphism iff L is onto (by Corollary 5.13)
iff range$(L) = \mathbb{R}^n$ iff span({columns of $\mathbf{A}$}) $= \mathbb{R}^n$
iff the n columns of $\mathbf{A}$ are linearly independent (by part (2) of Theorem 4.12).

(13) (a) Use Theorem 5.14.

(b) By Exercise 11(c) in Section 5.4, L is onto. Now, if L is not one-to-one, then $\ker(L)$ is nontrivial. Suppose $\mathbf{v} \in \ker(L)$ with $\mathbf{v} \neq \mathbf{0}$. If $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then there are $a_1, \ldots, a_n \in \mathbb{R}$ such that $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, where not every $a_i = 0$ (since $\mathbf{v} \neq \mathbf{0}$). Then

$$\mathbf{0}_{\mathcal{W}} = L(\mathbf{v}) = L(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1L(\mathbf{v}_1) + \cdots + a_nL(\mathbf{v}_n).$$

But since each $L(\mathbf{v}_j)$ is distinct, and some $a_i \neq 0$, this contradicts the linear independence of $L(B)$. Hence L is one-to-one.

(c) $T(\mathbf{e}_1) = [3,1]$, $T(\mathbf{e}_2) = [5,2]$, and $T(\mathbf{e}_3) = [3,1]$. Hence, $T(B) = \{[3,1],[5,2]\}$ (because identical elements of a set are not listed more than once). This set is clearly a basis for $\mathbb{R}^2$. T is not an isomorphism or else we would contradict Theorem 5.18, since $\dim(\mathbb{R}^3) \neq \dim(\mathbb{R}^2)$.

(d) Part (c) is not a counterexample to part (b) because in part (c) the images of the vectors in B are not distinct.

(14) Let $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, let $\mathbf{v} \in \mathcal{V}$, and let $a_1, \ldots, a_n \in \mathbb{R}$ such that $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$.
Then $[\mathbf{v}]_B = [a_1, \ldots, a_n]$. But

$$L(\mathbf{v}) = L(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1L(\mathbf{v}_1) + \cdots + a_nL(\mathbf{v}_n).$$

From Exercise 13(a), $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ is a basis for $L(B) = \mathcal{W}$ (since L is onto).
Hence, $[L(\mathbf{v})]_{L(B)} = [a_1, \ldots, a_n]$ as well.

(15) (Note: This exercise will be used as part of the proof of the Dimension Theorem, so we do not use the Dimension Theorem in this proof. Notice that we can use both Theorem 5.14 and part (a) of Exercise 11 because they were previously proven without invoking the Dimension Theorem.)

Because $\mathcal{Y}$ is a subspace of the finite dimensional vector space $\mathcal{V}$, $\mathcal{Y}$ is also finite dimensional (Theorem 4.13). Suppose $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ is a basis for $\mathcal{Y}$. Because L is one-to-one, the vectors $L(\mathbf{y}_1), \ldots, L(\mathbf{y}_k)$ are distinct. By part (1) of Theorem 5.14, $\{L(\mathbf{y}_1), \ldots, L(\mathbf{y}_k)\} = L(\{\mathbf{y}_1, \ldots, \mathbf{y}_k\})$ is linearly independent. Part (a) of Exercise 11 in Section 5.4 shows that

$$\text{span}(L(\{\mathbf{y}_1, \ldots, \mathbf{y}_k\})) = L(\text{span}(\{\mathbf{y}_1, \ldots, \mathbf{y}_k\})) = L(\mathcal{Y}).$$

Therefore, $\{L(\mathbf{y}_1), \ldots, L(\mathbf{y}_k)\}$ is a k-element basis for $L(\mathcal{Y})$. Hence,

$$\dim(L(\mathcal{Y})) = k = \dim(\mathcal{Y}).$$

(16) (a) Let $\mathbf{v} \in \ker(T)$. Then
$$(T_2 \circ T)(\mathbf{v}) = T_2(T(\mathbf{v})) = T_2(\mathbf{0}_{\mathcal{W}}) = \mathbf{0}_{\mathcal{Y}},$$
and so $\mathbf{v} \in \ker(T_2 \circ T)$. Hence, $\ker(T) \subseteq \ker(T_2 \circ T)$.

Next, suppose $\mathbf{v} \in \ker(T_2 \circ T)$. Then
$$\mathbf{0}_{\mathcal{Y}} = (T_2 \circ T)(\mathbf{v}) = T_2(T(\mathbf{v})),$$
which implies that $T(\mathbf{v}) \in \ker(T_2)$. But T_2 is one-to-one, and so $\ker(T_2) = \{\mathbf{0}_{\mathcal{W}}\}$. Therefore, $T(\mathbf{v}) = \mathbf{0}_{\mathcal{W}}$, implying that $\mathbf{v} \in \ker(T)$. Hence, $\ker(T_2 \circ T) \subseteq \ker(T)$.
Therefore, $\ker(T) = \ker(T_2 \circ T)$.

(b) Suppose $\mathbf{w} \in \text{range}(T \circ T_1)$. Then there is some $\mathbf{x} \in \mathcal{X}$ such that $(T \circ T_1)(\mathbf{x}) = \mathbf{w}$, implying $T(T_1(\mathbf{x})) = \mathbf{w}$. Hence $\mathbf{w}$ is the image under T of $T_1(\mathbf{x})$ and so $\mathbf{w} \in \text{range}(T)$. Thus, $\text{range}(T \circ T_1) \subseteq \text{range}(T)$. Similarly, if $\mathbf{w} \in \text{range}(T)$, there is some $\mathbf{v} \in \mathcal{V}$ such that $T(\mathbf{v}) = \mathbf{w}$. Since T_1^{-1} exists, $(T \circ T_1)(T_1^{-1}(\mathbf{v})) = \mathbf{w}$, and so $\mathbf{w} \in \text{range}(T \circ T_1)$. Thus, $\text{range}(T) \subseteq \text{range}(T \circ T_1)$, finishing the proof.

(c) Let $\mathbf{v} \in T_1(\ker(T \circ T_1))$. Thus, there is an $\mathbf{x} \in \ker(T \circ T_1)$ such that $T_1(\mathbf{x}) = \mathbf{v}$. Then,
$$T(\mathbf{v}) = T(T_1(\mathbf{x})) = (T \circ T_1)(\mathbf{x}) = \mathbf{0}_{\mathcal{W}}.$$
Thus, $\mathbf{v} \in \ker(T)$. Therefore, $T_1(\ker(T \circ T_1)) \subseteq \ker(T)$.

Now suppose that $\mathbf{v} \in \ker(T)$. Then $T(\mathbf{v}) = \mathbf{0}_{\mathcal{W}}$. Thus, $(T \circ T_1)(T_1^{-1}(\mathbf{v})) = \mathbf{0}_{\mathcal{W}}$ as well. Hence, $T_1^{-1}(\mathbf{v}) \in \ker(T \circ T_1)$. Therefore,
$$\mathbf{v} = T_1(T_1^{-1}(\mathbf{v})) \in T_1(\ker(T \circ T_1)),$$
implying $\ker(T) \subseteq T_1(\ker(T \circ T_1))$, completing the proof.

(d) From part (c), $\dim(\ker(T)) = \dim(T_1(\ker(T \circ T_1)))$. Apply Exercise 15 using T_1 as L and $\mathcal{Y} = \ker(T \circ T_1)$ to show that
$$\dim(T_1(\ker(T \circ T_1))) = \dim(\ker(T \circ T_1)).$$

(e) Suppose $\mathbf{y} \in \text{range}(T_2 \circ T)$. Hence, there is a $\mathbf{v} \in \mathcal{V}$ such that $(T_2 \circ T)(\mathbf{v}) = \mathbf{y}$. Thus, $\mathbf{y} = T_2(T(\mathbf{v}))$. Clearly, $T(\mathbf{v}) \in \text{range}(T)$, and so $\mathbf{y} \in T_2(\text{range}(T))$. This proves that
$$\text{range}(T_2 \circ T) \subseteq T_2(\text{range}(T)).$$

Next, suppose that $\mathbf{y} \in T_2(\text{range}(T))$. Then there is some $\mathbf{w} \in \text{range}(T)$ such that $T_2(\mathbf{w}) = \mathbf{y}$. Since $\mathbf{w} \in \text{range}(T)$, there is a $\mathbf{v} \in \mathcal{V}$ such that $T(\mathbf{v}) = \mathbf{w}$. Therefore,
$$(T_2 \circ T)(\mathbf{v}) = T_2(T(\mathbf{v})) = T_2(\mathbf{w}) = \mathbf{y}.$$
This establishes that $T_2(\text{range}(T)) \subseteq \text{range}(T_2 \circ T)$, finishing the proof.

(f) From part (e),
$$\dim(\text{range}(T_2 \circ T)) = \dim(T_2(\text{range}(T))).$$
Apply Exercise 15 using T_2 as L and $\mathcal{Y} = \text{range}(T)$ to show that
$$\dim(T_2(\text{range}(T))) = \dim(\text{range}(T)).$$

(17) (a) Part (c) of Exercise 16 states that $T_1(\ker(T \circ T_1)) = \ker(T)$. Substituting $T = L_2 \circ L$ and $T_1 = L_1^{-1}$ yields

$$L_1^{-1}\left(\ker\left(L_2 \circ L \circ L_1^{-1}\right)\right) = \ker(L_2 \circ L).$$

Now $M = L_2 \circ L \circ L_1^{-1}$. Hence, $L_1^{-1}(\ker(M)) = \ker(L_2 \circ L)$.

(b) With $L_2 = T_2$ and $L = T$, part (a) of Exercise 16 shows that $\ker(L_2 \circ L) = \ker(L)$. This, combined with part (a) of this exercise, shows that $L_1^{-1}(\ker(M)) = \ker(L)$.

(c) From part (b),

$$\dim(\ker(L)) = \dim(L_1^{-1}(\ker(M))).$$

Apply Exercise 15 with L replaced by L_1^{-1} and $\mathcal{Y} = \ker(M)$ to show that

$$\dim(L_1^{-1}(\ker(M))) = \dim(\ker(M)),$$

completing the proof.

(d) Part (e) of Exercise 16 states that $\text{range}(T_2 \circ T) = T_2(\text{range}(T))$. Substituting $L \circ L_1^{-1}$ for T and L_2 for T_2 yields

$$\text{range}\left(L_2 \circ L \circ L_1^{-1}\right) = L_2\left(\text{range}(L \circ L_1^{-1})\right).$$

Apply L_2^{-1} to both sides to produce

$$L_2^{-1}(\text{range}\left(L_2 \circ L \circ L_1^{-1}\right)) = \text{range}(L \circ L_1^{-1}).$$

Using $M = L_2 \circ L \circ L_1^{-1}$ gives $L_2^{-1}(\text{range}(M)) = \text{range}(L \circ L_1^{-1})$, the desired result.

(e) Part (b) of Exercise 16 states that $\text{range}(T \circ T_1) = \text{range}(T)$. Substituting L for T and L_1^{-1} for T_1 produces $\text{range}(L \circ L_1^{-1}) = \text{range}(L)$. Combining this with

$$L_2^{-1}(\text{range}(M)) = \text{range}(L \circ L_1^{-1})$$

from part (d) of this exercise yields $L_2^{-1}(\text{range}(M)) = \text{range}(L)$.

(f) From part (e), $\dim(L_2^{-1}(\text{range}(M))) = \dim(\text{range}(L))$. Apply Exercise 15 with L replaced by L_2^{-1} and $\mathcal{Y} = \text{range}(M)$ to show that

$$\dim(L_2^{-1}(\text{range}(M))) = \dim(\text{range}(M)),$$

completing the proof.

(18) (a) A nonsingular 2×2 matrix must have at least one of its first column entries nonzero. Then, use an appropriate equation from the two given in the problem.

(b) $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} kx \\ y \end{bmatrix}$, a contraction (if $k \le 1$) or dilation (if $k \ge 1$) along the x-coordinate.

Similarly, $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ ky \end{bmatrix}$, a contraction (if $k \le 1$) or dilation (if $k \ge 1$) along the y-coordinate.

(c) By the hint, multiplying by $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ first performs a dilation/contraction by part (a) followed by a reflection about the y-axis. Using $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -k \end{bmatrix}$ shows that multiplying by $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$ first performs a dilation/contraction by part (a) followed by a reflection about the x-axis.

(d) These matrices are defined to represent shears in Exercise 11 in Section 5.1.

(e) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$; that is, this matrix multiplication switches the coordinates of a vector in $\mathbb{R}^2$. That is precisely the result of a reflection through the line $y = x$.

(f) This is merely a summary of parts (a) through (e).

(19) $\begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ \frac{\sqrt{2}}{2} & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{2} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$,

where the matrices on the right side are, respectively, a contraction along the x-coordinate, a shear in the y-direction, a dilation along the y-coordinate, and a shear in the x-direction.

(20) (a) T (b) T (c) F (d) F (e) F (f) T (g) T (h) T

Section 5.6

(1) (a)

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = 2$	$\{[1,0]\}$	2	1

(b)

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = 3$	$\{[1,4]\}$	1	1
$\lambda_2 = 2$	$\{[0,1]\}$	1	1

(c)

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = 1$	$\{[2,1,1]\}$	1	1
$\lambda_2 = -1$	$\{[-1,0,1]\}$	1	1
$\lambda_3 = 2$	$\{[1,1,1]\}$	1	1

(d)

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = 2$	$\{[5,4,0],[3,0,2]\}$	2	2
$\lambda_2 = 3$	$\{[0,1,-1]\}$	1	1

(e)

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = -1$	$\{[-1,2,3]\}$	2	1
$\lambda_2 = 0$	$\{[-1,1,3]\}$	1	1

(f)

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = 0$	$\{[6,-1,1,4]\}$	2	1
$\lambda_2 = 2$	$\{[7,1,0,5]\}$	2	1

(2) (a) $C = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$; $B = ([1,1,0,0],[0,0,1,1],[1,-1,0,0],[0,0,1,-1])$;

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

(b) $C = (x^2, x, 1)$; $B = (x^2 - 2x + 1, -x + 1, 1)$;

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

(c) Not diagonalizable; $C = (x^2, x, 1)$; $\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 2 & -3 & 0 \\ 0 & 1 & -3 \end{bmatrix}$;

Eigenvalue	Basis for E_λ	Alg. Mult.	Geom. Mult.
$\lambda_1 = -1$	$\{2x^2 + 2x + 1\}$	1	1
$\lambda_2 = -3$	$\{1\}$	2	1

(d) $C = (x^2, x, 1)$; $B = (2x^2 - 8x + 9, x, 1)$;

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ -12 & -4 & 0 \\ 18 & 0 & -5 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -5 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 2 & 0 & 0 \\ -8 & 1 & 0 \\ 9 & 0 & 1 \end{bmatrix}$$

(e) Not diagonalizable; no eigenvalues; $\mathbf{A} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$

(f) $C = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$;

$B = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right)$;

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(g) $C = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$;

$B = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right)$;

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(h) $C = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$;

$B = \left(\begin{bmatrix} 3 & 0 \\ 5 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \right)$;

$$\mathbf{A} = \begin{bmatrix} -4 & 0 & 3 & 0 \\ 0 & -4 & 0 & 3 \\ -10 & 0 & 7 & 0 \\ 0 & -10 & 0 & 7 \end{bmatrix}; \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 5 & 0 & 2 & 0 \\ 0 & 5 & 0 & 2 \end{bmatrix}$$

(3) (a) $p_L(x) = \begin{vmatrix} (x-5) & -2 & 0 & -1 \\ 2 & (x-1) & 0 & 1 \\ -4 & -4 & (x-3) & -2 \\ -16 & 0 & 8 & (x+5) \end{vmatrix}$

$= (x-3)\begin{vmatrix} (x-5) & -2 & -1 \\ 2 & (x-1) & 1 \\ -16 & 0 & x+5 \end{vmatrix} - 8\begin{vmatrix} (x-5) & -2 & -1 \\ 2 & (x-1) & 1 \\ -4 & -4 & -2 \end{vmatrix}$

$= (x-3)\left(-16\begin{vmatrix} -2 & -1 \\ (x-1) & 1 \end{vmatrix} + (x+5)\begin{vmatrix} (x-5) & -2 \\ 2 & (x-1) \end{vmatrix}\right)$

$- 8\left((x-5)\begin{vmatrix} (x-1) & 1 \\ -4 & -2 \end{vmatrix} - 2\begin{vmatrix} -2 & -1 \\ -4 & -2 \end{vmatrix} - 4\begin{vmatrix} -2 & -1 \\ (x-1) & 1 \end{vmatrix}\right)$

$= (x-3)(-16(x-3) + (x+5)(x^2-6x+9)) - 8((x-5)(-2x+6) - 2(0) - 4(x-3))$
$= (x-3)(-16(x-3) + (x+5)(x-3)^2) - 8((-2)(x-5)(x-3) - 4(x-3))$
$= (x-3)^2(-16 + (x+5)(x-3)) + 16(x-3)((x-5)+2)$
$= (x-3)^2(x^2+2x-31) + 16(x-3)^2$
$= (x-3)^2(x^2+2x-15) = (x-3)^3(x+5)$.

You can check that this polynomial expands as claimed.

(b) $\begin{bmatrix} -10 & -2 & 0 & -1 \\ 2 & -6 & 0 & 1 \\ -4 & -4 & -8 & -2 \\ -16 & 0 & 8 & 0 \end{bmatrix}$ becomes $\begin{bmatrix} 1 & 0 & 0 & \frac{1}{8} \\ 0 & 1 & 0 & -\frac{1}{8} \\ 0 & 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 & 0 \end{bmatrix}$ when put into reduced row echelon form.

(4) For both parts, the only eigenvalue is $\lambda = 1$; $E_1 = \{1\}$.

(5) Since $\mathbf{A}$ is upper triangular, so is $x\mathbf{I}_n - \mathbf{A}$. Thus, by Theorem 3.2, $|x\mathbf{I}_n - \mathbf{A}|$ equals the product of the main diagonal entries of $x\mathbf{I}_n - \mathbf{A}$, which is $(x-a_{ii})^n$, since $a_{11} = a_{22} = \cdots = a_{nn}$. Thus $\mathbf{A}$ has exactly one eigenvalue λ with algebraic multiplicity n. Hence $\mathbf{A}$ is diagonalizable $\Longleftrightarrow \lambda$ has geometric multiplicity $n \Longleftrightarrow E_\lambda = \mathbb{R}^n \Longleftrightarrow \mathbf{Av} = \lambda\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n \Longleftrightarrow \mathbf{Av} = \lambda\mathbf{v}$ for all $\mathbf{v} \in \{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \Longleftrightarrow \mathbf{A} = \lambda\mathbf{I}_n$ (using Exercise 14(b) in Section 1.5).

(6) In both cases, the given linear operator is diagonalizable, but there are fewer than n distinct eigenvalues. This occurs because at least one of the eigenvalues has geometric multiplicity greater than 1.

(7) (a) $\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; eigenvalue $\lambda = 1$; basis for $E_1 = \{[1,0,0], [0,1,1]\}$; λ has algebraic multiplicity 3 and geometric multiplicity 2

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; eigenvalues $\lambda_1 = 1, \lambda_2 = 0$; basis for $E_1 = \{[1,0,0],\ [0,1,0]\}$; λ_1 has algebraic multiplicity 2 and geometric multiplicity 2

(8) (a) Zero is not an eigenvalue for L iff $\ker(L) = \{\mathbf{0}\}$. Now apply part (1) of Theorem 5.12 and Corollary 5.13.

(b) $L(\mathbf{v}) = \lambda\mathbf{v} \implies L^{-1}(L(\mathbf{v})) = L^{-1}(\lambda\mathbf{v}) \implies \mathbf{v} = \lambda L^{-1}(\mathbf{v}) \implies \frac{1}{\lambda}\mathbf{v} = L^{-1}(\mathbf{v})$ (since $\lambda \neq 0$ by part (a)).

(9) Let $\mathbf{P}$ be such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal. Let $C = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, where $[\mathbf{v}_i]_B = i$th column of $\mathbf{P}$. Then, by definition, $\mathbf{P}$ is the transition matrix from C to B. Then, for all $\mathbf{v} \in \mathcal{V}$,

$$\mathbf{P}^{-1}\mathbf{AP}[\mathbf{v}]_C = \mathbf{P}^{-1}\mathbf{A}[\mathbf{v}]_B = \mathbf{P}^{-1}[L(\mathbf{v})]_B = [L(\mathbf{v})]_C.$$

Hence, as in part (a), $\mathbf{P}^{-1}\mathbf{AP}$ is the matrix for L with respect to C. But $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

(10) If $\mathbf{P}$ is the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ is a diagonal matrix with $\lambda_1, \ldots, \lambda_n$ as the main diagonal entries of $\mathbf{D}$. Clearly $|\mathbf{D}| = \lambda_1\lambda_2\cdots\lambda_n$. But

$$|\mathbf{D}| = |\mathbf{P}^{-1}\mathbf{AP}| = |\mathbf{P}^{-1}||\mathbf{A}||\mathbf{P}| = |\mathbf{A}||\mathbf{P}^{-1}||\mathbf{P}| = |\mathbf{A}|.$$

(11) Clearly $n \geq \sum_{i=1}^{k}$(algebraic multiplicity of λ_i) $\geq \sum_{i=1}^{k}$(geometric multiplicity of λ_i), by Theorem 5.26.

(12) Proof by contrapositive: If $p_L(x)$ has a nonreal root, then the sum of the algebraic multiplicities of the eigenvalues of L is less than n. Apply Theorem 5.28.

(13) (a) $\mathbf{A}(\mathbf{Bv}) = (\mathbf{AB})\mathbf{v} = (\mathbf{BA})\mathbf{v} = \mathbf{B}(\mathbf{Av}) = \mathbf{B}(\lambda\mathbf{v}) = \lambda(\mathbf{Bv})$.

(b) The eigenspaces E_λ for $\mathbf{A}$ are one-dimensional. Let $\mathbf{v} \in E_\lambda$, with $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{Bv} \in E_\lambda$, by part (a). Hence $\mathbf{Bv} = c\mathbf{v}$ for some $c \in \mathbb{R}$, and so $\mathbf{v}$ is an eigenvector for $\mathbf{B}$. Repeating this for each eigenspace of $\mathbf{A}$ shows that $\mathbf{B}$ has n linearly independent eigenvectors. Now apply Theorem 5.22.

(14) (a) Suppose $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors for $\mathbf{A}$ corresponding to λ_1 and λ_2, respectively. Let $\mathbf{K}_1$ be the 2×2 matrix $[\mathbf{v}_1\ \mathbf{0}]$ (where the vectors represent columns). Similarly, let $\mathbf{K}_2 = [\mathbf{0}\ \mathbf{v}_1]$, $\mathbf{K}_3 = [\mathbf{v}_2\ \mathbf{0}]$, and $\mathbf{K}_4 = [\mathbf{0}\ \mathbf{v}_2]$. Then a little thought shows that $\{\mathbf{K}_1, \mathbf{K}_2\}$ is a basis for E_{λ_1} in $\mathcal{M}_{22}$, and $\{\mathbf{K}_3, \mathbf{K}_4\}$ is a basis for E_{λ_2} in $\mathcal{M}_{22}$.

(b) Generalize the construction in part (a).

(15) Suppose $\mathbf{v} \in B_1 \cap B_2$. Then $\mathbf{v} \in B_1 \Longrightarrow \mathbf{v} \in E_{\lambda_1} \Longrightarrow L(\mathbf{v}) = \lambda_1\mathbf{v}$.
Similarly, $\mathbf{v} \in B_2 \Longrightarrow \mathbf{v} \in E_{\lambda_2} \Longrightarrow L(\mathbf{v}) = \lambda_2\mathbf{v}$.
Hence, $\lambda_1\mathbf{v} = \lambda_2\mathbf{v}$, and so $(\lambda_1 - \lambda_2)\mathbf{v} = \mathbf{0}$. Since $\lambda_1 \neq \lambda_2, \mathbf{v} = \mathbf{0}$.
But $\mathbf{0}$ can not be an element of any basis, and so $B_1 \cap B_2$ is empty.

(16) (a) E_{λ_i} is a subspace, hence closed.

(b) First, substitute $\mathbf{u}_i$ for $\sum_{j=1}^{k_i} a_{ij}\mathbf{v}_{ij}$ in the given double sum equation. This proves $\sum_{i=1}^{n} \mathbf{u}_i = \mathbf{0}$. Now, the set of all nonzero $\mathbf{u}_i$'s is linearly independent by Theorem 5.23, since they are eigenvectors corresponding to distinct eigenvalues. But then, the nonzero terms in $\sum_{i=1}^{n} \mathbf{u}_i$ would give a nontrivial linear combination from a linearly independent set equal to the zero vector. This contradiction shows that all of the $\mathbf{u}_i$'s must equal $\mathbf{0}$.

(c) Using part (b) and the definition of $\mathbf{u}_i$, $\mathbf{0} = \sum_{j=1}^{k_i} a_{ij}\mathbf{v}_{ij}$ for each i. But $\{\mathbf{v}_{i1}, \ldots, \mathbf{v}_{ik_i}\}$ is linearly independent, since it is a basis for E_{λ_i}. Hence, for each i, $a_{i1} = \cdots = a_{ik_i} = 0$.

(d) Apply the definition of linear independence to B.

(17) Example 7 states that $p_{\mathbf{A}}(x) = x^3 - 4x^2 + x + 6$. So,

$$\begin{aligned} p_{\mathbf{A}}(\mathbf{A}) &= \mathbf{A}^3 - 4\mathbf{A}^2 + \mathbf{A} + 6\mathbf{I}_3 \\ &= \begin{bmatrix} -17 & 22 & 76 \\ -374 & 214 & 1178 \\ 54 & -27 & -163 \end{bmatrix} - 4\begin{bmatrix} 5 & 2 & -4 \\ -106 & 62 & 334 \\ 18 & -9 & -53 \end{bmatrix} \\ &\quad + \begin{bmatrix} 31 & -14 & -92 \\ -50 & 28 & 158 \\ 18 & -9 & -55 \end{bmatrix} + 6\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

verifying the Cayley-Hamilton Theorem for this matrix.

(18) (a) F (b) T (c) T (d) F (e) T (f) T (g) T (h) F (i) F (j) T

Chapter 5 Review Exercises

(1) (a) Not a linear transformation: $f([0,0,0]) \neq [0,0,0]$

(b) Linear transformation:

$$\begin{aligned} & g((ax^3 + bx^2 + cx + d) + (kx^3 + lx^2 + mx + n)) \\ = \;& g((a+k)x^3 + (b+l)x^2 + (c+m)x + (d+n)) \\ = \;& \begin{bmatrix} 4(b+l) - (c+m) & 3(d+n) - (a+k) \\ 2(d+n) + 3(a+k) & 4(c+m) \\ 5(a+k) + (c+m) + 2(d+n) & 2(b+l) - 3(d+n) \end{bmatrix} \\ = \;& \begin{bmatrix} 4b - c & 3d - a \\ 2d + 3a & 4c \\ 5a + c + 2d & 2b - 3d \end{bmatrix} + \begin{bmatrix} 4l - m & 3n - k \\ 2n + 3k & 4m \\ 5k + m + 2n & 2l - 3n \end{bmatrix} \\ = \;& g(ax^3 + bx^2 + cx + d) + g(kx^3 + lx^2 + mx + n). \end{aligned}$$

Also,

$$\begin{aligned} g(k(ax^3 + bx^2 + cx + d)) &= g((ka)x^3 + (kb)x^2 + (kc)x + (kd)) \\ &= \begin{bmatrix} 4(kb) - (kc) & 3(kd) - (ka) \\ 2(kd) + 3(ka) & 4(kc) \\ 5(ka) + (kc) + 2(kd) & 2(kb) - 3(kd) \end{bmatrix} \\ &= k\begin{bmatrix} 4b - c & 3d - a \\ 2d + 3a & 4c \\ 5a + c + 2d & 2b - 3d \end{bmatrix} \\ &= k(g(ax^3 + bx^2 + cx + d)). \end{aligned}$$

(c) Not a linear transformation: $h([1,0] + [0,1]) \neq h([1,0]) + h([0,1])$ since $h([1,1]) = [2,-3]$ and $h([1,0]) + h([0,1]) = [0,0] + [0,0] = [0,0]$.

(2) $[1.598, 3.232]$

(3) $f(\mathbf{A}_1) + f(\mathbf{A}_2) = \mathbf{C}\mathbf{A}_1\mathbf{B}^{-1} + \mathbf{C}\mathbf{A}_2\mathbf{B}^{-1} = \mathbf{C}(\mathbf{A}_1\mathbf{B}^{-1} + \mathbf{A}_2\mathbf{B}^{-1}) = \mathbf{C}(\mathbf{A}_1 + \mathbf{A}_2)\mathbf{B}^{-1} = f(\mathbf{A}_1 + \mathbf{A}_2)$; $f(k\mathbf{A}) = \mathbf{C}(k\mathbf{A})\mathbf{B}^{-1} = k\mathbf{C}\mathbf{A}\mathbf{B}^{-1} = kf(\mathbf{A})$.

(4) $L([6, 2, -7]) = [20, 10, 44]$;

$$L([x, y, z]) = \begin{bmatrix} -3 & 5 & -4 \\ 2 & -1 & 0 \\ 4 & 3 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = [-3x + 5y - 4z,\ 2x - y,\ 4x + 3y - 2z]$$

(5) (a) Use Theorem 5.2 and part (1) of Theorem 5.3.

(b) Use Theorem 5.2 and part (2) of Theorem 5.3.

(6) (a) $\mathbf{A}_{BC} = \begin{bmatrix} 29 & 32 & -2 \\ 43 & 42 & -6 \end{bmatrix}$ (b) $\mathbf{A}_{BC} = \begin{bmatrix} 113 & -58 & 51 & -58 \\ 566 & -291 & 255 & -283 \\ -1648 & 847 & -742 & 823 \end{bmatrix}$

(7) (a) $\mathbf{A}_{BC} = \begin{bmatrix} 2 & 1 & -3 & 0 \\ 1 & 3 & 0 & -4 \\ 0 & 0 & 1 & -2 \end{bmatrix}$; $\mathbf{A}_{DE} = \begin{bmatrix} -151 & 99 & 163 & 16 \\ 171 & -113 & -186 & -17 \\ 238 & -158 & -260 & -25 \end{bmatrix}$

(b) $\mathbf{A}_{BC} = \begin{bmatrix} 6 & -1 & -1 \\ 0 & 3 & 2 \\ 2 & 0 & -4 \\ 1 & -5 & 1 \end{bmatrix}$; $\mathbf{A}_{DE} = \begin{bmatrix} 115 & -45 & 59 \\ 374 & -146 & 190 \\ -46 & 15 & -25 \\ -271 & 108 & -137 \end{bmatrix}$

(8) $\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(9) (a) $p_{\mathbf{A}_{BB}}(x) = x^3 - x^2 - x + 1 = (x + 1)(x - 1)^2$

(b) Fundamental eigenvectors for $\lambda_1 = 1$: $\{[2, 1, 0],\ [2, 0, 3]\}$; for $\lambda_2 = -1$: $\{[3, -6, 2]\}$

(c) $C = ([2, 1, 0],\ [2, 0, 3],\ [3, -6, 2])$

(d) $\mathbf{A}_{CC} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. (Note: $\mathbf{P} = \begin{bmatrix} 2 & 2 & 3 \\ 1 & 0 & -6 \\ 0 & 3 & 2 \end{bmatrix}$ and $\mathbf{P}^{-1} = \frac{1}{41}\begin{bmatrix} 18 & 5 & -12 \\ -2 & 4 & 15 \\ 3 & -6 & -2 \end{bmatrix}$)

(e) This is a reflection through the plane spanned by $\{[2, 1, 0],\ [2, 0, 3]\}$. The equation for this plane is $3x - 6y + 2z = 0$.

(10) (a) Basis for $\ker(L) = \{[2, -3, 1, 0], [-3, 4, 0, 1]\}$; basis for $\text{range}(L) = \{[3, 2, 2, 1], [1, 1, 3, 4]\}$

(b) $2 + 2 = 4$

(c) $[-18, 26, -4, 2] \notin \ker(L)$ because $L([-18, 26, -4, 2]) = [-6, -2, 10, 20] \neq \mathbf{0}$;
$[-18, 26, -6, 2] \in \ker(L)$ because $L([-18, 26, -6, 2]) = \mathbf{0}$

(d) $[8, 3, -11, -23] \in \text{range}(L)$; row reduction shows $[8, 3, -11, -23] = 5[3, 2, 2, 1] - 7[1, 1, 3, 4]$, and so, $L([5, -7, 0, 0]) = [8, 3, -11, -23]$

(11) Matrix for L with respect to standard bases: $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;

basis for $\ker(L) = \left\{ \begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix} \right\}$;
basis for $\text{range}(L) = \{x^3, x^2, x\}$;
$\dim(\ker(L)) + \dim(\text{range}(L)) = 3 + 3 = 6 = \dim(\mathcal{M}_{32})$

(12) (a) If $\mathbf{v} \in \ker(L_1)$, then $L_1(\mathbf{v}) = \mathbf{0}_{\mathcal{W}}$. Thus, $(L_2 \circ L_1)(\mathbf{v}) = L_2(L_1(\mathbf{v})) = L_2(\mathbf{0}_{\mathcal{W}}) = \mathbf{0}_{\mathcal{X}}$, and so $\mathbf{v} \in \ker(L_2 \circ L_1)$. Therefore, $\ker(L_1) \subseteq \ker(L_2 \circ L_1)$. Hence, $\dim(\ker(L_1)) \leq \dim(\ker(L_2 \circ L_1))$.

(b) Let L_1 be the projection onto the x-axis ($L_1([x, y]) = [x, 0]$) and L_2 be the projection onto the y-axis ($L_2([x, y]) = [0, y]$). A basis for $\ker(L_1)$ is $\{[0, 1]\}$, while $\ker(L_2 \circ L_1) = \mathbb{R}^2$.

(13) (a) By part (2) of Theorem 5.9 applied to L and M,

$$\begin{aligned} \dim(\ker(L)) - \dim(\ker(M)) &= (n - \text{rank}(\mathbf{A})) - (m - \text{rank}(\mathbf{A}^T)) \\ &= (n - m) - (\text{rank}(\mathbf{A}) - \text{rank}(\mathbf{A}^T)) = n - m, \end{aligned}$$

by Corollary 5.11.

(b) Since L is onto, $\dim(\text{range}(L)) = m = \text{rank}(\mathbf{A})$ (by part (1) of Theorem 5.9) $= \text{rank}(\mathbf{A}^T)$ (by Corollary 5.11) $= \dim(\text{range}(M))$. Thus, by part (3) of Theorem 5.9, $\dim(\ker(M)) + \dim(\text{range}(M)) = m$, implying $\dim(\ker(M)) + m = m$, and so $\dim(\ker(M)) = 0$, and M is one-to-one (by part (1) of Theorem 5.12).

(c) Converse: M one-to-one $\Longrightarrow$ L onto. This is true.
M one-to-one $\Longrightarrow \dim(\ker(M)) = 0$
$\Longrightarrow \dim(\ker(L)) = n - m$ (by part (a))
$\Longrightarrow (n - m) + \dim(\text{range}(L)) = n$ (by part (3) of Theorem 5.9)
$\Longrightarrow \dim(\text{range}(L)) = m$
$\Longrightarrow$ L is onto (by Theorem 4.13, and part (2) of Theorem 5.12).

(14) (a) L is not one-to-one because $L(x^3 - x + 1) = \mathbf{O}_{22}$. Corollary 5.13 then shows that L is not onto.

(b) $\ker(L)$ has $\{x^3 - x + 1\}$ as a basis, so $\dim(\ker(L)) = 1$. Thus, $\dim(\text{range}(L)) = 3$ by the Dimension Theorem.

(15) In each part, let $\mathbf{A}$ represent the given matrix for L with respect to the standard bases.

(a) The reduced row echelon form of $\mathbf{A}$ is $\mathbf{I}_3$. Therefore, $\ker(L) = \{\mathbf{0}\}$, and so $\dim(\ker(L)) = 0$ and L is one-to-one. By Corollary 5.13, L is also onto. Hence $\dim(\text{range}(L)) = 3$.

(b) The reduced row echelon form of $\mathbf{A}$ is $\begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

$\dim(\ker(L)) = 4 - \text{rank}(\mathbf{A}) = 1$. L is not one-to-one.
$\dim(\text{range}(L)) = \text{rank}(\mathbf{A}) = 3$. L is onto.

(16) (a) $\dim(\ker(L)) = \dim(\mathcal{P}_3) - \dim(\text{range}(L)) \geq 4 - \dim(\mathbb{R}^3) = 1$, so L cannot be one-to-one.

(b) $\dim(\text{range}(L)) = \dim(\mathcal{P}_2) - \dim(\ker(L)) \leq \dim(\mathcal{P}_2) = 3 < \dim(\mathcal{M}_{22})$, so L cannot be onto.

(17) (a) $L(\mathbf{v}_1) = cL(\mathbf{v}_2) = L(c\mathbf{v}_2) \implies \mathbf{v}_1 = c\mathbf{v}_2$, because L is one-to-one. In this case, the set $\{L(\mathbf{v}_1), L(\mathbf{v}_2)\}$ is linearly dependent, since $L(\mathbf{v}_1)$ is a nonzero multiple of $L(\mathbf{v}_2)$. Part (1) of Theorem 5.14 then implies that $\{\mathbf{v}_1, \mathbf{v}_2\}$ must be linearly dependent, which is correct, since $\mathbf{v}_1$ is a nonzero multiple of $\mathbf{v}_2$.

(b) Consider the contrapositive: For L onto and $\mathbf{w} \in \mathcal{W}$, if $\{\mathbf{v}_1, \mathbf{v}_2\}$ spans $\mathcal{V}$, then $L(a\mathbf{v}_1 + b\mathbf{v}_2) = \mathbf{w}$ for some $a, b \in \mathbb{R}$.

Proof of the contrapositive: Suppose L is onto, $\mathbf{w} \in \mathcal{W}$, and $\{\mathbf{v}_1, \mathbf{v}_2\}$ spans $\mathcal{V}$. By part (2) of Theorem 5.14, $\{L(\mathbf{v}_1), L(\mathbf{v}_2)\}$ spans $\mathcal{W}$, so there exist $a, b \in \mathbb{R}$ such that

$$\mathbf{w} = aL(\mathbf{v}_1) + bL(\mathbf{v}_2) = L(a\mathbf{v}_1 + b\mathbf{v}_2).$$

(18) (a) L_1 and L_2 are isomorphisms (by Theorem 5.16) because their (given) matrices are nonsingular. Their inverses are given in part (b).

(b) Matrix for $L_2 \circ L_1 = \begin{bmatrix} 81 & 71 & -15 & 18 \\ 107 & 77 & -31 & 19 \\ 69 & 45 & -23 & 11 \\ -29 & -36 & -1 & -9 \end{bmatrix}$;

matrix for L_1^{-1}: $\frac{1}{5}\begin{bmatrix} 2 & -10 & 19 & 11 \\ 0 & 5 & -10 & -5 \\ 3 & -15 & 26 & 14 \\ -4 & 15 & -23 & -17 \end{bmatrix}$;

matrix for L_2^{-1}: $\frac{1}{2}\begin{bmatrix} -8 & 26 & -30 & 2 \\ 10 & -35 & 41 & -4 \\ 10 & -30 & 34 & -2 \\ -14 & 49 & -57 & 6 \end{bmatrix}$.

(c) Both computations yield $\frac{1}{10}\begin{bmatrix} -80 & 371 & -451 & 72 \\ 20 & -120 & 150 & -30 \\ -110 & 509 & -619 & 98 \\ 190 & -772 & 922 & -124 \end{bmatrix}$.

(19) (a) The determinant of the shear matrix given in Table 5.1 is 1, so this matrix is nonsingular. Therefore, the given mapping is an isomorphism by Theorem 5.16.

(b) The inverse isomorphism is $L\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & -k \\ 0 & 1 & -k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 - ka_3 \\ a_2 - ka_3 \\ a_3 \end{bmatrix}$. This represents a shear in the z-direction with factor $-k$.

(20) (a) $(\mathbf{B}^T\mathbf{A}\mathbf{B})^T = \mathbf{B}^T\mathbf{A}^T(\mathbf{B}^T)^T = \mathbf{B}^T\mathbf{A}\mathbf{B}$, since $\mathbf{A}$ is symmetric.

(b) Suppose $\mathbf{A}, \mathbf{C} \in \mathcal{W}$. Note that if $\mathbf{B}^T\mathbf{A}\mathbf{B} = \mathbf{B}^T\mathbf{C}\mathbf{B}$, then $\mathbf{A} = \mathbf{C}$ because $\mathbf{B}$ is nonsingular. Thus L is one-to-one, and so L is onto by Corollary 5.13.

An alternate approach is as follows: Suppose $\mathbf{C} \in \mathcal{W}$. If $\mathbf{A} = (\mathbf{B}^T)^{-1}\mathbf{C}\mathbf{B}^{-1}$, then $L(\mathbf{A}) = \mathbf{C}$. Thus L is onto, and so L is one-to-one by Corollary 5.13.

(21) (a) $L(ax^4 + bx^3 + cx^2) = 4ax^3 + (12a + 3b)x^2 + (6b + 2c)x + 2c$. Clearly $\ker(L) = \{\mathbf{0}\}$. Apply part (1) of Theorem 5.12.

(b) No. $\dim(\mathcal{W}) = 3 \neq \dim(\mathcal{P}_3)$.

(c) The polynomial $x \notin \text{range}(L)$ since it does not have the correct form for an image under L (as stated in part (a)).

(22) In parts (a), (b), and (c), let $\mathbf{A}$ represent the given matrix.

(a) (i) $p_{\mathbf{A}}(x) = x^3 - 3x^2 - x + 3 = (x-1)(x+1)(x-3)$;
eigenvalues for L: $\lambda_1 = 1$, $\lambda_2 = -1$, and $\lambda_3 = 3$;
basis for E_1: $\{[-1, 3, 4]\}$; basis for E_{-1}: $\{[-1, 4, 5]\}$; basis for E_3: $\{[-6, 20, 27]\}$
(ii) All algebraic and geometric multiplicities equal 1; L is diagonalizable.
(iii) $B = \{[-1, 3, 4], [-1, 4, 5], [-6, 20, 27]\}$;

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} -1 & -1 & -6 \\ 3 & 4 & 20 \\ 4 & 5 & 27 \end{bmatrix}. \text{ (Note that } \mathbf{P}^{-1} = \begin{bmatrix} -8 & 3 & -4 \\ 1 & 3 & -2 \\ 1 & -1 & 1 \end{bmatrix}\text{).}$$

(b) (i) $p_{\mathbf{A}}(x) = x^3 + 5x^2 + 20x + 28 = (x+2)(x^2 + 3x + 14)$;
eigenvalue for L: $\lambda = -2$ since $x^2 + 3x + 14$ has no real roots; basis for E_{-2}: $\{[0, 1, 1]\}$
(ii) For $\lambda = -2$: algebraic multiplicity = geometric multiplicity = 1; L is not diagonalizable.

(c) (i) $p_{\mathbf{A}}(x) = x^3 - 5x^2 + 3x + 9 = (x+1)(x-3)^2$;
eigenvalues for L: $\lambda_1 = -1$, and $\lambda_2 = 3$;
basis for E_{-1}: $\{[1, 3, 3]\}$; basis for E_3: $\{[1, 5, 0], [3, 0, 25]\}$
(ii) For $\lambda_1 = -1$: algebraic multiplicity = geometric multiplicity = 1;
For $\lambda_2 = 3$: algebraic multiplicity = geometric multiplicity = 2; L is diagonalizable.
(iii) $B = \{[1, 3, 3], [1, 5, 0], [3, 0, 25]\}$;

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} 1 & 1 & 3 \\ 3 & 5 & 0 \\ 3 & 0 & 25 \end{bmatrix}. \text{ (Note that } \mathbf{P}^{-1} = \tfrac{1}{5}\begin{bmatrix} 125 & -25 & -15 \\ -75 & 16 & 9 \\ -15 & 3 & 2 \end{bmatrix}\text{).}$$

(d) Matrix for L with respect to standard coordinates: $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 0 & 0 & 0 \\ 0 & -2 & -1 & 0 \\ 0 & 0 & -1 & -2 \end{bmatrix}$.

(i) $p_{\mathbf{A}}(x) = x^4 + 2x^3 - x^2 - 2x = x(x-1)(x+1)(x+2)$;
eigenvalues for L: $\lambda_1 = 1$, $\lambda_2 = 0$, $\lambda_3 = -1$, and $\lambda_4 = -2$;
basis for E_1: $\{-x^3 + 3x^2 - 3x + 1\}$ ($\{[-1, 3, -3, 1]\}$ in standard coordinates);
basis for E_0: $\{x^2 - 2x + 1\}$ ($\{[0, 1, -2, 1]\}$ in standard coordinates);
basis for E_{-1}: $\{-x + 1\}$ ($\{[0, 0, -1, 1]\}$ in standard coordinates);
basis for E_{-2}: $\{1\}$ ($\{[0, 0, 0, 1]\}$ in standard coordinates)
(ii) All algebraic and geometric multiplicities equal 1; L is diagonalizable.
(iii) $B = \{-x^3 + 3x^2 - 3x + 1, x^2 - 2x + 1, -x + 1, 1\}$
($\{[-1, 3, -3, 1], [0, 1, -2, 1], [0, 0, -1, 1], [0, 0, 0, 1]\}$ in standard coordinates);

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}; \mathbf{P} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -3 & -2 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \text{ (Note that } \mathbf{P}^{-1} = \mathbf{P}\text{).}$$

(23) Basis for E_1: $\{[a, b, c], [d, e, f]\}$; basis for E_{-1}: $\{[bf - ce, cd - af, ae - bd]\}$.
basis of eigenvectors: $\{[a, b, c], [d, e, f], [bf - ce, cd - af, ae - bd]\}$;

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(24) $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{A}^4 - 4\mathbf{A}^3 - 18\mathbf{A}^2 + 108\mathbf{A} - 135\mathbf{I}_4$

$$= \begin{bmatrix} 649 & 216 & -176 & -68 \\ -568 & -135 & 176 & 68 \\ 1136 & 432 & -271 & -136 \\ -1088 & 0 & 544 & 625 \end{bmatrix} - 4\begin{bmatrix} 97 & 54 & -8 & 19 \\ -70 & -27 & 8 & -19 \\ 140 & 108 & 11 & 38 \\ 304 & 0 & -152 & -125 \end{bmatrix}$$

$$- 18\begin{bmatrix} 37 & 12 & -8 & -2 \\ -28 & -3 & 8 & 2 \\ 56 & 24 & -7 & -4 \\ -32 & 0 & 16 & 25 \end{bmatrix} + 108\begin{bmatrix} 5 & 2 & 0 & 1 \\ -2 & 1 & 0 & -1 \\ 4 & 4 & 3 & 2 \\ 16 & 0 & -8 & -5 \end{bmatrix} - \begin{bmatrix} 135 & 0 & 0 & 0 \\ 0 & 135 & 0 & 0 \\ 0 & 0 & 135 & 0 \\ 0 & 0 & 0 & 135 \end{bmatrix}$$

$$= \mathbf{O}_{44}$$

(25) (a) T (b) F (c) T (d) T (e) T (f) F (g) F (h) T (i) T (j) F (k) F (l) T (m) F (n) T (o) T (p) F (q) F (r) T (s) F (t) T (u) T (v) F (w) T (x) T (y) T (z) T

Chapter 6

Section 6.1

(1) Orthogonal, not orthonormal: (a), (f);
Orthogonal and orthonormal: (b), (d), (e), (g);
Neither: (c)

(2) Orthogonal: (a), (d), (e);
Not orthogonal, columns not normalized: (b), (c)

(3) (a) $[\mathbf{v}]_B = \left[\frac{2\sqrt{3}+3}{2}, \frac{3\sqrt{3}-2}{2}\right]$ (b) $[\mathbf{v}]_B = [2, -1, 4]$ (c) $[\mathbf{v}]_B = [3, \frac{13\sqrt{3}}{3}, \frac{5\sqrt{6}}{3}, 4\sqrt{2}]$

(4) (a) $\{[5, -1, 2], [5, -3, -14]\}$
(b) $\{[2, -1, 3, 1], [-7, -1, 0, 13]\}$
(c) $\{[2, 1, 0, -1], [-1, 1, 3, -1], [5, -7, 5, 3]\}$
(d) $\{[0, 1, 3, -2], [2, 3, -1, 0], [-3, 2, 0, 1]\}$
(e) $\{[4, -1, -2, 2], [2, 0, 3, -1], [3, 8, -4, -6]\}$

(5) (a) $\{[2, 2, -3], [13, -4, 6], [0, 3, 2]\}$
(b) $\{[1, -4, 3], [25, 4, -3], [0, 3, 4]\}$
(c) $\{[1, -3, 1], [2, 5, 13], [4, 1, -1]\}$
(d) $\{[3, 1, -2], [5, -3, 6], [0, 2, 1]\}$
(e) $\{[2, 1, -2, 1], [3, -1, 2, -1], [0, 5, 2, -1], [0, 0, 1, 2]\}$
(f) $\{[2, 1, 0, -3], [0, 3, 2, 1], [5, -1, 0, 3], [0, 3, -5, 1]\}$

(6) Orthogonal basis for $\mathcal{W} = \{[-2, -1, 4, -2, 1], [4, -3, 0, -2, 1], [-1, -33, 15, 38, -19], [3, 3, 3, -2, -7]\}$

(7) (a) $[-1, 3, 3]$ (b) $[3, 3, 1]$ (c) $[5, 1, 1]$ (d) $[4, 3, 2]$

(8) (a) $(c_i\mathbf{v}_i) \cdot (c_j\mathbf{v}_j) = c_ic_j(\mathbf{v}_i \cdot \mathbf{v}_j) = c_ic_j0 = 0$, for $i \neq j$.
(b) No

(9) (a) Express $\mathbf{v}$ and $\mathbf{w}$ as linear combinations of $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ using Theorem 6.3. Then expand and simplify $\mathbf{v} \cdot \mathbf{w}$.
(b) Let $\mathbf{w} = \mathbf{v}$ in part (a).

(10) Follow the hint. Then use Exercise 9(b). Finally, drop some terms to get the inequality.

(11) (a) $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^{-1}$, since $\mathbf{A}$ is orthogonal.
(b) $\mathbf{AB}(\mathbf{AB})^T = \mathbf{AB}(\mathbf{B}^T\mathbf{A}^T) = \mathbf{A}(\mathbf{BB}^T)\mathbf{A}^T = \mathbf{AIA}^T = \mathbf{AA}^T = \mathbf{I}$. Hence $(\mathbf{AB})^T = (\mathbf{AB})^{-1}$.

(12) $\mathbf{A} = \mathbf{A}^T \implies \mathbf{A}^2 = \mathbf{AA}^T \implies \mathbf{I}_n = \mathbf{AA}^T \implies \mathbf{A}^T = \mathbf{A}^{-1}$.
Conversely, $\mathbf{A}^T = \mathbf{A}^{-1} \implies \mathbf{I}_n = \mathbf{AA}^T \implies \mathbf{A} = \mathbf{A}^2\mathbf{A}^T \implies \mathbf{A} = \mathbf{I}_n\mathbf{A}^T \implies \mathbf{A} = \mathbf{A}^T$.

(13) Suppose $\mathbf{A}$ is both orthogonal and skew-symmetric. Then $\mathbf{A}^2 = \mathbf{A}(-\mathbf{A}^T) = -\mathbf{AA}^T = -\mathbf{I}_n$. Hence $|\mathbf{A}|^2 = |-\mathbf{I}_n| = (-1)^n|\mathbf{I}_n|$ (by Corollary 3.4) $= -1$, if n is odd, a contradiction.

(14) $\mathbf{A}(\mathbf{A} + \mathbf{I}_n)^T = \mathbf{A}(\mathbf{A}^T + \mathbf{I}_n) = \mathbf{AA}^T + \mathbf{A} = \mathbf{I}_n + \mathbf{A}$ (since $\mathbf{A}$ is orthogonal). Hence $|\mathbf{I}_n + \mathbf{A}| = |\mathbf{A}|\,|(\mathbf{A} + \mathbf{I}_n)^T| = (-1)|\mathbf{A} + \mathbf{I}_n|$, implying $|\mathbf{A} + \mathbf{I}_n| = 0$. Thus $\mathbf{A} + \mathbf{I}_n$ is singular.

(15) Use part (2) of Theorem 6.7. To be a unit vector, the first column must be $[\pm 1, 0, 0]$. To be a unit vector orthogonal to the first column, the second column must be $[0, \pm 1, 0]$, etc.

(16) (a) By Theorem 6.5, we can expand the orthonormal set $\{\mathbf{u}\}$ to an orthonormal basis for $\mathbb{R}^n$. Form the matrix using these basis vectors as rows. Then use part (1) of Theorem 6.7.

(b) $\begin{bmatrix} \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{30}}{6} & -\frac{\sqrt{30}}{15} & -\frac{\sqrt{30}}{30} \\ 0 & \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \end{bmatrix}$ is one possible answer.

(17) Proof of the other half of part (1) of Theorem 6.7: (i, j) entry of $\mathbf{A}\mathbf{A}^T = (i\text{th row of } \mathbf{A}) \cdot (j\text{th row of } \mathbf{A})$ $= \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ (since the rows of $\mathbf{A}$ form an orthonormal set). Hence $\mathbf{A}\mathbf{A}^T = \mathbf{I}_n$, and $\mathbf{A}$ is orthogonal.

Proof of part (2): $\mathbf{A}$ is orthogonal iff $\mathbf{A}^T$ is orthogonal (by part (2) of Theorem 6.6) iff the rows of $\mathbf{A}^T$ form an orthonormal basis for $\mathbb{R}^n$ (by part (1) of Theorem 6.7) iff the columns of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$.

(18) (a) $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} = \mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{v}$ (by Theorem 6.9) $= \|\mathbf{A}\mathbf{v}\|^2$. Then take square roots.

(b) Using part (a) and Theorem 6.9, we have

$$\frac{\mathbf{v} \cdot \mathbf{w}}{(\|\mathbf{v}\|)(\|\mathbf{w}\|)} = \frac{(\mathbf{A}\mathbf{v}) \cdot (\mathbf{A}\mathbf{w})}{(\|\mathbf{A}\mathbf{v}\|)(\|\mathbf{A}\mathbf{w}\|)},$$

and so the cosines of the appropriate angles are equal.

(19) A detailed proof of this can be found in the beginning of the proof of Theorem 6.19 in Appendix A. (That portion of the proof uses no results beyond Section 6.1.)

(20) (a) Let $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be the matrices whose columns are the vectors in B and C, respectively. Then $\mathbf{Q}_1$ is orthogonal by part (2) of Theorem 6.7. Also, $\mathbf{Q}_1$ (respectively, $\mathbf{Q}_2$) is the transition matrix from B (respectively, C) to standard coordinates. By Theorem 4.19, $\mathbf{Q}_2^{-1}$ is the transition matrix from standard coordinates to C. Theorem 4.18 then implies $\mathbf{P} = \mathbf{Q}_2^{-1}\mathbf{Q}_1$. Hence, $\mathbf{Q}_2 = \mathbf{Q}_1\mathbf{P}^{-1}$. By parts (2) and (3) of Theorem 6.6, $\mathbf{Q}_2$ is orthogonal, since both $\mathbf{P}$ and $\mathbf{Q}_1$ are orthogonal. Thus C is an orthonormal basis by part (2) of Theorem 6.7.

(b) Let $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$, where $k = \dim(\mathcal{V})$. Then the ith column of the transition matrix from B to C is $[\mathbf{b}_i]_C$. Now, $[\mathbf{b}_i]_C \cdot [\mathbf{b}_j]_C = \mathbf{b}_i \cdot \mathbf{b}_j$ (by Exercise 19, since C is an orthonormal basis). This equals 0 if $i \neq j$ and equals 1 if $i = j$, since B is orthonormal. Hence, the columns of the transition matrix form an orthonormal set of k vectors in $\mathbb{R}^k$, and so this matrix is orthogonal by part (2) of Theorem 6.7.

(c) Let $C = (\mathbf{c}_1, \ldots, \mathbf{c}_k)$, where $k = \dim(\mathcal{V})$. Then $\mathbf{c}_i \cdot \mathbf{c}_j = [\mathbf{c}_i]_B \cdot [\mathbf{c}_j]_B$ (by Exercise 19 since B is orthonormal) $= (\mathbf{P}[\mathbf{c}_i]_B) \cdot (\mathbf{P}[\mathbf{c}_j]_B)$ (by Theorem 6.9 since $\mathbf{P}$ is orthogonal) $= [\mathbf{c}_i]_C \cdot [\mathbf{c}_j]_C$ (since $\mathbf{P}$ is the transition matrix from B to C) $= \mathbf{e}_i \cdot \mathbf{e}_j$, which equals 0 if $i \neq j$ and equals 1 if $i = j$. Hence C is orthonormal.

(21) First note that $\mathbf{A}^T\mathbf{A}$ is an $n \times n$ matrix. If $i \neq j$, the (i, j) entry of $\mathbf{A}^T\mathbf{A}$ equals (ith row of $\mathbf{A}^T$) $\cdot$ (jth column of $\mathbf{A}$) $=$ (ith column of $\mathbf{A}$) $\cdot$ (jth column of $\mathbf{A}$) $= 0$, because the columns of $\mathbf{A}$ are orthogonal to each other. Thus, $\mathbf{A}^T\mathbf{A}$ is a diagonal matrix. Finally, the (i, i) entry of $\mathbf{A}^T\mathbf{A}$ equals (ith row of $\mathbf{A}^T$) $\cdot$ (ith column of $\mathbf{A}$) $=$ (ith column of $\mathbf{A}$) $\cdot$ (ith column of $\mathbf{A}$) $= 1$, because the ith column of $\mathbf{A}$ is a unit vector.

(22) (a) F (b) T (c) T (d) F (e) T (f) T (g) T (h) F (i) T (j) T

Section 6.2

(1) (a) $\mathcal{W}^\perp = \text{span}(\{[2,3]\})$
(b) $\mathcal{W}^\perp = \text{span}(\{[5,2,-1],[0,1,2]\})$
(c) $\mathcal{W}^\perp = \text{span}(\{[2,3,7]\})$
(d) $\mathcal{W}^\perp = \text{span}(\{[3,-1,4]\})$
(e) $\mathcal{W}^\perp = \text{span}(\{[-2,5,-1]\})$
(f) $\mathcal{W}^\perp = \text{span}(\{[7,1,-2,-3],[0,4,-1,2]\})$
(g) $\mathcal{W}^\perp = \text{span}(\{[3,-2,4,1]\})$

(2) (a) $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v} = [-\frac{33}{35}, \frac{111}{35}, \frac{12}{7}]$; $\mathbf{w}_2 = [-\frac{2}{35}, -\frac{6}{35}, \frac{2}{7}]$

(b) $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v} = [-\frac{17}{9}, -\frac{10}{9}, \frac{14}{9}]$; $\mathbf{w}_2 = [\frac{26}{9}, -\frac{26}{9}, \frac{13}{9}]$

(c) $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v} = [\frac{1}{7}, -\frac{3}{7}, -\frac{2}{7}]$; $\mathbf{w}_2 = [\frac{13}{7}, \frac{17}{7}, -\frac{19}{7}]$

(d) $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v} = [-\frac{54}{35}, \frac{4}{35}, \frac{42}{35}, \frac{92}{35}]$; $\mathbf{w}_2 = [\frac{19}{35}, \frac{101}{35}, \frac{63}{35}, -\frac{22}{35}]$

(3) Use the orthonormal basis $\{\mathbf{i}, \mathbf{j}\}$ for $\mathcal{W}$.

(4) (a) $\frac{3\sqrt{129}}{43}$ (b) $\frac{\sqrt{3806}}{22}$ (c) $\frac{2\sqrt{247}}{13}$ (d) $\frac{8\sqrt{17}}{17}$

(5) (a) Orthogonal projection onto $3x + y + z = 0$
(b) Orthogonal reflection through $x - 2y - 2z = 0$
(c) Neither. $p_L(x) = (x-1)(x+1)^2$. $E_1 = \text{span}\{[-7,9,5]\}$, $E_{-1} = \text{span}\{[-7,2,0],\ [11,0,2]\}$. Eigenspaces have wrong dimensions.
(d) Neither. $p_L(x) = (x-1)^2(x+1)$. $E_1 = \text{span}\{[-1,4,0],\ [-7,0,4]\}$, $E_{-1} = \text{span}\{[2,1,3]\}$. Eigenspaces are not orthogonal.

(6) $\frac{1}{9}\begin{bmatrix} 5 & 2 & -4 \\ 2 & 8 & 2 \\ -4 & 2 & 5 \end{bmatrix}$

(7) $\frac{1}{7}\begin{bmatrix} -2 & 3 & -6 \\ 3 & 6 & 2 \\ -6 & 2 & 3 \end{bmatrix}$

(8) (a) $\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{5}{6} & -\frac{1}{6} \\ -\frac{1}{3} & -\frac{1}{6} & \frac{5}{6} \end{bmatrix}$
(b) This is straightforward.

(9) (a) Characteristic polynomial $= x^3 - 2x^2 + x$
(b) Characteristic polynomial $= x^3 - x^2$
(c) Characteristic polynomial $= x^3 - x^2 - x + 1$

(10) (a) $\frac{1}{59}\begin{bmatrix} 50 & -21 & -3 \\ -21 & 10 & -7 \\ -3 & -7 & 58 \end{bmatrix}$ (c) $\frac{1}{9}\begin{bmatrix} 2 & 2 & 3 & -1 \\ 2 & 8 & 0 & 2 \\ 3 & 0 & 6 & -3 \\ -1 & 2 & -3 & 2 \end{bmatrix}$

(b) $\frac{1}{17}\begin{bmatrix} 8 & 6 & -6 \\ 6 & 13 & 4 \\ -6 & 4 & 13 \end{bmatrix}$ (d) $\frac{1}{15}\begin{bmatrix} 9 & 3 & -3 & -6 \\ 3 & 1 & -1 & -2 \\ -3 & -1 & 1 & 2 \\ -6 & -2 & 2 & 4 \end{bmatrix}$

(11) $\mathcal{W}_1^\perp = \mathcal{W}_2^\perp \implies (\mathcal{W}_1^\perp)^\perp = (\mathcal{W}_2^\perp)^\perp \implies \mathcal{W}_1 = \mathcal{W}_2$ (by Corollary 6.14).

(12) Let $\mathbf{v} \in \mathcal{W}_2^\perp$ and let $\mathbf{w} \in \mathcal{W}_1$. Then $\mathbf{w} \in \mathcal{W}_2$, and so $\mathbf{v} \cdot \mathbf{w} = 0$, which implies $\mathbf{v} \in \mathcal{W}_1^\perp$.

(13) Both parts rely on the uniqueness assertion in Corollary 6.16, and the equation $\mathbf{v} = \mathbf{v} + \mathbf{0}$.

(14) $\mathbf{v} \in \mathcal{W}^\perp \implies \mathbf{proj}_{\mathcal{W}}\mathbf{v} = \mathbf{0}$ (see Exercise 13) $\implies$ minimum distance between P and $\mathcal{W} = \|\mathbf{v}\|$ (by Theorem 6.18).
Conversely, suppose $\|\mathbf{v}\|$ is the minimum distance from P to $\mathcal{W}$. Let $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v}$ and $\mathbf{w}_2 = \mathbf{v} - \mathbf{proj}_{\mathcal{W}}\mathbf{v}$. Then $\|\mathbf{v}\| = \|\mathbf{w}_2\|$, by Theorem 6.18. Hence

$$\begin{aligned} \|\mathbf{v}\|^2 &= \|\mathbf{w}_1 + \mathbf{w}_2\|^2 = (\mathbf{w}_1 + \mathbf{w}_2) \cdot (\mathbf{w}_1 + \mathbf{w}_2) \\ &= \mathbf{w}_1 \cdot \mathbf{w}_1 + 2\mathbf{w}_1 \cdot \mathbf{w}_2 + \mathbf{w}_2 \cdot \mathbf{w}_2 \\ &= \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \qquad \text{(since } \mathbf{w}_1 \perp \mathbf{w}_2\text{)}. \end{aligned}$$

Subtracting $\|\mathbf{v}\|^2 = \|\mathbf{w}_2\|^2$ from both sides yields $0 = \|\mathbf{w}_1\|^2$, implying $\mathbf{w}_1 = \mathbf{0}$. Hence $\mathbf{v} = \mathbf{w}_2 \in \mathcal{W}^\perp$.

(15) Let $\mathbf{A} \in \mathcal{V}$, $\mathbf{B} \in \mathcal{W}$. Then

$$\begin{aligned} \text{“}\mathbf{A}\cdot\mathbf{B}\text{”} &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij} \\ &= \sum_{i=2}^{n}\sum_{j=1}^{i-1} a_{ij}b_{ij} + \sum_{i=1}^{n} a_{ii}b_{ii} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_{ij}b_{ij} \\ &= \sum_{i=2}^{n}\sum_{j=1}^{i-1} a_{ij}b_{ij} + 0 + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_{ji}(-b_{ji}) \qquad (\text{since } b_{ii} = 0) \\ &= \sum_{i=2}^{n}\sum_{j=1}^{i-1} a_{ij}b_{ij} - \sum_{j=2}^{n}\sum_{i=1}^{j-1} a_{ji}b_{ji} = 0. \end{aligned}$$

Now use Exercise 14(b) in Section 4.6 and follow the hint in the text.

(16) Let $\mathbf{u} = \frac{\mathbf{a}}{\|\mathbf{a}\|}$. Then $\{\mathbf{u}\}$ is an orthonormal basis for $\mathcal{W}$. Hence $\mathbf{proj}_{\mathcal{W}}\mathbf{b} = (\mathbf{b}\cdot\mathbf{u})\mathbf{u} = (\mathbf{b}\cdot\frac{\mathbf{a}}{\|\mathbf{a}\|})\frac{\mathbf{a}}{\|\mathbf{a}\|} = (\frac{\mathbf{b}\cdot\mathbf{a}}{\|\mathbf{a}\|^2})\mathbf{a} = \mathbf{proj}_{\mathbf{a}}\mathbf{b}$, according to Section 1.2.

(17) Let S be a spanning set for $\mathcal{W}$. Clearly, if $\mathbf{v} \in \mathcal{W}^\perp$ and $\mathbf{u} \in S$ then $\mathbf{v}\cdot\mathbf{u} = 0$.
Conversely, suppose $\mathbf{v}\cdot\mathbf{u} = 0$ for all $\mathbf{u} \in S$. Let $\mathbf{w} \in \mathcal{W}$. Then $\mathbf{w} = a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n$ for some $\mathbf{u}_1, \ldots, \mathbf{u}_n \in S$. Hence, $\mathbf{v}\cdot\mathbf{w} = a_1(\mathbf{v}\cdot\mathbf{u}_1) + \cdots + a_n(\mathbf{v}\cdot\mathbf{u}_n) = 0 + \cdots + 0 = 0$. Thus $\mathbf{v} \in \mathcal{W}^\perp$.

(18) First, show $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$. Let $\mathbf{w} \in \mathcal{W}$. We need to show that $\mathbf{w} \in (\mathcal{W}^\perp)^\perp$. That is, we must prove that $\mathbf{w} \perp \mathbf{v}$ for all $\mathbf{v} \in \mathcal{W}^\perp$. But, by the definition of $\mathcal{W}^\perp$, $\mathbf{w} \cdot \mathbf{v} = 0$, since $\mathbf{w} \in \mathcal{W}$, which completes this half of the proof.
Next, prove $\mathcal{W} = (\mathcal{W}^\perp)^\perp$. By Corollary 6.13, $\dim(\mathcal{W}) = n - \dim(\mathcal{W}^\perp) = n - (n - \dim((\mathcal{W}^\perp)^\perp)) = \dim((\mathcal{W}^\perp)^\perp)$. Thus, by Theorem 4.13, $\mathcal{W} = (\mathcal{W}^\perp)^\perp$.

(19) The proof that L is a linear operator is straightforward.
Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthonormal basis for $\mathcal{W}$.
First, we prove that range$(L) = \mathcal{W}$. Note that if $\mathbf{w} \in \mathcal{W}$, then $\mathbf{w} = \mathbf{w} + \mathbf{0}$, where $\mathbf{w} \in \mathcal{W}$ and $\mathbf{0} \in \mathcal{W}^\perp$. Hence, by the uniqueness statement in Corollary 6.16, $L(\mathbf{w}) = \mathbf{proj}_{\mathcal{W}}\mathbf{w} = \mathbf{w}$. Therefore, every $\mathbf{w}$ in $\mathcal{W}$ is in the range of L. Similarly, Corollary 6.16 implies that $\mathbf{proj}_{\mathcal{W}}\mathbf{v} \in \mathcal{W}$ for every $\mathbf{v} \in \mathbb{R}^n$. Hence, range$(L) = \mathcal{W}$.
To show $\mathcal{W}^\perp = \ker(L)$, we first prove that $\mathcal{W}^\perp \subseteq \ker(L)$. If $\mathbf{v} \in \mathcal{W}^\perp$, then since $\mathbf{v} \cdot \mathbf{w} = 0$ for every $\mathbf{w} \in \mathcal{W}$, $L(\mathbf{v}) = \mathbf{proj}_{\mathcal{W}}(c\mathbf{v}) = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k = 0\mathbf{u}_1 + \cdots + 0\mathbf{u}_k = \mathbf{0}$. Hence, $\mathbf{v} \in \ker(L)$.
Finally, by the Dimension Theorem, we have

$$\dim(\ker(L)) = n - \dim(\text{range}(L)) = n - \dim(\mathcal{W}) = \dim(\mathcal{W}^\perp).$$

Since $\mathcal{W}^\perp \subseteq \ker(L)$, Theorem 4.13 implies that $\mathcal{W}^\perp = \ker(L)$.

(20) Since for each $\mathbf{v} \in \ker(L)$ we have $\mathbf{A}\mathbf{v} = \mathbf{0}$, each row of $\mathbf{A}$ is orthogonal to $\mathbf{v}$. Hence, $\ker(L) \subseteq (\text{row space of } \mathbf{A})^\perp$. Also, $\dim(\ker(L)) = n - \text{rank}(\mathbf{A})$ (by part (2) of Theorem 5.9) $= n - \dim(\text{row space of } \mathbf{A}) = \dim((\text{row space of } \mathbf{A})^\perp)$ (by Corollary 6.13). Then apply Theorem 4.13.

(21) Suppose $\mathbf{v} \in (\ker(L))^\perp$ and $T(\mathbf{v}) = \mathbf{0}$. Then $L(\mathbf{v}) = \mathbf{0}$, so $\mathbf{v} \in \ker(L)$. Apply Theorem 6.11 to show $\ker(T) = \{\mathbf{0}\}$.

(22) First, $\mathbf{a} \in \mathcal{W}^\perp$ by Corollary 6.16. Also, since $\mathbf{proj}_{\mathcal{W}}\mathbf{v} \in \mathcal{W}$ and $\mathbf{w} \in \mathcal{W}$ (since T is in $\mathcal{W}$), we have $\mathbf{b} = (\mathbf{proj}_{\mathcal{W}}\mathbf{v}) - \mathbf{w} \in \mathcal{W}$. Finally,

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= \|\mathbf{v} - \mathbf{proj}_{\mathcal{W}}\mathbf{v} + (\mathbf{proj}_{\mathcal{W}}\mathbf{v}) - \mathbf{w}\|^2 \\ &= \|\mathbf{a} + \mathbf{b}\|^2 = (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 \qquad (\text{since } \mathbf{a} \cdot \mathbf{b} = 0) \\ &\geq \|\mathbf{a}\|^2 \\ &= \|\mathbf{v} - \mathbf{proj}_{\mathcal{W}}\mathbf{v}\|^2. \end{aligned}$$

(23) (a) Mimic the proof of Theorem 6.11.
(b) Mimic the proofs of Theorem 6.12 and Corollary 6.13.

(24) Note that, for any $\mathbf{v} \in \mathbb{R}^k$, if we consider $\mathbf{v}$ to be an $k \times 1$ matrix, then $\mathbf{v}^T\mathbf{v}$ is the 1×1 matrix whose single entry is $\mathbf{v} \cdot \mathbf{v}$. Thus, we can equate $\mathbf{v} \cdot \mathbf{v}$ with $\mathbf{v}^T\mathbf{v}$.

(a) $\mathbf{v} \cdot (\mathbf{A}\mathbf{w}) = \mathbf{v}^T(\mathbf{A}\mathbf{w}) = (\mathbf{v}^T\mathbf{A})\mathbf{w} = (\mathbf{A}^T\mathbf{v})^T\mathbf{w} = (\mathbf{A}^T\mathbf{v}) \cdot \mathbf{w}$
(b) Let $\mathbf{v} \in \ker(L_2)$. We must show that $\mathbf{v}$ is orthogonal to every vector in range(L_1). An arbitrary element of range(L_1) is of the form $L_1(\mathbf{w})$, for some $\mathbf{w} \in \mathbb{R}^n$. Now,

$$\begin{aligned} \mathbf{v} \cdot L_1(\mathbf{w}) &= L_2(\mathbf{v}) \cdot \mathbf{w} && (\text{by part (a)}) \\ &= \mathbf{0} \cdot \mathbf{w} && (\text{because } \mathbf{v} \in \ker(L_2)) \\ &= 0. \end{aligned}$$

Hence, $\mathbf{v} \in (\text{range}(L_1))^\perp$.

(c) By Theorem 5.9, $\dim(\ker(L_2)) = m - \text{rank}(\mathbf{A}^T) = m - \text{rank}(\mathbf{A})$ (by Corollary 5.11). Also, $\dim(\text{range}(L_1)^\perp) = m - \dim(\text{range}(L_1))$ (by Corollary 6.13) $= m - \text{rank}(\mathbf{A})$ (by Theorem 5.9). Hence, $\dim(\ker(L_2)) = \dim(\text{range}(L_1)^\perp)$. This, together with part (b) and Theorem 4.13 completes the proof.

(d) The row space of $\mathbf{A}$ = the column space of $\mathbf{A}^T = \text{range}(L_2)$. Switching the roles of $\mathbf{A}$ and $\mathbf{A}^T$ in parts (a), (b) and (c), we see that $(\text{range}(L_2))^\perp = \ker(L_1)$. Taking the orthogonal complement of both sides of this equality and using Corollary 6.14 yields $\text{range}(L_2) = (\ker(L_1))^\perp$, which completes the proof.

(25) (a) T (b) T (c) F (d) F (e) T (f) F (g) T (h) T (i) T (j) F (k) T (l) T

Section 6.3

(1) Parts (a) and (g) are symmetric, because the matrix for L is symmetric.

Part (b) is not symmetric, because the matrix for L with respect to the standard basis is not symmetric.

Parts (c), (d), and (f) are symmetric, since L is orthogonally diagonalizable.

Part (e) is not symmetric, since L is not diagonalizable, and hence not orthogonally diagonalizable.

(2) (a) $\frac{1}{25}\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

(b) $\frac{1}{11}\begin{bmatrix} 11 & -6 & 6 \\ -6 & 18 & 0 \\ 6 & 0 & 4 \end{bmatrix}$

(c) $\frac{1}{49}\begin{bmatrix} 58 & -6 & -18 \\ -6 & 53 & 12 \\ -18 & 12 & 85 \end{bmatrix}$

(d) $\frac{1}{169}\begin{bmatrix} -119 & -72 & -96 & 0 \\ -72 & 119 & 0 & 96 \\ -96 & 0 & 119 & -72 \\ 0 & 96 & -72 & -119 \end{bmatrix}$

(3) (a) $B = (\frac{1}{13}[5, 12], \frac{1}{13}[-12, 5])$, $\mathbf{P} = \frac{1}{13}\begin{bmatrix} 5 & -12 \\ 12 & 5 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 169 \end{bmatrix}$

(b) $B = (\frac{1}{5}[4, 3], \frac{1}{5}[3, -4])$, $\mathbf{P} = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$

(c) $B = \left(\frac{1}{\sqrt{2}}[-1, 1, 0], \frac{1}{3\sqrt{2}}[1, 1, 4], \frac{1}{3}[-2, -2, 1]\right)$

(other bases are possible, since E_1 is two-dimensional),

$\mathbf{P} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Another possibility is

$B = (\frac{1}{3}[-1, 2, 2], \frac{1}{3}[2, -1, 2], \frac{1}{3}[2, 2, -1])$, $\mathbf{P} = \frac{1}{3}\begin{bmatrix} -1 & 2 & -2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

(d) $B = (\frac{1}{9}[-8, 1, 4], \frac{1}{9}[4, 4, 7], \frac{1}{9}[1, -8, 4])$,

$\mathbf{P} = \frac{1}{9}\begin{bmatrix} -8 & 4 & 1 \\ 1 & 4 & -8 \\ 4 & 7 & 4 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$

(e) $B = \left(\frac{1}{\sqrt{14}}[3,2,1,0], \frac{1}{\sqrt{14}}[-2,3,0,1], \frac{1}{\sqrt{14}}[1,0,-3,2], \frac{1}{\sqrt{14}}[0,-1,2,3]\right)$,

$$\mathbf{P} = \frac{1}{\sqrt{14}}\begin{bmatrix} 3 & -2 & 1 & 0 \\ 2 & 3 & 0 & -1 \\ 1 & 0 & -3 & 2 \\ 0 & 1 & 2 & 3 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

(f) $B = (\frac{1}{\sqrt{26}}[4,1,3], \frac{1}{5}[3,0,-4], \frac{1}{5\sqrt{26}}[4,-25,3])$

(other bases are possible, since E_{-13} is two-dimensional),

$$\mathbf{P} = \begin{bmatrix} \frac{4}{\sqrt{26}} & \frac{3}{5} & \frac{4}{5\sqrt{26}} \\ \frac{1}{\sqrt{26}} & 0 & -\frac{25}{5\sqrt{26}} \\ \frac{3}{\sqrt{26}} & -\frac{4}{5} & \frac{3}{5\sqrt{26}} \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 13 & 0 & 0 \\ 0 & -13 & 0 \\ 0 & 0 & -13 \end{bmatrix}$$

(g) $B = \left(\frac{1}{\sqrt{5}}[1,2,0], \frac{1}{\sqrt{6}}[-2,1,1], \frac{1}{\sqrt{30}}[2,-1,5]\right)$,

$$\mathbf{P} = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{6}} & \frac{2}{\sqrt{30}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{30}} \\ 0 & \frac{1}{\sqrt{6}} & \frac{5}{\sqrt{30}} \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & -15 \end{bmatrix}$$

(4) (a) $C = (\frac{1}{19}[-10,15,6], \frac{1}{19}[15,6,10])$, $B = (\frac{1}{19\sqrt{5}}[20,27,26], \frac{1}{19\sqrt{5}}[35,-24,-2])$,

$$\mathbf{A} = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}, \mathbf{P} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$$

(b) $C = (\frac{1}{2}[1,-1,1,1], \frac{1}{2}[-1,1,1,1], \frac{1}{2}[1,1,1,-1])$,
$B = (\frac{1}{2\sqrt{5}}[3,-1,3,1], \frac{1}{\sqrt{30}}[-2,4,3,1], \frac{1}{6}[-1,-1,0,2])$,

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & 2 \\ 2 & 2 & -1 \end{bmatrix}, \mathbf{P} = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} & -\frac{2}{\sqrt{6}} \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

(5) (a) $\frac{1}{25}\begin{bmatrix} 23 & -36 \\ -36 & 2 \end{bmatrix}$ (b) $\frac{1}{17}\begin{bmatrix} 353 & -120 \\ -120 & 514 \end{bmatrix}$ (c) $\frac{1}{3}\begin{bmatrix} 11 & 4 & -4 \\ 4 & 17 & -8 \\ -4 & -8 & 17 \end{bmatrix}$

(6) For example, the matrix $\mathbf{A}$ in Example 7 of Section 5.6 is diagonalizable but not symmetric and hence not orthogonally diagonalizable.

(7) $\frac{1}{2}\begin{bmatrix} a+c+\sqrt{(a-c)^2+4b^2} & 0 \\ 0 & a+c-\sqrt{(a-c)^2+4b^2} \end{bmatrix}$ (Note: You only need to compute the eigenvalues. You do not need the corresponding eigenvectors.)

(8) (a) Since L is diagonalizable and $\lambda = 1$ is the only eigenvalue, $E_1 = \mathbb{R}^n$. Hence, for every $\mathbf{v} \in \mathbb{R}^n$, $L(\mathbf{v}) = 1\mathbf{v} = \mathbf{v}$. Therefore, L is the identity operator.

(b) L must be the zero linear operator. Since L is diagonalizable, the eigenspace for 0 is all of $\mathcal{V}$.

(9) Let L_1 and L_2 be symmetric operators on $\mathbb{R}^n$. Then, note that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$(L_2 \circ L_1)(\mathbf{x}) \cdot \mathbf{y} = L_2(L_1(\mathbf{x})) \cdot \mathbf{y} = L_1(\mathbf{x}) \cdot L_2(\mathbf{y}) = \mathbf{x} \cdot (L_1 \circ L_2)(\mathbf{y}).$$

Now $L_2 \circ L_1$ is symmetric iff for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \cdot (L_1 \circ L_2)(\mathbf{y}) = (L_2 \circ L_1)(\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (L_2 \circ L_1)(\mathbf{y})$ iff for all $\mathbf{y} \in \mathbb{R}^n$, $(L_1 \circ L_2)(\mathbf{y}) = (L_2 \circ L_1)(\mathbf{y})$ iff $(L_1 \circ L_2) = (L_2 \circ L_1)$.

(This argument uses the fact that if for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{z}$, then $\mathbf{y} = \mathbf{z}$. Proof: Let $\mathbf{x} = \mathbf{y} - \mathbf{z}$. Then,

$$\begin{aligned}(\mathbf{y}-\mathbf{z}) \cdot \mathbf{y} = (\mathbf{y}-\mathbf{z}) \cdot \mathbf{z} \quad &\Longrightarrow \quad ((\mathbf{y}-\mathbf{z}) \cdot \mathbf{y}) - ((\mathbf{y}-\mathbf{z}) \cdot \mathbf{z}) = 0 \\ &\Longrightarrow \quad (\mathbf{y}-\mathbf{z}) \cdot (\mathbf{y}-\mathbf{z}) = 0 \\ &\Longrightarrow \quad \|\mathbf{y}-\mathbf{z}\|^2 = 0 \\ &\Longrightarrow \quad \mathbf{y}-\mathbf{z} = \mathbf{0} \quad \Longrightarrow \quad \mathbf{y} = \mathbf{z}.\ )\end{aligned}$$

(10) (i) $\Longrightarrow$ (ii): This is Exercise 6 in Section 3.4.

(ii) $\Longrightarrow$ (iii): Since $\mathbf{A}$ and $\mathbf{B}$ are symmetric, both are orthogonally similar to diagonal matrices $\mathbf{D}_1$ and $\mathbf{D}_2$, respectively, by Theorems 6.19 and 6.22. Now, since $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic polynomial, and hence the same eigenvalues with the same algebraic multiplicities (and geometric multiplicities, since $\mathbf{A}$ and $\mathbf{B}$ are diagonalizable), we can assume $\mathbf{D}_1 = \mathbf{D}_2$ (by listing the eigenvectors in an appropriate order when diagonalizing). Thus, if $\mathbf{D}_1 = \mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1$ and $\mathbf{D}_2 = \mathbf{P}_2^{-1}\mathbf{B}\mathbf{P}_2$, for some orthogonal matrices $\mathbf{P}_1$ and $\mathbf{P}_2$, then $(\mathbf{P}_1\mathbf{P}_2^{-1})^{-1}\mathbf{A}(\mathbf{P}_1\mathbf{P}_2^{-1}) = \mathbf{B}$. Note that $\mathbf{P}_1\mathbf{P}_2^{-1}$ is orthogonal by parts (2) and (3) of Theorem 6.6, so $\mathbf{A}$ and $\mathbf{B}$ are orthogonally similar.

(iii) $\Longrightarrow$ (i): Trivial.

(11) $L(\mathbf{v}_1) \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot L(\mathbf{v}_2) \Longrightarrow (\lambda_1\mathbf{v}_1) \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot (\lambda_2\mathbf{v}_2) \Longrightarrow (\lambda_2 - \lambda_1)(\mathbf{v}_1 \cdot \mathbf{v}_2) = 0 \Longrightarrow \mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, since $\lambda_2 - \lambda_1 \neq 0$.

(12) Suppose $\mathbf{A}$ is orthogonal, and λ is an eigenvalue for $\mathbf{A}$ with corresponding unit eigenvector $\mathbf{u}$. Then $1 = \mathbf{u} \cdot \mathbf{u} = \mathbf{A}\mathbf{u} \cdot \mathbf{A}\mathbf{u}$ (by Theorem 6.9) $= (\lambda\mathbf{u}) \cdot (\lambda\mathbf{u}) = \lambda^2(\mathbf{u} \cdot \mathbf{u}) = \lambda^2$. Hence $\lambda = \pm 1$.

Conversely, suppose that $\mathbf{A}$ is symmetric with all eigenvalues equal to ± 1. Let $\mathbf{P}$ be an orthogonal matrix with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, a diagonal matrix with all entries equal to ± 1 on the main diagonal. Note that $\mathbf{D}^2 = \mathbf{I}_n$. Then, since $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, $\mathbf{A}\mathbf{A}^T = \mathbf{A}\mathbf{A} = \mathbf{P}\mathbf{D}^2\mathbf{P}^{-1} = \mathbf{P}\mathbf{I}_n\mathbf{P}^{-1} = \mathbf{I}_n$. Hence $\mathbf{A}$ is orthogonal.

(13) (a) $-\mathbf{A}(\mathbf{A} - \mathbf{I}_n)^T = -\mathbf{A}(\mathbf{A}^T - \mathbf{I}_n^T) = -\mathbf{A}\mathbf{A}^T + \mathbf{A} = -\mathbf{I}_n + \mathbf{A} = \mathbf{A} - \mathbf{I}_n$. Therefore,

$$|\mathbf{A} - \mathbf{I}_n| = \left|-\mathbf{A}(\mathbf{A} - \mathbf{I}_n)^T\right| = |-\mathbf{A}|\left|(\mathbf{A} - \mathbf{I}_n)^T\right| = (-1)^n\,|\mathbf{A}|\,|(\mathbf{A} - \mathbf{I}_n)| = -\,|\mathbf{A} - \mathbf{I}_n|\,,$$

since n is odd and $|\mathbf{A}| = 1$. This implies that $|\mathbf{A} - \mathbf{I}_n| = 0$, and so $\mathbf{A} - \mathbf{I}_n$ is singular.

(b) By part (a), $\mathbf{A} - \mathbf{I}_n$ is singular. Hence, there is a nonzero vector $\mathbf{v}$ such that $(\mathbf{A} - \mathbf{I}_n)\mathbf{v} = \mathbf{0}$. Thus, $\mathbf{A}\mathbf{v} = \mathbf{I}_n\mathbf{v} = \mathbf{v}$, an so $\mathbf{v}$ is an eigenvector for the eigenvalue $\lambda = 1$.

(c) Suppose $\mathbf{v}$ is a unit eigenvector for $\mathbf{A}$ corresponding to $\lambda = 1$. Expand the set $\{\mathbf{v}\}$ to an orthonormal basis $B = \{\mathbf{w}, \mathbf{x}, \mathbf{v}\}$ for $\mathbb{R}^3$. Notice that we listed the vector $\mathbf{v}$ last. Let $\mathbf{Q}$ be the orthogonal matrix whose columns are $\mathbf{w}$, $\mathbf{x}$, and $\mathbf{v}$, in that order. Let $\mathbf{P} = \mathbf{Q}^T\mathbf{A}\mathbf{Q}$. Now, the last column of $\mathbf{P}$ equals $\mathbf{Q}^T\mathbf{A}\mathbf{v} = \mathbf{Q}^T\mathbf{v}$ (since $\mathbf{A}\mathbf{v} = \mathbf{v}$) $= \mathbf{e}_3$, since $\mathbf{v}$ is a unit vector orthogonal to the first two rows of $\mathbf{Q}^T$. Next, since $\mathbf{P}$ is the product of orthogonal matrices, $\mathbf{P}$ is orthogonal, and since $|\mathbf{Q}^T| = |\mathbf{Q}| = \pm 1$, we must have $|\mathbf{P}| = |\mathbf{A}| = 1$. Also, the first two columns of $\mathbf{P}$ are orthogonal to its last column, $\mathbf{e}_3$, making the $(3,1)$ and $(3,2)$ entries of $\mathbf{P}$ equal to 0. Since the first column of $\mathbf{P}$ is a unit vector with third coordinate 0, it can be expressed as $[\cos\theta, \sin\theta, 0]$, for some value of θ, with $0 \leq \theta < 2\pi$. The second column of $\mathbf{P}$ must be a unit vector with third coordinate 0, orthogonal to the first column. The only possibilities are $\pm[-\sin\theta, \cos\theta, 0]$. Choosing the minus sign makes $|\mathbf{P}| = -1$, so the second column must be $[-\sin\theta, \cos\theta, 0]$. Thus, $\mathbf{P} = \mathbf{Q}^T\mathbf{A}\mathbf{Q}$ has the desired form.

(d) The matrix for the linear operator L on $\mathbb{R}^3$ given by $L(\mathbf{v}) = \mathbf{A}\mathbf{v}$ with respect to the standard basis has a matrix that is orthogonal with determinant 1. But, by part (c), and the Orthogonal Diagonalization Method, the matrix for L with respect to the ordered orthonormal basis $B = (\mathbf{w}, \mathbf{x}, \mathbf{v})$ is $\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$. According to Table 5.1 in Section 5.2, this is a counterclockwise rotation around the z-axis through the angle θ. But since we are working in B-coordinates, the z-axis corresponds to the vector $\mathbf{v}$, the third vector in B. Because the basis B is orthonormal, the plane of the rotation is perpendicular to the axis of rotation.

(e) The axis of rotation is in the direction of the vector $[-1, 3, 3]$. The angle of rotation is $\theta \approx 278°$. This is a counterclockwise rotation about the vector $[-1, 3, 3]$ as you look down from the point $(-1, 3, 3)$ toward the plane through the origin spanned by $[6, 1, 1]$ and $[0, 1, -1]$.

(f) Let $\mathbf{G}$ be the matrix with respect to the standard basis for any chosen orthogonal reflection through a plane in $\mathbb{R}^3$. Then, $\mathbf{G}$ is orthogonal, $|\mathbf{G}| = -1$ (since it has eigenvalues $-1, 1, 1$), and $\mathbf{G}^2 = \mathbf{I}_3$. Let $\mathbf{C} = \mathbf{A}\mathbf{G}$. Then $\mathbf{C}$ is orthogonal since it is the product of orthogonal matrices. Since $|\mathbf{A}| = -1$, it follows that $|\mathbf{C}| = 1$. Thus, $\mathbf{A} = \mathbf{A}\mathbf{G}^2 = \mathbf{C}\mathbf{G}$, where $\mathbf{C}$ represents a rotation about some axis in $\mathbb{R}^3$ by part (d) of this exercise.

(14) (a) T (b) F (c) T (d) T (e) T (f) T

Chapter 6 Review Exercises

(1) (a) $[\mathbf{v}]_B = [-1, 4, 2]$ (b) $[\mathbf{v}]_B = [3, 2, 3]$

(2) (a) $\{[1, -1, -1, 1], [1, 1, 1, 1]\}$
(b) $\{[1, 3, 4, 3, 1], [-1, -1, 0, 1, 1], [-1, 3, -4, 3, -1]\}$

(3) $\{[6, 3, -6], [3, 6, 6], [2, -2, 1]\}$

(4) (a) $\{[4, 7, 0, 4], [2, 0, 1, -2], [29, -28, -18, 20], [0, 4, -14, -7]\}$
(b) $\{\frac{1}{9}[4, 7, 0, 4], \frac{1}{3}[2, 0, 1, -2], \frac{1}{9\sqrt{29}}[29, -28, -18, 20], \frac{1}{3\sqrt{29}}[0, 4, -14, -7]\}$
(c) $\begin{bmatrix} \frac{4}{9} & \frac{7}{9} & 0 & \frac{4}{9} \\ \frac{2}{3} & 0 & \frac{1}{3} & -\frac{2}{3} \\ \frac{\sqrt{29}}{9} & -\frac{28}{9\sqrt{29}} & -\frac{2}{\sqrt{29}} & \frac{20}{9\sqrt{29}} \\ 0 & \frac{4}{3\sqrt{29}} & -\frac{14}{3\sqrt{29}} & -\frac{7}{3\sqrt{29}} \end{bmatrix}$

(5) Using the hint, note that $\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot (\mathbf{A}^T\mathbf{A}\mathbf{w})$. Now, $\mathbf{A}^T\mathbf{A}\mathbf{e}_j$ is the jth column of $\mathbf{A}^T\mathbf{A}$, so $\mathbf{e}_i \cdot (\mathbf{A}^T\mathbf{A}\mathbf{e}_j)$ is the (i, j) entry of $\mathbf{A}^T\mathbf{A}$. Since this equals $\mathbf{e}_i \cdot \mathbf{e}_j$, we see that $\mathbf{A}^T\mathbf{A} = \mathbf{I}_n$. Hence, $\mathbf{A}^T = \mathbf{A}^{-1}$, and $\mathbf{A}$ is orthogonal.

(6) (a) $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v} = [0, -9, 18]$; $\mathbf{w}_2 = \mathbf{proj}_{\mathcal{W}^\perp}\mathbf{v} = [2, 16, 8]$

(b) $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v} = \left[\frac{7}{2}, \frac{23}{2}, \frac{5}{2}, -\frac{3}{2}\right]$; $\mathbf{w}_2 = \mathbf{proj}_{\mathcal{W}^\perp}\mathbf{v} = \left[-\frac{3}{2}, -\frac{3}{2}, \frac{9}{2}, -\frac{15}{2}\right]$

(7) (a) Distance ≈ 10.141294 (b) Distance ≈ 14.248050

(8) $\{[1, 0, -2, -2], [0, 1, -2, 2]\}$

(9) $\frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

(10) $\frac{1}{7}\begin{bmatrix} 3 & 6 & -2 \\ 6 & -2 & 3 \\ -2 & 3 & 6 \end{bmatrix}$

(11) (a) $p_L(x) = x(x-1)^2 = x^3 - 2x^2 + x$ (b) $p_L(x) = x^2(x-1) = x^3 - x^2$

(12) (a) Not a symmetric operator, since the matrix for L with respect to the standard basis is $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, which is not symmetric.

(b) Symmetric operator, since the matrix for L is orthogonally diagonalizable.

(c) Symmetric operator, since the matrix for L with respect to the standard basis is symmetric.

(13) (a) $B = \left(\frac{1}{\sqrt{6}}[-1, -2, 1], \frac{1}{\sqrt{30}}[1, 2, 5], \frac{1}{\sqrt{5}}[-2, 1, 0]\right)$; $\mathbf{P} = \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{6}} & \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{6}} & \frac{5}{\sqrt{30}} & 0 \end{bmatrix}$; $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(b) $B = \left(\frac{1}{\sqrt{3}}[1, -1, 1], \frac{1}{\sqrt{2}}[1, 1, 0], \frac{1}{\sqrt{6}}[-1, 1, 2]\right)$; $\mathbf{P} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$; $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

(14) Any diagonalizable 4×4 matrix that is not symmetric works. For example, chose any non-diagonal upper triangular 4×4 matrix with 4 distinct entries on the main diagonal.

(15) Because $\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\left(\mathbf{A}^T\right)^T = \mathbf{A}^T\mathbf{A}$, we see that $\mathbf{A}^T\mathbf{A}$ is symmetric. Therefore $\mathbf{A}^T\mathbf{A}$ is orthogonally diagonalizable by Corollary 6.23.

(16)

(a) T	(d) T	(g) T	(j) T	(m) F	(p) T	(s) T	(v) T
(b) T	(e) F	(h) T	(k) T	(n) T	(q) F	(t) T	
(c) T	(f) F	(i) T	(l) T	(o) T	(r) F	(u) F	

Chapter 7

Section 7.1

(1) (a) $[1+4i, 1+i, 6-i]$ (c) $[5+i, 2-i, 3i]$ (e) $1+28i$
(b) $[-12-32i, -7+30i, 53-29i]$ (d) $[-24-12i, -28-8i, -32i]$ (f) $3+77i$

(2) Let $\mathbf{z}_1 = [a_1, \ldots, a_n]$, and let $\mathbf{z}_2 = [b_1, \ldots, b_n]$.

(a) Part (1): $\mathbf{z}_1 \cdot \mathbf{z}_2 = \sum_{i=1}^n a_i\overline{b_i} = \sum_{i=1}^n \overline{\overline{(a_i(\overline{b_i}))}} = \sum_{i=1}^n \overline{(\overline{a_i}(\overline{\overline{b_i}}))} = \overline{\sum_{i=1}^n (\overline{a_i})b_i} = \overline{\sum_{i=1}^n b_i(\overline{a_i})} = \overline{\mathbf{z}_2 \cdot \mathbf{z}_1}$
Part (2): $\mathbf{z}_1 \cdot \mathbf{z}_1 = \sum_{i=1}^n a_i\overline{a_i} = \sum_{i=1}^n |a_i|^2 \geq 0$.
Also, $\mathbf{z}_1 \cdot \mathbf{z}_1 = \sum_{i=1}^n |a_i|^2 = \left(\sqrt{\sum_{i=1}^n |a_i|^2}\right)^2 = \|\mathbf{z}_1\|^2$

(b) $\overline{k}(\mathbf{z}_1 \cdot \mathbf{z}_2) = \overline{k}\sum_{i=1}^n a_i\overline{b_i} = \sum_{i=1}^n \overline{k}a_i\overline{b_i} = \sum_{i=1}^n a_i\overline{kb_i} = \mathbf{z}_1 \cdot (k\mathbf{z}_2)$

(3) (a) $\begin{bmatrix} 11+4i & -4-2i \\ 2-4i & 12 \end{bmatrix}$

(b) $\begin{bmatrix} 1-i & 2i & 6-4i \\ 0 & 3-i & 5 \\ 10i & 0 & 7+3i \end{bmatrix}$

(c) $\begin{bmatrix} 1-i & 0 & 10i \\ 2i & 3-i & 0 \\ 6-4i & 5 & 7+3i \end{bmatrix}$

(d) $\begin{bmatrix} -3-15i & -3 & 9i \\ 9-6i & 0 & 3+12i \end{bmatrix}$

(e) $\begin{bmatrix} -7+6i & -9-i \\ -3-3i & 4-6i \end{bmatrix}$

(f) $\begin{bmatrix} 1+40i & -4-14i \\ 13-50i & 23+21i \end{bmatrix}$

(g) $\begin{bmatrix} -12+6i & -16-3i & -11+41i \\ -15+23i & -20+5i & -61+15i \end{bmatrix}$

(h) $\begin{bmatrix} 86-33i & 6-39i \\ 61+36i & 13+9i \end{bmatrix}$

(i) $\begin{bmatrix} 4+36i & -5+39i \\ 1-7i & -6-4i \\ 5+40i & -7-5i \end{bmatrix}$

(j) $\begin{bmatrix} 40+58i & 50i & -20+80i \\ 4+8i & 8-6i & -20 \\ 56-10i & 50+10i & 80+102i \end{bmatrix}$

(4) (a) (i,j) entry of $(k\mathbf{Z})^* = (i,j)$ entry of $\overline{(k\mathbf{Z})}^T = (j,i)$ entry of $\overline{(k\mathbf{Z})} = \overline{kz_{ji}} = \overline{k}\,\overline{z_{ji}}$
$= \overline{k}((j,i)$ entry of $\overline{\mathbf{Z}}) = \overline{k}((i,j)$ entry of $\overline{\mathbf{Z}}^T) = \overline{k}((i,j)$ entry of $\mathbf{Z}^*)$

(b) Let $\mathbf{Z}$ be an $m \times n$ complex matrix and let $\mathbf{W}$ be an $n \times p$ matrix. Note that $(\mathbf{ZW})^*$ and $\mathbf{W}^*\mathbf{Z}^*$ are both $p \times m$ matrices. We present two methods of proof that $(\mathbf{ZW})^* = \mathbf{W}^*\mathbf{Z}^*$.

First method: Notice that the rule $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ holds for complex matrices, since matrix multiplication for complex matrices is defined in exactly the same manner as for real matrices, and hence the proof of Theorem 1.18 is valid for complex matrices as well. Thus we have:

$$(\mathbf{ZW})^* = \overline{(\mathbf{ZW})^T} = \overline{\mathbf{W}^T\mathbf{Z}^T} = \overline{(\mathbf{W}^T)}\,\overline{(\mathbf{Z}^T)} \text{ (by part (4) of Theorem 7.2)} = \mathbf{W}^*\mathbf{Z}^*.$$

Second method: We begin by computing the (i,j) entry of $(\mathbf{ZW})^*$. Now, the (j,i) entry of $\mathbf{ZW} = z_{j1}w_{1i} + \cdots + z_{jn}w_{ni}$. Hence,

$$\begin{aligned} (i,j) \text{ entry of } (\mathbf{ZW})^* &= \overline{((j,i) \text{ entry of } \mathbf{ZW})} \\ &= \overline{(z_{j1}w_{1i} + \cdots + z_{jn}w_{ni})} = \overline{z_{j1}w_{1i}} + \cdots + \overline{z_{jn}w_{ni}}. \end{aligned}$$

Next, we compute the (i,j) entry of $\mathbf{W}^*\mathbf{Z}^*$ to show that it equals the (i,j) entry of $(\mathbf{ZW})^*$. Let $\mathbf{A} = \mathbf{Z}^*$, an $n \times m$ matrix, and $\mathbf{B} = \mathbf{W}^*$, a $p \times n$ matrix. Then $a_{kl} = \overline{z_{lk}}$ and $b_{kl} = \overline{w_{lk}}$. Hence,

$$\begin{aligned}(i,j) \text{ entry of } \mathbf{W}^*\mathbf{Z}^* &= (i,j) \text{ entry of } \mathbf{BA} = b_{i1}a_{1j} + \cdots + b_{in}a_{nj} \\ &= \overline{w_{1i}}\,\overline{z_{j1}} + \cdots + \overline{w_{ni}}\,\overline{z_{jn}} = \overline{z_{j1}}\,\overline{w_{1i}} + \cdots + \overline{z_{jn}}\,\overline{w_{ni}}.\end{aligned}$$

Therefore, $(\mathbf{ZW})^* = \mathbf{W}^*\mathbf{Z}^*$.

(5) (a) Skew-Hermitian (b) Neither (c) Hermitian (d) Skew-Hermitian (e) Hermitian

(6) (a) $\mathbf{H}^* = (\frac{1}{2}(\mathbf{Z}+\mathbf{Z}^*))^* = \frac{1}{2}(\mathbf{Z}+\mathbf{Z}^*)^*$ (by part (3) of Theorem 7.2) $= \frac{1}{2}(\mathbf{Z}^*+\mathbf{Z})$ (by part (2) of Theorem 7.2) $= \mathbf{H}$. A similar calculation shows that $\mathbf{K}$ is skew-Hermitian.

(b) Part (a) proves existence for the decomposition by providing specific matrices $\mathbf{H}$ and $\mathbf{K}$. For uniqueness, suppose $\mathbf{Z} = \mathbf{H}+\mathbf{K}$, where $\mathbf{H}$ is Hermitian and $\mathbf{K}$ is skew-Hermitian. Our goal is to show that $\mathbf{H}$ and $\mathbf{K}$ must satisfy the formulas from part (a). Now $\mathbf{Z}^* = (\mathbf{H}+\mathbf{K})^* = \mathbf{H}^*+\mathbf{K}^* = \mathbf{H}-\mathbf{K}$. Hence $\mathbf{Z}+\mathbf{Z}^* = (\mathbf{H}+\mathbf{K})+(\mathbf{H}-\mathbf{K}) = 2\mathbf{H}$, which gives $\mathbf{H} = \frac{1}{2}(\mathbf{Z}+\mathbf{Z}^*)$. Similarly, $\mathbf{Z}-\mathbf{Z}^* = 2\mathbf{K}$, so $\mathbf{K} = \frac{1}{2}(\mathbf{Z}-\mathbf{Z}^*)$.

(7) (a) $\mathbf{HJ}$ is Hermitian iff $(\mathbf{HJ})^* = \mathbf{HJ}$ iff $\mathbf{J}^*\mathbf{H}^* = \mathbf{HJ}$ iff $\mathbf{JH} = \mathbf{HJ}$ (since $\mathbf{H}$ and $\mathbf{J}$ are Hermitian).

(b) Use induction on k.
Base Step ($k = 1$): $\mathbf{H}^1 = \mathbf{H}$ is given to be Hermitian.
Inductive Step: Assume $\mathbf{H}^k$ is Hermitian and prove that $\mathbf{H}^{k+1}$ is Hermitian. But $\mathbf{H}^k\mathbf{H} = \mathbf{H}^{k+1} = \mathbf{H}^{1+k} = \mathbf{H}\mathbf{H}^k$ (by part (1) of Theorem 1.17), and so by part (a), $\mathbf{H}^{k+1} = \mathbf{H}^k\mathbf{H}$ is Hermitian.

(c) $(\mathbf{P}^*\mathbf{HP})^* = \mathbf{P}^*\mathbf{H}^*(\mathbf{P}^*)^* = \mathbf{P}^*\mathbf{HP}$ (since $\mathbf{H}$ is Hermitian).

(8) $(\mathbf{AA}^*)^* = (\mathbf{A}^*)^*\mathbf{A}^* = \mathbf{AA}^*$ (and $\mathbf{A}^*\mathbf{A}$ is handled similarly).

(9) Let $\mathbf{Z}$ be Hermitian. Then $\mathbf{ZZ}^* = \mathbf{ZZ} = \mathbf{Z}^*\mathbf{Z}$. Similarly, if $\mathbf{Z}$ is skew-Hermitian,

$$\mathbf{ZZ}^* = \mathbf{Z}(-\mathbf{Z}) = -(\mathbf{ZZ}) = (-\mathbf{Z})\mathbf{Z} = \mathbf{Z}^*\mathbf{Z}.$$

(10) Let $\mathbf{Z}$ be normal. Consider $\mathbf{H}_1 = \frac{1}{2}(\mathbf{Z}+\mathbf{Z}^*)$ and $\mathbf{H}_2 = \frac{1}{2}(\mathbf{Z}-\mathbf{Z}^*)$. Note that $\mathbf{H}_1$ is Hermitian and $\mathbf{H}_2$ is skew-Hermitian (by Exercise 6(a)). Clearly $\mathbf{Z} = \mathbf{H}_1+\mathbf{H}_2$. Now,

$$\begin{aligned}\mathbf{H}_1\mathbf{H}_2 &= (\tfrac{1}{2})(\tfrac{1}{2})(\mathbf{Z}+\mathbf{Z}^*)(\mathbf{Z}-\mathbf{Z}^*) \\ &= \tfrac{1}{4}(\mathbf{Z}^2+\mathbf{Z}^*\mathbf{Z}-\mathbf{ZZ}^*-(\mathbf{Z}^*)^2) \\ &= \tfrac{1}{4}(\mathbf{Z}^2-(\mathbf{Z}^*)^2) \quad \text{(since } \mathbf{Z} \text{ is normal).}\end{aligned}$$

A similar computation shows that $\mathbf{H}_2\mathbf{H}_1$ is also equal to $\frac{1}{4}(\mathbf{Z}^2-(\mathbf{Z}^*)^2)$. Hence $\mathbf{H}_1\mathbf{H}_2 = \mathbf{H}_2\mathbf{H}_1$.

Conversely, let $\mathbf{H}_1$ be Hermitian and $\mathbf{H}_2$ be skew-Hermitian with $\mathbf{H}_1\mathbf{H}_2 = \mathbf{H}_2\mathbf{H}_1$. If $\mathbf{Z} = \mathbf{H}_1+\mathbf{H}_2$, then

$$\begin{aligned}\mathbf{ZZ}^* &= (\mathbf{H}_1+\mathbf{H}_2)(\mathbf{H}_1+\mathbf{H}_2)^* \\ &= (\mathbf{H}_1+\mathbf{H}_2)(\mathbf{H}_1^*+\mathbf{H}_2^*) \quad \text{(by part (2) of Theorem 7.2)} \\ &= (\mathbf{H}_1+\mathbf{H}_2)(\mathbf{H}_1-\mathbf{H}_2) \quad \text{(since } \mathbf{H}_1 \text{ is Hermitian and } \mathbf{H}_2 \text{ is skew-Hermitian)} \\ &= \mathbf{H}_1^2+\mathbf{H}_2\mathbf{H}_1-\mathbf{H}_1\mathbf{H}_2-\mathbf{H}_2^2 \\ &= \mathbf{H}_1^2-\mathbf{H}_2^2 \quad \text{(since } \mathbf{H}_1\mathbf{H}_2 = \mathbf{H}_2\mathbf{H}_1\text{).}\end{aligned}$$

A similar argument shows $\mathbf{Z}^*\mathbf{Z} = \mathbf{H}_1^2-\mathbf{H}_2^2$ also, and so $\mathbf{ZZ}^* = \mathbf{Z}^*\mathbf{Z}$, and $\mathbf{Z}$ is normal.

(11) (a) F (b) F (c) T (d) F (e) F (f) T

Section 7.2

(1) (a) $w = \frac{1}{5} + \frac{13}{5}i$; $z = \frac{28}{5} - \frac{3}{5}i$

(b) No solutions: After applying the row reduction method to the first 3 columns, the matrix changes to: $\begin{bmatrix} 1 & 1+i & 0 & 1-i \\ 0 & 0 & 1 & 3+4i \\ 0 & 0 & 0 & 4-2i \end{bmatrix}$

(c) Matrix row reduces to $\begin{bmatrix} 1 & 0 & 4-3i & 2+5i \\ 0 & 1 & -i & 5+2i \\ 0 & 0 & 0 & 0 \end{bmatrix}$;

Solution set $= \{(\ (2+5i) - (4-3i)c,\ (5+2i) + ic,\ c) \mid c \in \mathbb{C}\}$

(d) $w = 7 + i$; $z = -6 + 5i$

(e) No solutions: After applying the row reduction method to the first column, the matrix changes to: $\begin{bmatrix} 1 & 2+3i & -1+3i \\ 0 & 0 & 3+4i \end{bmatrix}$

(f) Matrix row reduces to $\begin{bmatrix} 1 & 0 & -3+2i & -i \\ 0 & 1 & 4+7i & 2-5i \end{bmatrix}$;

Solution set $= \{(\ (-i) + (3-2i)c,\ (2-5i) - (4+7i)c,\ c) \mid c \in \mathbb{C}\}$

(2) (a) $|\mathbf{A}| = 0$; $|\mathbf{A}^*| = 0 = \overline{|\mathbf{A}|}$

(b) $|\mathbf{A}| = -15 - 23i$; $|\mathbf{A}^*| = -15 + 23i = \overline{|\mathbf{A}|}$

(c) $|\mathbf{A}| = 3 - 5i$; $|\mathbf{A}^*| = 3 + 5i = \overline{|\mathbf{A}|}$

(3) Your computations may produce complex scalar multiples of the vectors given here, although they might not be immediately recognized as such.

(a) $p_{\mathbf{A}}(x) = x^2 + (1-i)x - i = (x-i)(x+1)$; eigenvalues $\lambda_1 = i$, $\lambda_2 = -1$; $E_i = \{c[1+i, 2] \mid c \in \mathbb{C}\}$; $E_{-1} = \{c[7+6i, 17] \mid c \in \mathbb{C}\}$.

(b) $p_{\mathbf{A}}(x) = x^3 - 11x^2 + 44x - 34$; eigenvalues: $\lambda_1 = 5+3i$, $\lambda_2 = 5-3i$, $\lambda_3 = 1$; $E_{\lambda_1} = \{c[1-3i, 5, 1-3i] \mid c \in \mathbb{C}\}$; $E_{\lambda_2} = \{c[1+3i, 5, 1+3i] \mid c \in \mathbb{C}\}$; $E_{\lambda_3} = \{c[1, 2, 2] \mid c \in \mathbb{C}\}$.

(c) $p_{\mathbf{A}}(x) = x^3 + (2-2i)x^2 - (1+4i)x - 2 = (x-i)^2(x+2)$; $\lambda_1 = i$, $\lambda_2 - 2$; $E_i = \{c[(-3-2i), 0, 2] + d[1, 1, 0] \mid c, d \in \mathbb{C}\}$; $E_{-2} = \{c[-1, i, 1] \mid c \in \mathbb{C}\}$.

(d) $p_{\mathbf{A}}(x) = x^3 - 2x^2 - x + 2 + i(-2x^2 + 4x) = (x-i)^2(x-2)$; eigenvalues: $\lambda_1 = i$, $\lambda_2 = 2$; $E_i = \{c[1, 1, 0] \mid c \in \mathbb{C}\}$; $E_2 = \{c[i, 0, 1] \mid c \in \mathbb{C}\}$. (Note that $\mathbf{A}$ is not diagonalizable.)

(4) (a) Let $\mathbf{A}$ be the matrix from Exercise 3(a). $\mathbf{A}$ is diagonalizable since $\mathbf{A}$ has 2 distinct eigenvalues. $\mathbf{P} = \begin{bmatrix} 1+i & 7+6i \\ 2 & 17 \end{bmatrix}$; $\mathbf{D} = \begin{bmatrix} i & 0 \\ 0 & -1 \end{bmatrix}$.

(b) Let $\mathbf{A}$ be the matrix from Exercise 3(d). The solution to Exercise 3(d) shows that the Diagonalization Method of Section 3.4 only produces two fundamental eigenvectors. Hence, $\mathbf{A}$ is not diagonalizable by Step 4 of that method.

(c) Let $\mathbf{A}$ be the matrix in Exercise 3(b). $p_{\mathbf{A}}(x) = x^3 - 11x^2 + 44x - 34$. There are 3 eigenvalues, $5+3i$, $5-3i$, and 1, each producing one complex fundamental eigenvector. Thus, there are three fundamental complex eigenvectors, and so $\mathbf{A}$ is diagonalizable as a complex matrix.

On the other hand, $\mathbf{A}$ has only one *real* eigenvalue $\lambda = 1$, and since λ produces only one real fundamental eigenvector, $\mathbf{A}$ is not diagonalizable as a real matrix.

(5) By the Fundamental Theorem of Algebra, the characteristic polynomial $p_{\mathbf{A}}(x)$ of a complex $n \times n$ matrix $\mathbf{A}$ must factor into n linear factors. Since each root of $p_{\mathbf{A}}(x)$ has multiplicity 1, there must be n distinct roots. Each of these roots is an eigenvalue, and each eigenvalue yields at least one fundamental eigenvector. Thus, we have a total of n fundamental eigenvectors, showing that $\mathbf{A}$ is diagonalizable.

(6) (a) T (b) F (c) F (d) F

Section 7.3

(1) (a) Mimic the proof in Example 6 in Section 4.1 of the textbook, restricting the discussion to polynomial functions of degree $\leq n$, and using complex-valued functions and complex scalars instead of their real counterparts.

(b) This is the complex analog of Theorem 1.12. Property (1) is proven for the real case just after the statement of Theorem 1.12 in Section 1.4 of the textbook. Generalize this proof to complex matrices. The other properties are similarly proved.

(2) (a) Linearly independent; dim = 2

(b) Not linearly independent; dim = 1. (Note: $[-3+6i, 3, 9i] = 3i[2+i, -i, 3]$.)

(c) Not linearly independent; dim = 2. (Note: Using the given vectors as columns, the matrix row reduces to $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$.)

(d) Not linearly independent, dim = 2. (Note: Using the given vectors as columns, the matrix row reduces to $\begin{bmatrix} 1 & 0 & i \\ 0 & 1 & -2i \\ 0 & 0 & 0 \end{bmatrix}$.)

(3) (a) Linearly independent; dim = 2. (Note: $[-i, 3+i, -1]$ is not a scalar multiple of $[2+i, -i, 3]$.)

(b) Linearly independent; dim = 2. (Note: $[-3+6i, 3, 9i]$ is not a *real* scalar multiple of $[2+i, -i, 3]$.)

(c) Not linearly independent; dim = 2.

(Note: $\begin{bmatrix} 3 & 1 & 1 \\ -1 & 1 & -3 \\ 1 & -2 & 5 \\ 2 & 0 & 2 \\ 0 & 4 & -8 \\ -1 & 1 & -3 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.)

(d) Linearly independent; dim = 3.

(Note: $\begin{bmatrix} 3 & 1 & 3 \\ -1 & 1 & 1 \\ 1 & -2 & -2 \\ 2 & 0 & 5 \\ 0 & 4 & 3 \\ -1 & 1 & -8 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.)

(4) (a) The 3×3 complex matrix whose columns (or rows) are the given vectors row reduces to $\mathbf{I}_3$.

(b) $[i, 1+i, -1]$

(5) Ordered basis $= ([1,0], [i,0], [0,1], [0,i])$; matrix $= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$

(6) Span: Let $\mathbf{v} \in \mathcal{V}$. Then there exists $c_1, \ldots, c_n \in \mathbb{C}$ such that $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Suppose, for each j, $c_j = a_j + b_j i$, for some $a_j, b_j \in \mathbb{R}$. Then

$$\mathbf{v} = (a_1 + b_1 i)\mathbf{v}_1 + \cdots + (a_n + b_n i)\mathbf{v}_n = a_1\mathbf{v}_1 + b_1(i\mathbf{v}_1) + \cdots + a_n\mathbf{v}_n + b_n(i\mathbf{v}_n),$$

which shows that $\mathbf{v} \in \operatorname{span}(\{\mathbf{v}_1, i\mathbf{v}_1, \ldots, \mathbf{v}_n, i\mathbf{v}_n\})$.

Linear Independence: Suppose $a_1\mathbf{v}_1 + b_1(i\mathbf{v}_1) + \cdots + a_n\mathbf{v}_n + b_n(i\mathbf{v}_n) = \mathbf{0}$. Then

$$(a_1 + b_1 i)\mathbf{v}_1 + \cdots + (a_n + b_n i)\mathbf{v}_n = \mathbf{0},$$

implying

$$(a_1 + b_1 i) = \cdots = (a_n + b_n i) = 0$$

(since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent). Hence, $a_1 = b_1 = a_2 = b_2 = \cdots = a_n = b_n = 0$.

(7) Exercise 6 clearly implies that if $\mathcal{V}$ is an n-dimensional complex vector space, then it is $2n$-dimensional when considered as a real vector space. Hence, the real dimension must be even. Therefore, $\mathbb{R}^3$ cannot be considered as a complex vector space since its real dimension is odd.

(8) $\begin{bmatrix} -3+i & -\frac{2}{5} - \frac{11}{5}i \\ \frac{1}{2} - \frac{3}{2}i & -i \\ -\frac{1}{2} + \frac{7}{2}i & -\frac{8}{5} - \frac{4}{5}i \end{bmatrix}$

(9) (a) T (b) F (c) T (d) F

Section 7.4

(1) (a) Not orthogonal; $[1+2i, -3-i] \cdot [4-2i, 3+i] = -10 + 10i$

(b) Not orthogonal; $[1-i, -1+i, 1-i] \cdot [i, -2i, 2i] = -5 - 5i$

(c) Orthogonal (d) Orthogonal

(2) $(c_i\mathbf{z}_i)\cdot(c_j\mathbf{z}_j) = c_i\overline{c_j}(\mathbf{z}_i\cdot\mathbf{z}_j)$ (by parts (4) and (5) of Theorem 7.1) $= c_i\overline{c_j}(0)$ (since $\mathbf{z}_i$ is orthogonal to $\mathbf{z}_j$) $= 0$, so $c_i\mathbf{z}_i$ is orthogonal to $c_j\mathbf{z}_j$. Also,

$$||c_i\mathbf{z}_i||^2 = (c_i\mathbf{z}_i)\cdot(c_i\mathbf{z}_i) = c_i\overline{c_i}(\mathbf{z}_i\cdot\mathbf{z}_i) = |c_i|^2||\mathbf{z}_i||^2 = 1,$$

since $|c_i| = 1$ and $\mathbf{z}_i$ is a unit vector. Hence, each $c_i\mathbf{z}_i$ is also a unit vector.

(3) (a) $\{[1+i, i, 1], [2, -1-i, -1+i], [0, 1, i]\}$

(b) $\begin{bmatrix} \frac{1+i}{2} & \frac{i}{2} & \frac{1}{2} \\ \frac{2}{\sqrt{8}} & \frac{-1-i}{\sqrt{8}} & \frac{-1+i}{\sqrt{8}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}$

(4) For any $n\times n$ matrix $\mathbf{B}$, $|\overline{\mathbf{B}}| = \overline{|\mathbf{B}|}$, by part (3) of Theorem 7.5. Therefore,

$$\Big|\,|\mathbf{A}|\,\Big|^2 = |\mathbf{A}|\,\overline{|\mathbf{A}|} = |\mathbf{A}|\,|\overline{\mathbf{A}}| = |\mathbf{A}|\,|(\overline{\mathbf{A}})^T| = |\mathbf{A}|\,|\mathbf{A}^*| = |\mathbf{A}\mathbf{A}^*| = |\mathbf{I}_n| = 1.$$

(5) (a) $\mathbf{A}$ unitary $\Rightarrow \mathbf{A}^* = \mathbf{A}^{-1} \Rightarrow (\mathbf{A}^*)^{-1} = \mathbf{A} \Rightarrow (\mathbf{A}^*)^{-1} = (\mathbf{A}^*)^* \Rightarrow \mathbf{A}^*$ is unitary.

(b) $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$ (by part (5) of Theorem 7.2) $= \mathbf{B}^{-1}\mathbf{A}^{-1} = (\mathbf{AB})^{-1}$.

(6) (a) $\mathbf{A}^* = \mathbf{A}^{-1}$ iff $(\overline{\mathbf{A}})^T = \mathbf{A}^{-1}$ iff $\overline{(\overline{\mathbf{A}})^T} = \overline{\mathbf{A}^{-1}}$ iff $(\overline{\mathbf{A}})^* = (\overline{\mathbf{A}})^{-1}$ (using the fact that $\overline{\mathbf{A}^{-1}} = (\overline{\mathbf{A}})^{-1}$, since $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n \implies \overline{\mathbf{A}\mathbf{A}^{-1}} = \overline{\mathbf{I}_n} = \mathbf{I}_n \implies (\overline{\mathbf{A}})(\overline{\mathbf{A}^{-1}}) = \mathbf{I}_n \implies \overline{\mathbf{A}^{-1}} = (\overline{\mathbf{A}})^{-1}$).

(b) $(\mathbf{A}^k)^* = (\overline{(\mathbf{A}^k)})^T = ((\overline{\mathbf{A}})^k)^T$ (by part (4) of Theorem 7.2) $= ((\overline{\mathbf{A}})^T)^k = (\mathbf{A}^*)^k = (\mathbf{A}^{-1})^k = (\mathbf{A}^k)^{-1}$.

(c) $\mathbf{A}^2 = \mathbf{I}_n$ iff $\mathbf{A} = \mathbf{A}^{-1}$ iff $\mathbf{A} = \mathbf{A}^*$ (since $\mathbf{A}$ is unitary) iff $\mathbf{A}$ is Hermitian.

(7) (a) $\mathbf{A}$ is unitary iff $\mathbf{A}^* = \mathbf{A}^{-1}$ iff $\overline{\mathbf{A}^T} = \mathbf{A}^{-1}$ iff $(\overline{\mathbf{A}^T})^T = (\mathbf{A}^{-1})^T$ iff $(\mathbf{A}^T)^* = (\mathbf{A}^T)^{-1}$ iff $\mathbf{A}^T$ is unitary.

(b) Follow the hint in the textbook.

(8) Modify the proof of Theorem 6.8.

(9) $\mathbf{Z}^* = \frac{1}{3}\begin{bmatrix} -1-3i & 2-2i & 2 \\ 2-2i & -2i & -2i \\ -2 & 2i & 1-4i \end{bmatrix}$ and $\mathbf{ZZ}^* = \mathbf{Z}^*\mathbf{Z} = \frac{1}{9}\begin{bmatrix} 22 & 8 & 10i \\ 8 & 16 & 2i \\ -10i & -2i & 25 \end{bmatrix}$.

(10) (a) Let $\mathbf{A}$ be the given matrix for L. $\mathbf{A}$ is normal since

$$\mathbf{A}^* = \begin{bmatrix} 1+6i & 2+10i \\ -10+2i & 5 \end{bmatrix} \quad \text{and} \quad \mathbf{AA}^* = \mathbf{A}^*\mathbf{A} = \begin{bmatrix} 141 & 12-12i \\ 12+12i & 129 \end{bmatrix}.$$

Hence $\mathbf{A}$ is unitarily diagonalizable by Theorem 7.9.

(b) $\mathbf{P} = \begin{bmatrix} \frac{-1+i}{\sqrt{6}} & \frac{1-i}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$; $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^*\mathbf{AP} = \begin{bmatrix} 9+6i & 0 \\ 0 & -3-12i \end{bmatrix}$

(11) (a) **A** is normal since $\mathbf{A}^* = \begin{bmatrix} -4-5i & 2-2i & 4-4i \\ 2-2i & -1-8i & -2+2i \\ 4-4i & -2+2i & -4-5i \end{bmatrix}$ and $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A} = 81\mathbf{I}_3$.

Hence **A** is unitarily diagonalizable by Theorem 7.9.

(b) $\mathbf{P} = \frac{1}{3}\begin{bmatrix} -2 & 1 & 2i \\ 1 & -2 & 2i \\ 2 & 2 & i \end{bmatrix}$ and $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^*\mathbf{AP} = \begin{bmatrix} -9 & 0 & 0 \\ 0 & 9i & 0 \\ 0 & 0 & 9i \end{bmatrix}$.

(12) (a) Note that

$$\begin{aligned} \mathbf{Az}\cdot\mathbf{Az} &= \lambda\mathbf{z}\cdot\lambda\mathbf{z} = \overline{\lambda}(\lambda\mathbf{z}\cdot\mathbf{z}) && \text{(by part (5) of Theorem 7.1)} \\ &= \overline{\lambda}\lambda(\mathbf{z}\cdot\mathbf{z}) && \text{(by part (4) of Theorem 7.1),} \end{aligned}$$

while

$$\begin{aligned} \mathbf{Az}\cdot\mathbf{Az} &= \mathbf{z}\cdot(\mathbf{A}^*(\mathbf{Az})) \quad \text{(by Theorem 7.3)} \\ &= \mathbf{z}\cdot((\mathbf{A}^*\mathbf{A})\mathbf{z}) \\ &= \mathbf{z}\cdot((\mathbf{A}^{-1}\mathbf{A})\mathbf{z}) \\ &= \mathbf{z}\cdot\mathbf{z}. \end{aligned}$$

Thus $\overline{\lambda}\lambda(\mathbf{z}\cdot\mathbf{z}) = \mathbf{z}\cdot\mathbf{z}$, and so $\overline{\lambda}\lambda = 1$, which gives $|\lambda|^2 = 1$, and hence $|\lambda| = 1$.

(b) Let **A** be unitary, and let λ be an eigenvalue of **A**. By part (a), $|\lambda| = 1$.
Now, suppose **A** is Hermitian. Then λ is real by Theorem 7.11. Hence $\lambda = \pm 1$.
Conversely, suppose every eigenvalue of **A** is ± 1. Since **A** is unitary, $\mathbf{A}^* = \mathbf{A}^{-1}$, and so $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A} = \mathbf{I}$, which implies that **A** is normal. Thus **A** is unitarily diagonalizable (by Theorem 7.9). Let $\mathbf{A} = \mathbf{PDP}^*$, where **P** is unitary and **D** is diagonal. But since the eigenvalues of **A** are ± 1 (real), $\mathbf{D}^* = \mathbf{D}$. Thus $\mathbf{A}^* = (\mathbf{PDP}^*)^* = \mathbf{PDP}^* = \mathbf{A}$, and **A** is Hermitian.

(13) The eigenvalues are -4, $2+\sqrt{6}$, and $2-\sqrt{6}$.

(14) (a) Since **A** is normal, **A** is unitarily diagonalizable (by Theorem 7.9). Hence $\mathbf{A} = \mathbf{PDP}^*$ for some diagonal matrix **D** and some unitary matrix **P**. Since all eigenvalues of **A** are real, the main diagonal elements of **D** are real, and so $\mathbf{D}^* = \mathbf{D}$. Thus,

$$\mathbf{A}^* = (\mathbf{PDP}^*)^* = \mathbf{PD}^*\mathbf{P}^* = \mathbf{PDP}^* = \mathbf{A},$$

and so **A** is Hermitian.

(b) Since **A** is normal, **A** is unitarily diagonalizable (by Theorem 7.9). Hence, $\mathbf{A} = \mathbf{PDP}^*$ for some diagonal matrix **D** and some unitary matrix **P**. Thus $\mathbf{PP}^* = \mathbf{I}$. Also note that if

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}, \text{ then } \mathbf{D}^* = \begin{bmatrix} \overline{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{\lambda_n} \end{bmatrix},$$

$$\text{and so } \mathbf{DD}^* = \begin{bmatrix} \lambda_1\overline{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\overline{\lambda_n} \end{bmatrix} = \begin{bmatrix} |\lambda_1|^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |\lambda_n|^2 \end{bmatrix}.$$

Since the main diagonal elements of $\mathbf{D}$ are the (not necessarily distinct) eigenvalues of $\mathbf{A}$, and these eigenvalues all have absolute value 1, it follows that $\mathbf{DD}^* = \mathbf{I}$. Then,

$$\begin{aligned}\mathbf{AA}^* &= (\mathbf{PDP}^*)(\mathbf{PDP}^*)^* = (\mathbf{PDP}^*)(\mathbf{PD}^*\mathbf{P}^*) = \mathbf{P}(\mathbf{D}(\mathbf{P}^*\mathbf{P})\mathbf{D}^*)\mathbf{P}^* \\ &= \mathbf{P}(\mathbf{DD}^*)\mathbf{P}^* = \mathbf{PP}^* = \mathbf{I}.\end{aligned}$$

Hence, $\mathbf{A}^* = \mathbf{A}^{-1}$, and $\mathbf{A}$ is unitary.

(c) Because $\mathbf{A}$ is unitary, $\mathbf{A}^{-1} = \mathbf{A}^*$. Hence, $\mathbf{AA}^* = \mathbf{I}_n = \mathbf{A}^*\mathbf{A}$. Therefore, $\mathbf{A}$ is normal.

(15) (a) F (b) T (c) T (d) T (e) F

Section 7.5

(1) (a) Note that $\langle \mathbf{x}, \mathbf{x} \rangle = (\mathbf{Ax}) \cdot (\mathbf{Ax}) = \|\mathbf{Ax}\|^2 \geq 0$.
However, $\|\mathbf{Ax}\|^2 = 0 \Longleftrightarrow \mathbf{Ax} = \mathbf{0} \Longleftrightarrow \mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{0} \Longleftrightarrow \mathbf{x} = \mathbf{0}$.
Also, $\langle \mathbf{x}, \mathbf{y} \rangle = (\mathbf{Ax}) \cdot (\mathbf{Ay}) = (\mathbf{Ay}) \cdot (\mathbf{Ax}) = \langle \mathbf{y}, \mathbf{x} \rangle$, and

$$\begin{aligned}\langle \mathbf{x}+\mathbf{y}, \mathbf{z} \rangle &= (\mathbf{A}(\mathbf{x}+\mathbf{y})) \cdot (\mathbf{Az}) = (\mathbf{Ax}+\mathbf{Ay}) \cdot \mathbf{Az} \\ &= ((\mathbf{Ax}) \cdot (\mathbf{Az})) + ((\mathbf{Ay}) \cdot (\mathbf{Az})) = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle.\end{aligned}$$

Finally, $\langle k\mathbf{x}, \mathbf{y} \rangle = (\mathbf{A}(k\mathbf{x})) \cdot (\mathbf{Ay}) = (k(\mathbf{Ax})) \cdot (\mathbf{Ay}) = k((\mathbf{Ax}) \cdot (\mathbf{Ay})) = k \langle \mathbf{x}, \mathbf{y} \rangle$.

(b) $\langle \mathbf{x}, \mathbf{y} \rangle = -183$, $\|\mathbf{x}\| = \sqrt{314}$

(2) Note that $\langle \mathbf{p}_1, \mathbf{p}_1 \rangle = a_n^2 + \cdots + a_1^2 + a_0^2 \geq 0$.
Clearly, $a_n^2 + \cdots + a_1^2 + a_0^2 = 0$ if and only if each $a_i = 0$.
Also, $\langle \mathbf{p}_1, \mathbf{p}_2 \rangle = a_nb_n + \cdots + a_1b_1 + a_0b_0 = b_na_n + \cdots + b_1a_1 + b_0a_0 = \langle \mathbf{p}_2, \mathbf{p}_1 \rangle$, and

$$\begin{aligned}\langle \mathbf{p}_1 + \mathbf{p}_2, \mathbf{p}_3 \rangle &= \langle (a_n+b_n)x^n + \cdots + (a_0+b_0),\ c_nx^n + \cdots + c_0 \rangle \\ &= (a_n+b_n)c_n + \cdots + (a_0+b_0)c_0 = (a_nc_n + b_nc_n) + \cdots + (a_0c_0 + b_0c_0) \\ &= (a_nc_n + \cdots + a_0c_0) + (b_nc_n + \cdots + b_0c_0) = \langle \mathbf{p}_1, \mathbf{p}_3 \rangle + \langle \mathbf{p}_2, \mathbf{p}_3 \rangle.\end{aligned}$$

Finally, $\langle k\mathbf{p}_1, \mathbf{p}_2 \rangle = (ka_n)b_n + \cdots + (ka_1)b_1 + (ka_0)b_0 = k(a_nb_n + \cdots + a_1b_1 + a_0b_0) = k \langle \mathbf{p}_1, \mathbf{p}_2 \rangle$.

(3) (a) Note that $\langle \mathbf{f}, \mathbf{f} \rangle = \int_a^b (\mathbf{f}(t))^2\, dt \geq 0$. Also, we know from calculus that a nonnegative continuous function on an interval has integral zero if and only if the function is the zero function.
Also, $\langle \mathbf{f}, \mathbf{g} \rangle = \int_a^b \mathbf{f}(t)\mathbf{g}(t)\, dt = \int_a^b \mathbf{g}(t)\mathbf{f}(t)\, dt = \langle \mathbf{g}, \mathbf{f} \rangle$, and

$$\begin{aligned}\langle \mathbf{f}+\mathbf{g}, \mathbf{h} \rangle &= \int_a^b (\mathbf{f}(t)+\mathbf{g}(t))\mathbf{h}(t)\, dt = \int_a^b (\mathbf{f}(t)\mathbf{h}(t) + \mathbf{g}(t)\mathbf{h}(t))\, dt \\ &= \int_a^b \mathbf{f}(t)\mathbf{h}(t)\, dt + \int_a^b \mathbf{g}(t)\mathbf{h}(t)\, dt = \langle \mathbf{f}, \mathbf{h} \rangle + \langle \mathbf{g}, \mathbf{h} \rangle.\end{aligned}$$

Finally, $\langle k\mathbf{f}, \mathbf{g} \rangle = \int_a^b (k\mathbf{f}(t))\mathbf{g}(t)\, dt = k\int_a^b \mathbf{f}(t)\mathbf{g}(t)\, dt = k \langle \mathbf{f}, \mathbf{g} \rangle$.

(b) $\langle \mathbf{f}, \mathbf{g} \rangle = \frac{1}{2}(e^{\pi} + 1)$; $\|\mathbf{f}\| = \sqrt{\frac{1}{2}(e^{2\pi} - 1)}$

(4) Note that $\langle \mathbf{A}, \mathbf{A} \rangle = \text{trace}(\mathbf{A}^T\mathbf{A})$, which by Exercise 28 in Section 1.5 is the sum of the squares of all the entries of $\mathbf{A}$, and hence is nonnegative. Also, this is clearly equal to zero if and only if $\mathbf{A} = \mathbf{O}_n$.
Next, $\langle \mathbf{A}, \mathbf{B} \rangle = \text{trace}(\mathbf{A}^T\mathbf{B}) = \text{trace}((\mathbf{A}^T\mathbf{B})^T)$ (by Exercise 13 in Section 1.4)
$= \text{trace}(\mathbf{B}^T\mathbf{A}) = \langle \mathbf{B}, \mathbf{A} \rangle$.
Also, $\langle \mathbf{A}+\mathbf{B}, \mathbf{C} \rangle = \text{trace}((\mathbf{A}+\mathbf{B})^T\mathbf{C}) = \text{trace}((\mathbf{A}^T+\mathbf{B}^T)\mathbf{C}) = \text{trace}(\mathbf{A}^T\mathbf{C} + \mathbf{B}^T\mathbf{C})$
$= \text{trace}(\mathbf{A}^T\mathbf{C}) + \text{trace}(\mathbf{B}^T\mathbf{C}) = \langle \mathbf{A}, \mathbf{C} \rangle + \langle \mathbf{B}, \mathbf{C} \rangle$.
Finally, $\langle k\mathbf{A}, \mathbf{B} \rangle = \text{trace}((k\mathbf{A})^T\mathbf{B}) = \text{trace}((k\mathbf{A}^T)\mathbf{B}) = \text{trace}(k(\mathbf{A}^T\mathbf{B}))$
$= k(\text{trace}(\mathbf{A}^T\mathbf{B})) = k \langle \mathbf{A}, \mathbf{B} \rangle$.

(5) (a) $\langle \mathbf{0}, \mathbf{x} \rangle = \langle \mathbf{0}+\mathbf{0}, \mathbf{x} \rangle = \langle \mathbf{0}, \mathbf{x} \rangle + \langle \mathbf{0}, \mathbf{x} \rangle$. Since $\langle \mathbf{0}, \mathbf{x} \rangle \in \mathbb{C}$, we know that "$-\langle \mathbf{0}, \mathbf{x} \rangle$" exists. Adding "$-\langle \mathbf{0}, \mathbf{x} \rangle$" to both sides gives $0 = \langle \mathbf{0}, \mathbf{x} \rangle$. Then, $\langle \mathbf{x}, \mathbf{0} \rangle = 0$, by property (3) in the definition of an inner product.

(b) For complex inner product spaces, $\langle \mathbf{x}, k\mathbf{y} \rangle = \overline{\langle k\mathbf{y}, \mathbf{x} \rangle}$ (by property (3) of an inner product space) $= \overline{k \langle \mathbf{y}, \mathbf{x} \rangle}$ (by property (5) of an inner product space) $= \overline{k}(\overline{\langle \mathbf{y}, \mathbf{x} \rangle}) = \overline{k} \langle \mathbf{x}, \mathbf{y} \rangle$ (by property (3) of an inner product space).

A similar proof works in real inner product spaces.

(6) $\|k\mathbf{x}\| = \sqrt{\langle k\mathbf{x}, k\mathbf{x} \rangle} = \sqrt{k \langle \mathbf{x}, k\mathbf{x} \rangle}$ (by property (5) of an inner product space) $= \sqrt{k\overline{k} \langle \mathbf{x}, \mathbf{x} \rangle}$ (by part (3) of Theorem 7.12) $= \sqrt{|k|^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |k|\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = |k|\, \|\mathbf{x}\|$.

(7) (a) We have

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y},\, \mathbf{x}+\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \end{aligned}$$

(by property (3) of an inner product space).

(b) Use part (a) and the fact that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal iff $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

(c) A proof similar to that in part (a) yields $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$.
Add this to the equation for $\|\mathbf{x}+\mathbf{y}\|^2$ in part (a).

(8) (a) Subtract the equation for $\|\mathbf{x}-\mathbf{y}\|^2$ in the answer to Exercise 7(c) from the equation for $\|\mathbf{x}+\mathbf{y}\|^2$ in Exercise 7(a). Then multiply by $\frac{1}{4}$.

(b) In the complex case,

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y},\, \mathbf{x}+\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{y}, \mathbf{y} \rangle. \\ \text{Similarly, } \|\mathbf{x}-\mathbf{y}\|^2 &= \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle - \overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{y}, \mathbf{y} \rangle, \\ \|\mathbf{x}+i\mathbf{y}\|^2 &= \langle \mathbf{x}, \mathbf{x} \rangle - i\langle \mathbf{x}, \mathbf{y} \rangle + i\overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{y}, \mathbf{y} \rangle, \\ \text{and } \|\mathbf{x}-i\mathbf{y}\|^2 &= \langle \mathbf{x}, \mathbf{x} \rangle + i\langle \mathbf{x}, \mathbf{y} \rangle - i\overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{y}, \mathbf{y} \rangle. \\ \text{Then, } \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 &= 2\langle \mathbf{x}, \mathbf{y} \rangle + 2\overline{\langle \mathbf{x}, \mathbf{y} \rangle}. \\ \text{Also, } i(\|\mathbf{x}+i\mathbf{y}\|^2 - \|\mathbf{x}-i\mathbf{y}\|^2) &= i(-2i\langle \mathbf{x}, \mathbf{y} \rangle + 2i\overline{\langle \mathbf{x}, \mathbf{y} \rangle}) \\ &= 2\langle \mathbf{x}, \mathbf{y} \rangle - 2\overline{\langle \mathbf{x}, \mathbf{y} \rangle}. \end{aligned}$$

Adding these and multiplying by $\frac{1}{4}$ produces the desired result.

(9) (a) $\sqrt{\frac{\pi^3}{3} - \frac{3\pi}{2}}$ (b) 0.941 radians, or, 53.9°

(10) (a) $\sqrt{174}$ (b) 0.586 radians, or, 33.6°

(11) (a) If either $\mathbf{x}$ or $\mathbf{y}$ is $\mathbf{0}$, then the theorem is obviously true. We only need to examine the case when both $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ are nonzero. We need to prove $-\|\mathbf{x}\|\,\|\mathbf{y}\| \le |\langle \mathbf{x}, \mathbf{y} \rangle| \le \|\mathbf{x}\|\,\|\mathbf{y}\|$. By property

(5) of an inner product space, and part (3) of Theorem 7.12, this is equivalent to proving

$$\left|\left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{y}}{\|\mathbf{y}\|} \right\rangle\right| \leq 1.$$

Hence, it is enough to show $|\langle \mathbf{a}, \mathbf{b}\rangle| \leq 1$ for any *unit* vectors $\mathbf{a}$ and $\mathbf{b}$.

Let $\langle \mathbf{a}, \mathbf{b}\rangle = z \in \mathbb{C}$. We need to show $|z| \leq 1$. If $z = 0$, we are done. If $z \neq 0$, then let $\mathbf{v} = \frac{|z|}{z}\mathbf{a}$. Then

$$\langle \mathbf{v}, \mathbf{b}\rangle = \left\langle \frac{|z|}{z}\mathbf{a}, \mathbf{b} \right\rangle = \frac{|z|}{z}\langle \mathbf{a}, \mathbf{b}\rangle = \frac{|z|}{z}z = |z| \in \mathbb{R}.$$

Also, since $\left|\frac{|z|}{z}\right| = 1$, $\mathbf{v} = \frac{|z|}{z}\mathbf{a}$ is a unit vector. Now,

$$\begin{aligned} 0 &\leq \|\mathbf{v}-\mathbf{b}\|^2 = \langle \mathbf{v}-\mathbf{b}, \mathbf{v}-\mathbf{b}\rangle \\ &= \|\mathbf{v}\|^2 - \langle \mathbf{v}, \mathbf{b}\rangle - \overline{\langle \mathbf{v}, \mathbf{b}\rangle} + \|\mathbf{b}\|^2 \\ &= 1 - |z| - \overline{|z|} + 1 \\ &= 2 - 2|z|. \end{aligned}$$

Hence $-2 \leq -2|z|$, or $1 \geq |z|$, completing the proof.

(b) Note that

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle \\ &= \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle \\ &= \|\mathbf{x}\|^2 + \langle \mathbf{x}, \mathbf{y}\rangle + \overline{\langle \mathbf{x}, \mathbf{y}\rangle} + \|\mathbf{y}\|^2 \\ &\qquad \text{(by property (3) of an inner product space)} \\ &= \|\mathbf{x}\|^2 + 2(\text{real part of } \langle \mathbf{x}, \mathbf{y}\rangle) + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2|\langle \mathbf{x}, \mathbf{y}\rangle| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &\qquad \text{(by the Cauchy-Schwarz Inequality)} \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

(12) The given inequality follows directly from the Cauchy-Schwarz Inequality.

(13) (i) $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \|(-1)(\mathbf{x}-\mathbf{y})\|$ (by Theorem 7.13) $= \|\mathbf{y}-\mathbf{x}\| = d(\mathbf{y}, \mathbf{x})$.

(ii) $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{\langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle} \geq 0$.
Also, $d(\mathbf{x}, \mathbf{y}) = 0$ iff $\|\mathbf{x}-\mathbf{y}\| = 0$ iff $\langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle = 0$ iff $\mathbf{x}-\mathbf{y} = \mathbf{0}$ (by property (2) of an inner product space) iff $\mathbf{x} = \mathbf{y}$.

(iii) $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \|(\mathbf{x}-\mathbf{z}) + (\mathbf{z}-\mathbf{y})\| \leq \|\mathbf{x}-\mathbf{z}\| + \|\mathbf{z}-\mathbf{y}\|$ (by the Triangle Inequality, part (2) of Theorem 7.14) $= d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

(14) (a) Orthogonal (b) Orthogonal (c) Not orthogonal (d) Orthogonal

(15) Suppose $\mathbf{x} \in T$, where $\mathbf{x}$ is a linear combination of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subseteq T$. Then $\mathbf{x} = a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n$. Consider

$$\begin{aligned} \langle \mathbf{x}, \mathbf{x}\rangle &= \langle \mathbf{x}, a_1\mathbf{x}_1\rangle + \cdots + \langle \mathbf{x}, a_n\mathbf{x}_n\rangle && \text{(by part (2) of Theorem 7.12)} \\ &= \overline{a_1}\langle \mathbf{x}, \mathbf{x}_1\rangle + \cdots + \overline{a_n}\langle \mathbf{x}, \mathbf{x}_n\rangle && \text{(by part (3) of Theorem 7.12)} \\ &= \overline{a_1}(0) + \cdots + \overline{a_n}(0) && \text{(since vectors in } T \text{ are orthogonal)} \\ &= 0, \end{aligned}$$

a contradiction, since $\mathbf{x}$ is assumed to be nonzero.

(16) (a) We have

$$\begin{aligned}\int_{-\pi}^{\pi} \cos mt\,dt &= \frac{1}{m}(\sin mt)|_{-\pi}^{\pi} = \frac{1}{m}(\sin m\pi - \sin(-m\pi)) \\ &= \frac{1}{m}(0) = 0 \quad \text{(since the sine of an integral multiple of } \pi \text{ is 0).}\end{aligned}$$

Similarly,

$$\begin{aligned}\int_{-\pi}^{\pi} \sin mt\,dt &= -\frac{1}{m}(\cos mt)|_{-\pi}^{\pi} = -\frac{1}{m}(\cos m\pi - \cos(-m\pi)) \\ &= -\frac{1}{m}(\cos m\pi - \cos m\pi) \quad \text{(since } \cos(-x) = \cos x) \\ &= -\frac{1}{m}(0) = 0.\end{aligned}$$

(b) Use the trigonometric identities

$$\begin{aligned}\cos mt \cos nt &= \frac{1}{2}(\cos(m-n)t + \cos(m+n)t) \\ \text{and} \quad \sin mt \sin nt &= \frac{1}{2}(\cos(m-n)t - \cos(m+n)t).\end{aligned}$$

Then,

$$\int_{-\pi}^{\pi} \cos mt \cos nt\,dt = \frac{1}{2}\int_{-\pi}^{\pi} \cos(m-n)t\,dt + \frac{1}{2}\int_{-\pi}^{\pi} \cos(m+n)t\,dt = \frac{1}{2}(0) + \frac{1}{2}(0) = 0,$$

by part (a), since $m \pm n$ is an integer. Also,

$$\begin{aligned}\int_{-\pi}^{\pi} \sin mt \sin nt\,dt &= \frac{1}{2}\int_{-\pi}^{\pi} \cos(m-n)t\,dt - \frac{1}{2}\int_{-\pi}^{\pi} \cos(m+n)t\,dt \\ &= \frac{1}{2}(0) - \frac{1}{2}(0) = 0, \quad \text{by part (a).}\end{aligned}$$

(c) Use the trigonometric identity $\sin nt \cos mt = \frac{1}{2}(\sin(n+m)t + \sin(n-m)t)$. Then,

$$\begin{aligned}\int_{-\pi}^{\pi} \cos mt \sin nt\,dt &= \frac{1}{2}\int_{-\pi}^{\pi} \sin(n+m)t\,dt + \frac{1}{2}\int_{-\pi}^{\pi} \sin(n-m)t\,dt \\ &= \frac{1}{2}(0) + \frac{1}{2}(0) = 0,\end{aligned}$$

by part (a), since $n \pm m$ is an integer.

(d) Obvious from parts (a), (b), and (c) with the real inner product of Example 5 (or Example 8) with $a = -\pi$, $b = \pi$: $\langle \mathbf{f}, \mathbf{g} \rangle = \int_{-\pi}^{\pi} f(t)g(t)\,dt$.

(17) Let $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ be an orthogonal ordered basis for a subspace $\mathcal{W}$ of $\mathcal{V}$, and let $\mathbf{v} \in \mathcal{W}$. Let $[\mathbf{v}]_B = [a_1, \ldots, a_k]$. The goal is to show that $a_i = \langle \mathbf{v}, \mathbf{v}_i \rangle / \|\mathbf{v}_i\|^2$. Now, $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. Hence,

$$\begin{aligned}
\langle \mathbf{v}, \mathbf{v}_i \rangle &= \langle a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k, \mathbf{v}_i \rangle \\
&= a_1 \langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + a_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle + \cdots + a_k \langle \mathbf{v}_k, \mathbf{v}_i \rangle \\
&\qquad \text{(by properties (4) and (5) of an inner product)} \\
&= a_1(0) + \cdots + a_{i-1}(0) + a_i\|\mathbf{v}_i\|^2 + a_{i+1}(0) + \cdots + a_k(0) \\
&\qquad \text{(since } B \text{ is orthogonal)} \\
&= a_i\|\mathbf{v}_i\|^2.
\end{aligned}$$

Hence, $a_i = \langle \mathbf{v}, \mathbf{v}_i \rangle / \|\mathbf{v}_i\|^2$. Also, if B is orthonormal, then each $\|\mathbf{v}_i\| = 1$, so $a_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$.

(18) Let $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$ and $\mathbf{w} = b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k$, where $a_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$ and $b_i = \langle \mathbf{w}, \mathbf{v}_i \rangle$ (by Theorem 7.16). Then

$$\begin{aligned}
\langle \mathbf{v}, \mathbf{w} \rangle &= \langle a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k,\ b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k \rangle \\
&= a_1\overline{b_1} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + a_2\overline{b_2} \langle \mathbf{v}_2, \mathbf{v}_2 \rangle + \cdots + a_k\overline{b_k} \langle \mathbf{v}_k, \mathbf{v}_k \rangle \\
&\qquad \text{(by property (5) of an inner product and part (3) of Theorem 7.12,} \\
&\qquad \text{and since } \langle \mathbf{v}_i, \mathbf{v}_j \rangle \text{ equals 0 when } i \text{ and } j \text{ are distinct)} \\
&= a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_k\overline{b_k} \qquad \text{(since } \|\mathbf{v}_i\| = 1, \text{ for } 1 \le i \le k) \\
&= \langle \mathbf{v}, \mathbf{v}_1 \rangle \overline{\langle \mathbf{w}, \mathbf{v}_1 \rangle} + \langle \mathbf{v}, \mathbf{v}_2 \rangle \overline{\langle \mathbf{w}, \mathbf{v}_2 \rangle} + \cdots + \langle \mathbf{v}, \mathbf{v}_k \rangle \overline{\langle \mathbf{w}, \mathbf{v}_k \rangle}.
\end{aligned}$$

(19) Using $\mathbf{w}_1 = t^2 - t + 1$, $\mathbf{w}_2 = 1$, and $\mathbf{w}_3 = t$ yields the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, with $\mathbf{v}_1 = t^2 - t + 1$, $\mathbf{v}_2 = -20t^2 + 20t + 13$, and $\mathbf{v}_3 = 15t^2 + 4t - 5$.

(20) $\{[-9, -4, 8], [27, 11, -22], [5, 3, -4]\}$

(21) The proof is totally analogous to the (long) proof of Theorem 6.4 in Section 6.1 of the textbook. Use Theorem 7.15 in the proof in place of Theorem 6.1.

(22) (a) Follow the proof of Theorem 6.11, using properties (4) and (5) of an inner product. In the last step, use the fact that $\langle \mathbf{w}, \mathbf{w} \rangle = 0 \implies \mathbf{w} = \mathbf{0}$ (by property (2) of an inner product).

(b) Prove part (4) of Theorem 7.19 first, using a proof similar to the proof of Theorem 6.12. (In that proof, use Theorem 7.16 in place of Theorem 6.3.) Then prove part (5) of Theorem 7.19 as follows:

Let $\mathcal{W}$ be a subspace of $\mathcal{V}$ of dimension k. By Theorem 7.17, $\mathcal{W}$ has an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Expand this basis to an orthogonal basis for all of $\mathcal{V}$. (That is, first expand to any basis for $\mathcal{V}$ by Theorem 4.15, then use the Gram-Schmidt Process. Since the first k vectors are already orthogonal, this expands $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ to an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for $\mathcal{V}$.) Then, by part (4) of Theorem 7.19, $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ is a basis for $\mathcal{W}^\perp$, and so $\dim(\mathcal{W}^\perp) = n - k$. Hence $\dim(\mathcal{W}) + \dim(\mathcal{W}^\perp) = n$.

(c) Since any vector in $\mathcal{W}$ is orthogonal to all vectors in $\mathcal{W}^\perp$, $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$.

(d) By part (3) of Theorem 7.19, $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$. Next, by part (5) of Theorem 7.19,

$$\dim(\mathcal{W}) = n - \dim(\mathcal{W}^\perp) = n - (n - \dim((\mathcal{W}^\perp)^\perp)) = \dim((\mathcal{W}^\perp)^\perp).$$

Thus, by Theorem 4.13, or its complex analog, $\mathcal{W} = (\mathcal{W}^\perp)^\perp$.

(23) $\mathcal{W}^{\perp} = \text{span}(\{t^3 - t^2,\, t+1\})$

(24) Using 1 and t as the vectors in the Gram-Schmidt Process, the basis obtained for $\mathcal{W}^{\perp}$ is $\{-5t^2 + 10t - 2, 15t^2 - 14t + 2\}$.

(25) As in the hint, let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be an orthonormal basis for $\mathcal{W}$. Now if $\mathbf{v} \in \mathcal{V}$, let

$$\mathbf{w}_1 = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{v}, \mathbf{v}_k \rangle \mathbf{v}_k$$

and $\mathbf{w}_2 = \mathbf{v} - \mathbf{w}_1$. Then, $\mathbf{w}_1 \in \mathcal{W}$ because $\mathbf{w}_1$ is a linear combination of basis vectors for $\mathcal{W}$. We claim that $\mathbf{w}_2 \in \mathcal{W}^{\perp}$. To see this, let $\mathbf{u} \in \mathcal{W}$. Then $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$ for some $a_1, \ldots, a_k$. Then

$$\begin{aligned} \langle \mathbf{u}, \mathbf{w}_2 \rangle &= \langle \mathbf{u}, \mathbf{v} - \mathbf{w}_1 \rangle = \langle a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k,\ \mathbf{v} - (\langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{v}, \mathbf{v}_k \rangle \mathbf{v}_k) \rangle \\ &= \langle a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k,\ \mathbf{v} \rangle - \langle a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k,\ \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{v}, \mathbf{v}_k \rangle \mathbf{v}_k \rangle \\ &= \sum_{i=1}^{k} a_i \langle \mathbf{v}_i, \mathbf{v} \rangle - \sum_{i=1}^{k}\sum_{j=1}^{k} a_i \overline{\langle \mathbf{v}, \mathbf{v}_j \rangle} \langle \mathbf{v}_i, \mathbf{v}_j \rangle . \end{aligned}$$

But $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ when $i \neq j$ and $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 1$ when $i = j$, since $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthonormal set. Hence,

$$\begin{aligned} \langle \mathbf{u}, \mathbf{w}_2 \rangle &= \sum_{i=1}^{k} a_i \langle \mathbf{v}_i, \mathbf{v} \rangle - \sum_{i=1}^{k} a_i \overline{\langle \mathbf{v}, \mathbf{v}_i \rangle} \\ &= \sum_{i=1}^{k} a_i \langle \mathbf{v}_i, \mathbf{v} \rangle - \sum_{i=1}^{k} a_i \langle \mathbf{v}_i, \mathbf{v} \rangle = 0. \end{aligned}$$

Since this is true for every $\mathbf{u} \in \mathcal{W}$, we conclude that $\mathbf{w}_2 \in \mathcal{W}^{\perp}$.

Finally, we want to show uniqueness of decomposition. Suppose that $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$ and $\mathbf{v} = \mathbf{w}_1' + \mathbf{w}_2'$, where $\mathbf{w}_1,\ \mathbf{w}_1' \in \mathcal{W}$ and $\mathbf{w}_2,\ \mathbf{w}_2' \in \mathcal{W}^{\perp}$. We want to show that $\mathbf{w}_1 = \mathbf{w}_1'$ and $\mathbf{w}_2 = \mathbf{w}_2'$. Now, $\mathbf{w}_1 - \mathbf{w}_1' = \mathbf{w}_2' - \mathbf{w}_2$. Also, $\mathbf{w}_1 - \mathbf{w}_1' \in \mathcal{W}$, but $\mathbf{w}_2' - \mathbf{w}_2 \in \mathcal{W}^{\perp}$. Thus, $\mathbf{w}_1 - \mathbf{w}_1' = \mathbf{w}_2' - \mathbf{w}_2 \in \mathcal{W} \cap \mathcal{W}^{\perp}$. By part (2) of Theorem 7.19, $\mathbf{w}_1 - \mathbf{w}_1' = \mathbf{w}_2' - \mathbf{w}_2 = \mathbf{0}$. Hence, $\mathbf{w}_1 = \mathbf{w}_1'$ and $\mathbf{w}_2 = \mathbf{w}_2'$.

(26) $\mathbf{w}_1 = \frac{1}{2\pi}(\sin t - \cos t)$, $\mathbf{w}_2 = \frac{1}{k}e^t - \frac{1}{2\pi}\sin t + \frac{1}{2\pi}\cos t$

(27) Orthonormal basis for $\mathcal{W} = \left\{\sqrt{\frac{15}{14}}(2t^2 - 1), \sqrt{\frac{3}{322}}(10t^2 + 7t + 2)\right\}$. Then $\mathbf{v} = 4t^2 - t + 3 = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 = \frac{1}{23}(32t^2 + 73t + 57) \in \mathcal{W}$ and $\mathbf{w}_2 = \frac{12}{23}(5t^2 - 8t + 1) \in \mathcal{W}^{\perp}$.

(28) Let $\mathcal{W} = \text{span}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$. Notice that $\mathbf{w}_1 = \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i$ (by Theorem 7.16, since $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthonormal basis for $\mathcal{W}$ by Theorem 7.15). Then, using the hint in the text yields

$$\begin{aligned} \|\mathbf{v}\|^2 &= \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_1 + \mathbf{w}_2 \rangle \\ &= \langle \mathbf{w}_1, \mathbf{w}_1 \rangle + \langle \mathbf{w}_2, \mathbf{w}_1 \rangle + \langle \mathbf{w}_1, \mathbf{w}_2 \rangle + \langle \mathbf{w}_2, \mathbf{w}_2 \rangle \\ &= \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \qquad \text{(since } \mathbf{w}_1 \in \mathcal{W} \text{ and } \mathbf{w}_2 \in \mathcal{W}^{\perp}). \end{aligned}$$

Hence,

$$\begin{aligned}
\|\mathbf{v}\|^2 &\geq \|\mathbf{w}_1\|^2 = \left\langle \left(\sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i\right), \left(\sum_{j=1}^{k} \langle \mathbf{v}, \mathbf{v}_j \rangle \mathbf{v}_j\right) \right\rangle \\
&= \sum_{i=1}^{k} \left\langle \left(\langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i\right), \left(\sum_{j=1}^{k} \langle \mathbf{v}, \mathbf{v}_j \rangle \mathbf{v}_j\right) \right\rangle \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k} \left\langle \left(\langle \mathbf{v}, \mathbf{v}_i \rangle \mathbf{v}_i\right), \left(\langle \mathbf{v}, \mathbf{v}_j \rangle \mathbf{v}_j\right) \right\rangle \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k} \left(\left(\langle \mathbf{v}, \mathbf{v}_i \rangle \ \langle \mathbf{v}, \mathbf{v}_j \rangle\right) \ \langle \mathbf{v}_i, \mathbf{v}_j \rangle\right) \\
&= \sum_{i=1}^{k} \left(\langle \mathbf{v}, \mathbf{v}_i \rangle \ \langle \mathbf{v}, \mathbf{v}_i \rangle\right) \qquad \text{(because } \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \text{ is an orthonormal set)} \\
&= \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{v}_i \rangle^2 .
\end{aligned}$$

(29) (a) Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be an orthonormal basis for $\mathcal{W}$. Let $\mathbf{x}, \mathbf{y} \in \mathcal{V}$. Then

$$\begin{aligned}
\mathbf{proj}_{\mathcal{W}}\mathbf{x} &= \langle \mathbf{x}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{x}, \mathbf{v}_k \rangle \mathbf{v}_k, \\
\mathbf{proj}_{\mathcal{W}}\mathbf{y} &= \langle \mathbf{y}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{y}, \mathbf{v}_k \rangle \mathbf{v}_k, \\
\mathbf{proj}_{\mathcal{W}}(\mathbf{x}+\mathbf{y}) &= \langle \mathbf{x}+\mathbf{y}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle \mathbf{x}+\mathbf{y}, \mathbf{v}_k \rangle \mathbf{v}_k, \quad \text{and} \\
\mathbf{proj}_{\mathcal{W}}(k\mathbf{x}) &= \langle k\mathbf{x}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots + \langle k\mathbf{x}, \mathbf{v}_k \rangle \mathbf{v}_k.
\end{aligned}$$

Clearly, by properties (4) and (5) of an inner product, $\mathbf{proj}_{\mathcal{W}}(\mathbf{x}+\mathbf{y}) = \mathbf{proj}_{\mathcal{W}}\mathbf{x} + \mathbf{proj}_{\mathcal{W}}\mathbf{y}$, and $\mathbf{proj}_{\mathcal{W}}(k\mathbf{x}) = k(\mathbf{proj}_{\mathcal{W}}\mathbf{x})$. Hence L is a linear transformation.

(b) $\ker(L) = \mathcal{W}^{\perp}$, $\text{range}(L) = \mathcal{W}$

(c) First, we establish that if $\mathbf{v} \in \mathcal{W}$, then $L(\mathbf{v}) = \mathbf{v}$. Note that for $\mathbf{v} \in \mathcal{W}$, $\mathbf{v} = \mathbf{v} + \mathbf{0}$, where $\mathbf{v} \in \mathcal{W}$ and $\mathbf{0} \in \mathcal{W}^{\perp}$. But by Theorem 7.20, any decomposition $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$ with $\mathbf{w}_1 \in \mathcal{W}$ and $\mathbf{w}_2 \in \mathcal{W}^{\perp}$ is unique. Hence, $\mathbf{w}_1 = \mathbf{v}$ and $\mathbf{w}_2 = \mathbf{0}$. Then, since $\mathbf{w}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{v}$, we get $L(\mathbf{v}) = \mathbf{v}$.

Finally, let $\mathbf{v} \in \mathcal{V}$ and let $\mathbf{x} = L(\mathbf{v}) = \mathbf{proj}_{\mathcal{W}}\mathbf{v} \in \mathcal{W}$. Then

$$(L \circ L)(\mathbf{v}) = L(L(\mathbf{v})) = L(\mathbf{x}) = \mathbf{proj}_{\mathcal{W}}\mathbf{x} = \mathbf{x} \text{ (since } \mathbf{x} \in \mathcal{W}) = L(\mathbf{v}).$$

Hence, $L \circ L = L$.

(30) (a) F (b) F (c) F (d) T (e) F

Chapter 7 Review Exercises

(1) (a) 0

(b) $(1+2i)(\mathbf{v} \cdot \mathbf{z}) = ((1+2i)\mathbf{v}) \cdot \mathbf{z} = -21 + 43i$, $\mathbf{v} \cdot ((1+2i)\mathbf{z}) = 47 - 9i$

(c) Since $(\mathbf{v} \cdot \mathbf{z}) = (13 + 17i)$,

$$(1 + 2i)(\mathbf{v} \cdot \mathbf{z}) = ((1 + 2i)\mathbf{v}) \cdot \mathbf{z} = (1 + 2i)(13 + 17i) = -21 + 43i,$$

while

$$\mathbf{v} \cdot ((1 + 2i)\mathbf{z}) = \overline{(1 + 2i)}(13 + 17i) = (1 - 2i)(13 + 17i) = 47 - 9i.$$

(d) $\mathbf{w} \cdot \mathbf{z} = -35 + 24i$; $\mathbf{w} \cdot (\mathbf{v} + \mathbf{z}) = -35 + 24i$, since $\mathbf{w} \cdot \mathbf{v} = 0$

(2) (a) $\mathbf{H} = \begin{bmatrix} 2 & 1+3i & 7-i \\ 1-3i & 34 & -10-10i \\ 7+i & -10+10i & 30 \end{bmatrix}$

(b) $\mathbf{A}\mathbf{A}^* = \begin{bmatrix} 32 & -2+14i \\ -2-14i & 34 \end{bmatrix}$, which is clearly Hermitian.

(3) $(\mathbf{A}\mathbf{z}) \cdot \mathbf{w} = \mathbf{z} \cdot (\mathbf{A}^*\mathbf{w})$ (by Theorem 7.3) $= \mathbf{z} \cdot (-\mathbf{A}\mathbf{w})$ (since $\mathbf{A}$ is skew-Hermitian) $= -\mathbf{z} \cdot (\mathbf{A}\mathbf{w})$ (by Theorem 7.1, part (5))

(4) (a) $w = 4 + 3i$, $z = -2i$

(b) $x = 2 + 7i$, $y = 0$, $z = 3 - 2i$

(c) No solutions

(d) $\{[(2 + i) - (3 - i)c,\ (7 + i) - ic,\ c] \mid c \in \mathbb{C}\} = \{[(2 - 3c) + (1 + c)i,\ 7 + (1 - c)i,\ c] \mid c \in \mathbb{C}\}$

(5) By part (3) of Theorem 7.5, $|\mathbf{A}^*| = \overline{|\mathbf{A}|}$. Hence, $|\mathbf{A}^*\mathbf{A}| = |\mathbf{A}^*| \cdot |\mathbf{A}| = \overline{|\mathbf{A}|} \cdot |\mathbf{A}| = \left| |\mathbf{A}| \right|^2$, which is real and nonnegative. This equals zero if and only if $|\mathbf{A}| = 0$, which occurs if and only if $\mathbf{A}$ is singular (by part (4) of Theorem 7.5).

(6) (a) $p_{\mathbf{A}}(x) = x^3 - x^2 + x - 1 = (x^2 + 1)(x - 1) = (x - i)(x + i)(x - 1)$;

$\mathbf{D} = \begin{bmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $\mathbf{P} = \begin{bmatrix} -2-i & -2+i & 0 \\ 1-i & 1+i & 2 \\ 1 & 1 & 1 \end{bmatrix}$

(b) $p_{\mathbf{A}}(x) = x^3 - 2ix^2 - x = x(x - i)^2$; not diagonalizable. Eigenspace for $\lambda = i$ is one-dimensional, with fundamental eigenvector $[1 - 3i,\ -1, 1]$. Fundamental eigenvector for $\lambda = 0$ is $[-i, 1, 1]$.

(7) (a) One possibility: Consider $L\colon \mathbb{C} \to \mathbb{C}$ given by $L(\mathbf{z}) = \overline{\mathbf{z}}$. Note that $L(\mathbf{v}+\mathbf{w}) = \overline{(\mathbf{v} + \mathbf{w})} = \overline{\mathbf{v}} + \overline{\mathbf{w}} = L(\mathbf{v}) + L(\mathbf{w})$. But L is not a linear operator on $\mathbb{C}$ because $L(i) = -i$, but $iL(1) = i(1) = i$, so the rule "$L(c\mathbf{v}) = cL(\mathbf{v})$" is not satisfied.

(b) The example given in part (a) *is* a linear operator on $\mathbb{C}$, thought of as a *real* vector space. In that case we may use only real scalars, and so, if $\mathbf{v} = a + bi$, then $L(c\mathbf{v}) = L(ca + cbi) = ca - cbi = c(a - bi) = cL(\mathbf{v})$.

Note: for any function L from any vector space to itself (real or complex), the rule "$L(\mathbf{v}+\mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$" implies that $L(c\mathbf{v}) = cL(\mathbf{v})$ for any *rational* scalar c. Thus, any example for real vector spaces for which $L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$ is satisfied, but $L(c\mathbf{v}) = cL(\mathbf{v})$ is not satisfied for some c, must involve an *irrational* value of c.

One possibility: Consider $\mathbb{R}$ as a vector space over the rationals (as the scalar field). Choose an uncountably infinite basis B for $\mathbb{R}$ over $\mathbb{Q}$ with $\mathbf{1} \in B$ and $\boldsymbol{\pi} \in B$. Define $L\colon \mathbb{R} \to \mathbb{R}$ by letting $L(\mathbf{1}) = \mathbf{1}$, but $L(\mathbf{v}) = \mathbf{0}$ for every $\mathbf{v} \in B$ with $\mathbf{v} \neq \mathbf{1}$. (Defining L on a basis uniquely determines

L.) This L is a linear operator on $\mathbb{R}$ over $\mathbb{Q}$, and so $L(\mathbf{v}+\mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$ is satisfied. But, by definition $L(\pi) = \mathbf{0}$, but $\pi L(\mathbf{1}) = \pi(\mathbf{1}) = \pi$. Thus, L is not a linear operator on the *real* vector space $\mathbb{R}$.

(8) (a) $B = \{[1, i, 1, -i],\ [4+5i,\ 7-4i,\ i,\ 1],[10,\ -2+6i,\ -8-i,\ -1+8i],\ [0,0,1,i]\}$

(b) $C = \{\frac{1}{2}[1, i, 1, -i], \frac{1}{6\sqrt{3}}[4+5i, 7-4i, i, 1], \frac{1}{3\sqrt{30}}[10, -2+6i, -8-i, -1+8i], \frac{1}{\sqrt{2}}[0,0,1,i]\}$

(c) $$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2}i & \frac{1}{2} & \frac{1}{2}i \\ \frac{1}{6\sqrt{3}}(4-5i) & \frac{1}{6\sqrt{3}}(7+4i) & -\frac{1}{6\sqrt{3}}i & \frac{1}{6\sqrt{3}} \\ \frac{10}{3\sqrt{30}} & \frac{1}{3\sqrt{30}}(-2-6i) & \frac{1}{3\sqrt{30}}(-8+i) & \frac{1}{3\sqrt{30}}(-1-8i) \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}i \end{bmatrix}$$

(9) (a) $p_{\mathbf{A}}(x) = x^2 - 5x + 4 = (x-1)(x-4)$; $\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$; $\mathbf{P} = \begin{bmatrix} \frac{1}{\sqrt{6}}(-1-i) & \frac{1}{\sqrt{3}}(1+i) \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$

(b) $p_{\mathbf{A}}(x) = x^3 + (-98+98i)x^2 - 4802ix = x(x-49+49i)^2$; $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 49-49i & 0 \\ 0 & 0 & 49-49i \end{bmatrix}$;

$\mathbf{P} = \begin{bmatrix} \frac{6}{7} & \frac{1}{\sqrt{5}}i & \frac{-4}{7\sqrt{5}} \\ \frac{3}{7}i & \frac{2}{\sqrt{5}} & \frac{-2i}{7\sqrt{5}} \\ \frac{2}{7} & 0 & \frac{15}{7\sqrt{5}} \end{bmatrix}$ is a probable answer. Another possibility: $\mathbf{P} = \frac{1}{7}\begin{bmatrix} 3i & -2 & -6i \\ 2 & 6i & 3 \\ 6i & 3 & 2i \end{bmatrix}$

(10) Both $\mathbf{AA}^*$ and $\mathbf{A}^*\mathbf{A}$ equal $\begin{bmatrix} 80 & 105+45i & 60 \\ 105-45i & 185 & 60-60i \\ 60 & 60+60i & 95 \end{bmatrix}$, and so $\mathbf{A}$ is normal. Hence, by Theorem 7.9, $\mathbf{A}$ is unitarily diagonalizable.

(11) Now, $\mathbf{A}^*\mathbf{A} =$

$$\begin{bmatrix} 459 & -46+459i & -459+46i & -11-925i & -83+2227i \\ -46-459i & 473 & 92+445i & -918+103i & 2248-139i \\ -459-46i & 92-445i & 473 & -81+932i & 305-2206i \\ -11+925i & -918-103i & -81-932i & 1871 & -4482-218i \\ -83-2227i & 2248+139i & 305+2206i & -4482+218i & 10843 \end{bmatrix},$$

but $\mathbf{AA}^* =$

$$\begin{bmatrix} 7759 & -120+2395i & 3850-1881i & -1230-3862i & -95-2905i \\ -120-2395i & 769 & -648-1165i & -1188+445i & -906+75i \\ 3850+1881i & -648+1165i & 2371 & 329-2220i & 660-1467i \\ -1230+3862i & -1188-445i & 329+2220i & 2127 & 1467+415i \\ -95+2905i & -906-75i & 660+1467i & 1467-415i & 1093 \end{bmatrix},$$

so $\mathbf{A}$ is not normal. Hence, $\mathbf{A}$ is not unitarily diagonalizable, by Theorem 7.9. (Note: $\mathbf{A}$ *is* diagonalizable (with eigenvalues 0, ± 1, and $\pm i$), but the eigenspaces are not orthogonal to each other.)

(12) If $\mathbf{U}$ is a unitary matrix, then $\mathbf{U}^* = \mathbf{U}^{-1}$. Hence, $\mathbf{U}^*\mathbf{U} = \mathbf{U}\mathbf{U}^* = \mathbf{I}$.

(13) Distance $= \sqrt{\frac{8}{105}} \approx 0.276$

(14) $\{[1,0,0],[4,3,0],[5,4,2]\}$

(15) Orthogonal basis for $\mathcal{W}$: $\{\sin x,\ x\cos x + \frac{1}{2}\sin x\}$; $\mathbf{w}_1 = 2\sin x - \frac{36}{4\pi^2+3}(x\cos x + \frac{1}{2}\sin x)$; $\mathbf{w}_2 = x - \mathbf{w}_1$

(16)

(a) T	(d) F	(g) F	(j) T	(m) F	(p) T	(s) T	(v) T
(b) F	(e) F	(h) T	(k) T	(n) T	(q) T	(t) T	
(c) T	(f) T	(i) F	(l) T	(o) T	(r) T	(u) T	(w) F

Chapter 8

Section 8.1

(1) Symmetric: (a), (b), (c), (d)

(a) G_1: $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$
(b) G_2 : $\begin{bmatrix} 1 & 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \end{bmatrix}$
(c) G_3: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(d) G_4: $\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$
(e) D_1: $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$
(f) D_2: $\begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(g) D_3: $\begin{bmatrix} 0 & 2 & 2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 2 & 0 \end{bmatrix}$
(h) D_4: $\begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

(2) All figures for this exercise appear on the next page.
C can be the adjacency matrix for a digraph (only) (see Figure 12).
F can be the adjacency matrix for either a graph or digraph (see Figure 13).
G can be the adjacency matrix for a digraph (only) (see Figure 14).
H can be the adjacency matrix for a digraph (only) (see Figure 15).
I can be the adjacency matrix for a graph or digraph (see Figure 16).
J can be the adjacency matrix for a graph or digraph (see Figure 17).
K can be the adjacency matrix for a graph or digraph (see Figure 18).

(3) The digraph is shown in Figure 19 (see the next page), and the adjacency matrix is

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}.$$

The transpose gives no new information. But it does suggest a different interpretation of the results: namely, the (i, j) entry of the transpose equals 1 if author j influences author i.

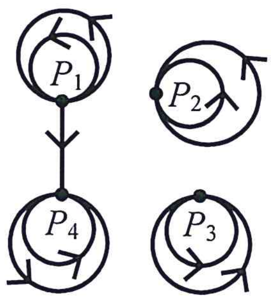


Figure 12: Digraph for **C**

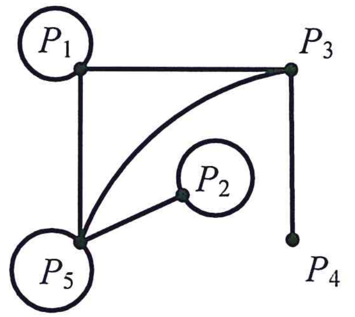


Figure 13: Graph for **F**

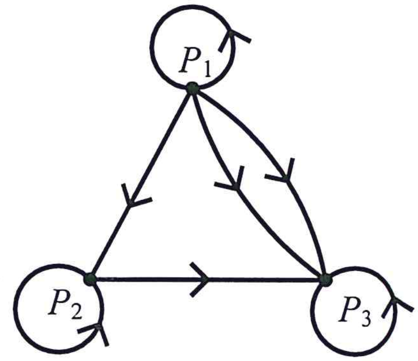


Figure 14: Digraph for **G**

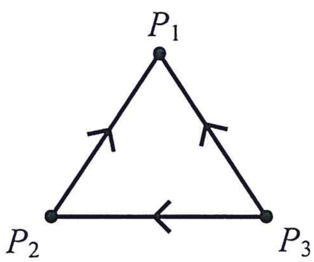


Figure 15: Digraph for **H**

Figure 16: Graph for **I**

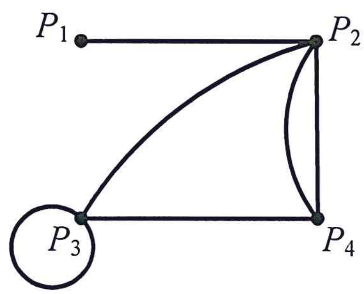


Figure 17: Graph for **J**

Figure 18: Graph for **K**

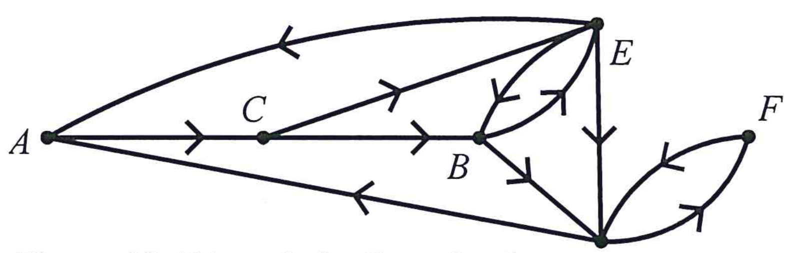


Figure 19: Digraph for Exercise 3

(4) (a) 15 (b) 56 (c) $74 = 1 + 2 + 15 + 56$ (d) $89 = 0 + 4 + 10 + 75$ (e) Length 2 (f) Length 2

(5) (a) 4 (b) 0 (c) $7 = 1 + 2 + 4$ (d) $18 = 2 + 0 + 8 + 8$ (e) No such path exists (f) Length 2

(6) (a) 8 (b) 114 (c) $92 = 0 + 6 + 12 + 74$

(7) (a) 3 (b) 10 (c) $19 = 1 + 2 + 4 + 12$

(8) (a) If the vertex is the ith vertex, then the ith row and ith column entries of the adjacency matrix all equal zero, except possibly for the (i, i) entry.
(b) If the vertex is the ith vertex, then the ith row entries of the adjacency matrix all equal zero, except possibly for the (i, i) entry. (Note: The ith column entries may be nonzero.)

(9) (a) The trace equals the total number of loops in the graph or digraph.
(b) The trace equals the total number of cycles of length k in the graph or digraph. (See Exercise 6 in the textbook for the definition of a cycle.)

(10) Connected: (b), (c); Disconnected: (a), (d)

(11) (a) Since G_2 has the same edges as G_1, along with one new loop at each vertex, G_2 has the same number of edges connecting any two distinct vertices as G_1. Thus, the entries off the main diagonal of the adjacency matrices for the two graphs are the same. But G_2 has one more loop at each vertex than G_1 has. Hence, the entries on the main diagonal of the adjacency matrix for G_2 are all 1 larger than the entries of $\mathbf{A}$, so the adjacency matrix for G_2 is found by adding $\mathbf{I}_n$ to $\mathbf{A}$. This does not change any entries off the main diagonal, but adds 1 to every entry on the main diagonal.
(b) Connectivity only involves the existence of a path between *distinct* vertices. If there is a path from P_i to P_j in G_1, for $i \neq j$, the same path connects these two vertices in G_2. If there is a path from P_i to P_j in G_2, then a similar path can be found from P_i to P_j in G_1 merely by deleting any loops that might appear in the path from P_i to P_j in G_2. Thus, P_i is connected by a path to P_j in G_1 if and only if P_i is connected by a path to P_j in G_2. Hence, G_1 is connected if and only if G_2 is connected.
(c) Suppose there is a path of length m from P_i to P_j in G_1, where $m \leq k$. Then, we can find a path of length k from P_i to P_j in G_2 by first following the known path of length m from P_i to P_j that exists in G_1 (and hence in G_2), and then going $(k - m)$ times around the newly added loop at P_j.
(d) In the discussion in the text for Theorem 8.3, we saw that a graph is connected if and only if every pair of distinct vertices are connected by a path of length $\leq (n-1)$. However, using an argument similar to that in part (c), if there is a path of length less than $(n - 1)$ connecting two vertices in G_2, then there is a path of length equal to $(n - 1)$ connecting those same two vertices. Hence, G_2 is connected if and only if each pair of distinct vertices is connected by a path of length $(n - 1)$. (Note: Because of the newly added loops, there is a path of every length ≥ 1 connecting each vertex in G_2 to itself. Thus, all of the main diagonal entries of any positive power of $(\mathbf{A} + \mathbf{I}_n)$ are nonzero.)
(e) By parts (a) and (d), including the note at the end of the answer for (d) above, and Theorem 8.1, G_2 is connected if and only if $(\mathbf{A} + \mathbf{I}_n)^{n-1}$ has no zero entries. But, by part (b), G_1 is connected if and only if G_2 is connected. Hence, G_1 is connected if and only if $(\mathbf{A} + \mathbf{I}_n)^{n-1}$ has no zero entries.

(f) Part (a): $(\mathbf{A}+\mathbf{I}_4)^3 = \begin{bmatrix} 41 & 0 & 32 & 40 \\ 0 & 8 & 0 & 0 \\ 32 & 0 & 25 & 32 \\ 40 & 0 & 32 & 37 \end{bmatrix}$; disconnected;

Part (b): $(\mathbf{A}+\mathbf{I}_4)^3 = \begin{bmatrix} 13 & 4 & 9 & 1 \\ 4 & 7 & 6 & 5 \\ 9 & 6 & 8 & 3 \\ 1 & 5 & 3 & 4 \end{bmatrix}$; connected;

Part (c): $(\mathbf{A}+\mathbf{I}_5)^4 = \begin{bmatrix} 9 & 12 & 8 & 12 & 6 \\ 12 & 93 & 120 & 28 & 120 \\ 8 & 120 & 252 & 42 & 168 \\ 12 & 28 & 42 & 21 & 24 \\ 6 & 120 & 168 & 24 & 172 \end{bmatrix}$; connected;

Part (d): $(\mathbf{A}+\mathbf{I}_5)^4 = \begin{bmatrix} 86 & 0 & 0 & 85 & 85 \\ 0 & 100 & 78 & 0 & 0 \\ 0 & 78 & 61 & 0 & 0 \\ 85 & 0 & 0 & 86 & 85 \\ 85 & 0 & 0 & 85 & 86 \end{bmatrix}$; disconnected.

(12) If the graph is connected, then vertex P_i must have at least one edge connecting it to some other distinct vertex P_j. (Otherwise there could be no paths at all connecting P_i to any other vertex.) But then, the path $P_i \to P_j \to P_i$ is a path of length 2 connecting P_i to itself. Hence, the (i,i) entry of $\mathbf{A}^2$ is at least 1. Since this is true for each i, all entries on the main diagonal of $\mathbf{A}^2$ are nonzero.

(13) (a) The digraph D_2 in Figure 8.2 is strongly connected since there is a path from any vertex to any other vertex. (This can be verified by inspection, or by using part (b) of this exercise – that is, letting $\mathbf{A}$ represent the adjacency matrix for this digraph and noting that $\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3$ has all nonzero entries off the main diagonal.) The digraph in Figure 8.7 is not strongly connected, since there is no path directed to P_5 from any other vertex.

(b) Use the fact that if a path exists between a given pair P_i, P_j of vertices, then there must be a path of length at most $n-1$ between P_i and P_j. (This is because any longer path would involve returning to a vertex P_k that was already visited earlier in the path since there are only n vertices. Hence, any portion of the longer path that involves travelling from P_k to itself could be removed from that path to yield a shorter path between P_i and P_j.) Then use Corollary 8.2.

(c) Create a new digraph by considering each directed edge of the given digraph and adding a corresponding directed edge going in the opposite direction. If we include loops in this process, the adjacency matrix for the new digraph will be $(\mathbf{A}+\mathbf{A}^T)$. Next, convert the new digraph into a graph by replacing each pair of directed edges (that is, each directed edge together with its opposite) by a single (undirected) edge. The entries of the adjacency matrix of this graph that are not on the main diagonal would agree with those of $(\mathbf{A}+\mathbf{A}^T)$, while the main diagonal entries of the adjacency matrix of this graph would agree with those of $\mathbf{A}$. (These actually represent the number of loops at each vertex in the original digraph.) But loops are irrelevant in determining connectivity since they do not connect distinct vertices. Hence, by Theorem 8.3, this new graph is connected if and only if $(\mathbf{A}+\mathbf{A}^T) + (\mathbf{A}+\mathbf{A}^T)^2 + (\mathbf{A}+\mathbf{A}^T)^3 + \cdots + (\mathbf{A}+\mathbf{A}^T)^{n-1}$ has all entries nonzero off the main diagonal.

(d) For the first adjacency matrix, let $\mathbf{C} = (\mathbf{A} + \mathbf{A}^T)$.Then

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 & 0 \\ 2 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 \end{bmatrix}, \quad \text{and,} \quad \mathbf{C} + \mathbf{C}^2 + \mathbf{C}^3 + \mathbf{C}^4 = \begin{bmatrix} 56 & 0 & 84 & 28 & 0 \\ 0 & 62 & 0 & 0 & 58 \\ 84 & 0 & 126 & 42 & 0 \\ 28 & 0 & 42 & 182 & 0 \\ 0 & 58 & 0 & 0 & 62 \end{bmatrix}.$$

However, some of the entries off the main diagonal of this sum are zero, so the digraph for $\mathbf{A}$ is not weakly connected.

For the second adjacency matrix, let $\mathbf{D} = (\mathbf{B} + \mathbf{B}^T)$. Then

$$\mathbf{D} = \begin{bmatrix} 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 2 \end{bmatrix}, \quad \text{and,} \quad \mathbf{D} + \mathbf{D}^2 + \mathbf{D}^3 + \mathbf{D}^4 = \begin{bmatrix} 80 & 20 & 24 & 20 & 32 \\ 20 & 6 & 4 & 6 & 4 \\ 24 & 4 & 52 & 44 & 32 \\ 20 & 6 & 44 & 46 & 12 \\ 32 & 4 & 32 & 12 & 52 \end{bmatrix}.$$

Since this sum has no zero entries off the main diagonal, the digraph for $\mathbf{B}$ is weakly connected.

(14) (a) Use part (c).

(b) Yes, it is a dominance digraph, because no tie games are possible and because each team plays every other team. Thus if P_i and P_j are two given teams, either P_i defeats P_j or vice versa.

(c) Since the entries of both $\mathbf{A}$ and $\mathbf{A}^T$ are all zeroes and ones, then, with $i \neq j$, the (i, j) entry of $\mathbf{A} + \mathbf{A}^T = a_{ij} + a_{ji} = 1$ iff $a_{ij} = 1$ or $a_{ji} = 1$, but not both. Also, the (i, i) entry of $\mathbf{A} + \mathbf{A}^T = 2a_{ii} = 0$ iff $a_{ii} = 0$. Hence, $\mathbf{A} + \mathbf{A}^T$ has the desired properties (all main diagonal entries equal to 0, and all other entries equal to 1) iff there is always exactly one directed edge between two distinct vertices ($a_{ij} = 1$ or $a_{ji} = 1$, but not both, for $i \neq j$), and the digraph has no loops ($a_{ii} = 0$). (This is the definition of a dominance digraph.)

(15) We prove Theorem 8.1 by induction on k. For the base step, $k = 1$. This case is true by the definition of the adjacency matrix. For the inductive step, let $\mathbf{B} = \mathbf{A}^k$. Then $\mathbf{A}^{k+1} = \mathbf{BA}$. Note that the (i, j) entry of $\mathbf{A}^{k+1}$

$$\begin{aligned} &= (i\text{th row of } \mathbf{B}) \cdot (j\text{th column of } \mathbf{A}) = \sum_{q=1}^{n} b_{iq} a_{qj} \\ &= \sum_{q=1}^{n} (\text{number of paths of length } k \text{ from } P_i \text{ to } P_q) \cdot (\text{number of paths of length 1 from } P_q \text{ to } P_j) \\ &= (\text{number of paths of length } k+1 \text{ from } P_i \text{ to } P_j). \end{aligned}$$

(16) (a) T (c) T (e) T (g) T (i) F (k) F (m) T
(b) F (d) F (f) T (h) F (j) F (l) F

Section 8.2

(1) (a) $I_1 = 8$, $I_2 = 5$, $I_3 = 3$

(b) $I_1 = 10$, $I_2 = 4$, $I_3 = 5$, $I_4 = 1$, $I_5 = 6$

(c) $I_1 = 12$, $I_2 = 5$, $I_3 = 3$, $I_4 = 2$, $I_5 = 2$, $I_6 = 7$

(d) $I_1 = 64$, $I_2 = 48$, $I_3 = 16$, $I_4 = 36$, $I_5 = 12$, $I_6 = 28$

(2) (a) T (b) T

Section 8.3

(1) (a) $y = -0.8x - 3.3$; $y \approx -7.3$ when $x = 5$
(b) $y = 1.12x + 1.05$; $y \approx 6.65$ when $x = 5$
(c) $y = -1.5x + 3.8$; $y \approx -3.7$ when $x = 5$

(2) (a) $y = 0.375x^2 + 0.35x + 3.60$
(b) $y = -0.83x^2 + 1.47x - 1.8$
(c) $y = -0.042x^2 + 0.633x + 0.266$

(3) (a) $y = \frac{1}{4}x^3 + \frac{25}{28}x^2 + \frac{25}{14}x + \frac{37}{35}$ (b) $y = -\frac{1}{6}x^3 - \frac{1}{14}x^2 - \frac{1}{3}x + \frac{117}{35}$

(4) (a) $y = 4.4286x^2 - 2.0571$
(b) $y = 0.7116x^2 + 1.6858x + 0.8989$
(c) $y = -0.1014x^2 + 0.9633x - 0.8534$
(d) $y = -0.1425x^3 + 0.9882x$
(e) $y = 0x^3 - 0.3954x^2 + 0.9706$

(5) (a) $y = 0.2x + 2.74$; the angle reaches 4.34° in the 8th month.
(b) The angle reaches 5.14° in the 12th month.
(c) $y = 0.007143x^2 + 0.1286x + 2.8614$; the tower will be leaning at 5.43° in the 12th month.
(d) The quadratic approximation will probably be more accurate because the amount of change in the angle from the vertical is generally increasing with each passing month.
(e) The angle reaches 3.82° after 5.4 months.
(f) The angle reaches 20° after 40.8 months.

(6) The answers to (a) and (b) assume that the years are renumbered as suggested in the textbook.

(a) $y = 25.81x + 150.97$; predicted population in 2005 is 292.93 million.
(b) Predicted population in 2030 is 357.45 million.
(c) $y = 0.97x^2 + 19x + 160.05$; predicted population in 2030 is 374.13 million.

(7) The least-squares polynomial is $y = \frac{4}{5}x^2 - \frac{2}{5}x + 2$, which is the *exact* quadratic through the given three points.

(8) When $t = 1$, the system in Theorem 8.4 is

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

which becomes $$\begin{bmatrix} n & \sum_{i=1}^n a_i \\ \sum_{i=1}^n a_i & \sum_{i=1}^n a_i^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n b_i \\ \sum_{i=1}^n a_i b_i \end{bmatrix},$$

the desired system.

(9) (a) $x_1 = \frac{230}{39}$, $x_2 = \frac{155}{39}$;

$$\begin{cases} 4x_1 & - & 3x_2 & = & 11\frac{2}{3}, & \text{which is almost } 12 \\ 2x_1 & + & 5x_2 & = & 31\frac{2}{3}, & \text{which is almost } 32 \\ 3x_1 & + & x_2 & = & 21\frac{2}{3}, & \text{which is close to } 21 \end{cases}$$

(b) $x_1 = \frac{46}{11}$, $x_2 = -\frac{12}{11}$, $x_3 = \frac{62}{33}$;

$$\begin{cases} 2x_1 & - & x_2 & + & x_3 & = & 11\frac{1}{3}, & \text{which is almost } 11 \\ -x_1 & + & 3x_2 & - & x_3 & = & -9\frac{1}{3}, & \text{which is almost } -9 \\ x_1 & - & 2x_2 & + & 3x_3 & = & 12 & \\ 3x_1 & - & 4x_2 & + & 2x_3 & = & 20\frac{2}{3}, & \text{which is almost } 21 \end{cases}$$

(10) (a) T (b) F (c) F (d) F

Section 8.4

(1) **A** is not stochastic, since **A** is not square; **A** is not regular, since **A** is not stochastic.
B is not stochastic, since the entries of column 2 do not sum to 1; **B** is not regular, since **B** is not stochastic.
C is stochastic; **C** is regular, since **C** is stochastic and has all nonzero entries.
D is stochastic; **D** is not regular, since every positive power of **D** is a matrix whose rows are the rows of **D** permuted in some order, and hence every such power contains zero entries.
E is not stochastic, since the entries of column 1 do not sum to 1; **E** is not regular, since **E** is not stochastic.
F is stochastic; **F** is not regular, since every positive power of **F** has all second row entries zero.
G is not stochastic, since **G** is not square; **G** is not regular, since **G** is not stochastic.
H is stochastic; **H** is regular, since

$$\mathbf{H}^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix},$$

which has no zero entries.

(2) (a) $\mathbf{p}_1 = [\frac{5}{18}, \frac{13}{18}]$, $\mathbf{p}_2 = [\frac{67}{216}, \frac{149}{216}]$

(b) $\mathbf{p}_1 = [\frac{2}{9}, \frac{13}{36}, \frac{5}{12}]$, $\mathbf{p}_2 = [\frac{25}{108}, \frac{97}{216}, \frac{23}{72}]$

(c) $\mathbf{p}_1 = [\frac{17}{48}, \frac{1}{3}, \frac{5}{16}]$, $\mathbf{p}_2 = [\frac{205}{576}, \frac{49}{144}, \frac{175}{576}]$

(3) (a) $[\frac{2}{5}, \frac{3}{5}]$ (b) $[\frac{18}{59}, \frac{20}{59}, \frac{21}{59}]$ (c) $[\frac{1}{4}, \frac{3}{10}, \frac{3}{10}, \frac{3}{20}]$

(4) (a) $[0.4, 0.6]$ (b) $[0.305, 0.339, 0.356]$

(5) (a) $[0.34, 0.175, 0.34, 0.145]$ in the next election;
$[0.3555, 0.1875, 0.2875, 0.1695]$ in the election after that

(b) The steady-state vector is $[0.36, 0.20, 0.24, 0.20]$; in a century, the votes would be 36% for Party A and 24% for Party C.

(6) (a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{6} & \frac{1}{5} & 0 \\ \frac{1}{8} & \frac{1}{2} & 0 & 0 & \frac{1}{5} \\ \frac{1}{8} & 0 & \frac{1}{2} & \frac{1}{10} & \frac{1}{10} \\ \frac{1}{4} & 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{5} \\ 0 & \frac{1}{3} & \frac{1}{6} & \frac{1}{5} & \frac{1}{2} \end{bmatrix}$

(b) If $\mathbf{M}$ is the stochastic matrix in part (a), then $\mathbf{M}^2 = \begin{bmatrix} \frac{41}{120} & \frac{1}{6} & \frac{1}{5} & \frac{13}{60} & \frac{9}{100} \\ \frac{1}{8} & \frac{27}{80} & \frac{13}{240} & \frac{13}{200} & \frac{1}{5} \\ \frac{3}{20} & \frac{13}{240} & \frac{73}{240} & \frac{29}{200} & \frac{3}{25} \\ \frac{13}{48} & \frac{13}{120} & \frac{29}{120} & \frac{107}{300} & \frac{13}{60} \\ \frac{9}{80} & \frac{1}{3} & \frac{1}{5} & \frac{13}{60} & \frac{28}{75} \end{bmatrix}$, which has all entries nonzero. Thus, $\mathbf{M}$ is regular.

(c) $\frac{29}{120}$, since the probability vector after two time intervals is $[\frac{1}{5}, \frac{13}{240}, \frac{73}{240}, \frac{29}{120}, \frac{1}{5}]$

(d) $[\frac{1}{5}, \frac{3}{20}, \frac{3}{20}, \frac{1}{4}, \frac{1}{4}]$; over time, the rat frequents rooms B and C the least and rooms D and E the most.

(7) The limit vector for any initial input is $[1, 0, 0]$. However, $\mathbf{M}$ is not regular, since every positive power of $\mathbf{M}$ is upper triangular, and hence is zero below the main diagonal. The unique fixed point is $[1, 0, 0]$.

(8) (a) Any steady-state vector is a solution of the system

$$\left(\begin{bmatrix} 1-a & b \\ a & 1-b \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which simplifies to $\begin{bmatrix} -a & b \\ a & -b \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Using the fact that $x_1 + x_2 = 1$ yields a larger system whose augmented matrix is $\left[\begin{array}{cc|c} -a & b & 0 \\ a & -b & 0 \\ 1 & 1 & 1 \end{array}\right]$. This reduces to $\left[\begin{array}{cc|c} 1 & 0 & \frac{b}{a+b} \\ 0 & 1 & \frac{a}{a+b} \\ 0 & 0 & 0 \end{array}\right]$, assuming a and b are not both zero. (Begin the row reduction with the row operation $\langle 1 \rangle \leftrightarrow \langle 3 \rangle$ in case $a = 0$.) Hence, $x_1 = \frac{b}{a+b}$, $x_2 = \frac{a}{a+b}$ is the unique steady-state vector.

(b) Here, $a = \frac{1}{2}$, $b = \frac{1}{3}$, and so the steady-state vector is $\frac{6}{5}[\frac{1}{3}, \frac{1}{2}] = [\frac{2}{5}, \frac{3}{5}]$, which agrees with the result from Exercise 3(a).

(9) It is enough to show that the product of any two stochastic matrices is stochastic. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ stochastic matrices. Then the entries of $\mathbf{AB}$ are clearly nonnegative, since the entries of both $\mathbf{A}$

and **B** are nonnegative. Furthermore, the sum of the entries in the jth column of $\mathbf{AB}$ =

$$\begin{aligned}\sum_{i=1}^{n}(\mathbf{AB})_{ij} &= \sum_{i=1}^{n}(i\text{th row of }\mathbf{A})\cdot(j\text{th column of }\mathbf{B})\\ &= \sum_{i=1}^{n}(a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj})\\ &= \left(\sum_{i=1}^{n}a_{i1}\right)b_{1j}+\left(\sum_{i=1}^{n}a_{i2}\right)b_{2j}+\cdots+\left(\sum_{i=1}^{n}a_{in}\right)b_{nj}\\ &= (1)b_{1j}+(1)b_{2j}+\cdots+(1)b_{nj}\\ &\quad\text{(since }\mathbf{A}\text{ is stochastic, and each of the summations}\\ &\quad\text{is a sum of the entries of a column of }\mathbf{A})\\ &= 1,\end{aligned}$$

since **B** is stochastic. Hence **AB** is stochastic.

(10) Suppose $\mathbf{M}$ is a $k \times k$ matrix.
Base Step ($n = 1$): ith entry in $\mathbf{Mp} = (i\text{th row of }\mathbf{M})\cdot\mathbf{p} = \sum_{j=1}^{k} m_{ij}p_j$

$$\begin{aligned}&= \sum_{j=1}^{k}(\text{probability of moving from state }S_j\text{ to }S_i)(\text{probability of being in state }S_j)\\ &= \text{probability of being in state }S_i\text{ after 1 step of the process}\\ &= i\text{th entry of }\mathbf{p}_1.\end{aligned}$$

Inductive Step: Assume $\mathbf{p}_k = \mathbf{M}^k\mathbf{p}$ is the probability vector after k steps of the process. Then, after an additional step of the process, the probability vector $\mathbf{p}_{k+1} = \mathbf{Mp}_k = \mathbf{M}(\mathbf{M}^k\mathbf{p}) = \mathbf{M}^{k+1}\mathbf{p}$.

(11) (a) F (b) T (c) T (d) T (e) F

Section 8.5

(1) (a)

−24	−46	−15	−30	10	16	39	62
26	42	51	84	24	37	−11	−23

(b)

97	177	146	96	169	143	113
201	171	93	168	133	175	311
254	175	312	256	238	(430)	(357)

(where "27" was used twice to pad the last vector)

(2) (a) HOMEWORK IS GOOD FOR THE SOUL
(b) WHO IS BURIED IN GRANT(')S TOMB–
(c) DOCTOR, I HAVE A CODE
(d) TO MAKE A SLOW HORSE FAST DON(')T FEED IT– –

(3) (a) T (b) T (c) F

Section 8.6

(1) (a) $\theta = \frac{1}{2}\arctan(-\frac{\sqrt{3}}{3}) = -\frac{\pi}{12}$; $\mathbf{P} = \begin{bmatrix} \frac{\sqrt{6}+\sqrt{2}}{4} & \frac{\sqrt{6}-\sqrt{2}}{4} \\ \frac{\sqrt{2}-\sqrt{6}}{4} & \frac{\sqrt{6}+\sqrt{2}}{4} \end{bmatrix}$;

equation in uv-coordinates: $u^2 - v^2 = 2$, or, $\frac{u^2}{2} - \frac{v^2}{2} = 1$;
center in uv-coordinates: $(0, 0)$; center in xy-coordinates: $(0, 0)$;
see Figures 20 and 21.

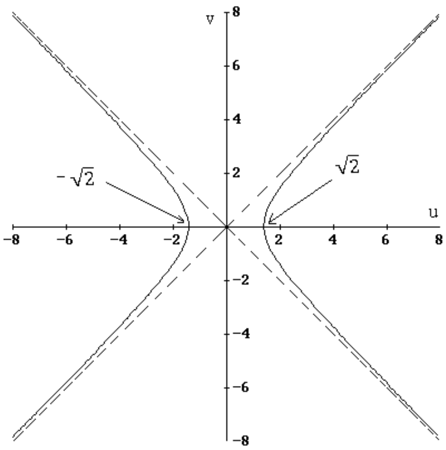


Figure 20: Rotated Hyperbola

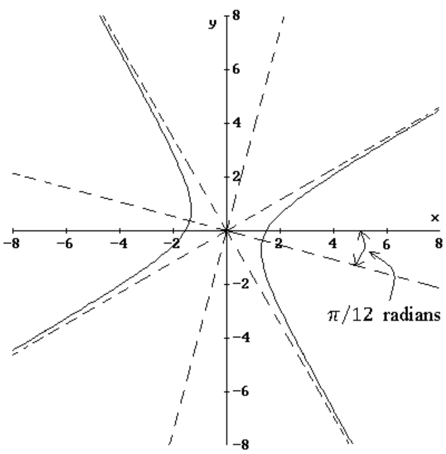


Figure 21: Original Hyperbola

(b) $\theta = \frac{\pi}{4}$, $\mathbf{P} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$;

equation in uv-coordinates: $4u^2 + 9v^2 - 8u = 32$, or, $\frac{(u-1)^2}{9} + \frac{v^2}{4} = 1$;
center in uv-coordinates: $(1, 0)$; center in xy-coordinates: $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$;
see Figures 22 and 23.

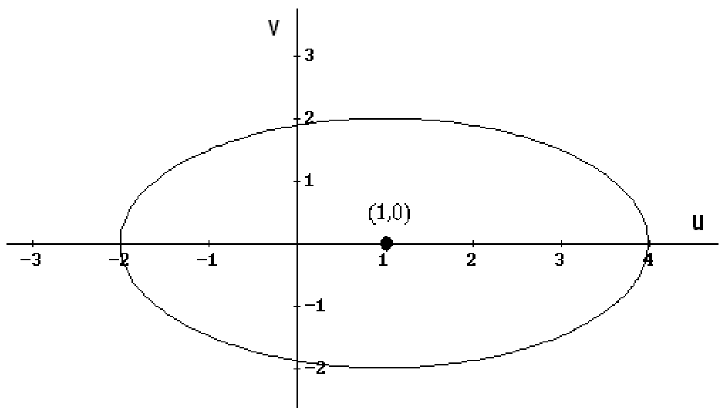


Figure 22: Rotated Ellipse

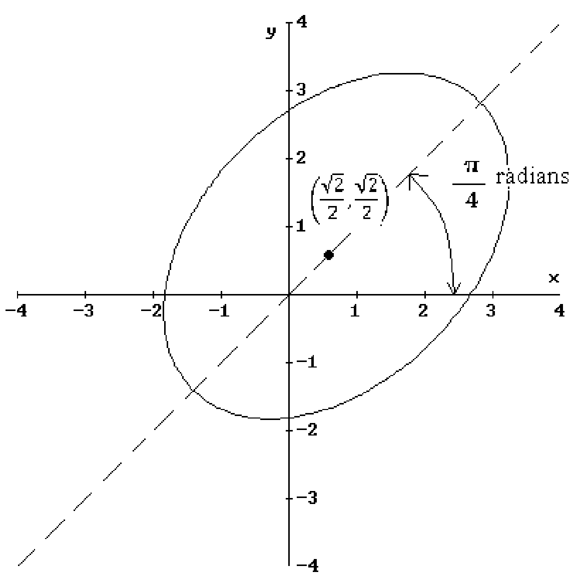


Figure 23: Original Ellipse

(c) $\theta = \frac{1}{2}\arctan(-\sqrt{3}) = -\frac{\pi}{6}$; $\mathbf{P} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$;

equation in uv-coordinates: $v = 2u^2 - 12u + 13$, or $(v + 5) = 2(u - 3)^2$;

vertex in uv-coordinates: $(3, -5)$; vertex in xy-coordinates: $(0.0981, -5.830)$; see Figures 24 and 25.

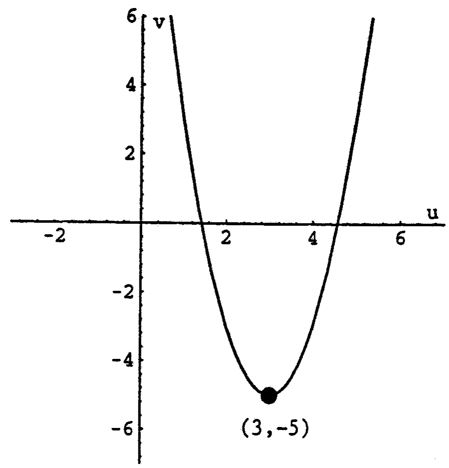


Figure 24: Rotated Parabola

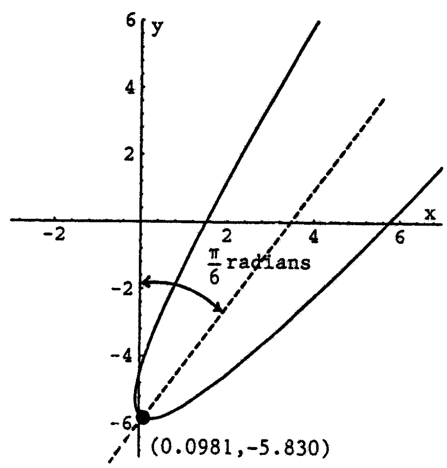


Figure 25: Original Parabola

(d) $\theta \approx 0.6435$ radians (about $36°52'$); $\mathbf{P} = \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}$;

equation in uv-coordinates: $4u^2 - 16u + 9v^2 + 18v = 11$, or, $\frac{(u-2)^2}{9} + \frac{(v+1)^2}{4} = 1$;
center in uv-coordinates: $(2, -1)$; center in xy-coordinates $= (\frac{11}{5}, \frac{2}{5})$;
see Figures 26 and 27.

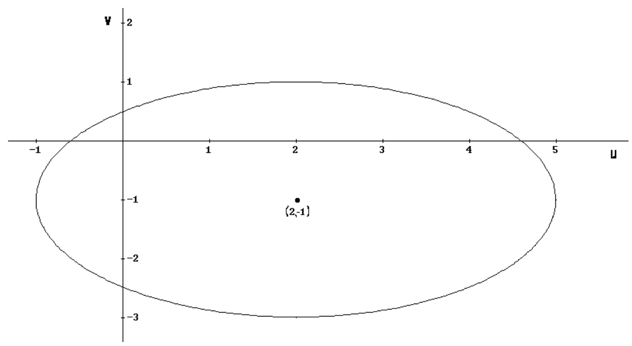


Figure 26: Rotated Ellipse

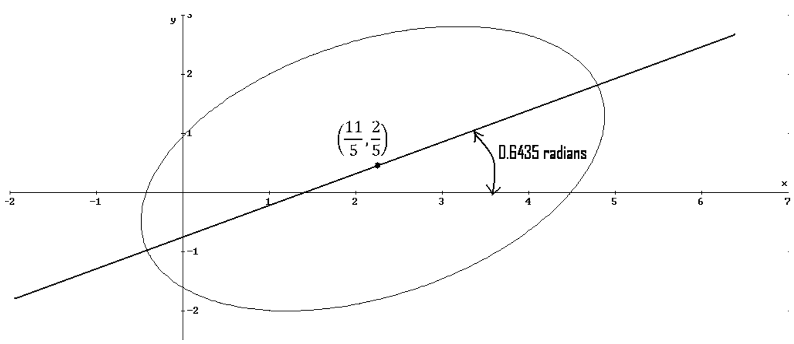


Figure 27: Original Ellipse

(e) $\theta \approx -0.6435$ radians (about $-36°52'$); $\mathbf{P} = \begin{bmatrix} \frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & \frac{4}{5} \end{bmatrix}$;

equation in uv-coordinates: $12v = u^2 + 12u$, or, $(v+3) = \frac{1}{12}(u+6)^2$;
vertex in uv-coordinates: $(-6,-3)$; vertex in xy-coordinates $= (-\frac{33}{5}, \frac{6}{5})$;
see Figures 28 and 29.

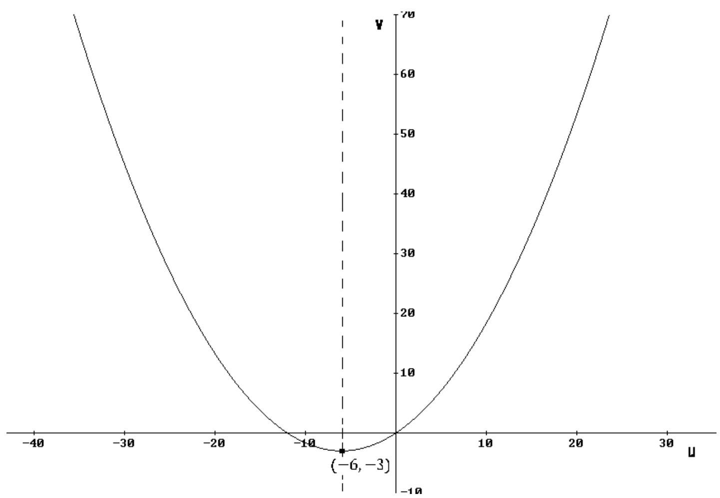


Figure 28: Rotated Parabola

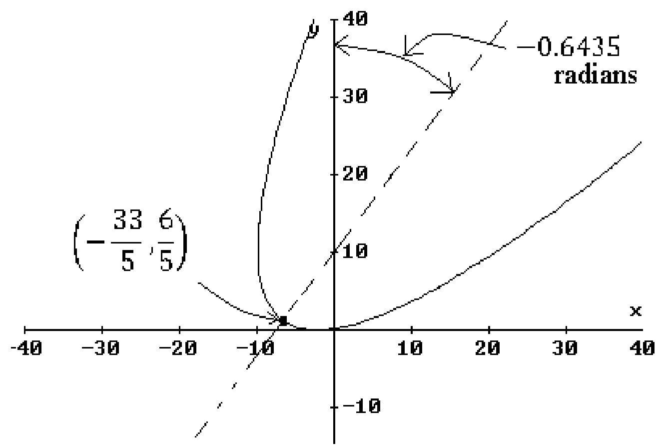


Figure 29: Original Parabola

(f) (All answers are rounded to four significant digits.)
$\theta \approx 0.4442$ radians (about $25°27'$); $\mathbf{P} = \begin{bmatrix} 0.9029 & -0.4298 \\ 0.4298 & 0.9029 \end{bmatrix}$;
equation in uv-coordinates: $\frac{u^2}{(1.770)^2} - \frac{v^2}{(2.050)^2} = 1$;
center in uv-coordinates: $(0,0)$; center in xy-coordinates $= (0,0)$;
see Figures 30 and 31.

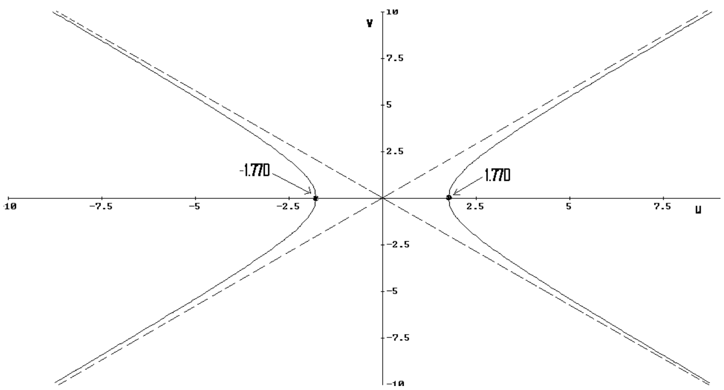


Figure 30: Rotated Hyperbola

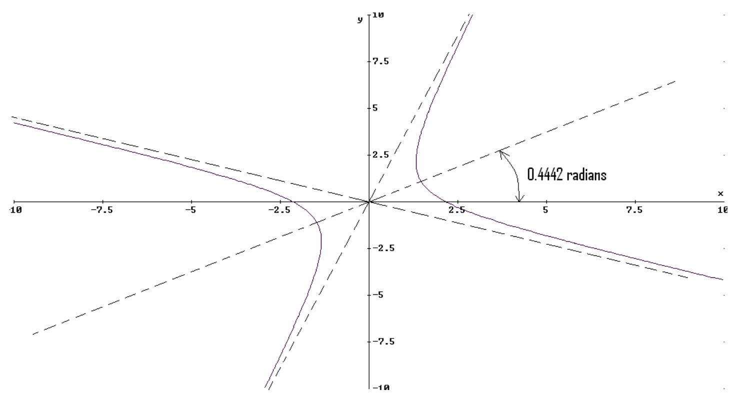


Figure 31: Original Hyperbola

(2) (a) T (b) F (c) T (d) F

Section 8.7

(1) (a) $(9,1), (9,5), (12,1), (12,5), (14,3)$
(b) $(3,5), (1,9), (5,7), (3,10), (6,9)$
(c) $(-2,5), (0,9), (-5,7), (-2,10), (-5,10)$
(d) $(20,6), (20,14), (32,6), (32,14), (40,10)$

(2) (a) $(3, 11), (5, 9), (7, 11), (11, 7), (15, 11)$; see Figure 32

(b) $(-8, 2), (-7, 5), (-10, 6), (-9, 11), (-14, 13)$; see Figure 33

(c) $(8, 1), (8, 4), (11, 4), (10, 10), (16, 11)$; see Figure 34

(d) $(3, 18), (4, 12), (5, 18), (7, 6), (9, 18)$; see Figure 35

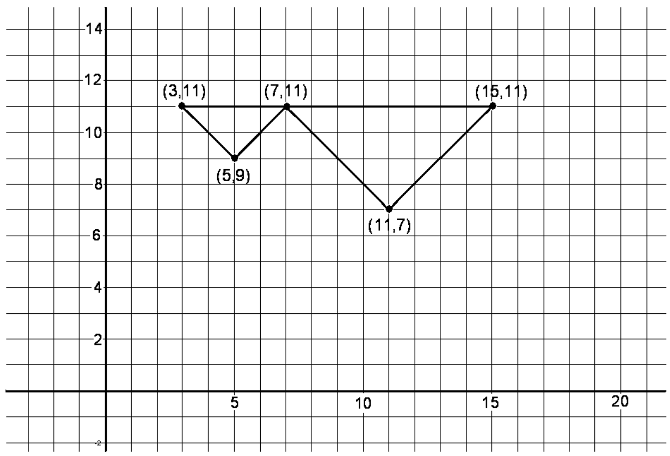


Figure 32

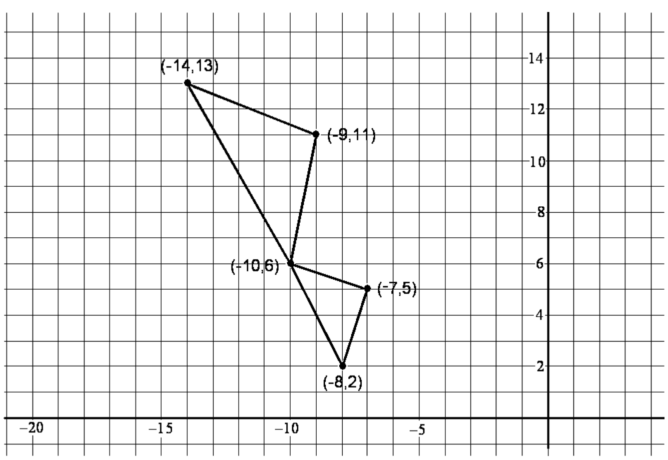


Figure 33

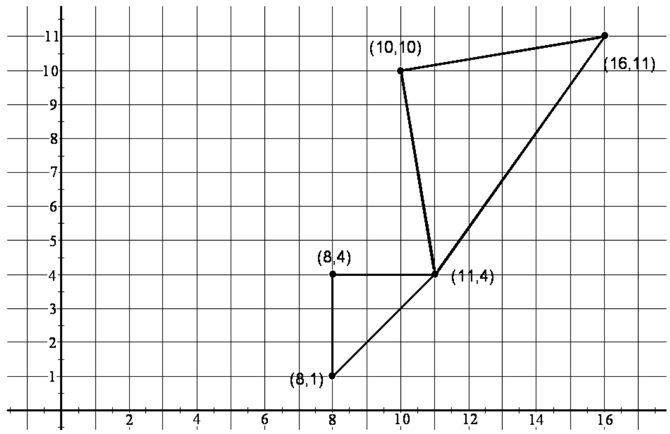


Figure 34

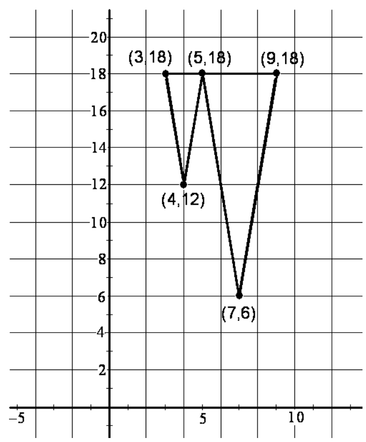


Figure 35

(3) (a) $(3,-4),(3,-10),(7,-6),(9,-9),(10,-3)$
(b) $(4,3),(10,3),(6,7),(9,9),(3,10)$
(c) $(-2,6),(0,8),(-8,17),(-10,22),(-16,25)$
(d) $(6,4),(11,7),(11,2),(14,1),(10,-3)$

(4) (a) $(14,9),(10,6),(11,11),(8,9),(6,8),(11,14)$
(b) $(-2,-3),(-6,-3),(-3,-6),(-6,-6),(-8,-6),(-2,-9)$
(c) $(2,4),(2,6),(8,5),(8,6)$ (twice), $(14,4)$

(5) (a) $(4,7),(6,9),(10,8),(9,2),(11,4)$
(b) $(0,5),(1,7),(0,11),(-5,8),(-4,10)$
(c) $(7,3),(7,12),(9,18),(10,3),(10,12)$

(6) (a) $(2,20),(3,17),(5,14),(6,19),(6,16),(9,14)$; see Figure 36
(b) $(17,17),(17,14),(18,10),(20,15),(19,12),(22,10)$; see Figure 37

(c) $(1, 18), (-3, 13), (-6, 8), (-2, 17), (-5, 12), (-6, 10)$; see Figure 38

(d) $(-19, 6), (-16, 7), (-15, 9), (-27, 7), (-21, 8), (-27, 10)$; see Figure 39

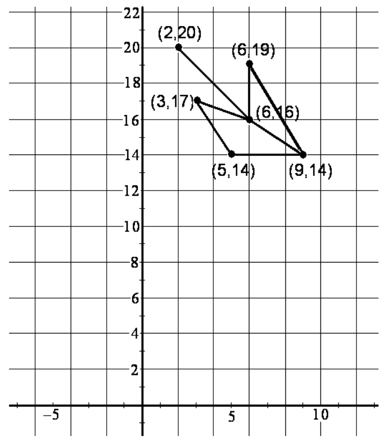


Figure 36

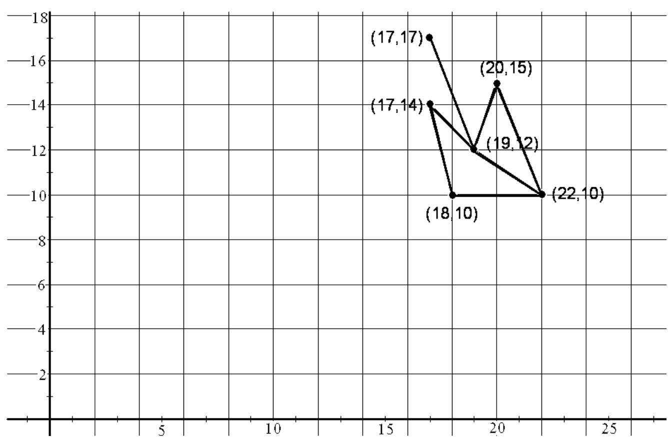


Figure 37

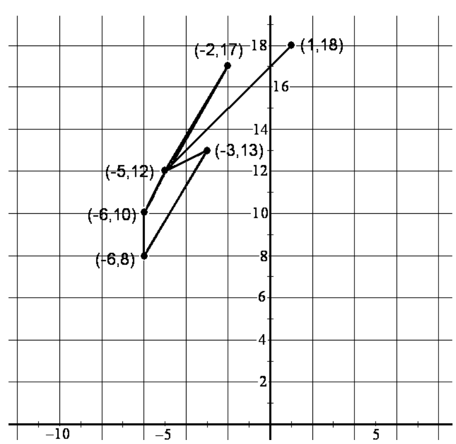


Figure 38

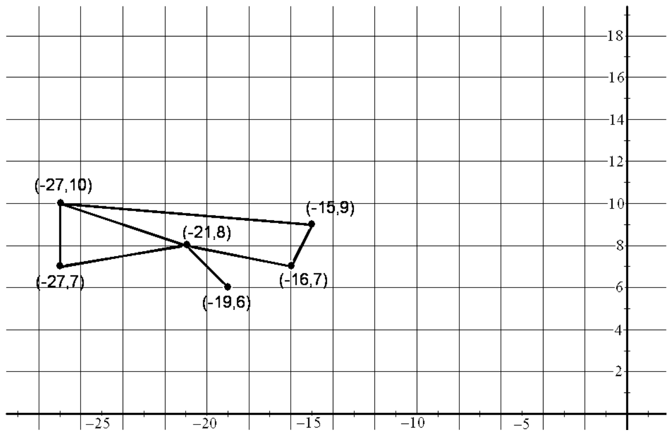


Figure 39

(7) (a) Show that the given matrix is equal to the product

$$\begin{bmatrix} 1 & 0 & r \\ 0 & 1 & s \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -r \\ 0 & 1 & -s \\ 0 & 0 & 1 \end{bmatrix}.$$

(b) Show that the given matrix is equal to the product

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \left(\frac{1}{1+m^2}\right) \begin{bmatrix} 1-m^2 & 2m & 0 \\ 2m & m^2-1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -b \\ 0 & 0 & 1 \end{bmatrix}.$$

(c) Show that the given matrix is equal to the product

$$\begin{bmatrix} 1 & 0 & r \\ 0 & 1 & s \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -r \\ 0 & 1 & -s \\ 0 & 0 & 1 \end{bmatrix}.$$

(8) Show that the given matrix is equal to the product

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

(9) (a) $(4,7), (6,9), (10,8), (9,2), (11,4)$

(b) $(0,5), (1,7), (0,11), (-5,8), (-4,10)$

(c) $(7,3), (7,12), (9,18), (10,3), (10,12)$

(10) (a) Multiplying the two matrices yields $\mathbf{I}_3$.

(b) The first matrix represents a translation along the vector $[a,b]$, while the second is the inverse translation along the vector $[-a,-b]$.

(c) See the *answer* for Exercise 7(a), and substitute a for r, and b for s, and then use part (a) of this Exercise.

(11) (a) Show that the matrices in parts (a) and (c) of Exercise 7 commute with each other when $c = d$. Both products equal

$$\begin{bmatrix} c(\cos\theta) & -c(\sin\theta) & r(1 - c(\cos\theta)) + sc(\sin\theta) \\ c(\sin\theta) & c(\cos\theta) & s(1 - c(\cos\theta)) - rc(\sin\theta) \\ 0 & 0 & 1 \end{bmatrix}.$$

(b) Consider the reflection about the y-axis, whose matrix is given in Section 8.7 of the textbook as $\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and a counterclockwise rotation of 90° about the origin, whose matrix is $\mathbf{B} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then $\mathbf{AB} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, but $\mathbf{BA} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and so $\mathbf{A}$ and $\mathbf{B}$ do not commute. In particular, starting from the point $(1, 0)$, performing the rotation and then the reflection yields $(0, 1)$. However, performing the reflection followed by the rotation produces $(0, -1)$.

(c) Consider a reflection about $y = -x$ and scaling by a factor of c in the x-direction and d in the y-direction, with $c \neq d$. Now,

$$(\text{reflection}) \circ (\text{scaling})([1, 0]) = (\text{reflection})([c, 0]) = [0, -c].$$

But,

$$(\text{scaling}) \circ (\text{reflection})([1, 0]) = (\text{scaling})([0, -1]) = [0, -d].$$

Hence,

$$(\text{reflection}) \circ (\text{scaling}) \neq (\text{scaling}) \circ (\text{reflection}).$$

(12) (a) The matrices are $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ and $\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$. If $\mathbf{A}$ represents either matrix, we can verify $\mathbf{A}$ is orthogonal by directly computing $\mathbf{AA}^T$.

(b) If $\mathbf{A} = \begin{bmatrix} 0 & -1 & 18 \\ 1 & 0 & -6 \\ 0 & 0 & 1 \end{bmatrix}$, then $\mathbf{AA}^T = \begin{bmatrix} 325 & -108 & 18 \\ -108 & 37 & -6 \\ 18 & -6 & 1 \end{bmatrix} \neq \mathbf{I}_3$.

(c) Let m be the slope of a nonvertical line. Then, the matrices are, respectively,

$$\begin{bmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} & 0 \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

For a vertical line, the matrices are $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. All of these matrices are obviously symmetric, and since a reflection is its own inverse, all of these matrices are their own inverses. Hence, if $\mathbf{A}$ is any one of these matrices, then $\mathbf{AA}^T = \mathbf{AA} = \mathbf{I}$, so $\mathbf{A}$ is orthogonal.

(13) (a) F (b) F (c) F (d) T (e) T (f) F

Section 8.8

(1) (a) $b_1e^t\begin{bmatrix}7\\3\end{bmatrix}+b_2e^{-t}\begin{bmatrix}2\\1\end{bmatrix}$ (b) $b_1e^{3t}\begin{bmatrix}1\\1\end{bmatrix}+b_2e^{-2t}\begin{bmatrix}3\\4\end{bmatrix}$

(c) $b_1\begin{bmatrix}0\\-1\\1\end{bmatrix}+b_2e^t\begin{bmatrix}1\\-1\\1\end{bmatrix}+b_3e^{3t}\begin{bmatrix}2\\0\\1\end{bmatrix}$

(d) $b_1e^t\begin{bmatrix}-1\\1\\0\end{bmatrix}+b_2e^t\begin{bmatrix}5\\0\\2\end{bmatrix}+b_3e^{4t}\begin{bmatrix}1\\1\\1\end{bmatrix}$

(There are other possible answers. For example, the first two vectors in the sum could be any basis for the two-dimensional eigenspace corresponding to the eigenvalue 1.)

(e) $b_1e^t\begin{bmatrix}-1\\-1\\1\\0\end{bmatrix}+b_2e^t\begin{bmatrix}1\\3\\0\\1\end{bmatrix}+b_3\begin{bmatrix}2\\3\\0\\1\end{bmatrix}+b_4e^{-3t}\begin{bmatrix}0\\1\\1\\1\end{bmatrix}$

(There are other possible answers. For example, the first two vectors in the sum could be any basis for the two-dimensional eigenspace corresponding to the eigenvalue 1.)

(2) (a) $y=b_1e^{2t}+b_2e^{-3t}$ (b) $y=b_1e^t+b_2e^{-t}+b_3e^{5t}$

(c) $y=b_1e^{2t}+b_2e^{-2t}+b_3e^{\sqrt{2}t}+b_4e^{-\sqrt{2}t}$

(3) Using the fact that $\mathbf{A}\mathbf{v}_i=\lambda_i\mathbf{v}_i$, it can be easily seen that, with $\mathbf{F}(t)=b_1e^{\lambda_1 t}\mathbf{v}_1+\cdots+b_ke^{\lambda_k t}\mathbf{v}_k$, both sides of $\mathbf{F}'(t)=\mathbf{A}\mathbf{F}(t)$ yield $\lambda_1b_1e^{\lambda_1 t}\mathbf{v}_1+\cdots+\lambda_kb_ke^{\lambda_k t}\mathbf{v}_k$.

(4) (a) Suppose $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ is an ordered basis for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$ corresponding to the eigenvalues $\lambda_1,\ldots,\lambda_n$. Then, by Theorem 8.9, all solutions to $\mathbf{F}'(t)=\mathbf{A}\mathbf{F}(t)$ are of the form $\mathbf{F}(t)=b_1e^{\lambda_1 t}\mathbf{v}_1+\cdots+b_ne^{\lambda_n t}\mathbf{v}_n$. Substituting $t=0$ yields $\mathbf{v}=\mathbf{F}(0)=b_1\mathbf{v}_1+\cdots+b_n\mathbf{v}_n$. But since $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ forms a basis for $\mathbb{R}^n$, $b_1,\ldots,b_n$ are uniquely determined (by Theorem 4.9). Thus, $\mathbf{F}(t)$ is uniquely determined.

(b) $\mathbf{F}(t)=2e^{5t}\begin{bmatrix}1\\0\\1\end{bmatrix}-e^t\begin{bmatrix}-1\\2\\0\end{bmatrix}-2e^{-t}\begin{bmatrix}1\\1\\1\end{bmatrix}$.

(5) (a) The equations $f_1(t)=y,\ f_2(t)=y',\ldots,f_n(t)=y^{(n-1)}$ translate into these $(n-1)$ equations: $f_1'(t)=f_2(t),\ f_2'(t)=f_3(t),\ldots,f_{n-1}'(t)=f_n(t)$. Also, $y^{(n)}+a_{n-1}y^{(n-1)}+\cdots+a_1y'+a_0y=0$ translates into $f_n'(t)=-a_0f_1(t)-a_1f_2(t)-\cdots-a_{n-1}f_n(t)$. These n equations taken together are easily seen to be represented by $\mathbf{F}'(t)=\mathbf{A}\mathbf{F}(t)$, where $\mathbf{A}$ and $\mathbf{F}$ are as given.

(b) Base Step ($n=1$): $\mathbf{A}=[-a_0]$, $x\mathbf{I}_1-\mathbf{A}=[x+a_0]$, and so $p_{\mathbf{A}}(x)=x+a_0$.
Inductive Step: Assume true for k. Prove true for $k+1$. For the case $k+1$,

$$\mathbf{A}=\begin{bmatrix}0&1&0&0&\cdots&0\\0&0&1&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&0&\cdots&1\\-a_0&-a_1&-a_2&-a_3&\cdots&-a_k\end{bmatrix}.$$

Then

$$(x\mathbf{I}_{k+1} - \mathbf{A}) = \begin{bmatrix} x & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & x & -1 \\ a_0 & a_1 & a_2 & a_3 & \cdots & a_{k-1} & x + a_k \end{bmatrix}.$$

Using a cofactor expansion on the first column yields

$$\begin{aligned} |x\mathbf{I}_{k+1} - \mathbf{A}| &= x \begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_1 & a_2 & a_3 & \cdots & a_{k-1} & x + a_k \end{vmatrix} \\ &\quad + (-1)^{(k+1)+1} a_0 \begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ x & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x & -1 \end{vmatrix}. \end{aligned}$$

The first determinant equals $|x\mathbf{I}_k - \mathbf{B}|$, where

$$\mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_1 & -a_2 & -a_3 & \cdots & -a_{k-1} & -a_k \end{bmatrix}.$$

Now, using the inductive hypothesis on the first determinant, and Theorems 3.2 and 3.10 on the second determinant (since its corresponding matrix is lower triangular) yields

$$p_{\mathbf{A}}(x) = x(x^k + a_k x^{k-1} + \cdots + a_2 x + a_1) + (-1)^{k+2} a_0 (-1)^k = x^{k+1} + a_k x^k + \cdots + a_1 x + a_0.$$

(6) (a) $[b_2, b_3, \ldots, b_n, -a_0 b_1 - a_1 b_2 - \cdots - a_{n-1} b_n]$

(b) Using Theorem 8.10,

$$0 = p_{\mathbf{A}}(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

Thus, $\lambda^n = -a_0 - a_1\lambda - \cdots - a_{n-1}\lambda^{n-1}$. Letting $\mathbf{x} = [1, \lambda, \lambda^2, \ldots, \lambda^{n-1}]$ and substituting $\mathbf{x}$ for $[b_1, \ldots, b_n]$ in part (a) gives

$$\begin{aligned} \mathbf{Ax} &= [\lambda, \lambda^2, \ldots, \lambda^{n-1}, -a_0 - a_1\lambda - \cdots - a_{n-1}\lambda^{n-1}] \\ &= [\lambda, \lambda^2, \ldots, \lambda^{n-1}, \lambda^n] \;=\; \lambda\mathbf{x}. \end{aligned}$$

(c) Let $\mathbf{v} = [c, v_2, v_3, \ldots, v_n]$. Then $\lambda\mathbf{v} = [c\lambda, v_2\lambda, \ldots, v_n\lambda]$. By part (a),

$$\mathbf{Av} = [v_2, v_3, \ldots, v_n, -a_0 c - \cdots - a_{n-1} v_n].$$

Equating the first $n-1$ coordinates of $\mathbf{Av}$ and $\lambda\mathbf{v}$ yields

$$v_2 = c\lambda, \;\; v_3 = v_2\lambda, \ldots, \;\; v_n = v_{n-1}\lambda.$$

Further recursive substitution gives

$$v_2 = c\lambda,\ v_3 = c\lambda^2, \ldots,\ v_n = c\lambda^{n-1}.$$

Hence $\mathbf{v} = c[1, \lambda, \ldots, \lambda^{n-1}]$.

(d) By parts (b) and (c), $\{[1, \lambda, \ldots, \lambda^{n-1}]\}$ is a basis for E_λ.

(7) (a) T (b) T (c) T (d) F

Section 8.9

(1) (a) Unique least-squares solution: $\mathbf{v} = \left[\frac{23}{30}, \frac{11}{10}\right]$; $||\mathbf{Av} - \mathbf{b}|| = \frac{\sqrt{6}}{6} \approx 0.408$; $||\mathbf{Az} - \mathbf{b}|| = 1$

(b) Unique least-squares solution: $\mathbf{v} = \left[\frac{229}{75}, \frac{1}{3}\right]$; $||\mathbf{Av} - \mathbf{b}|| = \frac{59\sqrt{3}}{15} \approx 6.813$; $||\mathbf{Az} - \mathbf{b}|| = 7$

(c) Infinite number of least-squares solutions, all of the form $\mathbf{v} = \left[7c + \frac{17}{3}, -13c - \frac{23}{3}, c\right]$; two particular least-squares solutions are $\left[\frac{17}{3}, -\frac{23}{3}, 0\right]$ and $\left[8, -12, \frac{1}{3}\right]$; with $\mathbf{v}$ as either of these vectors, $||\mathbf{Av} - \mathbf{b}|| = \frac{\sqrt{6}}{3} \approx 0.816$; $||\mathbf{Az} - \mathbf{b}|| = 3$

(d) Infinite number of least-squares solutions, all of the form $\mathbf{v} = \left[\frac{1}{2}c + \frac{5}{4}d - 1, -\frac{3}{2}c - \frac{11}{4}d + \frac{19}{3}, c, d\right]$; two particular least-squares solutions are $\left[-\frac{1}{2}, \frac{29}{6}, 1, 0\right]$ and $\left[\frac{1}{4}, \frac{43}{12}, 0, 1\right]$; with $\mathbf{v}$ as either of these vectors, $||\mathbf{Av} - \mathbf{b}|| = \frac{\sqrt{6}}{3} \approx 0.816$; $||\mathbf{Az} - \mathbf{b}|| = \sqrt{5} \approx 2.236$

(2) (a) Infinite number of least-squares solutions, all of the form $\mathbf{v} = \left[-\frac{4}{7}c + \frac{19}{42}, \frac{8}{7}c - \frac{5}{21}, c\right]$, with $\frac{5}{24} \leq c \leq \frac{19}{24}$.

(b) Infinite number of least-squares solutions, all of the form $\mathbf{v} = \left[-\frac{6}{5}c + \frac{19}{10}, -\frac{1}{5}c + \frac{19}{35}, c\right]$, with $0 \leq c \leq \frac{19}{12}$.

(3) (a) $\mathbf{v} \approx [1.58, -0.58]$; $(\lambda'\mathbf{I} - \mathbf{C})\mathbf{v} \approx [0.025, -0.015]$ (Actual nearest eigenvalue is $2 + \sqrt{3} \approx 3.732$)

(b) $\mathbf{v} \approx [0.46, -0.36, 0.90]$; $(\lambda'\mathbf{I} - \mathbf{C})\mathbf{v} \approx [0.03, -0.04, 0.07]$ (Actual nearest eigenvalue is $\sqrt{2} \approx 1.414$)

(c) $\mathbf{v} \approx [1.45, -0.05, -0.40]$; $(\lambda'\mathbf{I} - \mathbf{C})\mathbf{v} \approx [-0.09, -0.06, -0.15]$ (Actual nearest eigenvalue is $\sqrt[3]{12} \approx 2.289$)

(4) If $\mathbf{AX} = \mathbf{b}$ is consistent, then $\mathbf{b}$ is in the subspace $\mathcal{W}$ of Theorem 8.13. Thus, $\mathbf{proj}_{\mathcal{W}}\mathbf{b} = \mathbf{b}$. Finally, the fact that statement (3) of Theorem 8.13 is equivalent to statement (1) of Theorem 8.13 shows that the actual solutions are the same as the least-squares solutions in this case.

(5) Let $\mathbf{b} = \mathbf{Av}_1$. Then $\mathbf{A}^T\mathbf{Av}_1 = \mathbf{A}^T\mathbf{b}$. Also, $\mathbf{A}^T\mathbf{Av}_2 = \mathbf{A}^T\mathbf{Av}_1 = \mathbf{A}^T\mathbf{b}$. Hence $\mathbf{v}_1$ and $\mathbf{v}_2$ both satisfy part (3) of Theorem 8.13. Therefore, by part (1) of Theorem 8.13, if $\mathcal{W} = \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\}$, then $\mathbf{Av}_1 = \mathbf{proj}_{\mathcal{W}}\mathbf{b} = \mathbf{Av}_2$.

(6) Let $\mathbf{b} = [b_1, ..., b_n]$, let $\mathbf{A}$ be as in Theorem 8.4, and let $\mathcal{W} = \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\}$. Since $\mathbf{proj}_{\mathcal{W}}\mathbf{b} \in \mathcal{W}$ exists, there must be some $\mathbf{v} \in \mathbb{R}^n$ such that $\mathbf{Av} = \mathbf{proj}_{\mathcal{W}}\mathbf{b}$. Hence, $\mathbf{v}$ satisfies part (1) of Theorem 8.13, so $(\mathbf{A}^T\mathbf{A})\mathbf{v} = \mathbf{A}^T\mathbf{b}$ by part (3) of Theorem 8.13. This shows that the system $(\mathbf{A}^T\mathbf{A})\mathbf{X} = \mathbf{A}^T\mathbf{b}$ is consistent, which proves part (2) of Theorem 8.4.

Next, let $f(x) = d_0 + d_1x + \cdots + d_tx^t$ and $\mathbf{z} = [d_0, d_1, \ldots, d_t]$. A short computation shows that $||\mathbf{Az} - \mathbf{b}||^2 = S_f$, the sum of the squares of the vertical distances illustrated in Figure 8.10, just before the definition of a least-squares polynomial in Section 8.3 of the textbook. Hence, minimizing $||\mathbf{Az} - \mathbf{b}||$ over all possible $(t+1)$-vectors $\mathbf{z}$ gives the coefficients of a degree t least-squares polynomial for the given points $(a_1, b_1), \ldots, (a_n, b_n)$. However, parts (2) and (3) of Theorem 8.13 show that such a minimal solution is found by solving $(\mathbf{A}^T\mathbf{A})\mathbf{v} = \mathbf{A}^T\mathbf{b}$, thus proving part (1) of Theorem 8.4.

Finally, when $\mathbf{A}^T\mathbf{A}$ is row equivalent to $\mathbf{I}_{t+1}$, the uniqueness condition holds by Theorems 2.15 and 2.16, which proves part (3) of Theorem 8.4.

(7) (a) F (b) T (c) T (d) T (e) T

Section 8.10

(1) (a) $\mathbf{C} = \begin{bmatrix} 8 & 12 \\ 0 & -9 \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 8 & 6 \\ 6 & -9 \end{bmatrix}$

(b) $\mathbf{C} = \begin{bmatrix} 7 & -17 \\ 0 & 11 \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 7 & -\frac{17}{2} \\ -\frac{17}{2} & 11 \end{bmatrix}$

(c) $\mathbf{C} = \begin{bmatrix} 5 & 4 & -3 \\ 0 & -2 & 5 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 5 & 2 & -\frac{3}{2} \\ 2 & -2 & \frac{5}{2} \\ -\frac{3}{2} & \frac{5}{2} & 0 \end{bmatrix}$

(2) (a) $\mathbf{A} = \begin{bmatrix} 43 & -24 \\ -24 & 57 \end{bmatrix}$, $\mathbf{P} = \frac{1}{5}\begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 75 & 0 \\ 0 & 25 \end{bmatrix}$,

$B = \left(\frac{1}{5}[-3, 4], \frac{1}{5}[4, 3]\right)$, $[\mathbf{x}]_B = [-7, -4]$, $Q(\mathbf{x}) = 4075$

(b) $\mathbf{A} = \begin{bmatrix} -5 & 16 & 40 \\ 16 & 37 & 16 \\ 40 & 16 & 49 \end{bmatrix}$, $\mathbf{P} = \frac{1}{9}\begin{bmatrix} 1 & -8 & 4 \\ -8 & 1 & 4 \\ 4 & 4 & 7 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 27 & 0 & 0 \\ 0 & -27 & 0 \\ 0 & 0 & 81 \end{bmatrix}$,

$B = \left(\frac{1}{9}[1, -8, 4], \frac{1}{9}[-8, 1, 4], \frac{1}{9}[4, 4, 7]\right)$, $[\mathbf{x}]_B = [3, -6, 3]$, $Q(\mathbf{x}) = 0$

(c) $\mathbf{A} = \begin{bmatrix} 18 & 48 & -30 \\ 48 & -68 & 18 \\ -30 & 18 & 1 \end{bmatrix}$, $\mathbf{P} = \frac{1}{7}\begin{bmatrix} 2 & -6 & 3 \\ 3 & -2 & -6 \\ 6 & 3 & 2 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 49 & 0 \\ 0 & 0 & -98 \end{bmatrix}$,

$B = \left(\frac{1}{7}[2, 3, 6], \frac{1}{7}[-6, -2, 3], \frac{1}{7}[3, -6, 2]\right)$, $[\mathbf{x}]_B = [5, 0, 6]$, $Q(\mathbf{x}) = -3528$

(d) $\mathbf{A} = \begin{bmatrix} 1 & 0 & -12 & 12 \\ 0 & 5 & 60 & 60 \\ -12 & 60 & 864 & 576 \\ 12 & 60 & 576 & 864 \end{bmatrix}$, $\mathbf{P} = \frac{1}{17}\begin{bmatrix} 1 & 0 & 12 & -12 \\ 0 & 1 & -12 & -12 \\ -12 & 12 & 1 & 0 \\ 12 & 12 & 0 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 289 & 0 & 0 & 0 \\ 0 & 1445 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$,

$B = \left(\frac{1}{17}[1, 0, -12, 12], \frac{1}{17}[0, 1, 12, 12], \frac{1}{17}[12, -12, 1, 0], \frac{1}{17}[-12, -12, 0, 1]\right)$,

$[\mathbf{x}]_B = [1, -3, -3, -10]$, $Q(\mathbf{x}) = 13294$

(3) First, $a_{ii} = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_i = \mathbf{e}_i^T\mathbf{B}\mathbf{e}_i = b_{ii}$. Also, for $i \neq j$, let $\mathbf{x} = \mathbf{e}_i + \mathbf{e}_j$. Then

$$a_{ii} + a_{ij} + a_{ji} + a_{jj} = \mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{B}\mathbf{x} = b_{ii} + b_{ij} + b_{ji} + b_{jj}.$$

Using $a_{ii} = b_{ii}$ and $a_{jj} = b_{jj}$, we get $a_{ij} + a_{ji} = b_{ij} + b_{ji}$. Hence, since $\mathbf{A}$ and $\mathbf{B}$ are symmetric, $2a_{ij} = 2b_{ij}$, and so $a_{ij} = b_{ij}$.

(4) Yes; if $Q(\mathbf{x}) = \Sigma a_{ij}x_ix_j$, $1 \leq i \leq j \leq n$, then $\mathbf{x}^T\mathbf{C}_1\mathbf{x}$ and $\mathbf{C}_1$ upper triangular imply that the (i, j) entry for $\mathbf{C}_1$ is zero if $i > j$ and a_{ij} if $i \leq j$. A similar argument describes $\mathbf{C}_2$. Thus, $\mathbf{C}_1 = \mathbf{C}_2$.

(5) (a) Consider the expression $Q(\mathbf{x}) = d_{11}y_1^2 + \cdots + d_{nn}y_n^2$ in Step 3 of the Quadratic Form Method, where $d_{11}, \ldots, d_{nn}$ are the eigenvalues of $\mathbf{A}$, and $[\mathbf{x}]_B = [y_1, \ldots, y_n]$. Clearly, if all of $d_{11}, \ldots, d_{nn}$ are positive and $\mathbf{x} \neq \mathbf{0}$, then $Q(\mathbf{x})$ must be positive. (Obviously, $Q(\mathbf{0}) = \mathbf{0}$.)

Conversely, if $Q(\mathbf{x})$ is positive for all $\mathbf{x} \neq \mathbf{0}$, then choose $\mathbf{x}$ so that $[\mathbf{x}]_B = \mathbf{e}_i$ to yield $Q(\mathbf{x}) = d_{ii}$, thus proving that the eigenvalue d_{ii} is positive.

(b) Replace "positive" with "nonnegative" throughout the solution to part (a).

(6) (a) T (b) F (c) F (d) T (e) T

Chapter 9

Section 9.1

(1) (a) Solution to first system: $(602, 1500)$; solution to second system: $(302, 750)$. The systems are ill-conditioned, because a very small change in the coefficient of y leads to a very large change in the solution.

(b) Solution to first system: $(400, 800, 2000)$; solution to second system: $(336, 666\frac{2}{3}, 1600)$. Systems are ill-conditioned.

(2) Answers to this problem may differ significantly from the following, depending on where the rounding is performed in the algorithm:

(a) Without partial pivoting: $(3210, 0.765)$; with partial pivoting: $(3230, 0.767)$. Actual solution is $(3214, 0.765)$.

(b) Without partial pivoting: $(3040, 10.7, -5.21)$; with partial pivoting: $(3010, 10.1, -5.03)$. (Actual solution is $(3000, 10, -5)$.)

(c) Without partial pivoting: $(2.26, 1.01, -2.11)$; with partial pivoting: $(277, -327, 595)$. Actual solution is $(267, -315, 573)$.

(3) Answers to this problem may differ significantly from the following, depending on where the rounding is performed in the algorithm:

(a) Without partial pivoting: $(3214, 0.7651)$; with partial pivoting: $(3213, 0.7648)$. Actual solution is $(3214, 0.765)$.

(b) Without partial pivoting: $(3001, 10, -5)$; with partial pivoting: $(3000, 9.995, -4.999)$. (Actual solution is $(3000, 10, -5)$.)

(c) Without partial pivoting: $(-2.380, 8.801, -16.30)$; with partial pivoting: $(267.8, -315.9, 574.6)$. Actual solution is $(267, -315, 573)$.

(4) (a)

	x_1	x_2
Initial Values	0.000	0.000
After 1 Step	5.200	−6.000
After 2 Steps	6.400	−8.229
After 3 Steps	6.846	−8.743
After 4 Steps	6.949	−8.934
After 5 Steps	6.987	−8.978
After 6 Steps	6.996	−8.994
After 7 Steps	6.999	−8.998
After 8 Steps	7.000	−9.000
After 9 Steps	7.000	−9.000

(b)

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	−0.778	−4.375	2.000
After 2 Steps	−1.042	−4.819	2.866
After 3 Steps	−0.995	−4.994	2.971
After 4 Steps	−1.003	−4.995	2.998
After 5 Steps	−1.000	−5.000	2.999
After 6 Steps	−1.000	−5.000	3.000
After 7 Steps	−1.000	−5.000	3.000

(c)

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	−8.857	4.500	−4.333
After 2 Steps	−10.738	3.746	−8.036
After 3 Steps	−11.688	4.050	−8.537
After 4 Steps	−11.875	3.975	−8.904
After 5 Steps	−11.969	4.005	−8.954
After 6 Steps	−11.988	3.998	−8.991
After 7 Steps	−11.997	4.001	−8.996
After 8 Steps	−11.999	4.000	−8.999
After 9 Steps	−12.000	4.000	−9.000
After 10 Steps	−12.000	4.000	−9.000

(d)

	x_1	x_2	x_3	x_4
Initial Values	0.000	0.000	0.000	0.000
After 1 Step	0.900	−1.667	3.000	−2.077
After 2 Steps	1.874	−0.972	3.792	−2.044
After 3 Steps	1.960	−0.999	3.966	−2.004
After 4 Steps	1.993	−0.998	3.989	−1.999
After 5 Steps	1.997	−1.001	3.998	−2.000
After 6 Steps	2.000	−1.000	3.999	−2.000
After 7 Steps	2.000	−1.000	4.000	−2.000
After 8 Steps	2.000	−1.000	4.000	−2.000

(5) (a)

	x_1	x_2
Initial Values	0.000	0.000
After 1 Step	5.200	−8.229
After 2 Steps	6.846	−8.934
After 3 Steps	6.987	−8.994
After 4 Steps	6.999	−9.000
After 5 Steps	7.000	−9.000
After 6 Steps	7.000	−9.000

(b)

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	−0.778	−4.569	2.901
After 2 Steps	−0.963	−4.978	2.993
After 3 Steps	−0.998	−4.999	3.000
After 4 Steps	−1.000	−5.000	3.000
After 5 Steps	−1.000	−5.000	3.000

(c)

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	−8.857	3.024	−7.790
After 2 Steps	−11.515	3.879	−8.818
After 3 Steps	−11.931	3.981	−8.974
After 4 Steps	−11.990	3.997	−8.996
After 5 Steps	−11.998	4.000	−8.999
After 6 Steps	−12.000	4.000	−9.000
After 7 Steps	−12.000	4.000	−9.000

(d)

	x_1	x_2	x_3	x_4
Initial Values	0.000	0.000	0.000	0.000
After 1 Step	0.900	−1.767	3.510	−2.012
After 2 Steps	1.980	−1.050	4.003	−2.002
After 3 Steps	2.006	−1.000	4.002	−2.000
After 4 Steps	2.000	−1.000	4.000	−2.000
After 5 Steps	2.000	−1.000	4.000	−2.000

(6) Strictly diagonally dominant: (a), (c)

(7) (a) Put the third equation first, and move the other two down to get the following:

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	3.125	−0.481	1.461
After 2 Steps	2.517	−0.500	1.499
After 3 Steps	2.500	−0.500	1.500
After 4 Steps	2.500	−0.500	1.500

(b) Put the first equation last (and leave the other two alone) to get:

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	3.700	−6.856	4.965
After 2 Steps	3.889	−7.980	5.045
After 3 Steps	3.993	−8.009	5.004
After 4 Steps	4.000	−8.001	5.000
After 5 Steps	4.000	−8.000	5.000
After 6 Steps	4.000	−8.000	5.000

(c) Put the second equation first, the fourth equation second, the first equation third, and the third equation fourth to get the following:

	x_1	x_2	x_3	x_4
Initial Values	0.000	0.000	0.000	0.000
After 1 Step	5.444	−5.379	9.226	−10.447
After 2 Steps	8.826	−8.435	10.808	−11.698
After 3 Steps	9.820	−8.920	10.961	−11.954
After 4 Steps	9.973	−8.986	10.994	−11.993
After 5 Steps	9.995	−8.998	10.999	−11.999
After 6 Steps	9.999	−9.000	11.000	−12.000
After 7 Steps	10.000	−9.000	11.000	−12.000
After 8 Steps	10.000	−9.000	11.000	−12.000

(8) The Jacobi Method yields the following:

	x_1	x_2	x_3
Initial Values	0.0	0.0	0.0
After 1 Step	16.0	−13.0	12.0
After 2 Steps	−37.0	59.0	−87.0
After 3 Steps	224.0	−61.0	212.0
After 4 Steps	−77.0	907.0	−1495.0
After 5 Steps	3056.0	2515.0	−356.0
After 6 Steps	12235.0	19035.0	−23895.0

The Gauss-Seidel Method yields the following:

	x_1	x_2	x_3
Initial Values	0.0	0.0	0.0
After 1 Step	16.0	83.0	−183.0
After 2 Steps	248.0	1841.0	−3565.0
After 3 Steps	5656.0	41053.0	−80633.0
After 4 Steps	124648.0	909141.0	−1781665.0

The actual solution is $(2, -3, 1)$.

(9) (a)

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	3.500	2.250	1.625
After 2 Steps	1.563	2.406	2.516
After 3 Steps	1.039	2.223	2.869
After 4 Steps	0.954	2.089	2.979
After 5 Steps	0.966	2.028	3.003
After 6 Steps	0.985	2.006	3.005
After 7 Steps	0.995	2.000	3.003
After 8 Steps	0.999	1.999	3.001
After 9 Steps	1.000	2.000	3.000
After 10 Steps	1.000	2.000	3.000

(b)

	x_1	x_2	x_3
Initial Values	0.000	0.000	0.000
After 1 Step	3.500	4.000	4.500
After 2 Steps	−0.750	0.000	0.750
After 3 Steps	3.125	4.000	4.875
After 4 Steps	−0.938	0.000	0.938
After 5 Steps	3.031	4.000	4.969
After 6 Steps	−0.984	0.000	0.985
After 7 Steps	3.008	4.000	4.992
After 8 Steps	−0.996	0.000	0.996

(10) (a) T (b) F (c) F (d) T (e) F (f) F

Section 9.2

(1) (a) $\mathbf{LDU} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

(b) $\mathbf{LDU} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 1 \end{bmatrix}$

(c) $\mathbf{LDU} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & 4 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -4 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$

(d) $\mathbf{LDU} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{5}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{3}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$

(e) $\mathbf{LDU} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 & 0 \\ -2 & -\frac{3}{2} & 1 & 0 \\ \frac{2}{3} & -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ 0 & 1 & \frac{5}{2} & -\frac{11}{2} \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(f) $\mathbf{LDU} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -3 & -3 & 1 & 0 \\ -1 & 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & 4 & -2 & -3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(2) (a) With the given values of **L**, **D**, and **U**, $\mathbf{LDU} = \begin{bmatrix} x & xz \\ wx & wxz + y \end{bmatrix}$. The first row of **LDU** cannot equal $[0, 1]$, since this would give $x = 0$, forcing $xz = 0 \neq 1$.

(b) The matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ cannot be reduced to row echelon form using only Type (I) and lower Type (II) row operations.

(3) (a) $\mathbf{KU} = \begin{bmatrix} -1 & 0 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix}$; solution $= \{(4, -1)\}$.

(b) $\mathbf{KU} = \begin{bmatrix} 2 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & -3 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$; solution $= \{(5, -1, 2)\}$.

(c) $\mathbf{KU} = \begin{bmatrix} -1 & 0 & 0 \\ 4 & 3 & 0 \\ -2 & 5 & -2 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$; solution $= \{(2, -3, 1)\}$.

(d) $\mathbf{KU} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ -5 & -1 & 4 & 0 \\ 1 & 3 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & -5 & 2 & 2 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; solution $= \{(2, -2, 1, 3)\}$.

(4) (a) F (b) T (c) F (d) F

Section 9.3

(1) (a) After nine iterations, eigenvector $= [0.60, 0.80]$ and eigenvalue $= 50$.

(b) After five iterations, eigenvector $= [0.91, 0.42]$ and eigenvalue $= 5.31$.

(c) After seven iterations, eigenvector $= [0.41, 0.41, 0.82]$ and eigenvalue $= 3.0$.

(d) After five iterations, eigenvector $= [0.58, 0.58, 0.00, 0.58]$ and eigenvalue $= 6.00$.

(e) After fifteen iterations, eigenvector $= [0.346, 0.852, 0.185, 0.346]$ and eigenvalue $= 5.405$

(f) After six iterations, eigenvector $= [0.4455, -0.5649, -0.6842, -0.1193]$ and eigenvalue $= 5.7323$.

(2) For the matrices in both parts, the Power Method fails to converge, even after many iterations. The matrix in part (a) is not diagonalizable, with 1 as its only eigenvalue and $\dim(E_1) = 1$. The matrix in part (b) has eigenvalues 1, 3, and -3, none of which are strictly dominant.

(3) (a) $\mathbf{u}_i$ is derived from $\mathbf{u}_{i-1}$ as $\mathbf{u}_i = c_i\mathbf{A}\mathbf{u}_{i-1}$, where c_i is a normalizing constant. A proof by induction shows that $\mathbf{u}_i = k_i\mathbf{A}^i\mathbf{u}_0$ for some nonzero constant k_i. If $\mathbf{u}_0 = a_0\mathbf{v}_1 + b_0\mathbf{v}_2$, then

$$\mathbf{u}_i = k_i a_0 \mathbf{A}^i\mathbf{v}_1 + k_i b_0 \mathbf{A}^i \mathbf{v}_2 = k_i a_0 \lambda_1^i \mathbf{v}_1 + k_i b_0 \lambda_2^i \mathbf{v}_2.$$

Hence $a_i = k_i a_0 \lambda_1^i$ and $b_i = k_i b_0 \lambda_2^i$. Thus,

$$\frac{|a_i|}{|b_i|} = \left|\frac{k_i a_0 \lambda_1^i}{k_i b_0 \lambda_2^i}\right| = \left|\frac{\lambda_1}{\lambda_2}\right|^i \frac{|a_0|}{|b_0|}.$$

(b) Let $\lambda_1, \ldots \lambda_n$ be the eigenvalues of $\mathbf{A}$ with $|\lambda_1| > |\lambda_j|$, for $2 \le j \le n$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be as given in the exercise. Suppose the initial vector in the Power Method is $\mathbf{u}_0 = a_{01}\mathbf{v}_1 + \cdots + a_{0n}\mathbf{v}_n$ and the ith iteration yields $\mathbf{u}_i = a_{i1}\mathbf{v}_1 + \cdots + a_{in}\mathbf{v}_n$. As in part (a), a proof by induction shows that $\mathbf{u}_i = k_i\mathbf{A}^i\mathbf{u}_0$ for some nonzero constant k_i. Therefore,

$$\begin{aligned} \mathbf{u}_i &= k_i a_{01}\mathbf{A}^i\mathbf{v}_1 + k_i a_{02}\mathbf{A}^i\mathbf{v}_2 + \cdots + k_i a_{0n}\mathbf{A}^i\mathbf{v}_n \\ &= k_i a_{01}\lambda_1^i\mathbf{v}_1 + k_i a_{02}\lambda_2^i\mathbf{v}_2 + \cdots + k_i a_{0n}\lambda_n^i\mathbf{v}_n. \end{aligned}$$

Hence, $a_{ij} = k_i a_{0j}\lambda_j^i$. Thus, for $2 \le j \le n$, $\lambda_j \ne 0$, and $a_{0j} \ne 0$, we have

$$\frac{|a_{i1}|}{|a_{ij}|} = \frac{\left|k_i a_{01}\lambda_1^i\right|}{\left|k_i a_{0j}\lambda_j^i\right|} = \left|\frac{\lambda_1}{\lambda_j}\right|^i \frac{|a_{01}|}{|a_{0j}|}.$$

(4) (a) F (b) T (c) T (d) F

Section 9.4

(1) (a) $\mathbf{Q} = \frac{1}{3}\begin{bmatrix} 2 & 1 & -2 \\ -2 & 2 & -1 \\ 1 & 2 & 2 \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 3 & 6 & 3 \\ 0 & 6 & -9 \\ 0 & 0 & 3 \end{bmatrix}$

(b) $\mathbf{Q} = \frac{1}{11}\begin{bmatrix} 6 & -2 & -9 \\ 7 & -6 & 6 \\ 6 & 9 & 2 \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 11 & 22 & 11 \\ 0 & 11 & 22 \\ 0 & 0 & 11 \end{bmatrix}$

(c) $\mathbf{Q} = \begin{bmatrix} \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} & 0 \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} \sqrt{6} & 3\sqrt{6} & -\frac{2\sqrt{6}}{3} \\ 0 & 2\sqrt{3} & -\frac{10\sqrt{3}}{3} \\ 0 & 0 & \sqrt{2} \end{bmatrix}$

(d) $\mathbf{Q} = \frac{1}{3}\begin{bmatrix} 2 & 2 & 0 \\ 2 & -2 & 1 \\ 0 & -1 & -2 \\ 1 & 0 & -2 \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 6 & -3 & 9 \\ 0 & 9 & 12 \\ 0 & 0 & 15 \end{bmatrix}$

(e) $\mathbf{Q} = \frac{1}{105}\begin{bmatrix} 14 & 99 & -2 & 32 \\ 70 & 0 & -70 & -35 \\ 77 & -18 & 64 & 26 \\ 0 & 30 & 45 & -90 \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 105 & 105 & -105 & 210 \\ 0 & 210 & 105 & 315 \\ 0 & 0 & 105 & 420 \\ 0 & 0 & 0 & 210 \end{bmatrix}$

(2) (a) $\mathbf{Q} = \frac{1}{13}\begin{bmatrix} 3 & 4 \\ 4 & -12 \\ 12 & 3 \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 13 & 26 \\ 0 & 13 \end{bmatrix}$; $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{169}\begin{bmatrix} 940 \\ -362 \end{bmatrix} \approx \begin{bmatrix} 5.562 \\ -2.142 \end{bmatrix}$

(b) $\mathbf{Q} = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 0 & -2 \\ 0 & -1 & -2 \\ 2 & 2 & 0 \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 3 & 9 & 6 \\ 0 & 6 & 9 \\ 0 & 0 & 12 \end{bmatrix}$; $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{5}{72}\begin{bmatrix} 43 \\ 3 \\ 62 \end{bmatrix} \approx \begin{bmatrix} 2.986 \\ 0.208 \\ 4.306 \end{bmatrix}$

(c) $\mathbf{Q} = \frac{1}{9}\begin{bmatrix} 1 & 8 & 2\sqrt{2} \\ 4 & 0 & \frac{7}{2}\sqrt{2} \\ 0 & 4 & -\frac{9}{2}\sqrt{2} \\ -8 & 1 & 2\sqrt{2} \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 9 & -9 & 9 \\ 0 & 18 & -9 \\ 0 & 0 & 18\sqrt{2} \end{bmatrix}$; $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{108}\begin{bmatrix} -61 \\ 65 \\ 66 \end{bmatrix} \approx \begin{bmatrix} -0.565 \\ 0.602 \\ 0.611 \end{bmatrix}$

(d) $\mathbf{Q} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{19}\sqrt{19} \\ \frac{2}{9} & \frac{4}{9} & -\frac{1}{19}\sqrt{19} \\ \frac{4}{9} & \frac{2}{9} & -\frac{3}{19}\sqrt{19} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{1}{19}\sqrt{19} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{19}\sqrt{19} \end{bmatrix}$; $\mathbf{R} = \begin{bmatrix} 9 & -18 & 18 \\ 0 & 9 & 9 \\ 0 & 0 & 2\sqrt{19} \end{bmatrix}$; $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{1026}\begin{bmatrix} 2968 \\ 6651 \\ 3267 \end{bmatrix} \approx \begin{bmatrix} 2.893 \\ 6.482 \\ 3.184 \end{bmatrix}$

(3) We will show that the entries of **Q** and **R** are uniquely determined by the given requirements. We will proceed column by column, using a proof by induction on the column number m.

Base Step: $m = 1$. Here,

$$[\text{1st column of } \mathbf{A}] = \mathbf{Q} \cdot [\text{1st column of } \mathbf{R}] = [\text{1st column of } \mathbf{Q}](r_{11}),$$

because $\mathbf{R}$ is upper triangular. But, since the 1st column of $\mathbf{Q}$ is a unit vector and r_{11} is positive,

$$r_{11} = \|[\text{1st column of } \mathbf{A}]\|, \quad \text{and} \quad [\text{1st column of } \mathbf{Q}] = \frac{1}{r_{11}}[\text{1st column of } \mathbf{A}].$$

Hence, the first column in each of $\mathbf{Q}$ and $\mathbf{R}$ is uniquely determined.

Inductive Step: Assume $m > 1$ and that columns $1, \ldots, (m-1)$ of both $\mathbf{Q}$ and $\mathbf{R}$ are uniquely determined. We will show that the mth column in each of $\mathbf{Q}$ and $\mathbf{R}$ is uniquely determined. Now,

$$[m\text{th column of } \mathbf{A}] = \mathbf{Q} \cdot [m\text{th column of } \mathbf{R}] = \sum_{i=1}^{m} [i\text{th column of } \mathbf{Q}]\ (r_{im}).$$

Let $\mathbf{v} = \sum_{i=1}^{m-1} [i\text{th column of } \mathbf{Q}]\ (r_{im})$. By the Inductive Hypothesis, $\mathbf{v}$ is uniquely determined. Now,

$$[m\text{th column of } \mathbf{Q}]\ (r_{mm}) = [m\text{th column of } \mathbf{A}] - \mathbf{v}.$$

But since the mth column of $\mathbf{Q}$ is a unit vector and r_{mm} is positive,

$$r_{mm} = \|[m\text{th column of } \mathbf{A}] - \mathbf{v}\|, \quad \text{and} \quad [m\text{th column of } \mathbf{Q}] = \frac{1}{r_{mm}}\left([m\text{th column of } \mathbf{A}] - \mathbf{v}\right).$$

Hence, the mth column in each of $\mathbf{Q}$ and $\mathbf{R}$ is uniquely determined.

(4) (a) If $\mathbf{A}$ is square, then $\mathbf{A} = \mathbf{QR}$, where $\mathbf{Q}$ is an orthogonal matrix and $\mathbf{R}$ has nonnegative entries along its main diagonal. Setting $\mathbf{U} = \mathbf{R}$, we see that

$$\mathbf{A}^T\mathbf{A} = \mathbf{U}^T\mathbf{Q}^T\mathbf{Q}\mathbf{U} = \mathbf{U}^T(\mathbf{I}_n)\mathbf{U} = \mathbf{U}^T\mathbf{U}.$$

(b) Suppose $\mathbf{A} = \mathbf{QR}$ (where this is a $\mathbf{QR}$ factorization of $\mathbf{A}$) and $\mathbf{A}^T\mathbf{A} = \mathbf{U}^T\mathbf{U}$. Then $\mathbf{R}$ and $\mathbf{R}^T$ must be nonsingular. Also, $\mathbf{P} = \mathbf{Q}(\mathbf{R}^T)^{-1}\mathbf{U}^T$ is an orthogonal matrix, because

$$\begin{aligned}
\mathbf{P}\mathbf{P}^T &= \mathbf{Q}(\mathbf{R}^T)^{-1}\mathbf{U}^T\left(\mathbf{Q}(\mathbf{R}^T)^{-1}\mathbf{U}^T\right)^T \\
&= \mathbf{Q}(\mathbf{R}^T)^{-1}\mathbf{U}^T\mathbf{U}\mathbf{R}^{-1}\mathbf{Q}^T \\
&= \mathbf{Q}(\mathbf{R}^T)^{-1}\mathbf{A}^T\mathbf{A}\mathbf{R}^{-1}\mathbf{Q}^T \\
&= \mathbf{Q}(\mathbf{R}^T)^{-1}(\mathbf{R}^T\mathbf{Q}^T)(\mathbf{QR})\mathbf{R}^{-1}\mathbf{Q}^T \\
&= \mathbf{Q}((\mathbf{R}^T)^{-1}\mathbf{R}^T)(\mathbf{Q}^T\mathbf{Q})(\mathbf{R}\mathbf{R}^{-1})\mathbf{Q}^T \\
&= \mathbf{Q}\mathbf{I}_n\mathbf{I}_n\mathbf{I}_n\mathbf{Q}^T \\
&= \mathbf{Q}\mathbf{Q}^T \\
&= \mathbf{I}_n.
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathbf{PU} &= \mathbf{Q}(\mathbf{R}^T)^{-1}\mathbf{U}^T\mathbf{U} = (\mathbf{Q}^T)^{-1}(\mathbf{R}^T)^{-1}\mathbf{A}^T\mathbf{A} \\
&= (\mathbf{R}^T\mathbf{Q}^T)^{-1}\mathbf{A}^T\mathbf{A} = (\mathbf{A}^T)^{-1}\mathbf{A}^T\mathbf{A} = \mathbf{I}_n\mathbf{A} = \mathbf{A},
\end{aligned}$$

and so $\mathbf{PU}$ is a $\mathbf{QR}$ factorization of $\mathbf{A}$, which is unique by Exercise 3. Hence, $\mathbf{U}$ is uniquely determined.

(5) (a) T (b) T (c) T (d) F (e) T

Section 9.5

(1) For each part, one possibility is given.

(a) $\mathbf{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 2\sqrt{10} & 0 \\ 0 & \sqrt{10} \end{bmatrix}$, $\mathbf{V} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$

(b) $\mathbf{U} = \frac{1}{5}\begin{bmatrix} -3 & -4 \\ -4 & 3 \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 15\sqrt{2} & 0 \\ 0 & 10\sqrt{2} \end{bmatrix}$, $\mathbf{V} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$

(c) $\mathbf{U} = \frac{1}{\sqrt{10}}\begin{bmatrix} -3 & 1 \\ 1 & 3 \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 9\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$, $\mathbf{V} = \frac{1}{3}\begin{bmatrix} -1 & -2 & 2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}$

(d) $\mathbf{U} = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 & -2 \\ -2 & 1 \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 5\sqrt{5} & 0 & 0 \\ 0 & 5\sqrt{5} & 0 \end{bmatrix}$, $\mathbf{V} = \frac{1}{5}\begin{bmatrix} -3 & 0 & 4 \\ 4 & 0 & 3 \\ 0 & 5 & 0 \end{bmatrix}$

(e) $\mathbf{U} = \begin{bmatrix} \frac{2}{\sqrt{13}} & \frac{18}{7\sqrt{13}} & -\frac{3}{7} \\ \frac{3}{\sqrt{13}} & \frac{-12}{7\sqrt{13}} & \frac{2}{7} \\ 0 & \frac{13}{7\sqrt{13}} & \frac{6}{7} \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{V} = \mathbf{U}$

(Note: $\mathbf{A}$ represents the orthogonal projection onto the plane $-3x + 2y + 6z = 0$.)

(f) $\mathbf{U} = \frac{1}{7}\begin{bmatrix} 6 & 2 & 3 \\ 3 & -6 & -2 \\ 2 & 3 & -6 \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}$, $\mathbf{V} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

(g) $\mathbf{U} = \frac{1}{11}\begin{bmatrix} 2 & 6 & 9 \\ -9 & 6 & -2 \\ 6 & 7 & -6 \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$, $\mathbf{V} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(h) $\mathbf{U} = \begin{bmatrix} -\frac{2}{5}\sqrt{2} & -\frac{8}{15}\sqrt{2} & \frac{1}{3} \\ \frac{1}{2}\sqrt{2} & -\frac{1}{6}\sqrt{2} & \frac{2}{3} \\ \frac{3}{10}\sqrt{2} & -\frac{13}{30}\sqrt{2} & -\frac{2}{3} \end{bmatrix}$, $\mathbf{\Sigma} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\mathbf{V} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

(2) (a) $\mathbf{A}^+ = \frac{1}{2250}\begin{bmatrix} 104 & 70 & 122 \\ -158 & 110 & 31 \end{bmatrix}$, $\mathbf{v} = \frac{1}{2250}\begin{bmatrix} 5618 \\ 3364 \end{bmatrix}$, $\mathbf{A}^T\mathbf{A}\mathbf{v} = \mathbf{A}^T\mathbf{b} = \frac{1}{15}\begin{bmatrix} 6823 \\ 3874 \end{bmatrix}$

(b) $\mathbf{A}^+ = \frac{1}{450}\begin{bmatrix} 8 & 43 & -11 \\ -6 & 24 & -48 \end{bmatrix}$, $\mathbf{v} = \frac{1}{90}\begin{bmatrix} 173 \\ 264 \end{bmatrix}$, $\mathbf{A}^T\mathbf{A}\mathbf{v} = \mathbf{A}^T\mathbf{b} = \begin{bmatrix} 65 \\ 170 \end{bmatrix}$

(c) $\mathbf{A}^+ = \frac{1}{84}\begin{bmatrix} 36 & 24 & 12 & 0 \\ 12 & 36 & -24 & 0 \\ -31 & -23 & 41 & 49 \end{bmatrix}$, $\mathbf{v} = \frac{1}{14}\begin{bmatrix} 44 \\ -18 \\ 71 \end{bmatrix}$, $\mathbf{A}^T\mathbf{A}\mathbf{v} = \mathbf{A}^T\mathbf{b} = \frac{1}{7}\begin{bmatrix} 127 \\ -30 \\ 60 \end{bmatrix}$

(d) $\mathbf{A}^+ = \frac{1}{54}\begin{bmatrix} 2 & 0 & 4 & 4 \\ 8 & 6 & 4 & 1 \\ 4 & 6 & -4 & -7 \end{bmatrix}$, $\mathbf{v} = \frac{1}{54}\begin{bmatrix} 26 \\ 59 \\ 7 \end{bmatrix}$, $\mathbf{A}^T\mathbf{A}\mathbf{v} = \mathbf{A}^T\mathbf{b} = \begin{bmatrix} 13 \\ 24 \\ -2 \end{bmatrix}$

(3) (a) $\mathbf{A} = 2\sqrt{10}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}\right)\left(\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\end{bmatrix}\right) + \sqrt{10}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}\right)\left(\frac{1}{\sqrt{5}}\begin{bmatrix}-1 & 2\end{bmatrix}\right)$

(b) $\mathbf{A} = 9\sqrt{10}\left(\frac{1}{\sqrt{10}}\begin{bmatrix}-3\\1\end{bmatrix}\right)\left(\frac{1}{3}\begin{bmatrix}-1 & -2 & 2\end{bmatrix}\right) + 3\sqrt{10}\left(\frac{1}{\sqrt{10}}\begin{bmatrix}1\\3\end{bmatrix}\right)\left(\frac{1}{3}\begin{bmatrix}-2 & 2 & 1\end{bmatrix}\right)$

(c) $\mathbf{A} = 2\sqrt{2}\left(\frac{1}{7}\begin{bmatrix}6\\3\\2\end{bmatrix}\right)\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix}\right) + \sqrt{2}\left(\frac{1}{7}\begin{bmatrix}2\\-6\\3\end{bmatrix}\right)\left(\frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1\end{bmatrix}\right)$

(d) $\mathbf{A} = 3\left(\frac{1}{11}\begin{bmatrix}2\\-9\\6\end{bmatrix}\right)\begin{bmatrix}0 & 1\end{bmatrix} + 2\left(\frac{1}{11}\begin{bmatrix}6\\6\\7\end{bmatrix}\right)\begin{bmatrix}1 & 0\end{bmatrix}$

(e) $\mathbf{A} = 2\begin{bmatrix}-\frac{2}{5}\sqrt{2}\\ \frac{1}{2}\sqrt{2}\\ \frac{3}{10}\sqrt{2}\end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}0 & -1 & 1 & 0\end{bmatrix}\right) + 2\begin{bmatrix}-\frac{8}{15}\sqrt{2}\\ -\frac{1}{6}\sqrt{2}\\ -\frac{13}{30}\sqrt{2}\end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 0 & 0 & 1\end{bmatrix}\right)$

(4) If $\mathbf{A}$ is orthogonal, then $\mathbf{A}^T\mathbf{A} = \mathbf{I}_n$. Therefore, $\lambda = 1$ is the only eigenvalue for $\mathbf{A}^T\mathbf{A}$, and the eigenspace is all of $\mathbb{R}^n$. Therefore, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ can be any orthonormal basis for $\mathbb{R}^n$. If we take $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ to be the standard basis for $\mathbb{R}^n$, then the corresponding Singular Value Decomposition for $\mathbf{A}$ is $\mathbf{A}\mathbf{I}_n\mathbf{I}_n$. If, instead, we use the *rows* of $\mathbf{A}$, which form an orthonormal basis for $\mathbb{R}^n$, to represent $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then the corresponding Singular Value Decomposition for $\mathbf{A}$ is $\mathbf{I}_n\mathbf{I}_n\mathbf{A}$.

(5) If $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, then $\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^T$. For $i \le m$, let σ_i be the (i,i) entry of $\mathbf{\Sigma}$. Thus, for $i \le m$, $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^T\mathbf{v}_i = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{e}_i = \mathbf{V}\mathbf{\Sigma}^T(\sigma_i\mathbf{e}_i) = \mathbf{V}(\sigma_i^2\mathbf{e}_i) = \sigma_i^2\mathbf{v}_i$. (Note that in this proof, "$\mathbf{e}_i$" has been used to represent both a vector in $\mathbb{R}^n$ and a vector in $\mathbb{R}^m$.) If $i > m$, then $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^T\mathbf{v}_i = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{e}_i = \mathbf{V}\mathbf{\Sigma}^T(\mathbf{0}) = \mathbf{0}$. This proves the claims made in the exercise.

(6) Because $\mathbf{A}$ is symmetric, it is orthogonally diagonalizable. Let $\mathbf{D} = \mathbf{P}^T\mathbf{A}\mathbf{P}$ be an orthogonal diagonalization for $\mathbf{A}$, with the eigenvalues of $\mathbf{A}$ along the main diagonal of $\mathbf{D}$. Thus, $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T$. By Exercise 5, the columns of $\mathbf{P}$ form an orthonormal basis of eigenvectors for $\mathbf{A}^T\mathbf{A}$, and the corresponding eigenvalues are the squares of the diagonal entries in $\mathbf{D}$. Hence, the singular values of $\mathbf{A}$ are the square roots of these eigenvalues, which are the absolute values of the diagonal entries of $\mathbf{D}$. Since the diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$, this completes the proof.

(7) Express $\mathbf{v} \in \mathbb{R}^n$ as $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, where $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are right singular vectors for $\mathbf{A}$. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, $\|\mathbf{v}\| = \sqrt{a_1^2 + \cdots + a_n^2}$. Thus,

$$\begin{aligned}
\|\mathbf{A}\mathbf{v}\|^2 &= (\mathbf{A}\mathbf{v}) \cdot (\mathbf{A}\mathbf{v}) \\
&= (a_1\mathbf{A}\mathbf{v}_1 + \cdots + a_n\mathbf{A}\mathbf{v}_n) \cdot (a_1\mathbf{A}\mathbf{v}_1 + \cdots + a_n\mathbf{A}\mathbf{v}_n) \\
&= a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2 \quad \text{(by parts (2) and (3) of Lemma 9.4)} \\
&\le a_1^2\sigma_1^2 + a_2^2\sigma_1^2 + \cdots + a_n^2\sigma_1^2 \\
&= \sigma_1^2\left(a_1^2 + a_2^2 + \cdots + a_n^2\right) = \sigma_1^2\|\mathbf{v}\|^2.
\end{aligned}$$

Therefore, $\|\mathbf{A}\mathbf{v}\| \le \sigma_1\|\mathbf{v}\|$.

(8) In all parts, assume k represents the number of nonzero singular values.

(a) The ith column of $\mathbf{V}$ is the right singular vector $\mathbf{v}_i$, which is a unit eigenvector corresponding to the eigenvalue λ_i of $\mathbf{A}^T\mathbf{A}$. But $-\mathbf{v}_i$ is also a unit eigenvector corresponding to the eigenvalue λ_i

of $\mathbf{A}^T\mathbf{A}$. Thus, if any vector $\mathbf{v}_i$ is replaced with its opposite vector, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is still an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors for $\mathbf{A}^T\mathbf{A}$. Since the vectors are kept in the same order, the λ_i do not increase, and thus $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ fulfills all the necessary conditions to be a set of right singular vectors for $\mathbf{A}$. For $i \leq k$, the left singular vector $\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i$, so when we change the sign of $\mathbf{v}_i$, we must adjust $\mathbf{U}$ by changing the sign of $\mathbf{u}_i$ as well. For $i > k$, changing the sign of $\mathbf{v}_i$ has no effect on $\mathbf{U}$, but still produces a valid Singular Value Decomposition.

(b) If the eigenspace E_λ for $\mathbf{A}^T\mathbf{A}$ has dimension higher than 1, then the corresponding right singular vectors can be replaced by any orthonormal basis for E_λ, for which there is an infinite number of choices. Then the associated left singular vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ must be adjusted accordingly.

(c) Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of right singular vectors for $\mathbf{A}$. By Exercise 5, the ith diagonal entry of $\mathbf{\Sigma}$ must be the square root of an eigenvalue of $\mathbf{A}^T\mathbf{A}$ corresponding to the eigenvector $\mathbf{v}_i$. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of eigenvectors, it must correspond to a complete set of eigenvalues for $\mathbf{A}^T\mathbf{A}$. Also by Exercise 5, all of the vectors $\mathbf{v}_i$, for $i > m$, correspond to the eigenvalue 0. Hence, all of the square roots of the nonzero eigenvalues of $\mathbf{A}^T\mathbf{A}$ must lie on the diagonal of Σ. Thus, the values that appear on the diagonal of $\mathbf{\Sigma}$ are uniquely determined. Finally, the order in which these numbers appear is determined by the requirement for the Singular Value Decomposition that they appear in non-increasing order.

(d) By definition, for $1 \leq i \leq k$, $\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i$, and so these left singular vectors are completely determined by the choices made for $\mathbf{v}_1, \ldots, \mathbf{v}_k$, the first k columns of $\mathbf{V}$.

(e) Columns $k+1$ through m of $\mathbf{U}$ are the left singular vectors $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_m$, which can be any orthonormal basis for the orthogonal complement of the column space of $\mathbf{A}$ (by parts (2) and (3) of Theorem 9.5). If $m = k+1$, this orthogonal complement to the column space is one-dimensional, and so there are only two choices for $\mathbf{u}_{k+1}$, which are opposites of each other, because $\mathbf{u}_{k+1}$ must be a unit vector. If $m > k+1$, the orthogonal complement to the column space has dimension greater than 1, and there is an infinite number of choices for its orthonormal basis.

(9) (a) Each right singular vector $\mathbf{v}_i$, for $1 \leq i \leq k$, must be an eigenvector for $\mathbf{A}^T\mathbf{A}$. Performing the Gram-Schmidt Process on the rows of $\mathbf{A}$, eliminating zero vectors, and normalizing will produce an orthonormal basis for the row space of $\mathbf{A}$, but there is no guarantee that it will consist of eigenvectors for $\mathbf{A}^T\mathbf{A}$. For example, if $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, performing the Gram-Schmidt Process on the rows of $\mathbf{A}$ produces the two vectors $[1, 1]$ and $\left[-\frac{1}{2}, \frac{1}{2}\right]$, neither of which is an eigenvector for $\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

(b) The right singular vectors $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ form an orthonormal basis for the eigenspace E_0 of $\mathbf{A}^T\mathbf{A}$. Any orthonormal basis for E_0 will do. By part (5) of Theorem 9.5, E_0 equals the kernel of the linear transformation L whose matrix with respect to the standard bases is $\mathbf{A}$. A basis for $\ker(L)$ can be found by using the Kernel Method. That basis can be turned into an orthonormal basis for $\ker(L)$ by applying the Gram-Schmidt Process and normalizing.

(10) If $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, as given in the exercise, then $\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T$, and so

$$\mathbf{A}^+\mathbf{A} = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{V}\mathbf{\Sigma}^+\mathbf{\Sigma}\mathbf{V}^T.$$

Note that $\mathbf{\Sigma}^+\mathbf{\Sigma}$ is an $n \times n$ diagonal matrix whose first k diagonal entries equal 1, with the remaining diagonal entries equal to 0. Note also that since the columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\mathbf{V}$ are orthonormal,
$\mathbf{V}^T\mathbf{v}_i = \mathbf{e}_i$, for $1 \leq i \leq n$.

(a) If $1 \le i \le k$, then $\mathbf{A}^+\mathbf{A}\mathbf{v}_i = \mathbf{V}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{v}_i = \mathbf{V}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\mathbf{e}_i = \mathbf{V}\mathbf{e}_i = \mathbf{v}_i$.

(b) If $i > k$, then $\mathbf{A}^+\mathbf{A}\mathbf{v}_i = \mathbf{V}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{v}_i = \mathbf{V}\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}\mathbf{e}_i = \mathbf{V}(\mathbf{0}) = \mathbf{0}$.

(c) By parts (4) and (5) of Theorem 9.5, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthonormal basis for $(\ker(L))^{\perp}$, and $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $\ker(L)$. The orthogonal projection onto $(\ker(L))^{\perp}$ sends every vector in $(\ker(L))^{\perp}$ to itself, and every vector orthogonal to $(\ker(L))^{\perp}$ (that is, every vector in $\ker(L)$) to $\mathbf{0}$. Parts (a) and (b) prove that this is precisely the result obtained by multiplying $\mathbf{A}^+\mathbf{A}$ by the bases for $(\ker(L))^{\perp}$ and $\ker(L)$.

(d) Using part (a), if $1 \le i \le k$, then

$$\mathbf{A}\mathbf{A}^+\mathbf{A}\mathbf{v}_i = \mathbf{A}\left(\mathbf{A}^+\mathbf{A}\mathbf{v}_i\right) = \mathbf{A}\mathbf{v}_i.$$

By part (b), if $i > k$,

$$\mathbf{A}\mathbf{A}^+\mathbf{A}\mathbf{v}_i = \mathbf{A}\left(\mathbf{A}^+\mathbf{A}\mathbf{v}_i\right) = \mathbf{A}(\mathbf{0}) = \mathbf{0}.$$

But, if $i > k$, $\mathbf{v}_i \in \ker(L)$, and so $\mathbf{A}\mathbf{v}_i = \mathbf{0}$ as well. Hence, $\mathbf{A}\mathbf{A}^+\mathbf{A}$ and $\mathbf{A}$ represent linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ that agree on a basis for $\mathbb{R}^n$. Therefore, they are equal as matrices.

(e) By part (d), $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$. If $\mathbf{A}$ is nonsingular, we can multiply both sides by $\mathbf{A}^{-1}$ (on the left) to obtain $\mathbf{A}^+\mathbf{A} = \mathbf{I}_n$. Multiplying by $\mathbf{A}^{-1}$ again (on the right) yields $\mathbf{A}^+ = \mathbf{A}^{-1}$.

(11) If $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, as given in the exercise, then $\mathbf{A}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^T$. Note that since the columns $\mathbf{u}_1, \ldots, \mathbf{u}_m$ of $\mathbf{U}$ are orthonormal, $\mathbf{U}^T\mathbf{u}_i = \mathbf{e}_i$, for $1 \le i \le m$.

(a) If $1 \le i \le k$, then

$$\mathbf{A}^+\mathbf{u}_i = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^T\mathbf{u}_i = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{e}_i = \mathbf{V}\left(\frac{1}{\sigma_i}\mathbf{e}_i\right) = \frac{1}{\sigma_i}\mathbf{v}_i.$$

Thus,

$$\mathbf{A}\mathbf{A}^+\mathbf{u}_i = \mathbf{A}\left(\mathbf{A}^+\mathbf{u}_i\right) = \mathbf{A}\left(\frac{1}{\sigma_i}\mathbf{v}_i\right) = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i = \mathbf{u}_i.$$

(b) If $i > k$, then

$$\mathbf{A}^+\mathbf{u}_i = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^T\mathbf{u}_i = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{e}_i = \mathbf{V}(\mathbf{0}) = \mathbf{0}.$$

Thus,

$$\mathbf{A}\mathbf{A}^+\mathbf{u}_i = \mathbf{A}\left(\mathbf{A}^+\mathbf{u}_i\right) = \mathbf{A}(\mathbf{0}) = \mathbf{0}.$$

(c) By Theorem 9.5, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthonormal basis for $\text{range}(L)$, and $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_m\}$ is an orthonormal basis for $(\text{range}(L))^{\perp}$. The orthogonal projection onto $\text{range}(L)$ sends every vector in $\text{range}(L)$ to itself, and every vector in $(\text{range}(L))^{\perp}$ to $\mathbf{0}$. Parts (a) and (b) prove that this is precisely the action of multiplying by $\mathbf{A}\mathbf{A}^+$ by showing how it acts on bases for $\text{range}(L)$ and $(\text{range}(L))^{\perp}$.

(d) Using part (a), if $1 \le i \le k$, then

$$\mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{u}_i = \mathbf{A}^+\left(\mathbf{A}\mathbf{A}^+\mathbf{u}_i\right) = \mathbf{A}^+\mathbf{u}_i.$$

By part (b), if $i > k$,

$$\mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{u}_i = \mathbf{A}^+\left(\mathbf{A}\mathbf{A}^+\mathbf{u}_i\right) = \mathbf{A}^+(\mathbf{0}) = \mathbf{0}.$$

But, if $i > k$, $\mathbf{A}^+\mathbf{u}_i = \mathbf{0}$ as well. Hence, $\mathbf{A}^+\mathbf{A}\mathbf{A}^+$ and $\mathbf{A}^+$ represent linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^n$ that agree on a basis for $\mathbb{R}^m$. Therefore, they are equal as matrices.

(e) Let L_1: $(\ker(L))^\perp \to \text{range}(L)$ be given by $L_1(\mathbf{v}) = \mathbf{A}\mathbf{v}$. Parts (2) and (3) of Theorem 9.5 shows that $\dim((\ker(L))^\perp) = k = \dim(\text{range}(L))$. Part (a) of this exercise and part (2) of Theorem 9.5 shows that

$$T\colon\ \text{range}(L) \to \text{range}(L) \quad \text{given by} \quad T(\mathbf{u}) = L_1(\mathbf{A}^+\mathbf{u})$$

is the identity linear transformation. Hence, L_1 must be onto, and so by Corollary 5.13, L_1 is an isomorphism. Because T is the identity transformation, the inverse of L_1 has $\mathbf{A}^+$ as its matrix with respect to the standard basis. This shows that $\mathbf{A}^+\mathbf{u}$ is uniquely determined for $\mathbf{u} \in \text{range}(L)$. Thus, since the subspace range(L) depends only on $\mathbf{A}$, not on the Singular Value Decomposition of $\mathbf{A}$, $\mathbf{A}^+\mathbf{u}$ is uniquely determined on the subspace range(L), independently from which Singular Value Decomposition is used for $\mathbf{A}$ to compute $\mathbf{A}^+$. Similarly, part (3) of Theorem 9.5 and part (b) of this exercise shows that $\mathbf{A}^+\mathbf{u} = \mathbf{0}$ on a basis for $(\text{range}(L))^\perp$, and so $\mathbf{A}^+\mathbf{u} = \mathbf{0}$ for *all* vectors in $(\text{range}(L))^\perp$ — again, independent of which Singular Value Decomposition of $\mathbf{A}$ is used. Thus, $\mathbf{A}^+\mathbf{u}$ is uniquely determined for $\mathbf{u} \in \text{range}(L)$ and $\mathbf{u} \in (\text{range}(L))^\perp$, and thus, on a basis for $\mathbb{R}^m$ which can be chosen from these two subspaces. Thus, $\mathbf{A}^+$ is uniquely determined.

(12) By part (a) of Exercise 28 in Section 1.5, trace $\left(\mathbf{A}\mathbf{A}^T\right)$ equals the sum of the squares of the entries of $\mathbf{A}$. If $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is a Singular Value Decomposition of $\mathbf{A}$, then

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T.$$

Using part (c) of Exercise 28 in Section 1.5, we see that

$$\text{trace}\left(\mathbf{A}\mathbf{A}^T\right) = \text{trace}\left(\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T\right) = \text{trace}\left(\mathbf{U}\mathbf{U}^T\mathbf{\Sigma}\mathbf{\Sigma}^T\right) = \text{trace}\left(\mathbf{\Sigma}\mathbf{\Sigma}^T\right) = \sigma_1^2 + \cdots + \sigma_k^2.$$

(13) Let $\mathbf{V}_1$ be the $n \times n$ matrix whose columns are

$$\mathbf{v}_i, \ldots, \mathbf{v}_j, \mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_n,$$

in that order. Let $\mathbf{U}_1$ be the $m \times m$ matrix whose columns are

$$\mathbf{u}_i, \ldots, \mathbf{u}_j, \mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{j+1}, \ldots, \mathbf{u}_m,$$

in that order. (The notation used here assumes that $i > 1$, $j < n$, and $j < m$, but the construction of $\mathbf{V}_1$ and $\mathbf{U}_1$ can be adjusted in an obvious way if $i = 1$, $j = n$, or $j = m$.) Let $\mathbf{\Sigma}_1$ be the diagonal $m \times n$ matrix with $\sigma_i, \ldots, \sigma_j$ in the first $j - i + 1$ diagonal entries and zero on the remaining diagonal entries. We claim that $\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$ is a Singular Value Decomposition for $\mathbf{A}_{ij}$. To show this, because $\mathbf{U}_1$, $\mathbf{\Sigma}_1$, and $\mathbf{V}_1$ are of the proper form, it is enough to show that $\mathbf{A}_{ij} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$.

If $i \le l \le j$, then

$$\mathbf{A}_{ij}\mathbf{v}_l = \left(\sigma_i\mathbf{u}_i\mathbf{v}_i^T + \cdots + \sigma_j\mathbf{u}_j\mathbf{v}_j^T\right)\mathbf{v}_l = \sigma_l\mathbf{u}_l,$$

and

$$\left(\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T\right)\mathbf{v}_l = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{e}_{l-i+1} = \mathbf{U}_1\sigma_l\mathbf{e}_{l-i+1} = \sigma_l\mathbf{u}_l.$$

If $l < i$ or $j < l$, then

$$\mathbf{A}_{ij}\mathbf{v}_l = \left(\sigma_i\mathbf{u}_i\mathbf{v}_i^T + \cdots + \sigma_j\mathbf{u}_j\mathbf{v}_j^T\right)\mathbf{v}_l = \mathbf{0}.$$

If $l < i$,

$$\left(\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T\right)\mathbf{v}_l = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{e}_{l+j-i+1} = \mathbf{U}_1\mathbf{0} = \mathbf{0},$$

and if $j < l$,

$$\left(\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T\right)\mathbf{v}_l = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{e}_l = \mathbf{U}_1\mathbf{0} = \mathbf{0}.$$

Hence, $\mathbf{A}_{ij}\mathbf{v}_l = \left(\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T\right)\mathbf{v}_l$ for every l, and so $\mathbf{A}_{ij} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$.

Finally, since we know a Singular Value Decomposition for $\mathbf{A}_{ij}$, we know that $\sigma_i, \ldots, \sigma_j$ are the singular values for $\mathbf{A}_{ij}$, and by part (1) of Theorem 9.5, $\text{rank}(\mathbf{A}_{ij}) = j - i + 1$.

(14) (a) $\mathbf{A} = \begin{bmatrix} 40 & -5 & 15 & -15 & 5 & -30 \\ 1.8 & 3 & 1.2 & 1.2 & -0.6 & 1.8 \\ 50 & 5 & 45 & -45 & -5 & -60 \\ -2.4 & -1.5 & 0.9 & 0.9 & 3.3 & -2.4 \\ 42.5 & -2.5 & 60 & -60 & 2.5 & -37.5 \end{bmatrix}$

(b) $\mathbf{A}_1 = \begin{bmatrix} 25 & 0 & 25 & -25 & 0 & -25 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 50 & 0 & 50 & -50 & 0 & -50 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 50 & 0 & 50 & -50 & 0 & -50 \end{bmatrix}$, $\mathbf{A}_2 = \begin{bmatrix} 35 & 0 & 15 & -15 & 0 & -35 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 55 & 0 & 45 & -45 & 0 & -55 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 40 & 0 & 60 & -60 & 0 & -40 \end{bmatrix}$,

$\mathbf{A}_3 = \begin{bmatrix} 40 & -5 & 15 & -15 & 5 & -30 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 50 & 5 & 45 & -45 & -5 & -60 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 42.5 & -2.5 & 60 & -60 & 2.5 & -37.5 \end{bmatrix}$, $\mathbf{A}_4 = \begin{bmatrix} 40 & -5 & 15 & -15 & 5 & -30 \\ 1.8 & 1.8 & 0 & 0 & -1.8 & 1.8 \\ 50 & 5 & 45 & -45 & -5 & -60 \\ -2.4 & -2.4 & 0 & 0 & 2.4 & -2.4 \\ 42.5 & -2.5 & 60 & -60 & 2.5 & -37.5 \end{bmatrix}$

(c) $N(\mathbf{A}) \approx 153.85$; $N(\mathbf{A} - \mathbf{A}_1)/N(\mathbf{A}) \approx 0.2223$; $N(\mathbf{A} - \mathbf{A}_2)/N(\mathbf{A}) \approx 0.1068$; $N(\mathbf{A} - \mathbf{A}_3)/N(\mathbf{A}) \approx 0.0436$; $N(\mathbf{A} - \mathbf{A}_4)/N(\mathbf{A}) \approx 0.0195$

(d) The method described in the text for the compression of digital images takes the matrix describing the image and alters it by zeroing out some of the lower singular values. This exercise illustrates how the matrices $\mathbf{A}_i$ that use only the first i singular values for a matrix $\mathbf{A}$ get closer to approximating $\mathbf{A}$ as i increases. The matrix for a digital image is, of course, much larger than the 5×6 matrix considered in this exercise. Also, you can frequently get a very good approximation of the image using a small enough number of singular values so that less data needs to be saved. Using the outer product form of the Singular Value Decomposition, only the singular values and the relevant singular vectors are needed to construct each $\mathbf{A}_i$, so not all mn entries of $\mathbf{A}_i$ need to be kept in storage.

(15) These are the steps to process a digital image in MATLAB, as described in the textbook:

- Enter the command: `edit`
- Use the text editor to enter the following MATLAB program:

```
function totalmat = RevisedPic(U, S, V, k)
T = V'
totalmat = 0;
for i = 1:k
totalmat = totalmat + U(:,i)*T(i,:)*S(i,i);
end
```

- Save this program under the name `RevisedPic.m`
- Enter the command: `A = imread('picturefilename')`
  where "picturefilename" is the name of the file containing the picture, including its file extension (preferably .tif or .jpg).
- Enter the command: `ndims(A)`

– If the response is "3", then MATLAB is treating your picture as if it is in color, even if it appears to be black-and-white.
  If this is the case, enter the command: `B = A(:,:,1);`
  (to use only the first color of the three used for color pictures)
  followed by the command: `C = double(B);`
  (to convert the integer format to decimal)
  Otherwise, just enter: `C = double(A);`

– Enter the command: `[U,S,V] = svd(C);`
  (This computes the Singular Value Decomposition of C.)

– Enter the command: `W = RevisedPic(U,S,V,100);`
  (Here, the number "100" represents the number of singular values you are using.
  You may change this to any value you like.)

– Enter the command: `R = uint8(round(W));`
  (This converts the decimal output to the correct integer format.)

– Enter the command: `imwrite(R,'Revised100.tif','tif')`
  (This will write the revised picture out to a file named "Revised100.tif".
  Of course, you can use any file name you would like, or use the .jpg extension instead of .tif.)

– You may repeat the steps:

```
W=RevisedPic(U,S,V,k);
R = uint8(round(W));
imwrite(R,'Revisedk.tif','tif')
```

  with different values for k to see how the picture gets more refined as more singular values are used.

– Output can be viewed using any appropriate software you may have on your computer for viewing such files.

(16) (a) F (b) T (c) F (d) F (e) F (f) T (g) F (h) T (i) F (j) T (k) T

Appendices

Appendix B

(1) (a) Not a function; undefined for $x < 1$

(b) Function; range $= \{x \in \mathbb{R} \mid x \geq 0\}$; image of $2 = 1$; pre-images of $2 = \{-3, 5\}$

(c) Not a function; two values are assigned to each $x \neq 1$

(d) Function; range $= \{-4, -2, 0, 2, 4, \ldots\}$; image of $2 = -2$; pre-images of $2 = \{6, 7\}$

(e) Not a function (k undefined at $\theta = \frac{\pi}{2}$)

(f) Function; range = all prime numbers; image of $2 = 2$; pre-image of $2 = \{0, 1, 2\}$

(g) Not a function; 2 is assigned to two different values (-1 and 6)

(2) (a) $\{-15, -10, -5, 5, 10, 15\}$

(b) $\{-9, -8, -7, -6, -5, 5, 6, 7, 8, 9\}$

(c) $\{\ldots, -8, -6, -4, -2, 0, 2, 4, 6, 8, \ldots\}$

(3) $(g \circ f)(x) = \frac{1}{4}\sqrt{75x^2 - 30x + 35}$; $(f \circ g)(x) = \frac{1}{4}(5\sqrt{3x^2 + 2} - 1)$

(4) $(g \circ f)\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -8 & 24 \\ 2 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$; $(f \circ g)\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -12 & 8 \\ -4 & 12 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

(5) (a) Let $f\colon A \to B$ be given by $f(1) = 4$, $f(2) = 5$, $f(3) = 6$, and $g\colon B \to C$ be given by $g(4) = g(7) = 8$, $g(5) = 9$, $g(6) = 10$. Then $g \circ f$ is onto, but $f^{-1}(\{7\})$ is empty, so f is not onto.

(b) Use the example from part (a). Note that $g \circ f$ is one-to-one, but $g(4) = g(7)$, so g is not one-to-one.

(6) The function f is onto because, given any real number x, the $n \times n$ diagonal matrix $\mathbf{A}$ with $a_{11} = x$ and $a_{ii} = 1$ for $2 \leq i \leq n$ has determinant x. Also, f is not one-to-one because for every $x \neq 0$, any other matrix obtained by performing a Type (II) operation on $\mathbf{A}$ (for example, $< 1 > \leftarrow < 2 > + < 1 >$) also has determinant x.

(7) The function f is not onto because only symmetric 3×3 matrices are in the range of f ($f(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T = (\mathbf{A}^T + \mathbf{A})^T = (f(\mathbf{A}))^T$). Also, f is not one-to-one because if $\mathbf{A}$ is any nonsymmetric 3×3 matrix, then $f(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T = \mathbf{A}^T + (\mathbf{A}^T)^T = f(\mathbf{A}^T)$, but $\mathbf{A} \neq \mathbf{A}^T$.

(8) f is not one-to-one, because $f(x + 1) = f(x + 3) = 1$; f is not onto, because there is no pre-image for x^n. For $n \geq 3$, the pre-image of $\mathcal{P}_2$ is $\mathcal{P}_3$.

(9) If $f(x_1) = f(x_2)$, then $3x_1^3 - 5 = 3x_2^3 - 5 \implies 3x_1^3 = 3x_2^3 \implies x_1^3 = x_2^3 \implies x_1 = x_2$. So f is one-to-one. Also, if $b \in \mathbb{R}$, then $f\left(\left(\frac{b+5}{3}\right)^{\frac{1}{3}}\right) = b$, so f is onto. Inverse of $f = f^{-1}(x) = \left(\frac{x+5}{3}\right)^{\frac{1}{3}}$.

(10) The function f is one-to-one, because

$$\begin{aligned} &f(\mathbf{A}_1) = f(\mathbf{A}_2) \\ \implies &\mathbf{B}^{-1}\mathbf{A}_1\mathbf{B} = \mathbf{B}^{-1}\mathbf{A}_2\mathbf{B} \\ \implies &\mathbf{B}(\mathbf{B}^{-1}\mathbf{A}_1\mathbf{B})\mathbf{B}^{-1} = \mathbf{B}(\mathbf{B}^{-1}\mathbf{A}_2\mathbf{B})\mathbf{B}^{-1} \\ \implies &(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}_1(\mathbf{B}\mathbf{B}^{-1}) = (\mathbf{B}\mathbf{B}^{-1})\mathbf{A}_2(\mathbf{B}\mathbf{B}^{-1}) \end{aligned}$$

$\Longrightarrow \mathbf{I}_n\mathbf{A}_1\mathbf{I}_n = \mathbf{I}_n\mathbf{A}_2\mathbf{I}_n$
$\Longrightarrow \mathbf{A}_1 = \mathbf{A}_2$.
The function f is onto, because for any $\mathbf{C} \in \mathcal{M}_{nn}$, $f(\mathbf{BCB}^{-1}) = \mathbf{B}^{-1}(\mathbf{BCB}^{-1})\mathbf{B} = \mathbf{C}$. Also, $f^{-1}(\mathbf{A}) = \mathbf{BAB}^{-1}$.

(11) (a) Let $c \in C$. Since $g \circ f$ is onto, there is some $a \in A$ such that $(g \circ f)(a) = c$. Then $g(f(a)) = c$. Let $f(a) = b$. Then $g(b) = c$, so g is onto. (Exercise 5(a) shows that f is not necessarily onto.)

(b) Let $a_1, a_2 \in A$ such that $f(a_1) = f(a_2)$. Then $g(f(a_1)) = g(f(a_2))$, implying $(g \circ f)(a_1) = (g \circ f)(a_2)$. But since $g \circ f$ is one-to-one, $a_1 = a_2$. Hence, f is one-to-one. (Exercise 5(b) shows that g is not necessarily one-to-one.)

(12) (a) F (b) T (c) F (d) F (e) F (f) F (g) F (h) F

Appendix C

(1) (a) $11 - i$ (b) $24 - 32i$ (c) $20 - 12i$ (d) $18 - 9i$ (e) $9 + 19i$ (f) $2 + 42i$ (g) $-17 - 19i$ (h) $5 - 4i$ (i) $9 + 2i$ (j) -6 (k) $16 + 22i$ (l) $\sqrt{73}$ (m) $\sqrt{53}$ (n) 5

(2) (a) $\frac{3}{20} + \frac{1}{20}i$ (b) $\frac{3}{25} - \frac{4}{25}i$ (c) $-\frac{4}{17} - \frac{1}{17}i$ (d) $-\frac{5}{34} + \frac{3}{34}i$

(3) In all parts, let $z_1 = a_1 + b_1 i$ and $z_2 = a_2 + b_2 i$.

(a) Part (1):

$$\begin{aligned}
\overline{z_1 + z_2} &= \overline{(a_1 + b_1 i) + (a_2 + b_2 i)} \\
&= \overline{(a_1 + a_2) + (b_1 + b_2)i} \\
&= (a_1 + a_2) - (b_1 + b_2)i \\
&= (a_1 - b_1 i) + (a_2 - b_2 i) \\
&= \overline{z_1} + \overline{z_2}.
\end{aligned}$$

Part (2):

$$\begin{aligned}
\overline{(z_1 z_2)} &= \overline{(a_1 + b_1 i)(a_2 + b_2 i)} \\
&= \overline{(a_1a_2 - b_1b_2) + (a_1b_2 + a_2b_1)i} \\
&= (a_1a_2 - b_1b_2) - (a_1b_2 + a_2b_1)i \\
&= (a_1a_2 - (-b_1)(-b_2)) + (a_1(-b_2) + a_2(-b_1))i \\
&= (a_1 - b_1 i)(a_2 - b_2 i) \\
&= \overline{z_1}\,\overline{z_2}.
\end{aligned}$$

(b) If $z_1 \neq 0$, then $\frac{1}{z_1}$ exists. Hence, $\frac{1}{z_1}(z_1 z_2) = \frac{1}{z_1}(0)$, implying $z_2 = 0$.

(c) Part (4): $z_1 = \overline{z_1} \Longleftrightarrow a_1 + b_1 i = a_1 - b_1 i \Longleftrightarrow b_1 i = -b_1 i \Longleftrightarrow 2b_1 i = 0 \Longleftrightarrow b_1 = 0 \Longleftrightarrow z_1$ is real.
Part (5): $z_1 = -\overline{z_1} \Longleftrightarrow a_1 + b_1 i = -(a_1 - b_1 i) \Longleftrightarrow a_1 + b_1 i = -a_1 + b_1 i \Longleftrightarrow a_1 = -a_1 \Longleftrightarrow 2a_1 = 0 \Longleftrightarrow a_1 = 0 \Longleftrightarrow z_1$ is pure imaginary.

(4) In all parts, let $z_1 = a_1 + b_1 i$ and $z_2 = a_2 + b_2 i$. Also note that $z_1\overline{z_1} = |z_1|^2$ because $z_1\overline{z_1} = (a_1 + b_1 i)(a_1 - b_1 i) = a_1^2 + b_1^2$.

(a) $|z_1 z_2|^2 = (z_1 z_2)\overline{(z_1 z_2)}$ (by the above) $= (z_1 z_2)\left(\overline{z_1}\ \overline{z_2}\right)$ (by part (2) of Theorem C.1) $= \left(z_1\overline{z_1}\right)\left(z_2\overline{z_2}\right) = |z_1|^2|z_2|^2 = \left(|z_1|\ |z_2|\right)^2$. Now take square roots.

(b) We have

$$\begin{aligned}
\left|\frac{1}{z_1}\right| &= \left|\frac{\overline{z_1}}{|z_1|^2}\right| \\
&\quad \text{(by the boxed equation just before Theorem C.1 in Appendix C)} \\
&= \left|\frac{a_1}{a_1^2+b_1^2} - \frac{b_1}{a_1^2+b_1^2}i\right| \\
&= \sqrt{\left(\frac{a_1}{a_1^2+b_1^2}\right)^2 + \left(\frac{b_1}{a_1^2+b_1^2}\right)^2} \\
&= \sqrt{\left(\frac{1}{a_1^2+b_1^2}\right)^2 (a_1^2+b_1^2)} \\
&= \frac{1}{\sqrt{a_1^2+b_1^2}} = \frac{1}{|z_1|}.
\end{aligned}$$

(c) We have

$$\begin{aligned}
\overline{\left(\frac{z_1}{z_2}\right)} &= \overline{\left(\frac{z_1\overline{z_2}}{z_2\overline{z_2}}\right)} = \overline{\left(\frac{z_1\overline{z_2}}{|z_2|^2}\right)} = \overline{\left(\frac{(a_1+b_1 i)(a_2-b_2 i)}{a_2^2+b_2^2}\right)} \\
&= \overline{\frac{(a_1a_2-b_1b_2)}{a_2^2+b_2^2} + \frac{(b_1a_2-a_1b_2)}{a_2^2+b_2^2}i} = \frac{(a_1a_2-b_1b_2)}{a_2^2+b_2^2} - \frac{(b_1a_2-a_1b_2)}{a_2^2+b_2^2}i \\
&= \frac{(a_1a_2-b_1b_2)}{a_2^2+b_2^2} + \frac{(-b_1a_2+a_1b_2)}{a_2^2+b_2^2}i = \frac{(a_1-b_1 i)(a_2+b_2 i)}{a_2^2+b_2^2} \\
&= \frac{\overline{z_1}z_2}{z_2\overline{z_2}} = \frac{\overline{z_1}}{\overline{z_2}}.
\end{aligned}$$

(5) (a) F (b) F (c) T (d) T (e) F

Appendix D

(1) (a) (III): $\langle 2\rangle \leftrightarrow \langle 3\rangle$; inverse operation is (III): $\langle 2\rangle \leftrightarrow \langle 3\rangle$. The matrix is its own inverse.

(b) (I): $\langle 2\rangle \leftarrow -2\,\langle 2\rangle$; inverse operation is (I): $\langle 2\rangle \leftarrow -\frac{1}{2}\,\langle 2\rangle$.

The inverse matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(c) (II): $\langle 3\rangle \leftarrow -4\,\langle 1\rangle + \langle 3\rangle$; inverse operation is (II): $\langle 3\rangle \leftarrow 4\,\langle 1\rangle + \langle 3\rangle$.

The inverse matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$.

(d) (I): $\langle 2\rangle \leftarrow 6\langle 2\rangle$; inverse operation is (I): $\langle 2\rangle \leftarrow \frac{1}{6}\langle 2\rangle$.

The inverse matrix is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(e) (II): $\langle 3\rangle \leftarrow -2\langle 4\rangle + \langle 3\rangle$; inverse operation is (II): $\langle 3\rangle \leftarrow 2\langle 4\rangle + \langle 3\rangle$.

The inverse matrix is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(f) (III): $\langle 1\rangle \leftrightarrow \langle 4\rangle$; inverse operation is (III): $\langle 1\rangle \leftrightarrow \langle 4\rangle$.
The matrix is its own inverse.

(2) (a) $\begin{bmatrix} 4 & 9 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 & \frac{9}{4} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(b) Not possible, since the matrix is singular.

(c) The product of the following matrices in the order listed:

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

$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\frac{5}{3} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & \frac{2}{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(3) If $\mathbf{A}$ and $\mathbf{B}$ are row equivalent, then $\mathbf{B} = \mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}$ for some elementary matrices $\mathbf{E}_1, \ldots, \mathbf{E}_k$, by Theorem D.3. Hence $\mathbf{B} = \mathbf{PA}$, with $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1$. By Corollary D.4, $\mathbf{P}$ is nonsingular.

Conversely, if $\mathbf{B} = \mathbf{PA}$ for a nonsingular matrix $\mathbf{P}$, then by Corollary D.4, $\mathbf{P} = \mathbf{E}_k \cdots \mathbf{E}_1$ for some elementary matrices $\mathbf{E}_1, \ldots, \mathbf{E}_k$. Hence $\mathbf{A}$ and $\mathbf{B}$ are row equivalent by Theorem D.3.

(4) Follow the hint in the textbook. Because $\mathbf{U}$ is upper triangular with all nonzero diagonal entries, when row reducing $\mathbf{U}$ to $\mathbf{I}_n$, no Type (III) row operations are needed, and all Type (II) row operations used will be of the form $\langle i\rangle \leftarrow c\langle j\rangle + \langle i\rangle$, where $i < j$ (the pivots are all below the targets). Thus, none of the Type (II) row operations change a diagonal entry, since $u_{ji} = 0$ when $i < j$. Hence, only Type (I) row operations make changes on the main diagonal, and no main diagonal entry will be made zero. Also, the elementary matrices for the Type (II) row operations mentioned are upper triangular: nonzero on the main diagonal, and perhaps in the (i, j) entry, with $i < j$. Since the elementary matrices for Type (I) row operations are also upper triangular, we see that $\mathbf{U}^{-1}$, which is the product of all these elementary matrices, is the product of upper triangular matrices. Therefore, it is also upper triangular.

(5) For Type (I) and (III) operations, $\mathbf{E} = \mathbf{E}^T$. If $\mathbf{E}$ corresponds to a Type (II) operation $\langle i\rangle \leftarrow c\langle j\rangle + \langle i\rangle$, then $\mathbf{E}^T$ corresponds to $\langle j\rangle \leftarrow c\langle i\rangle + \langle j\rangle$, so $\mathbf{E}^T$ is also an elementary matrix.

(6) Note that $(\mathbf{AF}^T)^T = \mathbf{FA}^T$. Now, multiplying by $\mathbf{F}$ on the left performs a row operation on $\mathbf{A}^T$. Taking the transpose again shows that this is a corresponding column operation on $\mathbf{A}$. The first equation show that this is the same as multiplying by $\mathbf{F}^T$ on the right side of $\mathbf{A}$.

(7) $\mathbf{A}$ is nonsingular iff $\text{rank}(\mathbf{A}) = n$ (by Theorem 2.15) iff $\mathbf{A}$ is row equivalent to $\mathbf{I}_n$ (by the definition of rank) iff $\mathbf{A} = \mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{I}_n = \mathbf{E}_k \cdots \mathbf{E}_1$ for some elementary matrices $\mathbf{E}_1, \ldots, \mathbf{E}_k$ (by Theorem D.3).

(8) $\mathbf{AX} = \mathbf{O}$ has a nontrivial solution iff $\text{rank}(\mathbf{A}) < n$ (by Theorem 2.7) iff $\mathbf{A}$ is singular (by Theorem 2.15) iff $\mathbf{A}$ cannot be expressed as a product of elementary matrices (by Corollary D.4).

(9) (a) By Theorem D.3, $\mathbf{E}_k \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A}$ is row equivalent to $\mathbf{A}$, so both have the same reduced row echelon form, and thus have the same rank (by definition of rank).

(b) The reduced row echelon form of $\mathbf{A}$ cannot have more nonzero rows than $\mathbf{A}$.

(c) If $\mathbf{A}$ has k rows of zeroes, then $\text{rank}(\mathbf{A}) = m - k$. But $\mathbf{AB}$ has at least k rows of zeroes, so $\text{rank}(\mathbf{AB}) \leq m - k$.

(d) Let $\mathbf{A} = \mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{D}$, where $\mathbf{D}$ is in reduced row echelon form. Then

$$\begin{aligned} \text{rank}(\mathbf{AB}) &= \text{rank}(\mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{DB}) \\ &= \text{rank}(\mathbf{DB}) && \text{(by repeated use of part (a))} \\ &\leq \text{rank}(\mathbf{D}) && \text{(by part (c))} \\ &= \text{rank}(\mathbf{A}) && \text{(by definition of rank).} \end{aligned}$$

(e) Exercise 18 in Section 2.3 proves the same results as those in parts (a) through (d), except that it is phrased in terms of the underlying row operations rather than in terms of elementary matrices.

(10) (a) T (b) F (c) F (d) T (e) T

Chapter Tests

We include here three sample tests for each of Chapters 1 through 7. Answer keys for all tests follow the test collection itself. The answer keys also include optional hints that an instructor might want to divulge for some of the more complicated or time-consuming problems.

Instructors who are supervising independent study students can use these tests to measure their students' progress at regular intervals. However, in a typical classroom situation, we do not expect instructors to give a test immediately after every chapter is covered, since that would amount to six or seven tests throughout the semester. Rather, we envision these tests as being used as a test bank; that is, a supply of questions, or ideas for questions, to assist instructors in composing their own tests.

Note that most of the tests, as printed here, would take a student much more than an hour to complete, even with the use of software on a computer and/or calculator. Hence, we expect instructors to choose appropriate subsets of these tests to fulfill their classroom needs.

Test for Chapter 1 — Version A

(1) A rower can propel a boat 5 km/hr on a calm river. If the rower rows southeastward against a current of 2 km/hr northward, what is the net velocity of the boat? Also, what is the net speed of the boat?

(2) Use a calculator to find the angle θ (to the nearest degree) between the vectors $\mathbf{x} = [3, -2, 5]$ and $\mathbf{y} = [-4, 1, -1]$.

(3) Prove that, for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2).$$

(4) Let $\mathbf{x} = [-3, 4, 2]$ represent the force on an object in a three-dimensional coordinate system, and let $\mathbf{a} = [5, -1, 3]$ be a given vector. Use $\mathbf{proj_a x}$ to decompose $\mathbf{x}$ into two component forces in directions parallel and orthogonal to $\mathbf{a}$. Verify that your answer is correct.

(5) State the contrapositive, converse, and inverse of the following statement:

If $\|\mathbf{x}+\mathbf{y}\| \neq \|\mathbf{x}\| + \|\mathbf{y}\|$, then $\mathbf{x}$ is not parallel to $\mathbf{y}$.

Which one of these is logically equivalent to the original statement?

(6) Give the negation of the following statement:

$\|\mathbf{x}\| = \|\mathbf{y}\|$ or $(\mathbf{x}-\mathbf{y}) \cdot (\mathbf{x}+\mathbf{y}) \neq 0$.

(7) Use a proof by induction to show that if the $n \times n$ matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$ are diagonal, then $\sum_{i=1}^{k} \mathbf{A}_i$ is diagonal.

(8) Decompose $\mathbf{A} = \begin{bmatrix} -4 & 3 & -2 \\ 6 & -1 & 7 \\ 2 & 4 & -3 \end{bmatrix}$ into the sum of $\mathbf{S}$ and $\mathbf{V}$, where $\mathbf{S}$ is a symmetric matrix and $\mathbf{V}$ is a skew-symmetric matrix.

(9) Given the following information about the employees of a certain TV network, calculate the total amount of salaries and perks paid out by the network for each TV show:

	Actors	Writers	Directors
TV Show 1	12	4	2
TV Show 2	10	2	3
TV Show 3	6	3	5
TV Show 4	9	4	1

	Salary	Perks
Actor	\$50000	\$40000
Writer	\$60000	\$30000
Director	\$80000	\$25000

(10) Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric $n \times n$ matrices. Prove that $\mathbf{AB}$ is symmetric if and only if $\mathbf{A}$ and $\mathbf{B}$ commute.

Test for Chapter 1 — Version B

(1) Three people are competing in a three-way tug-of-war, using three ropes all attached to a brass ring. Player A is pulling due north with a force of 100 lbs. Player B is pulling due east, and Player C is pulling at an angle of 30 degrees off due south, toward the west side. The ring is not moving. At what magnitude of force are Players B and C pulling?

(2) Use a calculator to find the angle θ (to the nearest degree) between the vectors $\mathbf{x} = [2, -5]$ and $\mathbf{y} = [-6, 5]$.

(3) Prove that, for any vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$ with $\mathbf{a} \neq \mathbf{0}$, the vector $\mathbf{b} - \mathbf{proj_a b}$ is orthogonal to $\mathbf{a}$.

(4) Let $\mathbf{x} = [5, -2, -4]$ represent the force on an object in a three-dimensional coordinate system, and let $\mathbf{a} = [-2, 3, -3]$ be a given vector. Use $\mathbf{proj_a x}$ to decompose $\mathbf{x}$ into two component forces in directions parallel and orthogonal to $\mathbf{a}$. Verify that your answer is correct.

(5) State the contrapositive, converse, and inverse of the following statement:

If $\|\mathbf{x}\|^2 + 2(\mathbf{x} \cdot \mathbf{y}) > 0$, then $\|\mathbf{x} + \mathbf{y}\| > \|\mathbf{y}\|$.

Which one of these is logically equivalent to the original statement?

(6) Give the negation of the following statement:

There is a unit vector $\mathbf{x}$ in $\mathbb{R}^3$ such that $\mathbf{x}$ is parallel to $[-2, 3, 1]$.

(7) Use a proof by induction to show that if the $n \times n$ matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$ are lower triangular, then $\sum_{i=1}^{k} \mathbf{A}_i$ is lower triangular.

(8) Decompose $\mathbf{A} = \begin{bmatrix} 5 & 8 & -2 \\ -6 & 3 & -3 \\ 9 & 4 & 1 \end{bmatrix}$ into the sum of $\mathbf{S}$ and $\mathbf{V}$, where $\mathbf{S}$ is a symmetric matrix and $\mathbf{V}$ is a skew-symmetric matrix.

(9) Given the following information about the amount of foods (in ounces) eaten by three cats each week, and the percentage of certain nutrients in each food type, find the total intake of each type of nutrient for each cat each week:

	Food A	Food B	Food C	Food D
Cat 1	9	4	7	8
Cat 2	6	3	10	4
Cat 3	4	6	9	7

	Nutrient 1	Nutrient 2	Nutrient 3
Food A	5%	4%	8%
Food B	2%	3%	2%
Food C	9%	2%	6%
Food D	6%	0%	8%

(10) Prove that if $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ skew-symmetric matrices, then $(\mathbf{AB})^T = \mathbf{BA}$.

Test for Chapter 1 — Version C

(1) Using Newton's Second Law of Motion, find the acceleration vector on a 10 kg object in a three-dimensional coordinate system when the following forces are simultaneously applied:

- a force of 6 newtons in the direction of the vector $[-4, 0, 5]$,
- a force of 4 newtons in the direction of the vector $[2, 3, -6]$, and
- a force of 8 newtons in the direction of the vector $[-1, 4, -2]$.

(2) Use a calculator to find the angle θ (to the nearest degree) between the vectors $\mathbf{x} = [-2, 4, -3]$ and $\mathbf{y} = [8, -3, -2]$.

(3) Without using the Triangle Inequality, prove that, for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\|\mathbf{x} + \mathbf{y}\|^2 \le (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.$$

(Hint: Use the Cauchy-Schwarz Inequality.)

(4) Let $\mathbf{x} = [-4, 2, -7]$ represent the force on an object in a three-dimensional coordinate system, and let $\mathbf{a} = [6, -5, 1]$ be a given vector. Use $\mathbf{proj_a x}$ to decompose $\mathbf{x}$ into two component forces in directions parallel and orthogonal to $\mathbf{a}$. Verify that your answer is correct.

(5) Use a proof by contrapositive to prove the following statement:

If $\mathbf{y} \neq \mathbf{proj_x y}$, then $\mathbf{y} \neq c\mathbf{x}$ for all $c \in \mathbb{R}$.

(6) Give the negation of the following statement:

For every vector $\mathbf{x} \in \mathbb{R}^n$, there is some $\mathbf{y} \in \mathbb{R}^n$ such that $\|\mathbf{y}\| > \|\mathbf{x}\|$.

(7) Decompose $\mathbf{A} = \begin{bmatrix} -3 & 6 & 3 \\ 7 & 5 & 2 \\ -2 & 4 & -4 \end{bmatrix}$ into the sum of $\mathbf{S}$ and $\mathbf{V}$, where $\mathbf{S}$ is a symmetric matrix and $\mathbf{V}$ is a skew-symmetric matrix.

(8) Given $\mathbf{A} = \begin{bmatrix} -4 & 3 & 1 & 5 \\ 6 & -9 & 2 & -4 \\ 8 & 7 & -1 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 6 & 4 & -2 \\ -1 & -3 & 4 \\ 2 & 3 & 9 \\ -2 & 5 & -8 \end{bmatrix}$, calculate, if possible, the third row of $\mathbf{AB}$ and the second column of $\mathbf{BA}$.

(9) Use a proof by induction to show that if the $n \times n$ matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$ are diagonal, then the product $\mathbf{A}_1 \cdots \mathbf{A}_k$ is diagonal.

(10) Prove that if $\mathbf{A}$ is a skew-symmetric matrix, then $\mathbf{A}^3$ is also skew-symmetric.

Test for Chapter 2 — Version A

(1) Use Gaussian Elimination to find the quadratic equation $y = ax^2 + bx + c$ that goes through the points $(3,7)$, $(-2,12)$, and $(-4,56)$.

(2) Use Gaussian Elimination or the Gauss-Jordan Method to solve the following linear system. Give the full solution set.

$$\begin{cases} 5x_1 + 10x_2 + 2x_3 + 14x_4 + 2x_5 = -13 \\ -7x_1 - 14x_2 - 2x_3 - 22x_4 + 4x_5 = 47 \\ 3x_1 + 6x_2 + x_3 + 9x_4 = -13 \end{cases}$$

(3) Solve the following homogeneous system using the Gauss-Jordan Method. Give the full solution set, expressing the vectors in it as linear combinations of particular solutions.

$$\begin{cases} 2x_1 + 5x_2 - 16x_3 - 9x_4 = 0 \\ x_1 + x_2 - 2x_3 - 2x_4 = 0 \\ -3x_1 + 2x_2 - 14x_3 - 2x_4 = 0 \end{cases}$$

(4) Find the rank of $\mathbf{A} = \begin{bmatrix} 2 & -1 & 8 \\ -2 & 2 & -10 \\ -5 & 3 & -21 \end{bmatrix}$. Is $\mathbf{A}$ row equivalent to $\mathbf{I}_3$?

(5) Solve the following two systems simultaneously:

$$\begin{cases} 2x_1 + 5x_2 + 11x_3 = 8 \\ 2x_1 + 7x_2 + 14x_3 = 6 \\ 3x_1 + 11x_2 + 22x_3 = 9 \end{cases} \quad \text{and} \quad \begin{cases} 2x_1 + 5x_2 + 11x_3 = 25 \\ 2x_1 + 7x_2 + 14x_3 = 30 \\ 3x_1 + 11x_2 + 22x_3 = 47 \end{cases}.$$

(6) Let $\mathbf{A}$ be a $m \times n$ matrix, $\mathbf{B}$ be an $n \times p$ matrix, and let R be the Type (I) row operation $R : \langle i \rangle \leftarrow c\langle i \rangle$, for some scalar c and some i with $1 \le i \le m$. Prove that $R(\mathbf{AB}) = R(\mathbf{A})\mathbf{B}$. (Since you are proving part of Theorem 2.1 in the textbook, you may not use that theorem in your proof. However, you may use the fact from Chapter 1 that (kth row of $(\mathbf{AB})$) = (kth row of $\mathbf{A}$)$\mathbf{B}$.)

(7) Determine whether or not $[-15, 10, -23]$ is in the row space of

$$\mathbf{A} = \begin{bmatrix} 5 & -4 & 8 \\ 2 & 1 & 2 \\ -1 & -5 & 1 \end{bmatrix}.$$

(8) Find the inverse of $\mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ 8 & -9 & 3 \\ 4 & -3 & 2 \end{bmatrix}$.

(9) Without using row reduction, find the inverse of the matrix $\mathbf{A} = \begin{bmatrix} 6 & -8 \\ 5 & -9 \end{bmatrix}$.

(10) Let $\mathbf{A}$ and $\mathbf{B}$ be nonsingular $n \times n$ matrices. Prove that $\mathbf{A}$ and $\mathbf{B}$ commute if and only if $(\mathbf{AB})^2 = \mathbf{A}^2\mathbf{B}^2$.

Test for Chapter 2 — Version B

(1) Use Gaussian Elimination to find the circle $x^2 + y^2 + ax + by + c = 0$ that goes through the points $(5, -1)$, $(6, -2)$, and $(1, -7)$.

(2) Use Gaussian Elimination or the Gauss-Jordan Method to solve the following linear system. Give the full solution set.

$$\begin{cases} 3x_1 - 2x_2 + 19x_3 + 11x_4 + 9x_5 - 27x_6 = 31 \\ -2x_1 + x_2 - 12x_3 - 7x_4 - 8x_5 + 21x_6 = -22 \\ 2x_1 - 2x_2 + 14x_3 + 8x_4 + 3x_5 - 14x_6 = 19 \end{cases}$$

(3) Solve the following homogeneous system using the Gauss-Jordan Method. Give the full solution set, expressing the vectors in it as linear combinations of particular solutions.

$$\begin{cases} -3x_1 + 9x_2 + 8x_3 + 4x_4 = 0 \\ 5x_1 - 15x_2 + x_3 + 22x_4 = 0 \\ 4x_1 - 12x_2 + 3x_3 + 22x_4 = 0 \end{cases}$$

(4) Find the rank of $\mathbf{A} = \begin{bmatrix} 3 & -9 & 7 \\ 1 & -2 & 2 \\ -15 & 41 & -34 \end{bmatrix}$. Is $\mathbf{A}$ row equivalent to $\mathbf{I}_3$?

(5) Solve the following two systems simultaneously:

$$\begin{cases} 4x_1 - 3x_2 - x_3 = -13 \\ -4x_1 - 2x_2 - 3x_3 = 14 \\ 5x_1 - x_2 + x_3 = -17 \end{cases} \quad \text{and} \quad \begin{cases} 4x_1 - 3x_2 - x_3 = 13 \\ -4x_1 - 2x_2 - 3x_3 = -16 \\ 5x_1 - x_2 + x_3 = 18 \end{cases}.$$

(6) Let $\mathbf{A}$ be a $m \times n$ matrix, $\mathbf{B}$ be an $n \times p$ matrix, and let R be the Type (II) row operation $R : \langle i \rangle \leftarrow c\langle j \rangle + \langle i \rangle$, for some scalar c and some i, j with $1 \le i, j \le m$. Prove that $R(\mathbf{AB}) = R(\mathbf{A})\mathbf{B}$. (Since you are proving part of Theorem 2.1 in the textbook, you may not use that theorem in your proof. However, you may use the fact from Chapter 1 that (kth row of $(\mathbf{AB})$) = (kth row of $\mathbf{A}$)$\mathbf{B}$.)

(7) Determine whether $[-2, 3, 1]$ is in the row space of $\mathbf{A} = \begin{bmatrix} 2 & -1 & -5 \\ -4 & 1 & 9 \\ -1 & 1 & 3 \end{bmatrix}$.

(8) Find the inverse of $\mathbf{A} = \begin{bmatrix} 3 & 3 & 1 \\ 2 & 1 & 0 \\ 10 & 3 & -1 \end{bmatrix}$.

(9) Without using row reduction, solve the linear system

$$\begin{cases} -5x_1 + 9x_2 + 4x_3 = -91 \\ 7x_1 - 14x_2 - 10x_3 = 146 \\ -7x_1 + 12x_2 + 4x_3 = -120 \end{cases}, \quad \text{where} \quad \begin{bmatrix} 32 & 6 & -17 \\ 21 & 4 & -11 \\ -7 & -\frac{3}{2} & \frac{7}{2} \end{bmatrix}$$

is the inverse of the coefficient matrix.

(10) (a) Prove that if $\mathbf{A}$ is a nonsingular matrix and $c \neq 0$, then $(c\mathbf{A})^{-1} = (\frac{1}{c})\mathbf{A}^{-1}$.

(b) Use the result from part (a) to prove that if $\mathbf{A}$ is a nonsingular skew-symmetric matrix, then $\mathbf{A}^{-1}$ is also skew-symmetric.

Test for Chapter 2 — Version C

(1) Use Gaussian Elimination to find the values of A, B, C, and D that solve the following partial fractions problem:

$$\frac{5x^2-18x+1}{(x-2)^2(x+3)} = \frac{A}{x-2} + \frac{Bx+C}{(x-2)^2} + \frac{D}{x+3}.$$

(2) Use Gaussian Elimination or the Gauss-Jordan Method to solve the following linear system. Give the full solution set.

$$\begin{cases} 9x_1 + 36x_2 - 11x_3 + 3x_4 + 34x_5 - 15x_6 = -6 \\ -10x_1 - 40x_2 + 9x_3 - 2x_4 - 38x_5 - 9x_6 = -89 \\ 4x_1 + 16x_2 - 4x_3 + x_4 + 15x_5 = 22 \\ 11x_1 + 44x_2 - 12x_3 + 2x_4 + 47x_5 + 3x_6 = 77 \end{cases}$$

(3) Solve the following homogeneous system using the Gauss-Jordan Method. Give the full solution set, expressing the vectors in it as linear combinations of particular solutions.

$$\begin{cases} -2x_1 - 2x_2 + 14x_3 + x_4 - 7x_5 = 0 \\ 6x_1 + 9x_2 - 27x_3 - 2x_4 + 21x_5 = 0 \\ -3x_1 - 4x_2 + 16x_3 + x_4 - 10x_5 = 0 \end{cases}$$

(4) Find the rank of $\mathbf{A} = \begin{bmatrix} 5 & 6 & -2 & 23 \\ 2 & 4 & -1 & 13 \\ -6 & -9 & 1 & -27 \\ 4 & 6 & -1 & 19 \end{bmatrix}$. Is $\mathbf{A}$ row equivalent to $\mathbf{I}_4$?

(5) Find the reduced row echelon form matrix $\mathbf{B}$ for the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 3 \\ -3 & 2 & 1 \\ 5 & 1 & 3 \end{bmatrix},$$

and list a series of row operations that converts $\mathbf{B}$ to $\mathbf{A}$.

(6) Determine whether $[25, 12, -19]$ is a linear combination of $\mathbf{a} = [7, 3, 5]$, $\mathbf{b} = [3, 3, 4]$, and $\mathbf{c} = [3, 2, 3]$.

(7) Find the inverse of $\mathbf{A} = \begin{bmatrix} -2 & 0 & 1 & -1 \\ 4 & -1 & -1 & 3 \\ 3 & 1 & -1 & -2 \\ 3 & 7 & -2 & -16 \end{bmatrix}$.

(8) Without using row reduction, solve the linear system

$$\begin{cases} -4x_1 + 2x_2 - 5x_3 = -3 \\ -4x_1 + x_2 - 3x_3 = -7 \\ 6x_1 + 3x_2 - 4x_3 = 26 \end{cases}, \quad \text{where} \quad \begin{bmatrix} \frac{5}{2} & -\frac{7}{2} & -\frac{1}{2} \\ -17 & 23 & 4 \\ -9 & 12 & 2 \end{bmatrix}$$

is the inverse of the coefficient matrix.

(9) Prove by induction that if $\mathbf{A}_1, \ldots, \mathbf{A}_k$ are nonsingular $n \times n$ matrices, then

$$(\mathbf{A}_1 \cdots \mathbf{A}_k)^{-1} = (\mathbf{A}_k)^{-1} \cdots (\mathbf{A}_1)^{-1}.$$

(10) Suppose that $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices, and that $R_1, \ldots, R_k$ are row operations such that $R_1(R_2(\cdots(R_k(\mathbf{AB}))\cdots)) = \mathbf{I}_n$. Prove that $R_1(R_2(\cdots(R_k(\mathbf{A}))\cdots)) = \mathbf{B}^{-1}$.

Test for Chapter 3 — Version A

(1) Calculate the area of the parallelogram determined by the vectors $\mathbf{x} = [-2, 5]$ and $\mathbf{y} = [6, -3]$.

(2) If $\mathbf{A}$ is a 5×5 matrix with determinant 4, what is $|6\mathbf{A}|$? Why?

(3) Calculate the determinant of $\mathbf{A} = \begin{bmatrix} 5 & -6 & -3 \\ 4 & 6 & 1 \\ 5 & 10 & 2 \end{bmatrix}$ by row reducing $\mathbf{A}$ to upper triangular form.

(4) Prove that if $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices, then $|\mathbf{A}\mathbf{B}^T| = |\mathbf{A}^T\mathbf{B}|$.

(5) Use cofactor expansion along any row or column to find the determinant of

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & -4 & 6 \\ 4 & -2 & 5 & 2 \\ -3 & 3 & 0 & -5 \\ 7 & 0 & 0 & -8 \end{bmatrix}.$$

Be sure to use cofactor expansion to find any 3×3 determinants needed as well.

(6) Prove that if $\mathbf{A}$ and $\mathbf{B}^{-1}$ are similar $n \times n$ matrices, then $|\mathbf{A}||\mathbf{B}| = 1$.

(7) Use Cramer's Rule to solve the following system:

$$\begin{cases} 5x_1 + x_2 + 2x_3 = 3 \\ -8x_1 + 2x_2 - x_3 = 17 \\ 6x_1 - x_2 + x_3 = -10 \end{cases}.$$

(8) Let $\mathbf{A} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$. Find a nonsingular matrix $\mathbf{P}$ having all integer entries, and a diagonal matrix $\mathbf{D}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

(9) Prove that an $n \times n$ matrix $\mathbf{A}$ is singular if and only if $\lambda = 0$ is an eigenvalue for $\mathbf{A}$.

(10) Suppose that $\lambda_1 = 2$ is an eigenvalue for an $n \times n$ matrix $\mathbf{A}$. Prove that $\lambda_2 = 8$ is an eigenvalue for $\mathbf{A}^3$.

Test for Chapter 3 — Version B

(1) Calculate the volume of the parallelepiped determined by the vectors $\mathbf{x} = [2, 7, -1]$, $\mathbf{y} = [4, 2, -3]$, and $\mathbf{z} = [-5, 6, 2]$.

(2) If $\mathbf{A}$ is a 4×4 matrix with determinant 7, what is $|5\mathbf{A}|$? Why?

(3) Calculate the determinant of

$$\mathbf{A} = \begin{bmatrix} 2 & -4 & -12 & 15 \\ 6 & 11 & 25 & -32 \\ 4 & 13 & 32 & -41 \\ 3 & -16 & -45 & 57 \end{bmatrix}$$

by row reducing $\mathbf{A}$ to upper triangular form.

(4) Prove that if $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices with $\mathbf{AB} = -\mathbf{BA}$, and n is odd, then either $\mathbf{A}$ or $\mathbf{B}$ is singular.

(5) Use cofactor expansion along any row or column to find the determinant of

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 5 & 2 \\ 4 & 3 & -1 & 0 \\ -6 & 8 & 0 & 0 \\ 1 & 7 & 0 & -3 \end{bmatrix}.$$

Be sure to use cofactor expansion to find any 3×3 determinants needed as well.

(6) Suppose $\mathbf{A}$ is an $m \times n$ matrix, R is the Type (II) row operation $\langle i \rangle \leftarrow b \langle j \rangle + \langle i \rangle$, and C is the Type (II) *column* operation $\langle \text{col. } i \rangle \leftarrow b \langle \text{col. } j \rangle + \langle \text{col. } i \rangle$. Prove that $(C(\mathbf{A}))^T = R(\mathbf{A}^T)$ by showing that the kth row of $(C(\mathbf{A}))^T = k$th row of $R(\mathbf{A}^T)$, first when $k \neq i$, and then when $k = i$.

(7) Use Cramer's Rule to solve the following system:

$$\begin{cases} 4x_1 - 4x_2 - 3x_3 = -10 \\ 6x_1 - 5x_2 - 10x_3 = -28 \\ -2x_1 + 2x_2 + 2x_3 = 6 \end{cases}.$$

(8) Let $\mathbf{A} = \begin{bmatrix} 2 & 1 & -3 \\ 1 & 2 & -3 \\ 1 & 1 & -2 \end{bmatrix}$. Find a nonsingular matrix $\mathbf{P}$ having all integer entries, and a diagonal matrix $\mathbf{D}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.

(9) Consider the matrix $\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

(a) Show that $\mathbf{A}$ has only one eigenvalue. What is it?

(b) Show that Step 3 of the Diagonalization Method of Section 3.4 produces only one fundamental eigenvector for $\mathbf{A}$.

(10) Consider the matrix $\mathbf{A} = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}$. It can be shown that, for every nonzero vector $\mathbf{X}$, the vectors $\mathbf{X}$ and $(\mathbf{AX})$ form an angle with each other measuring $\theta = \arccos(\frac{3}{5}) \approx 53^\circ$. (You may assume this fact.) Show why $\mathbf{A}$ is not diagonalizable, first from an algebraic perspective, and then from a geometric perspective.

Test for Chapter 3 — Version C

(1) Calculate the volume of the parallelepiped determined by the vectors $\mathbf{x} = [5, -2, 3]$, $\mathbf{y} = [-3, 1, -2]$, and $\mathbf{z} = [4, -5, 1]$.

(2) Calculate the determinant of

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 1 & 2 \\ 1 & 9 & 0 & 2 \\ 8 & 3 & 2 & -2 \\ 4 & 3 & 1 & 1 \end{bmatrix}$$

by row reducing $\mathbf{A}$ to upper triangular form.

(3) Use a cofactor expansion along any row or column to find the determinant of

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 5 \\ 0 & 12 & 0 & 9 & 4 \\ 0 & 0 & 0 & 8 & 3 \\ 0 & 14 & 11 & 7 & 2 \\ 15 & 13 & 10 & 6 & 1 \end{bmatrix}.$$

Be sure to use cofactor expansion on each 4×4, 3×3, and 2×2 determinant you need to calculate as well.

(4) Prove that if $\mathbf{A}$ is an orthogonal matrix (that is, $\mathbf{A}^T = \mathbf{A}^{-1}$), then $|\mathbf{A}| = \pm 1$.

(5) Suppose that $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{B}$ is an $n \times k$ matrix, and C is a Type (I), (II), or (III) *column* operation. Prove that $C(\mathbf{AB}) = \mathbf{A}(C(\mathbf{B}))$. (Hint: You can assume the fact that if R is the *row* operation corresponding to the *column* operation C, then, for any matrix $\mathbf{D}$, $(C(\mathbf{D})) = \left(R\left(\mathbf{D}^T\right)\right)^T$.)

(6) Perform (only) the Base Step in the proof by induction of the following statement, which is part of Theorem 3.3: Let $\mathbf{A}$ be an $n \times n$ matrix, with $n \geq 2$, and let R be the Type (II) row operation $\langle j \rangle \leftarrow c\langle i \rangle + \langle j \rangle$. Then $|\mathbf{A}| = |R(\mathbf{A})|$.

(7) Use Cramer's Rule to solve the following system:

$$\begin{cases} -9x_1 + 6x_2 + 2x_3 = -41 \\ 5x_1 - 3x_2 - x_3 = 22 \\ -8x_1 + 6x_2 + x_3 = -40 \end{cases}.$$

(8) Use diagonalization to calculate $\mathbf{A}^9$ if $\mathbf{A} = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$.

(9) Let $\mathbf{A} = \begin{bmatrix} -4 & 6 & -6 \\ 0 & 2 & 0 \\ 3 & -3 & 5 \end{bmatrix}$. Find a nonsingular matrix $\mathbf{P}$ having all integer entries, and a diagonal matrix $\mathbf{D}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.

(10) Let λ be an eigenvalue for an $n \times n$ matrix $\mathbf{A}$ having algebraic multiplicity k. Prove that λ is also an eigenvalue with algebraic multiplicity k for $\mathbf{A}^T$.

Test for Chapter 4 — Version A

(1) Prove that $\mathbb{R}$ is a vector space using the operations $\oplus$ and $\odot$ given by $\mathbf{x} \oplus \mathbf{y} = (x^5 + y^5)^{1/5}$ and $a \odot \mathbf{x} = a^{1/5}x$.

(2) Prove that the set of $n \times n$ skew-symmetric matrices is a subspace of $\mathcal{M}_{nn}$.

(3) Prove that the set $\{[5, -2, -1], [2, -1, -1], [8, -2, 1]\}$ spans $\mathbb{R}^3$.

(4) Find a simplified general form for all the vectors in span(S) in $\mathcal{P}_3$ if

$$S = \{4x^3 - 20x^2 - x - 11,\ 6x^3 - 30x^2 - 2x - 18,\ -5x^3 + 25x^2 + 2x + 16\}.$$

(5) Prove that the set $\{[2, -1, -11, -4], [9, -1, 2, 2], [4, 0, 21, 8], [-2, 1, 8, 3]\}$ in $\mathbb{R}^4$ is linearly independent.

(6) Determine whether

$$\left\{ \begin{bmatrix} 4 & 13 \\ -19 & 1 \end{bmatrix}, \begin{bmatrix} 3 & -15 \\ 19 & -1 \end{bmatrix}, \begin{bmatrix} 2 & -12 \\ 15 & -1 \end{bmatrix}, \begin{bmatrix} -4 & 7 \\ -5 & 2 \end{bmatrix} \right\}$$

is a basis for $\mathcal{M}_{22}$.

(7) Find a subset of

$$\begin{aligned} S \;=\; & \{3x^3 - 2x^2 + 3x - 1,\ -6x^3 + 4x^2 - 6x + 2,\ x^3 - 3x^2 + 2x - 1, \\ & 7x^3 + 5x - 1,\ 11x^3 - 12x^2 + 13x + 3\} \end{aligned}$$

in $\mathcal{P}_3$ that is a basis for span(S). What is dim(span(S))?

(8) Prove that the columns of a nonsingular $n \times n$ matrix $\mathbf{A}$ span $\mathbb{R}^n$.

(9) Consider the matrix $\mathbf{A} = \begin{bmatrix} 5 & -6 & 5 \\ 1 & 0 & 0 \\ -2 & 4 & -3 \end{bmatrix}$, which has $\lambda = 2$ as an eigenvalue. Find a basis B for $\mathbb{R}^3$ that contains a basis for the eigenspace E_2 for $\mathbf{A}$.

(10) (a) Find the transition matrix from B-coordinates to C-coordinates if

$$\begin{aligned} B &= ([-11, -19, 12],\ [7, 16, -2],\ [-8. - 12, 11]) \quad \text{and} \\ C &= ([2, 4, -1],\ [-1, -2, 1],\ [1, 1, -2]) \end{aligned}$$

are ordered bases for $\mathbb{R}^3$.

(b) Given $\mathbf{v} = [-63, -113, 62] \in \mathbb{R}^3$, find $[\mathbf{v}]_B$, and use your answer to part (a) to find $[\mathbf{v}]_C$.

Test for Chapter 4 — Version B

(1) Prove that $\mathbb{R}$ with the usual scalar multiplication, but with addition given by $\mathbf{x} \oplus \mathbf{y} = 3(x+y)$ is *not* a vector space.

(2) Prove that the set of all polynomials $\mathbf{p}$ in $\mathcal{P}_4$ for which the coefficient of the second-degree term equals the coefficient of the fourth-degree term is a subspace of $\mathcal{P}_4$.

(3) Let $\mathcal{V}$ be a vector space with subspaces $\mathcal{W}_1$ and $\mathcal{W}_2$. Prove that the intersection of $\mathcal{W}_1$ and $\mathcal{W}_2$ is a subspace of $\mathcal{V}$.

(4) Let $S = \{[5,-15,-4], [-2,6,1], [-9,27,7]\}$. Determine whether $[2,-4,3]$ is in span(S).

(5) Find a simplified general form for all the vectors in span(S) in $\mathcal{M}_{22}$ if

$$S = \left\{ \begin{bmatrix} 2 & -6 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 5 & -15 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 8 & -24 \\ -2 & 1 \end{bmatrix}, \begin{bmatrix} 3 & -9 \\ 1 & 2 \end{bmatrix} \right\}.$$

(6) Let S be the set

$$\{2x^3 - 11x^2 - 12x - 1, 7x^3 + 35x^2 + 16x - 7, \\ -5x^3 + 29x^2 + 33x + 2, -5x^3 - 26x^2 - 12x + 5\}$$

in $\mathcal{P}_3$ and let $\mathbf{p}(x) \in \mathcal{P}_3$. Prove that there is exactly one way to express $\mathbf{p}(x)$ as a linear combination of the elements of S.

(7) Find a subset of

$$S = \{[2,-3,4,-1], [-6,9,-12,3], [3,1,-2,2], [2,8,-12,3], [7,6,-10,4]\}$$

that is a basis for span(S) in $\mathbb{R}^4$. Does S span $\mathbb{R}^4$?

(8) Let $\mathbf{A}$ be a nonsingular $n \times n$ matrix. Prove that the rows of $\mathbf{A}$ are linearly independent.

(9) Consider the matrix $\mathbf{A} = \begin{bmatrix} -1 & -2 & 3 & 1 \\ 2 & -6 & 6 & 7 \\ 0 & 0 & -2 & 0 \\ 2 & -4 & 6 & 5 \end{bmatrix}$, which has $\lambda = -2$ as an eigenvalue. Find a basis B for $\mathbb{R}^4$ that contains a basis for the eigenspace E_{-2} for $\mathbf{A}$.

(10) (a) Find the transition matrix from B-coordinates to C-coordinates if

$$\begin{aligned} B &= (-21x^2 + 14x - 38,\ 17x^2 - 10x + 32,\ -10x^2 + 4x - 23) \text{ and} \\ C &= (-7x^2 + 4x - 14,\ 2x^2 - x + 4,\ x^2 - x + 1) \end{aligned}$$

are ordered bases for $\mathcal{P}_2$.

(b) Given $\mathbf{v} = -100x^2 + 64x - 181 \in \mathcal{P}_3$, find $[\mathbf{v}]_B$, and use your answer to part (a) to find $[\mathbf{v}]_C$.

Test for Chapter 4 — Version C

(1) The set $\mathbb{R}^2$ with operations $[x,y] \oplus [w,z] = [x+w+3, y+z-4]$, and $a \odot [x,y] = [ax+3a-3, ay-4a+4]$ is a vector space. Find the zero vector $\mathbf{0}$ and the additive inverse of $\mathbf{v} = [x,y]$ for this vector space.

(2) Prove that the set of all 3-vectors orthogonal to $[2,-3,1]$ forms a subspace of $\mathbb{R}^3$.

(3) Determine whether $[-3,6,5]$ is in span(S) in $\mathbb{R}^3$, if

$$S = \{[3,-6,-1],[4,-8,-3],[5,-10,-1]\}.$$

(4) Find a simplified form for all the vectors in span(S) in $\mathcal{P}_3$ if

$$S = \{4x^3 - x^2 - 11x + 18, -3x^3 + x^2 + 9x - 14, 5x^3 - x^2 - 13x + 22\}.$$

(5) Prove that the set

$$\left\{ \begin{bmatrix} -3 & 63 \\ 9 & -46 \end{bmatrix}, \begin{bmatrix} -5 & 56 \\ 14 & -41 \end{bmatrix}, \begin{bmatrix} -4 & 44 \\ 13 & -32 \end{bmatrix}, \begin{bmatrix} 5 & 14 \\ -13 & -10 \end{bmatrix} \right\}$$

in $\mathcal{M}_{22}$ is linearly independent.

(6) Find a subset of

$$S = \left\{ \begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}, \begin{bmatrix} -6 & 2 \\ 4 & -8 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -3 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 4 & -3 \\ -2 & 2 \end{bmatrix} \right\}$$

in $\mathcal{M}_{22}$ that is a basis for span(S). What is dim(span(S))?

(7) Let $\mathbf{A}$ be an $n \times n$ singular matrix and let S be the set of columns of $\mathbf{A}$. Prove that dim(span(S)) $< n$.

(8) Enlarge the linearly independent set

$$T = \{2x^3 - 3x^2 + 3x + 1, -x^3 + 4x^2 - 6x - 2\}$$

of $\mathcal{P}_3$ to a basis for $\mathcal{P}_3$.

(9) Let $\mathbf{A} \neq \mathbf{I}_n$ be an $n \times n$ matrix having eigenvalue $\lambda = 1$. Prove that $\dim(E_1) < n$.

(10) (a) Find the transition matrix from B-coordinates to C-coordinates if

$$\begin{aligned} B &= ([10,-17,8],[-4,10,-5],[29,-36,16]) \quad \text{and} \\ C &= ([12,-12,5],[-5,3,-1],[1,-2,1]) \end{aligned}$$

are ordered bases for $\mathbb{R}^3$.

(b) Given $\mathbf{v} = [-109, 155, -71] \in \mathbb{R}^3$, find $[\mathbf{v}]_B$, and use your answer to part (a) to find $[\mathbf{v}]_C$.

Test for Chapter 5 — Version A

(1) Prove that the mapping $f\colon \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $f(\mathbf{A}) = \mathbf{BAB}^{-1}$, where $\mathbf{B}$ is some fixed nonsingular $n \times n$ matrix, is a linear operator.

(2) Suppose $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation, and $L([-2,5,2]) = [2,1,-1]$, $L([0,1,1]) = [1,-1,0]$, and $L([1,-2,-1]) = [0,3,-1]$. Find $L([2,-3,1])$.

(3) Find the matrix $\mathbf{A}_{BC}$ for $L\colon \mathbb{R}^2 \to \mathcal{P}_2$ given by

$$L([a,b]) = (-2a+b)x^2 - (2b)x + (a+2b),$$

for the ordered bases $B = ([3,-7],[2,-5])$ for $\mathbb{R}^2$,
and $C = (9x^2+20x-21,\ 4x^2+9x-9,\ 10x^2+22x-23)$ for $\mathcal{P}_2$.

(4) For the linear transformation $L\colon \mathbb{R}^4 \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 4 & -8 & -1 & -7 \\ -3 & 6 & 1 & 6 \\ -4 & 8 & 2 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

find a basis for $\ker(L)$, a basis for $\operatorname{range}(L)$, and verify that the Dimension Theorem holds for L.

(5) (a) Consider the linear transformation $L\colon \mathcal{P}_3 \to \mathbb{R}^3$ given by

$$L(\mathbf{p}(x)) = \left[\mathbf{p}(3),\ \mathbf{p}'(1),\ \int_0^1 \mathbf{p}(x)\,dx\right].$$

Prove that $\ker(L)$ is nontrivial.

(b) Use part (a) to prove that there is a nonzero polynomial $\mathbf{p} \in \mathcal{P}_3$ such that $\mathbf{p}(3) = 0$, $\mathbf{p}'(1) = 0$, and $\int_0^1 \mathbf{p}(x)\,dx = 0$.

(6) Prove that the mapping $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$L([x,y,z]) = \begin{bmatrix} -5 & 2 & 1 \\ 6 & -3 & -2 \\ 10 & -3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

is one-to-one. Is L an isomorphism?

(7) Show that $L\colon \mathcal{P}_n \to \mathcal{P}_n$ given by $L(\mathbf{p}) = \mathbf{p} - \mathbf{p}'$ is an isomorphism.

(8) Consider the diagonalizable operator $L\colon \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by $L(\mathbf{K}) = \mathbf{K} - \mathbf{K}^T$. Let $\mathbf{A}$ be the matrix representation of L with respect to the standard basis C for $\mathcal{M}_{22}$.

(a) Find an ordered basis B of $\mathcal{M}_{22}$ consisting of fundamental eigenvectors for L, and the diagonal matrix $\mathbf{D}$ that is the matrix representation of L with respect to B.

(b) Calculate the transition matrix $\mathbf{P}$ from B to C, and verify that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.

(9) Indicate whether each given statement is true or false.

(a) If L is a linear operator on a nontrivial finite dimensional vector space for which every root of $p_L(x)$ is real, then L is diagonalizable.

(b) Let $L\colon \mathcal{V} \to \mathcal{V}$ be an isomorphism with eigenvalue λ. Then $\lambda \neq 0$ and $1/\lambda$ is an eigenvalue for L^{-1}.

(10) Prove that the linear operator on $\mathcal{P}_3$ given by $L(\mathbf{p}(x)) = \mathbf{p}(x+1)$ is not diagonalizable.

Test for Chapter 5 — Version B

(1) Prove that the mapping $f\colon \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $f(\mathbf{A}) = \mathbf{A} + \mathbf{A}^T$ is a linear operator.

(2) Suppose $L\colon \mathbb{R}^3 \longrightarrow \mathcal{P}_3$ is a linear transformation, and $L([1,-1,1]) = 2x^3 - x^2 + 3$, $L([2,5,6]) = -11x^3 + x^2 + 4x - 2$, and $L([1,4,4]) = -x^3 - 3x^2 + 2x + 10$. Find $L([6,-4,7])$.

(3) Find the matrix $\mathbf{A}_{BC}$ for $L\colon \mathcal{M}_{22} \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = [-2a + c - d, a - 2b + d, 2a - b + c],$$

for the ordered bases

$$B = \left(\begin{bmatrix} 1 & -2 \\ -4 & -1 \end{bmatrix}, \begin{bmatrix} -4 & 6 \\ 11 & 3 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 4 \\ 5 & 1 \end{bmatrix}\right)$$

for $\mathcal{M}_{22}$ and $C = ([-1,0,2], [2,2,-1], [1,3,2])$ for $\mathbb{R}^3$.

(4) For the linear transformation $L\colon \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 2 & 5 & 4 & 8 \\ 1 & -1 & -5 & -1 \\ -4 & 2 & 16 & 1 \\ 4 & 1 & -10 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

find a basis for $\ker(L)$, a basis for $\text{range}(L)$, and verify that the Dimension Theorem holds for L.

(5) Let $\mathbf{A}$ be a fixed $m \times n$ matrix, with $m \neq n$. Let $L_1\colon \mathbb{R}^n \to \mathbb{R}^m$ be given by $L_1(\mathbf{x}) = \mathbf{A}\mathbf{x}$, and let $L_2\colon \mathbb{R}^m \to \mathbb{R}^n$ be given by $L_2(\mathbf{y}) = \mathbf{A}^T\mathbf{y}$.

(a) Prove or disprove: $\dim(\text{range}(L_1)) = \dim(\text{range}(L_2))$.

(b) Prove or disprove: $\dim(\ker(L_1)) = \dim(\ker(L_2))$.

(6) Prove that the mapping $L\colon \mathcal{M}_{22} \to \mathcal{M}_{22}$ given by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 3a - 2b - 3c - 4d & -a - c - d \\ -a - d & 3a - b - c - d \end{bmatrix}$$

is one-to-one. Is L an isomorphism?

(7) Let $\mathbf{A}$ be a fixed nonsingular $n \times n$ matrix. Show that $L\colon \mathcal{M}_{nn} \to \mathcal{M}_{nn}$ given by $L(\mathbf{B}) = \mathbf{B}\mathbf{A}^{-1}$ is an isomorphism.

(8) Consider the diagonalizable operator L: $\mathcal{M}_{22} \to \mathcal{M}_{22}$ given by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} -7 & 12 \\ -4 & 7 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Let $\mathbf{A}$ be the matrix representation of L with respect to the standard basis C for $\mathcal{M}_{22}$.

(a) Find an ordered basis B of $\mathcal{M}_{22}$ consisting of fundamental eigenvectors for L, and the diagonal matrix $\mathbf{D}$ that is the matrix representation of L with respect to B.

(b) Calculate the transition matrix $\mathbf{P}$ from B to C, and verify that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

(9) Indicate whether each given statement is true or false.

(a) If L: $\mathcal{V} \to \mathcal{V}$ is an isomorphism on a nontrivial finite dimensional vector space $\mathcal{V}$ and λ is an eigenvalue for L, then $\lambda = \pm 1$.

(b) A linear operator L on an n-dimensional vector space is diagonalizable if and only if L has n distinct eigenvalues.

(10) Let L be the linear operator on $\mathbb{R}^3$ representing a counterclockwise rotation through an angle of $\pi/6$ radians around an axis through the origin that is parallel to $[4, -2, 3]$. Explain in words why L is not diagonalizable.

Test for Chapter 5 — Version C

(1) Let $\mathbf{A}$ be a fixed $n \times n$ matrix. Prove that the function L: $\mathcal{P}_n \rightarrow \mathcal{M}_{nn}$ given by

$$L(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0)$$
$$= a_n \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1 \mathbf{A} + a_0 \mathbf{I}_n$$

is a linear transformation.

(2) Suppose L: $\mathbb{R}^3 \rightarrow \mathcal{M}_{22}$ is a linear transformation such that $L([6,-1,1]) = \begin{bmatrix} 35 & -24 \\ 17 & 9 \end{bmatrix}$, $L([4,-2,-1]) = \begin{bmatrix} 21 & -18 \\ 5 & 11 \end{bmatrix}$, and $L([-7,3,1]) = \begin{bmatrix} -38 & 31 \\ -11 & -17 \end{bmatrix}$. Find $L([-3,1,-4])$.

(3) Find the matrix $\mathbf{A}_{BC}$ for L: $\mathbb{R}^3 \longrightarrow \mathbb{R}^4$ given by

$$L([x,y,z]) = [3x-y,\ x-z,\ 2x+2y-z,\ -x-2y],$$

for the ordered bases

$$\begin{aligned} B &= ([-5,1,-2],[10,-1,3],[41,-12,20]) \text{ for } \mathbb{R}^3, \text{ and} \\ C &= ([1,0,-1,-1],[0,1,-2,1],[-1,0,2,1],[0,-1,2,0]) \text{ for } \mathbb{R}^4. \end{aligned}$$

(4) Consider L: $\mathcal{M}_{22} \rightarrow \mathcal{P}_2$ given by

$$L\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = (a_{11}+a_{22})x^2 + (a_{11}-a_{21})x + a_{12}.$$

What is ker(L)? What is range(L)? Verify that the Dimension Theorem holds for L.

(5) Suppose that L_1: $\mathcal{V} \rightarrow \mathcal{W}$ and L_2: $\mathcal{W} \rightarrow \mathcal{Y}$ are linear transformations, where $\mathcal{V}$, $\mathcal{W}$, and $\mathcal{Y}$ are finite dimensional vector spaces.

(a) Prove that $\ker(L_1) \subseteq \ker(L_2 \circ L_1)$.

(b) Prove that $\dim(\text{range}(L_1)) \geq \dim(\text{range}(L_2 \circ L_1))$.

(6) Prove that the mapping L: $\mathcal{P}_2 \rightarrow \mathbb{R}^4$ given by

$$L(ax^2+bx+c) = [-5a+b-c,\ -7a-c,\ 2a-b-c,\ -2a+5b+c]$$

is one-to-one. Is L an isomorphism?

(7) Let $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an ordered basis for an n-dimensional vector space $\mathcal{V}$. Define L: $\mathbb{R}^n \rightarrow \mathcal{V}$ by $L([a_1, \ldots, a_n]) = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Prove that L is an isomorphism.

(8) Consider the diagonalizable operator L: $\mathcal{P}_2 \to \mathcal{P}_2$ given by

$$L(ax^2 + bx + c) \;=\; (3a + b - c)x^2 + (-2a + c)x + (4a + 2b - c).$$

Let $\mathbf{A}$ be the matrix representation of L with respect to the standard basis C for $\mathcal{P}_2$.

(a) Find an ordered basis B of $\mathcal{P}_2$ consisting of fundamental eigenvectors for L, and the diagonal matrix $\mathbf{D}$ that is the matrix representation of L with respect to B.

(b) Calculate the transition matrix $\mathbf{P}$ from B to C, and verify that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$.

(9) Indicate whether each given statement is true or false.

(a) If L is a diagonalizable linear operator on a nontrivial finite dimensional vector space, then the algebraic multiplicity of λ equals the geometric multiplicity of λ for every eigenvalue λ of L.

(b) A linear operator L on a nontrivial finite dimensional vector space $\mathcal{V}$ is one-to-one if and only if 0 is not an eigenvalue for L.

(10) Prove that the linear operator on $\mathcal{P}_n$, for $n > 0$, defined by $L(\mathbf{p}) = \mathbf{p}'$ is not diagonalizable.

Test for Chapter 6 — Version A

(1) Use the Gram-Schmidt Process to enlarge the following set to an orthogonal basis for $\mathbb{R}^3$: $\{[4, -3, 2]\}$.

(2) Let L: $\mathbb{R}^3 \to \mathbb{R}^3$ be the orthogonal reflection through the plane $x - 2y + 2z = 0$. Use eigenvalues and eigenvectors to find the matrix representation of L with respect to the standard basis for $\mathbb{R}^3$. (Hint: $[1, -2, 2]$ is orthogonal to the given plane.)

(3) For the subspace $\mathcal{W} = \text{span}(\{[2, -3, 1], [-3, 2, 2]\})$ of $\mathbb{R}^3$, find a basis for $\mathcal{W}^\perp$. What is $\dim(\mathcal{W}^\perp)$? (Hint: The two given vectors are *not* orthogonal.)

(4) Suppose $\mathbf{A}$ is an $n \times n$ orthogonal matrix. Prove that for every $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{Ax}\| = \|\mathbf{x}\|$.

(5) Prove the following part of Corollary 6.14: Let $\mathcal{W}$ be a subspace of $\mathbb{R}^n$. Then $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$.

(6) Let L: $\mathbb{R}^4 \longrightarrow \mathbb{R}^4$ be the orthogonal projection onto the subspace $\mathcal{W} = \text{span}(\{[2, 1, -3, -1], [5, 2, 3, 3]\})$ of $\mathbb{R}^4$. Find the characteristic polynomial of L.

(7) Explain why L: $\mathbb{R}^3 \to \mathbb{R}^3$ given by the orthogonal projection onto the plane $3x + 5y - 6z = 0$ is orthogonally diagonalizable.

(8) Consider L: $\mathbb{R}^3 \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Find an ordered orthonormal basis B of fundamental eigenvectors for L, and the diagonal matrix $\mathbf{D}$ that is the matrix for L with respect to B.

(9) Use orthogonal diagonalization to find a symmetric matrix $\mathbf{A}$ such that

$$\mathbf{A}^3 = \frac{1}{5}\begin{bmatrix} 31 & 18 \\ 18 & 4 \end{bmatrix}.$$

(10) Let L be a symmetric operator on $\mathbb{R}^n$ and let λ_1 and λ_2 be distinct eigenvalues for L with corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$. Prove that $\mathbf{v}_1 \perp \mathbf{v}_2$.

Test for Chapter 6 — Version B

(1) Use the Gram-Schmidt Process to find an orthogonal basis for $\mathbb{R}^3$, starting with the linearly independent set $\{[2,-3,1],[4,-2,-1]\}$. (Hint: The two given vectors are *not* orthogonal.)

(2) Let $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the orthogonal projection onto the plane $3x + 4z = 0$. Use eigenvalues and eigenvectors to find the matrix representation of L with respect to the standard basis for $\mathbb{R}^3$. (Hint: $[3,0,4]$ is orthogonal to the given plane.)

(3) Let $\mathcal{W} = \text{span}(\{[2,2,-3],[3,0,-1]\})$ in $\mathbb{R}^3$. Decompose $\mathbf{v} = [-12,-36,-43]$ into $\mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 \in \mathcal{W}$ and $\mathbf{w}_2 \in \mathcal{W}^\perp$. (Hint: The two given vectors spanning $\mathcal{W}$ are *not* orthogonal.)

(4) Suppose that $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ orthogonal matrices. Prove that $\mathbf{AB}$ is an orthogonal matrix.

(5) Prove the following part of Corollary 6.14: Let $\mathcal{W}$ be a subspace of $\mathbb{R}^n$. Assuming $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$ has already been shown, prove that $\mathcal{W} = (\mathcal{W}^\perp)^\perp$.

(6) Let $\mathbf{v} = [-3,5,1,0]$, and let $\mathcal{W}$ be the subspace of $\mathbb{R}^4$ spanned by $S = \{[0,3,6,-2],\ [-6,2,0,3]\}$. Find $\mathbf{proj}_{\mathcal{W}^\perp}\mathbf{v}$.

(7) Is $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 3 & 7 \\ 7 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

orthogonally diagonalizable? Explain.

(8) Consider $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 40 & 18 & 6 \\ 18 & 13 & -12 \\ 6 & -12 & 45 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Find an ordered orthonormal basis B of fundamental eigenvectors for L, and the diagonal matrix $\mathbf{D}$ that is the matrix for L with respect to B.

(9) Use orthogonal diagonalization to find a symmetric matrix $\mathbf{A}$ such that

$$\mathbf{A}^2 = \begin{bmatrix} 10 & -6 \\ -6 & 10 \end{bmatrix}.$$

(10) Let $\mathbf{A}$ be an $n \times n$ symmetric matrix. Prove that if all eigenvalues for $\mathbf{A}$ are ± 1, then $\mathbf{A}$ is orthogonal.

Test for Chapter 6 — Version C

(1) Use the Gram-Schmidt Process to enlarge the following set to an orthogonal basis for $\mathbb{R}^4$:

$$\{[-3, 1, -1, 2], [2, 4, 0, 1]\}.$$

(2) Let L: $\mathbb{R}^3 \to \mathbb{R}^3$ be the orthogonal projection onto the plane $2x - 2y - z = 0$. Use eigenvalues and eigenvectors to find the matrix representation of L with respect to the standard basis for $\mathbb{R}^3$. (Hint: $[2, -2, -1]$ is orthogonal to the given plane.)

(3) For the subspace $\mathcal{W} = \text{span}(\{[5, 1, -2], [-8, -2, 9]\})$ of $\mathbb{R}^3$, find a basis for $\mathcal{W}^\perp$. What is $\dim(\mathcal{W}^\perp)$? (Hint: The two given vectors are *not* orthogonal.)

(4) Let $\mathbf{A}$ be an $n \times n$ orthogonal skew-symmetric matrix. Show that n is even. (Hint: Assume n is odd and calculate $|\mathbf{A}^2|$ to reach a contradiction.)

(5) Prove that if $\mathcal{W}_1$ and $\mathcal{W}_2$ are two subspaces of $\mathbb{R}^n$ such that $\mathcal{W}_1 \subseteq \mathcal{W}_2$, then $\mathcal{W}_2^\perp \subseteq \mathcal{W}_1^\perp$.

(6) Let $P = (5, 1, 0, -4)$, and let

$$\mathcal{W} = \text{span}(\{[2, 1, 0, 2], [-2, 0, 1, 2], [0, 2, 2, -1]\}).$$

Find the minimum distance from P to $\mathcal{W}$.

(7) Explain why the matrix for L: $\mathbb{R}^3 \to \mathbb{R}^3$ given by the orthogonal reflection through the plane $4x - 3y - z = 0$ with respect to the standard basis for $\mathbb{R}^3$ is symmetric.

(8) Consider L: $\mathbb{R}^3 \to \mathbb{R}^3$ given by

$$L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 49 & -84 & -72 \\ -84 & 23 & -84 \\ -72 & -84 & 49 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Find an ordered orthonormal basis B of fundamental eigenvectors for L, and the diagonal matrix $\mathbf{D}$ that is the matrix for L with respect to B.

(9) Use orthogonal diagonalization to find a symmetric matrix $\mathbf{A}$ such that

$$\mathbf{A}^2 = \begin{bmatrix} 10 & 30 \\ 30 & 90 \end{bmatrix}.$$

(10) Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ symmetric matrices such that $p_{\mathbf{A}}(x) = p_{\mathbf{B}}(x)$. Prove that there is an orthogonal matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PBP}^{-1}$.

Test for Chapter 7 — Version A

Note that there are more than 10 problems in this test, with at least three problems from each section of Chapter 7.

(1) (Section 7.1) Verify that $\mathbf{a} \cdot \mathbf{b} = \overline{\mathbf{b} \cdot \mathbf{a}}$ for the complex vectors $\mathbf{a} = [2, i, 2+i]$ and $\mathbf{b} = [1+i, 1-i, 2+i]$.

(2) (Section 7.1) Suppose $\mathbf{H}$ and $\mathbf{P}$ are $n \times n$ complex matrices and that $\mathbf{H}$ is Hermitian. Prove that $\mathbf{P}^*\mathbf{H}\mathbf{P}$ is Hermitian.

(3) (Section 7.1) Indicate whether each given statement is true or false.

(a) If $\mathbf{Z}$ and $\mathbf{W}$ are $n \times n$ complex matrices, then $\mathbf{ZW} = \overline{(\mathbf{WZ})}$.

(b) If an $n \times n$ complex matrix is normal, then it is either Hermitian or skew-Hermitian.

(4) (Section 7.2) Use Gaussian Elimination to solve the following system of linear equations:

$$\left\{ \begin{array}{rcrcr} iz_1 & + & (1+3i)z_2 & = & 6+2i \\ (2+3i)z_1 & + & (7+7i)z_2 & = & 20-2i \end{array} \right. .$$

(5) (Section 7.2) Give an example of a matrix having all real entries that is not diagonalizable when thought of as a real matrix, but is diagonalizable when thought of as a complex matrix. Prove that your example works.

(6) (Section 7.2) Find all eigenvalues and a basis of fundamental eigenvectors for each eigenspace for $L\colon \mathbb{C}^3 \to \mathbb{C}^3$ given by

$$L\left(\left[\begin{array}{c} z_1 \\ z_2 \\ z_3 \end{array}\right]\right) = \left[\begin{array}{ccc} 1 & 0 & -1-i \\ -1 & i & i \\ -i & 0 & -1+i \end{array}\right]\left[\begin{array}{c} z_1 \\ z_2 \\ z_3 \end{array}\right].$$

Is L diagonalizable?

(7) (Section 7.3) Prove or disprove: For $n \geq 1$, the set of polynomials in $\mathcal{P}_n^{\mathbb{C}}$ with real coefficients is a complex subspace of $\mathcal{P}_n^{\mathbb{C}}$.

(8) (Section 7.3) Find a basis B for span(S) that is a subset of

$$S = \{[1+3i, 2+i, 1-2i], [-2+4i, 1+3i, 3-i], [1, 2i, 3+i], [5+6i, 3, -1-2i]\}.$$

(9) (Section 7.3) Find the matrix with respect to the standard bases for the complex linear transformation $L\colon \mathcal{M}_{22}^{\mathbb{C}} \to \mathbb{C}^2$ given by

$$L\left(\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]\right) = [a+bi,\ c+di].$$

Also, find the matrix for L with respect to the standard bases when thought of as a linear transformation between real vector spaces.

(10) (Section 7.4) Use the Gram-Schmidt Process to find an orthogonal basis for the complex vector space $\mathbb{C}^3$ starting with the linearly independent set $\{[1, i, 1+i], [3, -i, -1-i]\}$.

(11) (Section 7.4) Find a unitary matrix $\mathbf{P}$ such that

$$\mathbf{P}^* \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \mathbf{P}$$

is diagonal.

(12) (Section 7.4) Show that if $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ unitary matrices, then $\mathbf{AB}$ is unitary and $\left| |\mathbf{A}| \right| = 1$.

(13) (Section 7.5) Prove that

$$\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}(0)\mathbf{g}(0) + \mathbf{f}(1)\mathbf{g}(1) + \mathbf{f}(2)\mathbf{g}(2)$$

is a real inner product on $\mathcal{P}_2$. For $\mathbf{f}(x) = 2x + 1$ and $\mathbf{g}(x) = x^2 - 1$, calculate $\langle \mathbf{f}, \mathbf{g} \rangle$ and $\|\mathbf{f}\|$.

(14) (Section 7.5) Find the distance between $\mathbf{x} = [4, 1, 2]$ and $\mathbf{y} = [3, 3, 5]$ in the real inner product space consisting of $\mathbb{R}^3$ with inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{Ax} \cdot \mathbf{Ay}, \quad \text{where } \mathbf{A} = \begin{bmatrix} 4 & 0 & -1 \\ 3 & 3 & 1 \\ 1 & 2 & 1 \end{bmatrix}.$$

(15) (Section 7.5) Prove that in a real inner product space, for any vectors $\mathbf{x}$ and $\mathbf{y}$,

$$\|\mathbf{x}\| = \|\mathbf{y}\| \text{ if and only if } \langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0.$$

(16) (Section 7.5) Use the Generalized Gram-Schmidt Process to find an orthogonal basis for $\mathcal{P}_2$ starting with the linearly independent set $\{x^2, x\}$ using the real inner product given by

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_{-1}^{1} f(t)g(t)\, dt.$$

(17) (Section 7.5) Decompose $\mathbf{v} = [0, 1, -3]$ in $\mathbb{R}^3$ as $\mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 \in \mathcal{W} = \text{span}(\{[2, 1, -4], [4, 3, -8]\})$ and $\mathbf{w}_2 \in \mathcal{W}^\perp$, using the real inner product on $\mathbb{R}^3$ given by $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{Ax} \cdot \mathbf{Ay}$, with

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

Test for Chapter 7 — Version B

Note that there are more than 10 problems in this test, with at least three problems from each section of Chapter 7.

(1) (Section 7.1) Calculate $\mathbf{A}^*\mathbf{B}$ for the complex matrices

$$\mathbf{A} = \begin{bmatrix} i & 3+2i \\ 5 & -2i \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1-i & 3-2i \\ 3i & 7 \end{bmatrix}.$$

(2) (Section 7.1) Prove that if $\mathbf{H}$ is a Hermitian matrix, then $i\mathbf{H}$ is a skew-Hermitian matrix.

(3) (Section 7.1) Indicate whether each given statement is true or false.

(a) Every skew-Hermitian matrix must have all zeroes on its main diagonal.

(b) If $\mathbf{x}$, $\mathbf{y} \in \mathbb{C}^n$, but each entry of $\mathbf{x}$ and $\mathbf{y}$ is real, then $\mathbf{x} \cdot \mathbf{y}$ produces the same answer if the dot product is thought of as taking place in $\mathbb{C}^n$ as it would if the dot product were considered to be taking place in $\mathbb{R}^n$.

(4) (Section 7.2) Use the an inverse matrix to solve the given linear system:

$$\begin{cases} (4+i)z_1 & + & (2+14i)z_2 & = & 21+4i \\ (2-i)z_1 & + & (6+5i)z_2 & = & 10-6i \end{cases}.$$

(5) (Section 7.2) Explain why the sum of the algebraic multiplicities of a complex $n \times n$ matrix must equal n.

(6) (Section 7.2) Find all eigenvalues and a basis of fundamental eigenvectors for each eigenspace for $L: \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$L\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = \begin{bmatrix} -i & 1 \\ 4 & 3i \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Is L diagonalizable?

(7) (Section 7.3) Prove or disprove: The set of $n \times n$ Hermitian matrices is a complex subspace of $\mathcal{M}_{nn}^{\mathbb{C}}$. If it is a subspace, compute its dimension.

(8) (Section 7.3) If

$$S = \{[i, -1, -2+i], [2-i, 2+2i, 5+2i], [0, 2, 2-2i], [1+3i, -2+2i, -3+5i]\},$$

find a basis B for span(S) that uses vectors having a simpler form than those in S.

(9) (Section 7.3) Show that L: $\mathcal{M}_{22}^{\mathbb{C}} \to \mathcal{M}_{22}^{\mathbb{C}}$ given by $L(\mathbf{Z}) = \mathbf{Z}^*$ is not a complex linear transformation.

(10) (Section 7.4) Use the Gram-Schmidt Process to find an orthogonal basis for the complex vector space $\mathbb{C}^3$ containing $\{[1+i, 1-i, 2]\}$.

(11) (Section 7.4) Prove that $\mathbf{Z} = \begin{bmatrix} 1+i & -1+i \\ 1-i & 1+i \end{bmatrix}$ is unitarily diagonalizable.

(12) (Section 7.4) Let $\mathbf{A}$ be an $n \times n$ unitary matrix. Prove that $\mathbf{A}^2 = \mathbf{I}_n$ if and only if $\mathbf{A}$ is Hermitian.

(13) (Section 7.5) Let $\mathbf{A}$ be the 3×3 nonsingular matrix

$$\begin{bmatrix} 4 & 3 & 1 \\ -1 & 4 & 3 \\ -1 & 1 & 1 \end{bmatrix}.$$

Prove that $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{Ax} \cdot \mathbf{Ay}$ is a real inner product on $\mathbb{R}^3$. For $\mathbf{x} = [2, -1, 1]$ and $\mathbf{y} = [1, 4, 2]$, calculate $\langle \mathbf{x}, \mathbf{y} \rangle$ and $\|\mathbf{x}\|$ for this inner product.

(14) (Section 7.5) Find the distance between $\mathbf{w} = [4+i, 5, 4+i, 2+2i]$ and $\mathbf{z} = [1+i, 2i, -2i, 2+3i]$ in the complex inner product space $\mathbb{C}^3$ using the usual complex dot product as an inner product.

(15) (Section 7.5) Prove that in a real inner product space, for any vectors $\mathbf{x}$ and $\mathbf{y}$,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \tfrac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2).$$

(16) (Section 7.5) Use the Generalized Gram-Schmidt Process to find an orthogonal basis for $\mathcal{P}_2$ containing x^2 using the real inner product given by

$$\langle \mathbf{f}, \mathbf{g} \rangle \;=\; \mathbf{f}(0)\mathbf{g}(0) + \mathbf{f}(1)\mathbf{g}(1) + \mathbf{f}(2)\mathbf{g}(2).$$

(17) (Section 7.5) Decompose $\mathbf{v} = x + 3$ in $\mathcal{P}_2$ as $\mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 \in \mathcal{W} = \text{span}(\{x^2\})$ and $\mathbf{w}_2 \in \mathcal{W}^{\perp}$, using the real inner product on $\mathcal{P}_2$ given by

$$\langle \mathbf{f}, \mathbf{g} \rangle \;=\; \int_0^1 f(t)g(t)\,dt.$$

Test for Chapter 7 — Version C

Note that there are more than 10 problems in this test, with at least three problems from each section of Chapter 7.

(1) (Section 7.1) Calculate $\mathbf{A}^T\mathbf{B}^*$ for the complex matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 3-i \\ 1+2i & -i \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 4+i & 3i \\ 2 & 1-2i \end{bmatrix}.$$

(2) (Section 7.1) Let $\mathbf{H}$ and $\mathbf{J}$ be $n\times n$ Hermitian matrices. Prove that if $\mathbf{HJ}$ is Hermitian, then $\mathbf{HJ} = \mathbf{JH}$.

(3) (Section 7.1) Indicate whether each given statement is true or false.

(a) If $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ with $\mathbf{x}\cdot\mathbf{y} \in \mathbb{R}$, then $\mathbf{x}\cdot\mathbf{y} = \mathbf{y}\cdot\mathbf{x}$.

(b) If $\mathbf{A}$ is an $n \times n$ Hermitian matrix, and $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, then $\mathbf{Ax}\cdot\mathbf{y} = \mathbf{x}\cdot\mathbf{Ay}$.

(4) (Section 7.2) Use Cramer's Rule to solve the following system of linear equations:

$$\begin{cases} -iz_1 + (1-i)z_2 = -3-2i \\ (1-2i)z_1 + (4-i)z_2 = -5-6i \end{cases}.$$

(5) (Section 7.2) Show that a complex $n \times n$ matrix that is not diagonalizable must have an eigenvalue λ whose algebraic multiplicity is strictly greater than its geometric multiplicity.

(6) (Section 7.2) Find all eigenvalues and a basis of fundamental eigenvectors for each eigenspace for $L\colon \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ given by

$$L\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & 2 \\ -2i & 2+2i \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Is L diagonalizable?

(7) (Section 7.3) Prove that the set of normal 2×2 matrices is not a complex subspace of $\mathcal{M}_{22}^{\mathbb{C}}$ by showing that it is not closed under matrix addition.

(8) (Section 7.3) Find a basis B for $\mathbb{C}^4$ that contains

$$\{[1, 3+i, i, 1-i], [1+i, 3+4i, -1+i, 2]\}.$$

(9) (Section 7.3) Let $\mathbf{w} \in \mathbb{C}^n$ be a fixed nonzero vector. Show that $L\colon \mathbb{C}^n \to \mathbb{C}$ given by $L(\mathbf{z}) = \mathbf{w}\cdot\mathbf{z}$ is not a complex linear transformation.

(10) (Section 7.4) Use the Gram-Schmidt Process to find an orthogonal basis for the complex vector space $\mathbb{C}^3$, starting with the basis $\{[2-i, 1+2i, 1], [0, 1, 0], [-1, 0, 2+i]\}$.

(11) (Section 7.4) Let $\mathbf{Z}$ be an $n \times n$ complex matrix whose rows form an orthonormal basis for $\mathbb{C}^n$. Prove that the columns of $\mathbf{Z}$ also form an orthonormal basis for $\mathbb{C}^n$.

(12) (Section 7.4)

(a) Show that $\mathbf{A} = \begin{bmatrix} -23i & 36 \\ -36 & -2i \end{bmatrix}$ is normal, and hence unitarily diagonalizable.

(b) Find a unitary matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is diagonal.

(13) (Section 7.5) For $\mathbf{x} = [x_1, x_2]$ and $\mathbf{y} = [y_1, y_2]$ in $\mathbb{R}^2$, prove that $\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$ is a real inner product on $\mathbb{R}^2$. For $\mathbf{x} = [2, 3]$ and $\mathbf{y} = [4, -4]$, calculate $\langle \mathbf{x}, \mathbf{y} \rangle$ and $\|\mathbf{x}\|$ for this inner product.

(14) (Section 7.5) Find the distance between $\mathbf{f}(x) = \sin^2 x$ and $\mathbf{g}(x) = -\cos^2 x$ in the real inner product space consisting of the set of all real-valued continuous functions defined on the interval $[0, \pi]$ with inner product

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^{\pi} f(t)g(t)\,dt.$$

(15) (Section 7.5) Prove that in a real inner product space, vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if and only if $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

(16) (Section 7.5) Use the Generalized Gram-Schmidt Process to find an orthogonal basis for $\mathbb{R}^3$, starting with the linearly independent set $\{[-1, 1, 1], [3, -4, -1]\}$ using the real inner product given by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{Ax} \cdot \mathbf{Ay}, \quad \text{where} \quad \mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}.$$

(17) (Section 7.5) Decompose $\mathbf{v} = x^2$ in $\mathcal{P}_2$ as $\mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1 \in \mathcal{W} = \text{span}(\{x - 5, 6x - 5\})$ and $\mathbf{w}_2 \in \mathcal{W}^{\perp}$, using the real inner product on $\mathcal{P}_2$ given by

$$\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}(0)\mathbf{g}(0) + \mathbf{f}(1)\mathbf{g}(1) + \mathbf{f}(2)\mathbf{g}(2).$$

Answers to Test for Chapter 1 — Version A

(1) The net velocity vector is $[\frac{5\sqrt{2}}{2}, 2 - \frac{5\sqrt{2}}{2}] \approx [3.54, -1.54]$ km/hr. The net speed is approximately 3.85 km/hr.

(2) The angle (to the nearest degree) is 137°.

(3)
$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) + (\mathbf{x}-\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})\\ &= \mathbf{x}\cdot(\mathbf{x}+\mathbf{y}) + \mathbf{y}\cdot(\mathbf{x}+\mathbf{y}) + \mathbf{x}\cdot(\mathbf{x}-\mathbf{y}) - \mathbf{y}\cdot(\mathbf{x}-\mathbf{y})\\ &= \mathbf{x}\cdot\mathbf{x} + \mathbf{x}\cdot\mathbf{y} + \mathbf{y}\cdot\mathbf{x} + \mathbf{y}\cdot\mathbf{y} + \mathbf{x}\cdot\mathbf{x} - \mathbf{x}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{x} + \mathbf{y}\cdot\mathbf{y}\\ &= 2(\mathbf{x}\cdot\mathbf{x}) + 2(\mathbf{y}\cdot\mathbf{y}) = 2\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right)\end{aligned}$$

(4) The vector $\mathbf{x} = [-3, 4, 2]$ can be expressed as $[-\frac{65}{35}, \frac{13}{35}, -\frac{39}{35}] + [-\frac{40}{35}, \frac{127}{35}, \frac{109}{35}]$, where the first vector ($\mathbf{proj_a x}$) of the sum is parallel to $\mathbf{a}$ (since it equals $-\frac{13}{35}\mathbf{a}$), and the second vector ($\mathbf{x} - \mathbf{proj_a x}$) of the sum is easily seen to be orthogonal to $\mathbf{a}$.

(5) Contrapositive: If $\mathbf{x}$ is parallel to $\mathbf{y}$, then $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$.
Converse: If $\mathbf{x}$ is not parallel to $\mathbf{y}$, then $\|\mathbf{x}+\mathbf{y}\| \neq \|\mathbf{x}\| + \|\mathbf{y}\|$.
Inverse: If $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$, then $\mathbf{x}$ is parallel to $\mathbf{y}$.
Only the contrapositive is logically equivalent to the original statement.

(6) The negation is: $\|\mathbf{x}\| \neq \|\mathbf{y}\|$ and $(\mathbf{x}-\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) = 0$.

(7) Base Step: Assume $\mathbf{A}$ and $\mathbf{B}$ are any two diagonal $n \times n$ matrices. Then, $\mathbf{C} = \mathbf{A} + \mathbf{B}$ is also diagonal because for $i \neq j$, $c_{ij} = a_{ij} + b_{ij} = 0 + 0 = 0$, since $\mathbf{A}$ and $\mathbf{B}$ are both diagonal.

Inductive Step: Assume that the sum of any k diagonal $n \times n$ matrices is diagonal. We must show for any diagonal $n \times n$ matrices $\mathbf{A}_1, \ldots, \mathbf{A}_{k+1}$ that $\sum_{i=1}^{k+1} \mathbf{A}_i$ is diagonal. Now, $\sum_{i=1}^{k+1} \mathbf{A}_i = (\sum_{i=1}^{k} \mathbf{A}_i) + \mathbf{A}_{k+1}$. Let $\mathbf{B} = \sum_{i=1}^{k} \mathbf{A}_i$. Then $\mathbf{B}$ is diagonal by the inductive hypothesis. Hence, $\sum_{i=1}^{k+1} \mathbf{A}_i = \mathbf{B} + \mathbf{A}_{k+1}$ is the sum of two diagonal matrices, and so is diagonal by the Base Step.

(8) $\mathbf{A} = \mathbf{S} + \mathbf{V}$, where $\mathbf{S} = \begin{bmatrix} -4 & \frac{9}{2} & 0 \\ \frac{9}{2} & -1 & \frac{11}{2} \\ 0 & \frac{11}{2} & -3 \end{bmatrix}$, $\mathbf{V} = \begin{bmatrix} 0 & -\frac{3}{2} & -2 \\ \frac{3}{2} & 0 & \frac{3}{2} \\ 2 & -\frac{3}{2} & 0 \end{bmatrix}$, $\mathbf{S}$ is symmetric, and $\mathbf{V}$ is skew-symmetric.

(9)

	Salary	Perks
TV Show 1	\$1000000	\$650000
TV Show 2	\$860000	\$535000
TV Show 3	\$880000	\$455000
TV Show 4	\$770000	\$505000

(10) Assume $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ symmetric matrices. Then $\mathbf{AB}$ is symmetric if and only if $(\mathbf{AB})^T = \mathbf{AB}$ if and only if $\mathbf{B}^T\mathbf{A}^T = \mathbf{AB}$ (by Theorem 1.18) if and only if $\mathbf{BA} = \mathbf{AB}$ (since $\mathbf{A}$ and $\mathbf{B}$ are symmetric) if and only if $\mathbf{A}$ and $\mathbf{B}$ commute.

Answers to Test for Chapter 1 — Version B

(1) Player B is pulling with a force of $100/\sqrt{3}$ lbs, or about 57.735 lbs. Player C is pulling with a force of $200/\sqrt{3}$ lbs, or about 115.47 lbs.

(2) The angle (to the nearest degree) is 152°.

(3) See the proof in the textbook directly before Theorem 1.11 in Section 1.2.

(4) The vector $\mathbf{x} = [5, -2, -4]$ can be expressed as $[\frac{8}{22}, -\frac{12}{22}, \frac{12}{22}] + [\frac{102}{22}, -\frac{32}{22}, -\frac{100}{22}]$, where the first vector ($\mathbf{proj_a x}$) of the sum is parallel to $\mathbf{a}$ (since it equals $-\frac{4}{22}\mathbf{a}$), and the second vector ($\mathbf{x} - \mathbf{proj_a x}$) of the sum is easily seen to be orthogonal to $\mathbf{a}$.

(5) Contrapositive: If $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{y}\|$, then $\|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) \leq 0$.
Converse: If $\|\mathbf{x}+\mathbf{y}\| > \|\mathbf{y}\|$, then $\|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) > 0$.
Inverse: If $\|\mathbf{x}\|^2 + 2(\mathbf{x}\cdot\mathbf{y}) \leq 0$, then $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{y}\|$.
Only the contrapositive is logically equivalent to the original statement.

(6) The negation is: No unit vector $\mathbf{x} \in \mathbb{R}^3$ is parallel to $[-2, 3, 1]$.

(7) Base Step: Assume $\mathbf{A}$ and $\mathbf{B}$ are any two lower triangular $n \times n$ matrices. Then, $\mathbf{C} = \mathbf{A} + \mathbf{B}$ is also lower triangular because for $i < j$, $c_{ij} = a_{ij} + b_{ij} = 0 + 0 = 0$, since $\mathbf{A}$ and $\mathbf{B}$ are both lower triangular.

Inductive Step: Assume that the sum of any k lower triangular $n \times n$ matrices is lower triangular. We must show for any lower triangular $n \times n$ matrices $\mathbf{A}_1, \ldots, \mathbf{A}_{k+1}$ that $\sum_{i=1}^{k+1} \mathbf{A}_i$ is lower triangular. Now, $\sum_{i=1}^{k+1} \mathbf{A}_i = (\sum_{i=1}^{k} \mathbf{A}_i) + \mathbf{A}_{k+1}$. Let $\mathbf{B} = \sum_{i=1}^{k} \mathbf{A}_i$. Then $\mathbf{B}$ is lower triangular by the inductive hypothesis. Hence, $\sum_{i=1}^{k+1} \mathbf{A}_i = \mathbf{B} + \mathbf{A}_{k+1}$ is the sum of two lower triangular matrices, and so is lower triangular by the Base Step.

(8) $\mathbf{A} = \mathbf{S} + \mathbf{V}$, where $\mathbf{S} = \begin{bmatrix} 5 & 1 & \frac{7}{2} \\ 1 & 3 & \frac{1}{2} \\ \frac{7}{2} & \frac{1}{2} & 1 \end{bmatrix}$, $\mathbf{V} = \begin{bmatrix} 0 & 7 & -\frac{11}{2} \\ -7 & 0 & -\frac{7}{2} \\ \frac{11}{2} & \frac{7}{2} & 0 \end{bmatrix}$, $\mathbf{S}$ is symmetric, and $\mathbf{V}$ is skew-symmetric.

(9)

	Nutrient 1	Nutrient 2	Nutrient 3
Cat 1	1.64	0.62	1.86
Cat 2	1.50	0.53	1.46
Cat 3	1.55	0.52	1.54

(All figures are in ounces.)

(10) Assume $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ skew-symmetric matrices. Then $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ (by Theorem 1.18) $= (-\mathbf{B})(-\mathbf{A})$ (since $\mathbf{A}$, $\mathbf{B}$ are skew-symmetric) $= \mathbf{BA}$ (by part (4) of Theorem 1.16).

Answers to Test for Chapter 1 — Version C

(1) The acceleration vector is approximately $[-0.435, 0.870, -0.223]$ m/sec^2. This can also be expressed as $1.202[-0.362, 0.724, -0.186]$ m/sec^2, where the latter vector is a unit vector.

(2) The angle (to the nearest degree) is 118°.

(3) See the proof of Theorem 1.8 in Section 1.2 of the textbook.

(4) The vector $\mathbf{x} = [-4, 2, -7]$ can be expressed as $\left[-\frac{246}{62}, \frac{205}{62}, -\frac{41}{62}\right] + \left[-\frac{2}{62}, -\frac{81}{62}, -\frac{393}{62}\right]$, where the first vector ($\mathbf{proj_a x}$) of the sum is parallel to $\mathbf{a}$ (since it equals $-\frac{41}{62}\mathbf{a}$) and the second vector ($\mathbf{x} - \mathbf{proj_a x}$) of the sum is easily seen to be orthogonal to $\mathbf{a}$.

(5) Assume $\mathbf{y} = c\mathbf{x}$, for some $c \in \mathbb{R}$. We must prove that $\mathbf{y} = \mathbf{proj_x y}$. Now, if $\mathbf{y} = c\mathbf{x}$, then $\mathbf{proj_x y} = \left(\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\right)\mathbf{x} = \left(\frac{\mathbf{x}\cdot(c\mathbf{x})}{\|\mathbf{x}\|^2}\right)\mathbf{x} = \left(\frac{c(\mathbf{x}\cdot\mathbf{x})}{\|\mathbf{x}\|^2}\right)\mathbf{x}$ (by part (4) of Theorem 1.5) $= c\mathbf{x}$ (by part (2) of Theorem 1.5) $= \mathbf{y}$.

(6) The negation is: For some vector $\mathbf{x} \in \mathbb{R}^n$, there is no vector $\mathbf{y} \in \mathbb{R}^n$ such that $\|\mathbf{y}\| > \|\mathbf{x}\|$.

(7) $\mathbf{A} = \mathbf{S} + \mathbf{V}$, where $\mathbf{S} = \begin{bmatrix} -3 & \frac{13}{2} & \frac{1}{2} \\ \frac{13}{2} & 5 & 3 \\ \frac{1}{2} & 3 & -4 \end{bmatrix}$, $\mathbf{V} = \begin{bmatrix} 0 & -\frac{1}{2} & \frac{5}{2} \\ \frac{1}{2} & 0 & -1 \\ -\frac{5}{2} & 1 & 0 \end{bmatrix}$, $\mathbf{S}$ is symmetric, and $\mathbf{V}$ is skew-symmetric.

(8) The third row of $\mathbf{AB}$ is $[35, 18, -13]$, and the second column of $\mathbf{BA}$ is $[-32, 52, 42, -107]$.

(9) Base Step: Suppose $\mathbf{A}$ and $\mathbf{B}$ are two $n \times n$ diagonal matrices. Let $\mathbf{C} = \mathbf{AB}$. Then, $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. But if $k \neq i$, then $a_{ik} = 0$. Thus, $c_{ij} = a_{ii}b_{ij}$. Then, if $i \neq j$, $b_{ij} = 0$, hence $c_{ij} = a_{ii}0 = 0$. Hence $\mathbf{C}$ is diagonal.

Inductive Step: Assume that the product of any k diagonal $n \times n$ matrices is diagonal. We must show that for any diagonal $n \times n$ matrices $\mathbf{A}_1, \ldots, \mathbf{A}_{k+1}$ that the product $\mathbf{A}_1 \cdots \mathbf{A}_{k+1}$ is diagonal. Now, $\mathbf{A}_1 \cdots \mathbf{A}_{k+1} = (\mathbf{A}_1 \cdots \mathbf{A}_k)\mathbf{A}_{k+1}$. Let $\mathbf{B} = \mathbf{A}_1 \cdots \mathbf{A}_k$. Then $\mathbf{B}$ is diagonal by the inductive hypothesis. Hence, $\mathbf{A}_1 \cdots \mathbf{A}_k\mathbf{A}_{k+1} = \mathbf{B}\mathbf{A}_{k+1}$ is the product of two diagonal matrices, and so is diagonal by the Base Step.

(10) Assume that $\mathbf{A}$ is skew-symmetric. Then $(\mathbf{A}^3)^T = (\mathbf{A}^2\mathbf{A})^T = \mathbf{A}^T(\mathbf{A}^2)^T$ (by Theorem 1.18) $= \mathbf{A}^T(\mathbf{AA})^T = \mathbf{A}^T\mathbf{A}^T\mathbf{A}^T$ (again by Theorem 1.18) $= (-\mathbf{A})(-\mathbf{A})(-\mathbf{A})$ (since $\mathbf{A}$ is skew-symmetric) $= -(\mathbf{A}^3)$ (by part (4) of Theorem 1.16). Hence, $\mathbf{A}^3$ is skew-symmetric.

Answers to Test for Chapter 2 — Version A

(1) The quadratic equation is $y = 3x^2 - 4x - 8$.

(2) The solution set is $\{(-2b - 4d - 5,\, b,\, 3d + 2,\, d,\, 4) \mid b, d \in \mathbb{R}\}$.

(3) The solution set is $\{c(-2, 4, 1, 0) \mid c \in \mathbb{R}\}$.

(4) The rank of $\mathbf{A}$ is 2, and so $\mathbf{A}$ is not row equivalent to $\mathbf{I}_3$.

(5) The solution sets are, respectively, $\{(3, -4, 2)\}$ and $\{(1, -2, 3)\}$.

(6) First, ith row of $R(\mathbf{AB})$
$= c(i\text{th row of } (\mathbf{AB}))$
$= c(i\text{th row of } \mathbf{A})\mathbf{B}$ (by the hint)
$= (i\text{th row of } R(\mathbf{A}))\mathbf{B}$
$= i\text{th row of } (R(\mathbf{A})\mathbf{B})$ (by the hint).
Now, if $k \neq i$, kth row of $R(\mathbf{AB})$
$= k\text{th row of } (\mathbf{AB})$
$= (k\text{th row of } \mathbf{A})\mathbf{B}$ (by the hint)
$= (k\text{th row of } R(\mathbf{A}))\mathbf{B}$
$= k\text{th row of } (R(\mathbf{A})\mathbf{B})$ (by the hint).
Hence, $R(\mathbf{AB}) = R(\mathbf{A})\mathbf{B}$, since they are equal on each row.

(7) Yes: $[-15, 10, -23] = (-2)\,(\text{row } 1) + (-3)\,(\text{row } 2) + (-1)\,(\text{row } 3)$.

(8) The inverse of $\mathbf{A}$ is $\begin{bmatrix} -9 & -4 & 15 \\ -4 & -2 & 7 \\ 12 & 5 & -19 \end{bmatrix}$.

(9) The inverse of $\mathbf{A}$ is $-\frac{1}{14}\begin{bmatrix} -9 & 8 \\ -5 & 6 \end{bmatrix} = \begin{bmatrix} \frac{9}{14} & -\frac{4}{7} \\ \frac{5}{14} & -\frac{3}{7} \end{bmatrix}$.

(10) Suppose $\mathbf{A}$ and $\mathbf{B}$ are nonsingular $n \times n$ matrices. If $\mathbf{A}$ and $\mathbf{B}$ commute, then $\mathbf{AB} = \mathbf{BA}$. Hence, $\mathbf{A}(\mathbf{AB})\mathbf{B} = \mathbf{A}(\mathbf{BA})\mathbf{B}$, and so $\mathbf{A}^2\mathbf{B}^2 = (\mathbf{AB})^2$.

Conversely, suppose $\mathbf{A}^2\mathbf{B}^2 = (\mathbf{AB})^2$. Since both $\mathbf{A}$ and $\mathbf{B}$ are nonsingular, we know that $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ exist. Multiply both sides of $\mathbf{A}^2\mathbf{B}^2 = \mathbf{ABAB}$ by $\mathbf{A}^{-1}$ on the left and by $\mathbf{B}^{-1}$ on the right to obtain $\mathbf{AB} = \mathbf{BA}$. Hence, $\mathbf{A}$ and $\mathbf{B}$ commute.

Answers to Test for Chapter 2 — Version B

(1) The equation of the circle is $x^2 + y^2 - 6x + 8y + 12 = 0$, or, $(x-3)^2 + (y+4)^2 = 13$.

(2) The solution set is $\{(-5c - 3d + f + 6, 2c + d - 3f - 2, c, d, 2f + 1, f) \mid c, d, f \in \mathbb{R}\}$.

(3) The solution set is $\{b(3,1,0,0) + d(-4,0,-2,1) \mid b, d \in \mathbb{R}\}$.

(4) The rank of $\mathbf{A}$ is 3, and so $\mathbf{A}$ is row equivalent to $\mathbf{I}_3$.

(5) The solution sets are, respectively, $\{(-2,3,-4)\}$ and $\{(3,-1,2)\}$.

(6) First, ith row of $R(\mathbf{AB})$
$= c(j\text{th row of }(\mathbf{AB})) + (i\text{th row of }(\mathbf{AB}))$
$= c(j\text{th row of }\mathbf{A})\mathbf{B} + (i\text{th row of }\mathbf{A})\mathbf{B}$ (by the hint)
$= (c(j\text{th row of }\mathbf{A}) + (i\text{th row of }\mathbf{A}))\mathbf{B}$
$= (i\text{th row of }R(\mathbf{A}))\mathbf{B}$
$= i\text{th row of }(R(\mathbf{A})\mathbf{B})$ (by the hint).
Now, if $k \neq i$, kth row of $R(\mathbf{AB})$
$= k\text{th row of }(\mathbf{AB})$
$= (k\text{th row of }\mathbf{A})\mathbf{B}$ (by the hint)
$= (k\text{th row of }R(\mathbf{A}))\mathbf{B}$
$= k\text{th row of }(R(\mathbf{A})\mathbf{B})$ (by the hint).
Hence, $R(\mathbf{AB}) = R(\mathbf{A})\mathbf{B}$, since they are equal on each row.

(7) The vector $[-2,3,1]$ is *not* in the row space of $\mathbf{A}$.

(8) The inverse of $\mathbf{A}$ is $\begin{bmatrix} 1 & -6 & 1 \\ -2 & 13 & -2 \\ 4 & -21 & 3 \end{bmatrix}$.

(9) The solution set is $\{(4,-7,-2)\}$.

(10) (a) To prove that the inverse of $(c\mathbf{A})$ is $(\frac{1}{c})\mathbf{A}^{-1}$, simply note that the product $(c\mathbf{A})$ times $(\frac{1}{c})\mathbf{A}^{-1}$ equals $\mathbf{I}$ (by part (4) of Theorem 1.16).

(b) Now, assume $\mathbf{A}$ is a nonsingular skew-symmetric matrix. Then,

$$\begin{aligned} (\mathbf{A}^{-1})^T &= (\mathbf{A}^T)^{-1} && \text{by part (4) of Theorem 2.12} \\ &= (-1\mathbf{A})^{-1} && \text{since } \mathbf{A} \text{ is skew-symmetric} \\ &= (\tfrac{1}{-1})\mathbf{A}^{-1} && \text{by part (a)} \\ &= -(\mathbf{A}^{-1}). \end{aligned}$$

Hence, $\mathbf{A}^{-1}$ is skew-symmetric.

Answers to Test for Chapter 2 — Version C

(1) $A = 3$, $B = -2$, $C = 1$, and $D = 4$.

(2) The solution set is $\{(-4b - 3e + 3, b, 2e - 3, 5e - 2, e, 4) \,|\, b, e \in \mathbb{R}\}$.

(3) The solution set is $\{c(12, -5, 1, 0, 0) + e(-1, -1, 0, 3, 1) \,|\, c, e \in \mathbb{R}\}$.

(4) The rank of $\mathbf{A}$ is 3, and so $\mathbf{A}$ is not row equivalent to $\mathbf{I}_4$.

(5) The reduced row echelon form matrix for $\mathbf{A}$ is $\mathbf{B} = \mathbf{I}_3$. One possible sequence of row operations converting $\mathbf{B}$ to $\mathbf{A}$ is:

(II): $< 2 > \leftarrow \frac{11}{10} < 3 > + < 2 >$
(II): $< 1 > \leftarrow \frac{2}{5} < 3 > + < 1 >$
(I): $< 3 > \leftarrow -\frac{1}{10} < 3 >$
(II): $< 3 > \leftarrow -4 < 2 > + < 3 >$
(II): $< 1 > \leftarrow 1 < 2 > + < 1 >$
(I): $< 2 > \leftarrow 5 < 2 >$
(II): $< 3 > \leftarrow 5 < 1 > + < 3 >$
(II): $< 2 > \leftarrow -3 < 1 > + < 2 >$
(I): $< 1 > \leftarrow 2 < 1 >$

(6) Yes: $[25, 12, -19] = 4[7, 3, 5] + 2[3, 3, 4] - 3[3, 2, 3]$.

(7) The inverse of $\mathbf{A}$ is $\begin{bmatrix} 1 & 3 & -4 & 1 \\ 3 & 21 & -34 & 8 \\ 4 & 14 & -21 & 5 \\ 1 & 8 & -13 & 3 \end{bmatrix}$.

(8) The solution set is $\{(4, -6, -5)\}$.

(9) Base Step: Assume $\mathbf{A}$ and $\mathbf{B}$ are two $n \times n$ nonsingular matrices. Then $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, by part (3) of Theorem 2.12.

Inductive Step: Assume that for any set of k nonsingular $n \times n$ matrices, the inverse of their product is found by multiplying the inverses of the matrices in reverse order. We must show that if $\mathbf{A}_1, \ldots, \mathbf{A}_{k+1}$ are nonsingular $n \times n$ matrices, then

$$(\mathbf{A}_1 \cdots \mathbf{A}_{k+1})^{-1} = (\mathbf{A}_{k+1})^{-1} \cdots (\mathbf{A}_1)^{-1}.$$

Now, let $\mathbf{B} = \mathbf{A}_1 \cdots \mathbf{A}_k$. Then by the Base Step, or by part (3) of Theorem 2.12,

$$\begin{aligned} (\mathbf{A}_1 \cdots \mathbf{A}_{k+1})^{-1} &= (\mathbf{B}\mathbf{A}_{k+1})^{-1} = (\mathbf{A}_{k+1})^{-1}\mathbf{B}^{-1} \\ &= (\mathbf{A}_{k+1})^{-1}(\mathbf{A}_1 \cdots \mathbf{A}_k)^{-1} \\ &= (\mathbf{A}_{k+1})^{-1}(\mathbf{A}_k)^{-1} \cdots (\mathbf{A}_1)^{-1} \end{aligned}$$

by the inductive hypothesis.

(10) Now, $(R_1(R_2(\cdots(R_k(\mathbf{A}))\cdots)))\mathbf{B} = R_1(R_2(\cdots(R_k(\mathbf{AB}))\cdots))$ (by part (2) of Theorem 2.1) $= \mathbf{I}_n$ (given). Hence $R_1(R_2(\cdots(R_k(\mathbf{A}))\cdots)) = \mathbf{B}^{-1}$.

Answers for Test for Chapter 3 — Version A

(1) The area of the parallelogram is 24 square units.

(2) The determinant of $6\mathbf{A}$ is $6^5(4) = 31104$.

(3) The determinant of $\mathbf{A}$ is -2.

(4) Note that $|\mathbf{AB}^T| = |\mathbf{A}||\mathbf{B}^T|$ (by Theorem 3.7) $= |\mathbf{A}^T||\mathbf{B}|$ (by Theorem 3.10) $= |\mathbf{A}^T\mathbf{B}|$ (again by Theorem 3.7).

(5) The determinant of $\mathbf{A}$ is 1070.

(6) If $\mathbf{A}$ and $\mathbf{B}^{-1}$ are similar square matrices, there is some nonsingular matrix $\mathbf{P}$ such that $\mathbf{B}^{-1} = \mathbf{P}^{-1}\mathbf{AP}$. Then, $|\mathbf{B}^{-1}| = |\mathbf{P}^{-1}\mathbf{AP}|$. Since

$$|\mathbf{B}^{-1}| = \frac{1}{|\mathbf{B}|},$$

by Corollary 3.8, and since

$$|\mathbf{P}^{-1}\mathbf{AP}| = |\mathbf{P}^{-1}||\mathbf{A}||\mathbf{P}| \text{ (by Theorem 3.7) } = \frac{1}{|\mathbf{P}|}|\mathbf{A}||\mathbf{P}| = (\frac{1}{|\mathbf{P}|}|\mathbf{P}|)|\mathbf{A}| = |\mathbf{A}|,$$

we have

$$\frac{1}{|\mathbf{B}|} = |\mathbf{A}|,$$

which means $|\mathbf{A}||\mathbf{B}| = 1$.

(7) The solution set is $\{(-2, 3, 5)\}$.

(8) (Optional hint: $p_{\mathbf{A}}(x) = x^3 - x$.) Answer: $\mathbf{P} = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(9) $\lambda = 0$ is an eigenvalue for $\mathbf{A} \iff p_{\mathbf{A}}(0) = 0 \iff |0\mathbf{I}_n - \mathbf{A}| = 0 \iff |-\mathbf{A}| = 0 \iff (-1)^n|\mathbf{A}| = 0 \iff |\mathbf{A}| = 0 \iff \mathbf{A}$ is singular.

(10) Let $\mathbf{X}$ be an eigenvector for $\mathbf{A}$ corresponding to the eigenvalue $\lambda_1 = 2$. Hence, $\mathbf{AX} = 2\mathbf{X}$. Therefore,

$$\mathbf{A}^3\mathbf{X} = \mathbf{A}^2(\mathbf{AX}) = \mathbf{A}^2(2\mathbf{X}) = 2(\mathbf{A}^2\mathbf{X}) = 2\mathbf{A}(\mathbf{AX}) = 2\mathbf{A}(2\mathbf{X}) = 4\mathbf{AX} = 4(2\mathbf{X}) = 8\mathbf{X}.$$

This shows that $\mathbf{X}$ is an eigenvector corresponding to the eigenvalue $\lambda_2 = 8$ for $\mathbf{A}^3$.

Answers for Test for Chapter 3 — Version B

(1) The volume of the parallelepiped is 59 cubic units.

(2) The determinant of $5\mathbf{A}$ is $5^4(7) = 4375$.

(3) The determinant of $\mathbf{A}$ is -1.

(4) Assume $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices and that $\mathbf{AB} = -\mathbf{BA}$, with n odd. Then, $|\mathbf{AB}| = |-\mathbf{BA}|$. Now, $|-\mathbf{BA}| = (-1)^n|\mathbf{BA}|$ (by Corollary 3.4) $= -|\mathbf{BA}|$, since n is odd. Hence, $|\mathbf{AB}| = -|\mathbf{BA}|$. By Theorem 3.7, this means that $|\mathbf{A}| \cdot |\mathbf{B}| = -|\mathbf{B}| \cdot |\mathbf{A}|$. Since $|\mathbf{A}|$ and $|\mathbf{B}|$ are real numbers, this can only be true if either $|\mathbf{A}|$ or $|\mathbf{B}|$ equals zero. Hence, either $\mathbf{A}$ or $\mathbf{B}$ is singular (by Theorem 3.5).

(5) The determinant of $\mathbf{A}$ is -916.

(6) Case 1: Suppose $k \neq i$. Then, the (kth row of $(C(\mathbf{A}))^T$) $=$ kth column of $C(\mathbf{A})$ $=$ kth column of $\mathbf{A}$ (since $k \neq i$, $\mathbf{C}$ does not affect the kth column of any matrix) $=$ kth row of $\mathbf{A}^T$ $=$ kth row of $R\left(\mathbf{A}^T\right)$ (since $k \neq i$, R does not affect the kth row of any matrix).
Case 2: Consider the ith row of $(C(\mathbf{A}))^T$. The ith row of $(C(\mathbf{A}))^T$ $=$ ith column of $C(\mathbf{A})$ $= b$(jth column of $\mathbf{A}$) $+$ (ith column of $\mathbf{A}$) $= b$(jth row of $\mathbf{A}^T$) $+$ (ith row of $\mathbf{A}^T$) $=$ ith row of $R\left(\mathbf{A}^T\right)$.
Hence, every row of $(C(\mathbf{A}))^T$ equals the corresponding row of $R\left(\mathbf{A}^T\right)$, and so the matrices are equal.

(7) The solution set is $\{(-3, -2, 2)\}$.

(8) Optional hint: $p_{\mathbf{A}}(x) = x^3 - 2x^2 + x$. Answer: $\mathbf{P} = \begin{bmatrix} -1 & 3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(9) (a) Because $\mathbf{A}$ is upper triangular, we can easily see that $p_{\mathbf{A}}(x) = (x-2)^4$. Hence $\lambda = 2$ is the only eigenvalue for $\mathbf{A}$.

(b) The reduced row echelon form for $2\mathbf{I}_4 - \mathbf{A}$ is $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. There is only 1 non-pivot column (column 1), and so Step 3 of the Diagonalization Method will produce only 1 fundamental eigenvector. That fundamental eigenvector is $[1, 0, 0, 0]$.

(10) From the algebraic perspective, $p_{\mathbf{A}}(x) = x^2 - \frac{6}{5}x + 1$, which has no roots, since its discriminant is $-\frac{64}{25}$. Hence, $\mathbf{A}$ has no eigenvalues, and so $\mathbf{A}$ can not be diagonalized. From the geometric perspective, for any nonzero vector $\mathbf{X}$, $\mathbf{AX}$ will be rotated away from $\mathbf{X}$ through an angle θ. Hence, $\mathbf{X}$ and $\mathbf{AX}$ can not be parallel. Thus, $\mathbf{A}$ can not have any eigenvectors. This makes $\mathbf{A}$ nondiagonalizable.

Answers for Test for Chapter 3 — Version C

(1) The volume of the parallelepiped is 2 cubic units.

(2) The determinant of $\mathbf{A}$ is 3.

(3) One possibility: $|\mathbf{A}| = 5\begin{vmatrix} 0 & 12 & 0 & 9 \\ 0 & 0 & 0 & 8 \\ 0 & 14 & 11 & 7 \\ 15 & 13 & 10 & 6 \end{vmatrix} = 5(8)\begin{vmatrix} 0 & 12 & 0 \\ 0 & 14 & 11 \\ 15 & 13 & 10 \end{vmatrix} = 5(8)(-12)\begin{vmatrix} 0 & 11 \\ 15 & 10 \end{vmatrix}$ $= 5(8)(-12)(-11)\,|15| = 79200$. Cofactor expansions were used along the first row in each case, except for the 4×4 matrix, where the cofactor expansion was along the second row.

(4) Assume that $\mathbf{A}^T = \mathbf{A}^{-1}$. Then $|\mathbf{A}^T| = |\mathbf{A}^{-1}|$. By Corollary 3.8 and Theorem 3.10, we have $|\mathbf{A}| = 1/|\mathbf{A}|$. Hence, $|\mathbf{A}|^2 = 1$, and so $|\mathbf{A}| = \pm 1$.

(5) By the hint, $C(\mathbf{AB}) = \left(R\left((\mathbf{AB})^T\right)\right)^T = \left(R\left(\mathbf{B}^T\mathbf{A}^T\right)\right)^T$ (by Theorem 1.18) $= \left(R\left(\mathbf{B}^T\right)\mathbf{A}^T\right)^T$ (by part (1) of Theorem 2.1) $= \mathbf{A}\left(R\left(\mathbf{B}^T\right)\right)^T$ (by Theorem 1.18) $= \mathbf{A}\,(C(\mathbf{B}))$ (by the hint).

(6) Base Step ($n = 2$): Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Note that $|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$. Since $\mathbf{A}$ has only two rows, there are only two possible forms for the row operation R.

Case 1: $R = \langle 1\rangle \leftarrow c\langle 2\rangle + \langle 1\rangle$. Then $R(\mathbf{A}) = \begin{bmatrix} a_{11} + ca_{21} & a_{12} + ca_{22} \\ a_{21} & a_{22} \end{bmatrix}$. Hence $|R(\mathbf{A})| = (a_{11} + ca_{21})a_{22} - (a_{12} + ca_{22})a_{21} = a_{11}a_{22} + ca_{21}a_{22} - a_{12}a_{21} - ca_{22}a_{21} = a_{11}a_{22} - a_{12}a_{21} = |\mathbf{A}|$.

Case 2: $R = \langle 2\rangle \leftarrow c\langle 1\rangle + \langle 2\rangle$. Then $R(\mathbf{A}) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} + ca_{11} & a_{22} + ca_{12} \end{bmatrix}$. Hence $|R(\mathbf{A})| = a_{11}(a_{22} + ca_{12}) - a_{12}(a_{21} + ca_{11}) = a_{11}a_{22} + a_{11}ca_{12} - a_{12}a_{21} - a_{12}ca_{11} = a_{11}a_{22} - a_{12}a_{21} = |\mathbf{A}|$.

Therefore $|\mathbf{A}| = |R(\mathbf{A})|$ in all possible cases when $n = 2$, thus completing the Base Step.

(7) The solution set is $\{(3, -3, 2)\}$.

(8) If $\mathbf{P} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$, then $\mathbf{A} = \mathbf{P}\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}\mathbf{P}^{-1}$. Thus,

$$\mathbf{A}^9 = \mathbf{P}\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^9\mathbf{P}^{-1} = \mathbf{P}\begin{bmatrix} -1 & 0 \\ 0 & 512 \end{bmatrix}\mathbf{P}^{-1} = \begin{bmatrix} 1025 & -1026 \\ 513 & -514 \end{bmatrix}.$$

(9) (Optional hint: $p_{\mathbf{A}}(x) = (x+1)(x-2)^2$.) Answer: $\mathbf{P} = \begin{bmatrix} -2 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

(10) We will show that $p_{\mathbf{A}}(x) = p_{\mathbf{A}^T}(x)$. (Then the linear factor $(x - \lambda)$ will appear exactly k times in $p_{\mathbf{A}^T}(x)$ since it appears exactly k times in $p_{\mathbf{A}}(x)$.) Now, $p_{\mathbf{A}^T}(x) = |x\mathbf{I}_n - \mathbf{A}^T| = |(x\mathbf{I}_n)^T - \mathbf{A}^T|$ (since $(x\mathbf{I}_n)$ is diagonal, hence symmetric) $= |(x\mathbf{I}_n - \mathbf{A})^T|$ (by part (2) of Theorem 1.13) $= |x\mathbf{I}_n - \mathbf{A}|$ (by Theorem 3.10) $= p_{\mathbf{A}}(x)$.

Answers for Test for Chapter 4 — Version A

(1) Clearly the operation $\oplus$ is commutative. Also, $\oplus$ is associative because

$$\begin{aligned}(\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z} &= (x^5 + y^5)^{1/5} \oplus \mathbf{z} = (((x^5 + y^5)^{1/5})^5 + z^5)^{1/5} \\ &= ((x^5 + y^5) + z^5)^{1/5} = (x^5 + (y^5 + z^5))^{1/5} \\ &= (x^5 + ((y^5 + z^5)^{1/5})^5)^{1/5} = \mathbf{x} \oplus (y^5 + z^5)^{1/5} \\ &= \mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z}).\end{aligned}$$

Clearly, the real number 0 acts as the additive identity, and, for any real number x, the additive inverse of x is $-x$, because $(x^5 + (-x)^5)^{1/5} = 0^{1/5} = 0$, the additive identity.

Also, the first distributive law holds because

$$\begin{aligned}a \odot (\mathbf{x} \oplus \mathbf{y}) &= a \odot (x^5 + y^5)^{1/5} = a^{1/5}(x^5 + y^5)^{1/5} \\ &= (a(x^5 + y^5))^{1/5} = (ax^5 + ay^5)^{1/5} \\ &= ((a^{1/5}x)^5 + (a^{1/5}y)^5)^{1/5} = (a^{1/5}x) \oplus (a^{1/5}y) \\ &= (a \odot \mathbf{x}) \oplus (a \odot \mathbf{y}).\end{aligned}$$

Similarly, the other distributive law holds because

$$\begin{aligned}(a + b) \odot \mathbf{x} &= (a + b)^{1/5}x = (a + b)^{1/5}(x^5)^{1/5} \\ &= (ax^5 + bx^5)^{1/5} = ((a^{1/5}x)^5 + (b^{1/5}x)^5)^{1/5} \\ &= (a^{1/5}x) \oplus (b^{1/5}x) = (a \odot \mathbf{x}) \oplus (b \odot \mathbf{x}).\end{aligned}$$

Associativity of scalar multiplication holds because

$$(ab) \odot \mathbf{x} = (ab)^{1/5}x = a^{1/5}b^{1/5}x = a^{1/5}(b^{1/5}x) = a \odot (b^{1/5}x) = a \odot (b \odot \mathbf{x}).$$

Finally, $1 \odot \mathbf{x} = 1^{1/5}x = 1x = \mathbf{x}$.

(2) Since the set of $n \times n$ skew-symmetric matrices is nonempty, for any $n \geq 1$, it is enough to prove closure under addition and scalar multiplication. That is, we must show that if $\mathbf{A}$ and $\mathbf{B}$ are any two $n \times n$ skew-symmetric matrices, and if c is any real number, then $\mathbf{A} + \mathbf{B}$ is skew-symmetric, and $c\mathbf{A}$ is skew-symmetric. However, $\mathbf{A} + \mathbf{B}$ is skew-symmetric since $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$ (by part (2) of Theorem 1.13) $= -\mathbf{A} - \mathbf{B}$ (since $\mathbf{A}$ and $\mathbf{B}$ are skew-symmetric) $= -(\mathbf{A} + \mathbf{B})$. Similarly, $c\mathbf{A}$ is skew-symmetric because $(c\mathbf{A})^T = c\mathbf{A}^T$ (by part (3) of Theorem 1.13) $= c(-\mathbf{A}) = -c\mathbf{A}$.

(3) Form the matrix $\mathbf{A}$ whose rows are the given vectors: $\begin{bmatrix} 5 & -2 & -1 \\ 2 & -1 & -1 \\ 8 & -2 & 1 \end{bmatrix}$. It is easily shown that $\mathbf{A}$ row reduces to $\mathbf{I}_3$. Hence, the given vectors span $\mathbb{R}^3$. (Alternatively, the vectors span $\mathbb{R}^3$ since $|\mathbf{A}|$ is nonzero.)

(4) A simplified form for the vectors in span(S) is:

$$\{a(x^3 - 5x^2 - 2) + b(x + 3) \,|\, a, b \in \mathbb{R}\} = \{ax^3 - 5ax^2 + bx + (-2a + 3b) \,|\, a, b \in \mathbb{R}\}.$$

(5) Use the Independence Test Method (in Section 4.4). Note that the matrix whose columns are the given vectors row reduces to $\mathbf{I}_4$. Thus, there is a pivot in every column, and the vectors are linearly independent.

(6) The given set is a basis for $\mathcal{M}_{22}$. Since the set contains four vectors, it is only necessary to check either that it spans $\mathcal{M}_{22}$ or that it is linearly independent (see Theorem 4.12). To check for linear independence, form the matrix whose columns are the entries of the given matrices. (Be sure to take the entries from each matrix in the same order.) Since this matrix row reduces to $\mathbf{I}_4$, the set of vectors is linearly independent, hence a basis for $\mathcal{M}_{22}$.

(7) Using the Independence Test Method, as explained in Section 4.6, produces the basis

$$B = \{3x^3 - 2x^2 + 3x - 1, x^3 - 3x^2 + 2x - 1, 11x^3 - 12x^2 + 13x + 3\}$$

for span(S). Since B has 3 elements, $\dim(\text{span}(S)) = 3$.

(8) Let S be the set of columns of $\mathbf{A}$. Because $\mathbf{A}$ is nonsingular, it row reduces to $\mathbf{I}_n$. Hence, the Independence Test Method applied to S shows that S is linearly independent. Thus, S itself is a basis for span(S). Since S has n elements, $\dim(\text{span}(S)) = n$, and so span$(S) = \mathbb{R}^n$, by Theorem 4.13. Hence S spans $\mathbb{R}^n$.

(9) One possibility is $B = \{[2, 1, 0], [1, 0, 0], [0, 0, 1]\}$.

(10) (a) The transition matrix from B to C is $\begin{bmatrix} -2 & 3 & -1 \\ 4 & -3 & 2 \\ -3 & -2 & -4 \end{bmatrix}$.

(b) For $\mathbf{v} = [-63, -113, 62]$, $[\mathbf{v}]_B = [3, -2, 2]$, and $[\mathbf{v}]_C = [-14, 22, -13]$.

Answers for Test for Chapter 4 — Version B

(1) The operation $\oplus$ is not associative in general because

$$(\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z} = (3(x+y)) \oplus \mathbf{z} = 3(3(x+y)) + z) = 9x + 9y + 3z,$$

while

$$\mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z}) = \mathbf{x} \oplus (3(y+z)) = 3(x + 3(y+z)) = 3x + 9y + 9z$$

instead. For a particular counterexample, $(1 \oplus 2) \oplus 3 = 9 \oplus 3 = 30$, but $1 \oplus (2 \oplus 3) = 1 \oplus 15 = 48$.

(2) The set of all polynomials in $\mathcal{P}_4$ for which the coefficient of the second-degree term equals the coefficient of the fourth-degree term has the form $\{ax^4 + bx^3 + ax^2 + cx + d\}$. To show this set is a subspace of $\mathcal{P}_4$, it is enough to show that it is closed under addition and scalar multiplication, as it is clearly nonempty. Clearly, this set is closed under addition because

$$\begin{aligned} & (ax^4 + bx^3 + ax^2 + cx + d) + (ex^4 + fx^3 + ex^2 + gx + h) \\ = \; & (a+e)x^4 + (b+f)x^3 + (a+e)x^2 + (c+g)x + (d+h), \end{aligned}$$

which has the correct form because this latter polynomial has the coefficients of its second-degree and fourth-degree terms equal. Similarly,

$$k(ax^4 + bx^3 + ax^2 + cx + d) = (ka)x^4 + (kb)x^3 + (ka)x^2 + (kc)x + (kd),$$

which also has the coefficients of its second-degree and fourth-degree terms equal.

(3) Let $\mathcal{V}$ be a vector space with subspaces $\mathcal{W}_1$ and $\mathcal{W}_2$. Let $\mathcal{W} = \mathcal{W}_1 \cap \mathcal{W}_2$. First, $\mathcal{W}$ is clearly nonempty because $\mathbf{0} \in \mathcal{W}$ since $\mathbf{0}$ must be in both $\mathcal{W}_1$ and $\mathcal{W}_2$ due to the fact that they are subspaces. Thus, to show that $\mathcal{W}$ is a subspace of $\mathcal{V}$, it is enough to show that $\mathcal{W}$ is closed under addition and scalar multiplication. Let $\mathbf{x}$ and $\mathbf{y}$ be elements of $\mathcal{W}$, and let c be any real number. Then, $\mathbf{x}$ and $\mathbf{y}$ are elements of both $\mathcal{W}_1$ and $\mathcal{W}_2$, and since $\mathcal{W}_1$ and $\mathcal{W}_2$ are both subspaces of $\mathcal{V}$, hence closed under addition, we know that $\mathbf{x}+\mathbf{y}$ is in both $\mathcal{W}_1$ and $\mathcal{W}_2$. Thus $\mathbf{x}+\mathbf{y}$ is in $\mathcal{W}$. Similarly, since $\mathcal{W}_1$ and $\mathcal{W}_2$ are both subspaces of $\mathcal{V}$, then $c\mathbf{x}$ is in both $\mathcal{W}_1$ and $\mathcal{W}_2$, and hence $c\mathbf{x}$ is in $\mathcal{W}$.

(4) The vector $[2, -4, 3]$ is *not* in span(S). Note that

$$\left[\begin{array}{rrr|r} 5 & -2 & -9 & 2 \\ -15 & 6 & 27 & -4 \\ -4 & 1 & 7 & 3 \end{array}\right] \text{ row reduces to } \left[\begin{array}{rrr|r} 1 & 0 & -\frac{5}{3} & -\frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{23}{3} \\ 0 & 0 & 0 & 2 \end{array}\right].$$

(5) A simplified form for the vectors in span(S) is:

$$\left\{ a\begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \;\middle|\; a, b, c \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} a & -3a \\ b & c \end{bmatrix} \;\middle|\; a, b, c \in \mathbb{R} \right\}.$$

(6) Form the matrix $\mathbf{A}$ whose columns are the coefficients of the polynomials in S. Row reducing $\mathbf{A}$ shows that there is a pivot in every column. Hence S is linearly independent by the Independence Test Method in Section 4.4. Hence, since S has four elements and $\dim(\mathcal{P}_3) = 4$, Theorem 4.12 shows that S is a basis for $\mathcal{P}_3$. Therefore, S spans $\mathcal{P}_3$, showing that $\mathbf{p}(x)$ can be expressed as a linear combination of the elements in S. Finally, since S is linearly independent, the uniqueness assertion is true by Theorem 4.9.

(7) A basis for span(S) is
$$\{[2,-3,4,-1],[3,1,-2,2],[2,8,-12,3]\}.$$
Hence, $\dim(\text{span}(S)) = 3$. Since $\dim(\mathbb{R}^4) = 4$, S does not span $\mathbb{R}^4$.

(8) Since $\mathbf{A}$ is nonsingular, $\mathbf{A}^T$ is also nonsingular (by part (4) of Theorem 2.12). Therefore $\mathbf{A}^T$ row reduces to $\mathbf{I}_n$, giving a pivot in every column. The Independence Test Method from Section 4.4 now shows the columns of $\mathbf{A}^T$, and hence the rows of $\mathbf{A}$, are linearly independent.

(9) One possible answer: $B = \{[2,1,0,0], [-3,0,1,0], [1,0,0,0], [0,0,0,1]\}$

(10) (a) The transition matrix from B to C is $\begin{bmatrix} 3 & -1 & 1 \\ 2 & 4 & -3 \\ -4 & 2 & 3 \end{bmatrix}$.

(b) For $\mathbf{v} = -100x^2 + 64x - 181$, $[\mathbf{v}]_B = [2,-4,-1]$, and $[\mathbf{v}]_C = [9,-9,-19]$.

Answers for Test for Chapter 4 — Version C

(1) From part (2) of Theorem 4.1, we know $0\mathbf{v} = \mathbf{0}$ in any general vector space. Thus,

$$0 \odot [x, y] = [0x + 3(0) - 3, 0y - 4(0) + 4] = [-3, 4]$$

is the additive identity $\mathbf{0}$ in this given vector space. Similarly, by part (3) of Theorem 4.1, $(-1)\mathbf{v}$ is the additive inverse of $\mathbf{v}$. Hence, the additive inverse of $[x, y]$ is

$$(-1) \odot [x, y] = [-x - 3 - 3, -y + 4 + 4] = [-x - 6, -y + 8].$$

(2) Let $\mathcal{W}$ be the set of all 3-vectors orthogonal to $[2, -3, 1]$. First, $[0, 0, 0] \in \mathcal{W}$ because

$$[0, 0, 0] \cdot [2, -3, 1] = 0.$$

Hence $\mathcal{W}$ is nonempty. Thus, to show that $\mathcal{W}$ is a subspace of $\mathbb{R}^3$, it is enough to show that $\mathcal{W}$ is closed under addition and scalar multiplication. Let $\mathbf{x}, \mathbf{y} \in \mathcal{W}$; that is, let $\mathbf{x}$ and $\mathbf{y}$ both be orthogonal to $[2, -3, 1]$. Then, $\mathbf{x} + \mathbf{y}$ is also orthogonal to $[2, -3, 1]$, because

$$(\mathbf{x} + \mathbf{y}) \cdot [2, -3, 1] = (\mathbf{x} \cdot [2, -3, 1]) + (\mathbf{y} \cdot [2, -3, 1]) = 0 + 0 = 0.$$

Similarly, if c is any real number, and $\mathbf{x}$ is in $\mathcal{W}$, then $c\mathbf{x}$ is also in $\mathcal{W}$ because

$$\begin{aligned} (c\mathbf{x}) \cdot [2, -3, 1] &= c(\mathbf{x} \cdot [2, -3, 1]) \quad \text{(by part (4) of Theorem 1.5)} \\ &= c(0) = 0. \end{aligned}$$

(3) The vector $[-3, 6, 5]$ is in span(S) since it equals

$$\tfrac{11}{5}[3, -6, -1] - \tfrac{12}{5}[4, -8, -3] + 0[5, -10, -1].$$

(4) A simplified form for the vectors in span(S) is:

$$\{a(x^3 - 2x + 4) + b(x^2 + 3x - 2) \,|\, a, b \in \mathbb{R}\} = \{ax^3 + bx^2 + (-2a + 3b)x + (4a - 2b) \,|\, a, b \in \mathbb{R}\}.$$

(5) Form the matrix $\mathbf{A}$ whose columns are the entries of each given matrix (taking the entries in the same order each time). Noting that every column in the reduced row echelon form for $\mathbf{A}$ contains a pivot, the Independence Test Method from Section 4.4 shows that the given set is linearly independent.

(6) Using the Independence Test Method, as explained in Section 4.6, produces the basis

$$B = \left\{ \begin{bmatrix} 3 & -1 \\ -2 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}, \begin{bmatrix} 4 & -3 \\ -2 & 2 \end{bmatrix} \right\}.$$

for span(S). Since B has 3 elements, dim(span(S)) = 3.

(7) Note that span(S) is the row space of $\mathbf{A}^T$. Now $\mathbf{A}^T$ is also singular (use Theorem 2.12), and so rank$(\mathbf{A}^T) < n$ by Theorem 2.15. Hence, the Simplified Span Method as explained in Section 4.6 for finding a basis for span(S) will produce fewer than n basis elements (the nonzero rows of the reduced row echelon form of $\mathbf{A}^T$), showing that dim(span(S)) $< n$.

(8) One possible answer is $\{2x^3 - 3x^2 + 3x + 1, -x^3 + 4x^2 - 6x - 2, x^3, x\}$. (This is obtained by enlarging T with the standard basis for $\mathcal{P}_3$ and then eliminating x^2 and 1 for redundancy.)

(9) First, by Theorem 4.4, E_1 is a subspace of $\mathbb{R}^n$, and so by Theorem 4.13, $\dim(E_1) \leq \dim(\mathbb{R}^n) = n$. But, if $\dim(E_1) = n$, then Theorem 4.13 shows that $E_1 = \mathbb{R}^n$. Hence, for every vector $\mathbf{X} \in \mathbb{R}^n$, $\mathbf{AX} = \mathbf{X}$. In particular, this is true for each vector $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Using these vectors as columns forms the matrix $\mathbf{I}_n$. Thus we see that $\mathbf{AI}_n = \mathbf{I}_n$. implying $\mathbf{A} = \mathbf{I}_n$. This contradicts the given condition $\mathbf{A} \neq \mathbf{I}_n$. Hence $\dim(E_1) \neq n$, and so $\dim(E_1) < n$.

(10) (a) The transition matrix from B to C is $\begin{bmatrix} 2 & -1 & 3 \\ 3 & -2 & 2 \\ 1 & -2 & 3 \end{bmatrix}$.

(b) For $\mathbf{v} = [-109, 155, -71]$, $[\mathbf{v}]_B = [-1, 3, -3]$, and $[\mathbf{v}]_C = [-14, -15, -16]$.

Answers to Test for Chapter 5 — Version A

(1) To show f is a linear operator, we must show that $f(\mathbf{A}_1+\mathbf{A}_2) = f(\mathbf{A}_1)+f(\mathbf{A}_2)$, for all $\mathbf{A}_1, \mathbf{A}_2 \in \mathcal{M}_{nn}$, and that $f(c\mathbf{A}) = cf(\mathbf{A})$, for all $c \in \mathbb{R}$ and all $\mathbf{A} \in \mathcal{M}_{nn}$. However, the first equation holds since

$$\mathbf{B}(\mathbf{A}_1 + \mathbf{A}_2)\mathbf{B}^{-1} = \mathbf{B}\mathbf{A}_1\mathbf{B}^{-1} + \mathbf{B}\mathbf{A}_2\mathbf{B}^{-1}$$

by the distributive laws of matrix multiplication over addition. Similarly, the second equation holds true since

$$\mathbf{B}(c\mathbf{A})\mathbf{B}^{-1} = c\mathbf{B}\mathbf{A}\mathbf{B}^{-1}$$

by part (4) of Theorem 1.16.

(2) $L([2, -3, 1]) = [-1, -11, 4]$.

(3) The matrix $\mathbf{A}_{BC} = \begin{bmatrix} 167 & 117 \\ 46 & 32 \\ -170 & -119 \end{bmatrix}$.

(4) The kernel of $L = \{[2b + d, b, -3d, d] \,|\, b, d \in \mathbb{R}\}$. Hence, a basis for $\ker(L) = \{[2, 1, 0, 0], [1, 0, -3, 1]\}$. A basis for range$(L)$ is the set $\{[4, -3, 4], [-1, 1, 2]\}$, the first and third columns of the original matrix. Thus, $\dim(\ker(L)) + \dim(\text{range}(L)) = 2 + 2 = 4 = \dim(\mathbb{R}^4)$, and the Dimension Theorem is verified.

(5) (a) Now, $\dim(\text{range}(L)) \leq \dim(\mathbb{R}^3) = 3$. Hence $\dim(\ker(L)) = \dim(\mathcal{P}_3) - \dim(\text{range}(L))$ (by the Dimension Theorem) $= 4 - \dim(\text{range}(L)) \geq 4 - 3 = 1$. Thus, $\ker(L)$ is nontrivial.

(b) Any nonzero polynomial $\mathbf{p}$ in $\ker(L)$ satisfies the given conditions.

(6) The 3×3 matrix given in the definition of L is clearly the matrix for L with respect to the standard basis for $\mathbb{R}^3$. Since the determinant of this matrix is -1, it is nonsingular. Theorem 5.16 thus shows that L is an isomorphism, implying also that it is one-to-one.

(7) First, L is a linear operator, because

$$L(\mathbf{p}_1 + \mathbf{p}_2) = (\mathbf{p}_1 + \mathbf{p}_2) - (\mathbf{p}_1 - \mathbf{p}_2)' = (\mathbf{p}_1 - \mathbf{p}_1{}') + (\mathbf{p}_2 - \mathbf{p}_2{}') = L(\mathbf{p}_1) + L(\mathbf{p}_2),$$

and because

$$L(c\mathbf{p}) = c\mathbf{p} - (c\mathbf{p})' = c\mathbf{p} - c\mathbf{p}' = c(\mathbf{p} - \mathbf{p}') = cL(\mathbf{p}).$$

Since the domain and codomain of L have the same dimension, in order to show L is an isomorphism, it is only necessary to show either L is one-to-one, or L is onto (by Corollary 5.13). We show L is one-to-one. Suppose $L(\mathbf{p}) = \mathbf{0}$ (the zero polynomial). Then, $\mathbf{p} - \mathbf{p}' = \mathbf{0}$, and so $\mathbf{p} = \mathbf{p}'$. But these polynomials have different degrees unless $\mathbf{p}$ is constant, in which case $\mathbf{p}' = \mathbf{0}$. Therefore, $\mathbf{p} = \mathbf{p}' = \mathbf{0}$. Hence, $\ker(L) = \{\mathbf{0}\}$, and L is one-to-one.

(8) $B = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right)$, $\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

Other answers for B and $\mathbf{P}$ are possible since the eigenspace E_0 is three-dimensional.

(9) (a) False

(b) True

(10) The matrix for L with respect to the standard basis for $\mathcal{P}_3$ is lower triangular with all 1's on the main diagonal. (The ith column of this matrix consists of the coefficients of $(x+1)^{4-i}$ in order of descending degree.) Thus, $p_L(x) = (x-1)^4$, and $\lambda = 1$ is the only eigenvalue for L. Now, if L were diagonalizable, the geometric multiplicity of $\lambda = 1$ would have to be 4, and we would have $E_1 = \mathcal{P}_3$. This would imply that $L(\mathbf{p}) = \mathbf{p}$ for all $\mathbf{p} \in \mathcal{P}_3$. But this is clearly false since the image of the polynomial x is $x + 1$.

Answers to Test for Chapter 5 — Version B

(1) To show f is a linear operator, we must show that $f(\mathbf{A}_1+\mathbf{A}_2) = f(\mathbf{A}_1)+f(\mathbf{A}_2)$, for all $\mathbf{A}_1, \mathbf{A}_2 \in \mathcal{M}_{nn}$, and that $f(c\mathbf{A}) = cf(\mathbf{A})$, for all $c \in \mathbb{R}$ and all $\mathbf{A} \in \mathcal{M}_{nn}$. However, the first equation holds since

$$\begin{aligned}(\mathbf{A}_1+\mathbf{A}_2)+(\mathbf{A}_1+\mathbf{A}_2)^T &= \mathbf{A}_1+\mathbf{A}_2+\mathbf{A}_1{}^T+\mathbf{A}_2{}^T \text{ (by part (2) of Theorem 1.13)}\\ &= (\mathbf{A}_1+\mathbf{A}_2)+(\mathbf{A}_1+\mathbf{A}_2)^T = \mathbf{A}_1+\mathbf{A}_2+\mathbf{A}_1{}^T+\mathbf{A}_2{}^T\\ &= (\mathbf{A}_1+\mathbf{A}_1{}^T)+(\mathbf{A}_2+\mathbf{A}_2{}^T) \text{ (by commutativity of matrix addition).}\end{aligned}$$

Similarly,

$$f(c\mathbf{A}) = (c\mathbf{A})+(c\mathbf{A})^T = c\mathbf{A}+c\mathbf{A}^T = c(\mathbf{A}+\mathbf{A}^T) = cf(\mathbf{A}).$$

(2) $L([6,-4,7]) = -x^2+2x+3$

(3) The matrix $\mathbf{A}_{BC} = \begin{bmatrix} -55 & 165 & 49 & 56 \\ -46 & 139 & 42 & 45 \\ 32 & -97 & -29 & -32 \end{bmatrix}$.

(4) The kernel of $L = \{[3c,-2c,c,0] \mid c \in \mathbb{R}\}$. Hence, a basis for $\ker(L) = \{[3,-2,1,0]\}$. The range of L is spanned by columns 1, 2, and 4 of the original matrix. Since these columns are linearly independent, a basis for $\text{range}(L) = \{[2,1,-4,4], [5,-1,2,1], [8.-1,1,2]\}$. The Dimension Theorem is verified because $\dim(\ker(L)) = 1$, $\dim(\text{range}(L)) = 3$, and the sum of these dimensions equals the dimension of $\mathbb{R}^4$, the domain of L.

(5) (a) This statement is true. For, $\dim(\text{range}(L_1)) = \text{rank}(\mathbf{A})$ (by part (1) of Theorem 5.9) $= \text{rank}(\mathbf{A}^T)$ (by Corollary 5.11) $= \dim(\text{range}(L_2))$ (by part (1) of Theorem 5.9).

(b) This statement is false. For a particular counterexample, use $\mathbf{A} = \mathbf{O}_{mn}$, in which case

$$\dim(\ker(L_1)) = n \neq m = \dim(\ker(L_2)).$$

In general, using the Dimension Theorem,

$$\dim(\ker(L_1)) = n - \dim(\text{range}(L_1)),$$

and

$$\dim(\ker(L_2)) = m - \dim(\text{range}(L_1)).$$

But, by part (a),

$$\dim(\text{range}(L_1)) = \dim(\text{range}(L_2)),$$

and since $m \neq n$, $\dim(\ker(L_1))$ can never equal $\dim(\ker(L_2))$.

(6) Solving the appropriate homogeneous system to find the kernel of L gives $\ker(L) = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}$. Hence, L is one-to-one by part (1) of Theorem 5.12. Since the domain and codomain of L have the same dimension, L is also an isomorphism by Corollary 5.13.

(7) First, L is a linear operator, since

$$L(\mathbf{B}_1+\mathbf{B}_2) = (\mathbf{B}_1+\mathbf{B}_2)\mathbf{A}^{-1} = \mathbf{B}_1\mathbf{A}^{-1}+\mathbf{B}_2\mathbf{A}^{-1} = L(\mathbf{B}_1)+L(\mathbf{B}_2),$$

and since

$$L(c\mathbf{B}) = (c\mathbf{B})\mathbf{A}^{-1} = c(\mathbf{B}\mathbf{A}^{-1}) = cL(\mathbf{B}).$$

Since the domain and codomain of L have the same dimension, in order to show that L is an isomorphism, by Corollary 5.13 it is enough to show either that L is one-to-one, or that L is onto. We show L is one-to-one. Suppose $L(\mathbf{B}_1) = L(\mathbf{B}_2)$. Then $\mathbf{B}_1\mathbf{A}^{-1} = \mathbf{B}_2\mathbf{A}^{-1}$. Multiplying both sides on the right by $\mathbf{A}$ gives $\mathbf{B}_1 = \mathbf{B}_2$. Hence, L is one-to-one.

(8) (Optional Hint: $p_L(x) = x^4 - 2x^2 + 1 = (x-1)^2(x+1)^2$.)
Answer:

$$B = \left(\begin{bmatrix} 3 & 0 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \right),$$

$$\mathbf{A} = \begin{bmatrix} -7 & 0 & 12 & 0 \\ 0 & -7 & 0 & 12 \\ -4 & 0 & 7 & 0 \\ 0 & -4 & 0 & 7 \end{bmatrix}, \mathbf{P} = \begin{bmatrix} 3 & 0 & 2 & 0 \\ 0 & 3 & 0 & 2 \\ 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Other answers for B and $\mathbf{P}$ are possible since the eigenspaces E_1 and E_{-1} are both two-dimensional.

(9) (a) False

(b) False

(10) All vectors parallel to $[4, -2, 3]$ are mapped to themselves under L, and so $\lambda = 1$ is an eigenvalue for L. Now, all other nonzero vectors in $\mathbb{R}^3$ are rotated around the axis through the origin parallel to $[4, -2, 3]$, and so their images under L are not parallel to themselves. Hence E_1 is the one-dimensional subspace spanned by $[4, -2, 3]$, and there are no other eigenvalues. Therefore, the sum of the geometric multiplicities of all eigenvalues is 1, which is less than 3, the dimension of $\mathbb{R}^3$. Thus Theorem 5.28 shows that L is not diagonalizable.

Answers to Test for Chapter 5 — Version C

(1) To show that L is a linear transformation, we must prove that $L(\mathbf{p}+\mathbf{q}) = L(\mathbf{p})+L(\mathbf{q})$ and that $L(c\mathbf{p}) = cL(\mathbf{p})$ for all $\mathbf{p},\mathbf{q} \in \mathcal{P}_n$ and all $c \in \mathbb{R}$. Let $\mathbf{p} = a_nx^n + \cdots + a_1x + a_0$ and let $\mathbf{q} = b_nx^n + \cdots + b_1x + b_0$. Then

$$\begin{aligned} L(\mathbf{p}+\mathbf{q}) &= L(a_nx^n + \cdots + a_1x + a_0 + b_nx^n + \cdots + b_1x + b_0) \\ &= L((a_n+b_n)x^n + \cdots + (a_1+b_1)x + (a_0+b_0)) \\ &= (a_n+b_n)\mathbf{A}^n + \cdots + (a_1+b_1)\mathbf{A} + (a_0+b_0)\mathbf{I}_n \\ &= a_n\mathbf{A}^n + \cdots + a_1\mathbf{A} + a_0\mathbf{I}_n + b_n\mathbf{A}^n + \cdots + b_1\mathbf{A} + b_0\mathbf{I}_n \\ &= L(\mathbf{p}) + L(\mathbf{q}). \end{aligned}$$

Similarly,

$$\begin{aligned} L(c\mathbf{p}) &= L(c(a_nx^n + \cdots + a_1x + a_0)) \\ &= L(ca_nx^n + \cdots + ca_1x + ca_0) \\ &= ca_n\mathbf{A}^n + \cdots + ca_1\mathbf{A} + ca_0\mathbf{I}_n \\ &= c(a_n\mathbf{A}^n + \cdots + a_1\mathbf{A} + a_0\mathbf{I}_n) \\ &= cL(\mathbf{p}). \end{aligned}$$

(2) $L([-3,1,-4]) = \begin{bmatrix} -29 & 21 \\ -2 & 6 \end{bmatrix}$.

(3) The matrix $\mathbf{A}_{BC} = \begin{bmatrix} -44 & 91 & 350 \\ -13 & 23 & 118 \\ -28 & 60 & 215 \\ -10 & 16 & 97 \end{bmatrix}$.

(4) The matrix for L with respect to the standard bases for $\mathcal{M}_{22}$ and $\mathcal{P}_2$ is $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, which row reduces to $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$. Hence $\ker(L) = \left\{ d \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} \middle| \, d \in \mathbb{R} \right\}$, which has dimension 1. A basis for range(L) consists of the first three "columns" of $\mathbf{A}$, and hence dim(range$(L)) = 3$, implying range$(L) = \mathcal{P}_2$. Note that dim(ker(L)) + dim(range(L)) $= 1 + 3 = 4 = \dim(\mathcal{M}_{22})$, thus verifying the Dimension Theorem.

(5) (a) $\mathbf{v} \in \ker(L_1) \Longrightarrow L_1(\mathbf{v}) = \mathbf{0}_{\mathcal{W}} \Longrightarrow L_2(L_1(\mathbf{v})) = L_2(\mathbf{0}_{\mathcal{W}}) \Longrightarrow (L_2 \circ L_1)(\mathbf{v}) = \mathbf{0}_{\mathcal{Y}}$ $\Longrightarrow \mathbf{v} \in \ker(L_2 \circ L_1)$. Hence, $\ker(L_1) \subseteq \ker(L_2 \circ L_1)$.

(b) By part (a) and Theorem 4.13, $\dim(\ker(L_1)) \leq \dim(\ker(L_2 \circ L_1))$. Hence, by the Dimension Theorem, $\dim(\text{range}(L_1)) = \dim(\mathcal{V}) - \dim(\ker(L_1)) \geq \dim(\mathcal{V}) - \dim(\ker(L_2 \circ L_1))$ $= \dim(\text{range}(L_2 \circ L_1))$.

(6) Solving the appropriate homogeneous system to find the kernel of L shows that $\ker(L)$ consists only of the zero polynomial. Hence, L is one-to-one by part (1) of Theorem 5.12. Also, $\dim(\ker(L)) = 0$. However, L is not an isomorphism, because by the Dimension Theorem, $\dim(\text{range}(L)) = 3$, which is less than the dimension of the codomain $\mathbb{R}^4$.

(7) First, we must show that L is a linear transformation. Now

$$\begin{aligned} L([a_1,\ldots,a_n]+[b_1,\ldots,b_n]) &= L([a_1+b_1,\ldots,a_n+b_n]) \\ &= (a_1+b_1)\mathbf{v}_1+\cdots+(a_n+b_n)\mathbf{v}_n \\ &= a_1\mathbf{v}_1+b_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n+b_n\mathbf{v}_n \\ &= a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n+b_1\mathbf{v}_1+\cdots+b_n\mathbf{v}_n \\ &= L([a_1,\ldots,a_n])+L([b_1,\ldots,b_n]). \end{aligned}$$

Also,

$$\begin{aligned} L(c[a_1,\ldots,a_n]) &= L([ca_1,\ldots,ca_n]) \\ &= ca_1\mathbf{v}_1+\cdots+ca_n\mathbf{v}_n \\ &= c(a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n) \\ &= cL([a_1,\ldots,a_n]). \end{aligned}$$

In order to prove that L is an isomorphism, it is enough to show that L is one-to-one or that L is onto (by Corollary 5.13), since the dimensions of the domain and the codomain are the same. We show that L is one-to-one. Suppose $L([a_1,\ldots,a_n]) = \mathbf{0}_{\mathcal{V}}$. Then $a_1\mathbf{v}_1+a_2\mathbf{v}_2+\cdots+a_n\mathbf{v}_n = \mathbf{0}_{\mathcal{V}}$, implying $a_1 = a_2 = \cdots = a_n = 0$ because B is a linearly independent set (since B is a basis for $\mathcal{V}$). Hence $[a_1,\ldots,a_n] = [0,\ldots,0]$. Therefore, $\ker(L) = \{[0,\ldots,0]\}$, and L is one-to-one by part (1) of Theorem 5.12.

(8) $B = (x^2-x+2, x^2-2x, x^2+2)$, $\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Other answers for B and $\mathbf{P}$ are possible since the eigenspace E_1 is two-dimensional.

(9) (a) True

(b) True

(10) If $\mathbf{p}$ is a nonconstant polynomial, then $\mathbf{p}'$ is nonzero and has a degree lower than that of $\mathbf{p}$. Hence $\mathbf{p}' \neq c\mathbf{p}$, for any $c \in \mathbb{R}$. Hence, nonconstant polynomials can not be eigenvectors. All constant polynomials, however, are in the eigenspace E_0 of $\lambda = 0$. Thus, E_0 is one-dimensional. Therefore, the sum of the geometric multiplicities of all eigenvalues is 1, which is less than $\dim(\mathcal{P}_n) = n+1$, since $n > 0$. Hence L is not diagonalizable by Theorem 5.28.

Answers for Test for Chapter 6 — Version A

(1) Performing the Gram-Schmidt Process as outlined in the text and using appropriate multiples to avoid fractions leads to the following basis: $\{[4,-3,2],[13,12,-8],[0,2,3]\}$.

(2) Matrix for $L = \frac{1}{9}\begin{bmatrix} 7 & 4 & -4 \\ 4 & 1 & 8 \\ -4 & 8 & 1 \end{bmatrix}$.

(3) Performing the Gram-Schmidt Process as outlined in the text and using appropriate multiples to avoid fractions leads to the following basis for $\mathcal{W}$: $\{[2,-3,1],[-11,-1,19]\}$. Continuing the process leads to the following basis for $\mathcal{W}^\perp$: $\{[8,7,5]\}$. Also, $\dim(\mathcal{W}^\perp) = 1$.

(4) $\|\mathbf{Ax}\|^2 = \mathbf{Ax}\cdot\mathbf{Ax} = \mathbf{x}\cdot\mathbf{x}$ (by Theorem 6.9) $= \|\mathbf{x}\|^2$. Taking square roots yields $\|\mathbf{Ax}\| = \|\mathbf{x}\|$.

(5) Let $\mathbf{w} \in \mathcal{W}$. We need to show that $\mathbf{w} \in (\mathcal{W}^\perp)^\perp$. To do this, we need only prove that $\mathbf{w}\perp\mathbf{v}$ for all $\mathbf{v} \in \mathcal{W}^\perp$. But, by the definition of $\mathcal{W}^\perp$, $\mathbf{w}\cdot\mathbf{v} = 0$, since $\mathbf{w} \in \mathcal{W}$, which completes the proof.

(6) The characteristic polynomial is $p_L(x) = x^2(x-1)^2 = x^4 - 2x^3 + x^2$. (This is the characteristic polynomial for every orthogonal projection onto a 2-dimensional subspace of $\mathbb{R}^4$.)

(7) If $\mathbf{v}_1$ and $\mathbf{v}_2$ are unit vectors in the given plane with $\mathbf{v}_1 \perp \mathbf{v}_2$, then the matrix for L with respect to the ordered orthonormal basis

$$\left(\frac{1}{\sqrt{70}}[3,5,-6], \mathbf{v}_1, \mathbf{v}_2\right) \quad \text{is} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(8) (Optional Hint: $p_L(x) = x^3 - 18x^2 + 81x$.)

Answer: $B = (\frac{1}{3}[2,-2,1], \frac{1}{3}[1,2,2], \frac{1}{3}[2,1,-2])$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

Other answers for B are possible since the eigenspace E_9 is two-dimensional. Another likely answer for B is $(\frac{1}{3}[2,-2,1], \frac{1}{\sqrt{2}}[1,1,0], \frac{1}{\sqrt{18}}[-1,1,4])$.

(9) $\mathbf{A}^3 = \left(\frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}\right)\begin{bmatrix} -1 & 0 \\ 0 & 8 \end{bmatrix}\left(\frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}\right)$.

Hence, one possible answer is

$$\mathbf{A} = \left(\frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}\right)\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}\left(\frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}\right) = \frac{1}{5}\begin{bmatrix} 7 & 6 \\ 6 & -2 \end{bmatrix}.$$

(10) (Optional Hint: Use the definition of a symmetric operator to show that $(\lambda_2 - \lambda_1)(\mathbf{v}_1\cdot\mathbf{v}_2) = 0$.)
Answer: L symmetric $\Longrightarrow \mathbf{v}_1\cdot L(\mathbf{v}_2) = L(\mathbf{v}_1)\cdot\mathbf{v}_2 \Longrightarrow \mathbf{v}_1\cdot(\lambda_2\mathbf{v}_2) = (\lambda_1\mathbf{v}_1)\cdot\mathbf{v}_2 \Longrightarrow \lambda_2(\mathbf{v}_1\cdot\mathbf{v}_2) = \lambda_1(\mathbf{v}_1\cdot\mathbf{v}_2)$ $\Longrightarrow (\lambda_2 - \lambda_1)(\mathbf{v}_1\cdot\mathbf{v}_2) = 0 \Longrightarrow (\mathbf{v}_1\cdot\mathbf{v}_2) = 0$ (since $\lambda_2 \neq \lambda_1$) $\Longrightarrow \mathbf{v}_1\perp\mathbf{v}_2$.

Answers for Test for Chapter 6 — Version B

(1) Performing the Gram-Schmidt Process as outlined in the text and using appropriate multiples to avoid fractions leads to the following basis: $\{[2,-3,1],[30,11,-27],[5,6,8]\}$.

(2) Matrix for $L = \frac{1}{25}\begin{bmatrix} 16 & 0 & -12 \\ 0 & 25 & 0 \\ -12 & 0 & 9 \end{bmatrix}$.

(3) Performing the Gram-Schmidt Process as outlined in the text and using appropriate multiples to avoid fractions leads to the following basis for $\mathcal{W}$: $\{[2,2,-3],[33,-18,10]\}$. Continuing the process leads to the following basis for $\mathcal{W}^\perp$: $\{[2,7,6]\}$. Using this information, we can show that $[-12,-36,-43] = [0,6,7]+[-12,-42,-36]$, where the first vector of this sum is in $\mathcal{W}$, and the second is in $\mathcal{W}^\perp$.

(4) Because $\mathbf{A}$ and $\mathbf{B}$ are orthogonal, $\mathbf{A}\mathbf{A}^T = \mathbf{B}\mathbf{B}^T = \mathbf{I}_n$. Hence,

$$\begin{aligned}(\mathbf{AB})(\mathbf{AB})^T &= (\mathbf{AB})(\mathbf{B}^T\mathbf{A}^T)\\ &= \mathbf{A}(\mathbf{B}\mathbf{B}^T)\mathbf{A}^T\\ &= \mathbf{A}\mathbf{I}_n\mathbf{A}^T = \mathbf{A}\mathbf{A}^T = \mathbf{I}_n.\end{aligned}$$

Thus, $\mathbf{AB}$ is orthogonal.

(5) By Corollary 6.13, $\dim(\mathcal{W}) = n - \dim(\mathcal{W}^\perp) = n - (n - \dim((\mathcal{W}^\perp)^\perp)) = \dim((\mathcal{W}^\perp)^\perp)$. Thus, by Theorem 4.13, $\mathcal{W} = (\mathcal{W}^\perp)^\perp$.

(6) Now S is an orthogonal basis for $\mathcal{W}$, and so

$$\mathbf{proj}_{\mathcal{W}}\mathbf{v} = \frac{\mathbf{v}\cdot[0,3,6,-2]}{[0,3,6,-2]\cdot[0,3,6,-2]}[0,3,6,-2] + \frac{\mathbf{v}\cdot[-6,2,0,3]}{[-6,2,0,3]\cdot[-6,2,0,3]}[-6,2,0,3] = \tfrac{1}{7}[-24,17,18,6].$$

Thus, $\mathbf{proj}_{\mathcal{W}^\perp}\mathbf{v} = \mathbf{v} - \mathbf{proj}_{\mathcal{W}}\mathbf{v} = \frac{1}{7}[3,18,-11,-6]$.

(7) Yes, by Theorems 6.19 and 6.22, because the matrix for L with respect to the standard basis is symmetric.

(8) (Optional Hint: $\lambda_1 = 0$, $\lambda_2 = 49$)

Answer: $B = (\frac{1}{7}[-3,6,2], \frac{1}{7}[6,2,3], \frac{1}{7}[2,3,-6])$, $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 49 & 0 \\ 0 & 0 & 49 \end{bmatrix}$.

Other answers for B are possible since the eigenspace E_{49} is two-dimensional. Another likely answer for B is $(\frac{1}{7}[-3,6,2], \frac{1}{\sqrt{5}}[2,1,0], \frac{1}{\sqrt{245}}[2,-4,15])$.

(9) $\mathbf{A}^2 = \left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right)\begin{bmatrix} 4 & 0 \\ 0 & 16 \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\right)$.

Hence, one possible answer is

$$\mathbf{A} = \left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right)\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\right) = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}.$$

(10) (Optional Hint: Consider the orthogonal diagonalization of $\mathbf{A}$ and the fact that $\mathbf{A}^2 = \mathbf{A}\mathbf{A}^T$.)
Answer: Since $\mathbf{A}$ is symmetric, it is orthogonally diagonalizable. So there is an orthogonal matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, or equivalently,

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}\mathbf{P}^T.$$

Since all eigenvalues of $\mathbf{A}$ are ± 1, $\mathbf{D}$ has only these values on its main diagonal, and hence $\mathbf{D}^2 = \mathbf{I}_n$. Therefore,

$$\begin{aligned}
\mathbf{A}\mathbf{A}^T &= \mathbf{P}\mathbf{D}\mathbf{P}^T(\mathbf{P}\mathbf{D}\mathbf{P}^T)^T \\
&= \mathbf{P}\mathbf{D}\mathbf{P}^T((\mathbf{P}^T)^T\mathbf{D}^T\mathbf{P}^T) \\
&= \mathbf{P}\mathbf{D}\mathbf{P}^T(\mathbf{P}\mathbf{D}\mathbf{P}^T) \\
&= \mathbf{P}\mathbf{D}(\mathbf{P}^T\mathbf{P})\mathbf{D}\mathbf{P}^T \\
&= \mathbf{P}\mathbf{D}\mathbf{I}_n\mathbf{D}\mathbf{P}^T \\
&= \mathbf{P}\mathbf{D}^2\mathbf{P}^T \\
&= \mathbf{P}\mathbf{I}_n\mathbf{P}^T \\
&= \mathbf{P}\mathbf{P}^T \\
&= \mathbf{I}_n.
\end{aligned}$$

Thus, $\mathbf{A}^T = \mathbf{A}^{-1}$, and $\mathbf{A}$ is orthogonal.

Answers for Test for Chapter 6 — Version C

(1) (Optional Hint: The two given vectors are orthogonal. Therefore, only two additional vectors need to be found.) Performing the Gram-Schmidt Process as outlined in the text and using appropriate multiples to avoid fractions leads to the following basis:

$$\{[-3,1,-1,2],[2,4,0,1],[22,-19,-21,32],[0,1,-7,-4]\}.$$

(2) Matrix for $L = \frac{1}{9}\begin{bmatrix} 5 & -4 & 2 \\ -4 & 5 & 2 \\ 2 & 2 & 8 \end{bmatrix}$.

(3) Performing the Gram-Schmidt Process as outlined in the text and using appropriate multiples to avoid fractions leads to the following basis for $\mathcal{W}$: $\{[5,1,-2],\ [2,0,5]\}$. Continuing the process leads to the following basis for $\mathcal{W}^{\perp}$: $\{[5,-29,-2]\}$. Also, $\dim(\mathcal{W}^{\perp}) = 1$.

(4) Suppose n is odd. Then

$$|\mathbf{A}|^2 = |\mathbf{A}^2| = |\mathbf{A}\mathbf{A}| = |\mathbf{A}(-\mathbf{A}^T)| = |-\mathbf{A}\mathbf{A}^T| = (-1)^n|\mathbf{A}\mathbf{A}^T| = (-1)|\mathbf{I}_n| = -1.$$

But $|\mathbf{A}|^2$ can not be negative. This contradiction shows that n is even.

(5) Let $\mathbf{w}_2 \in \mathcal{W}_2^{\perp}$. We must show that $\mathbf{w}_2 \in \mathcal{W}_1^{\perp}$. To do this, we show that $\mathbf{w}_2 \cdot \mathbf{w}_1 = 0$ for all $\mathbf{w}_1 \in \mathcal{W}_1$. So, let $\mathbf{w}_1 \in \mathcal{W}_1 \subseteq \mathcal{W}_2$. Thus, $\mathbf{w}_1 \in \mathcal{W}_2$. Hence, $\mathbf{w}_1 \cdot \mathbf{w}_2 = 0$, because $\mathbf{w}_2 \in \mathcal{W}_2^{\perp}$. This completes the proof.

(6) Let $\mathbf{v} = [5,1,0,-4]$. Then, $\mathbf{proj}_{\mathcal{W}}\mathbf{v} = \frac{1}{3}[14,5,-2,-12]$; $\mathbf{proj}_{\mathcal{W}^{\perp}}\mathbf{v} = \frac{1}{3}[1,-2,2,0]$; minimum distance $= \left\|\frac{1}{3}[1,-2,2,0]\right\| = 1$.

(7) If $\mathbf{v}_1$ and $\mathbf{v}_2$ are unit vectors in the given plane with $\mathbf{v}_1 \perp \mathbf{v}_2$, then the matrix for L with respect to the ordered orthonormal basis

$$\left(\frac{1}{\sqrt{26}}[4,-3,-1],\mathbf{v}_1,\mathbf{v}_2\right) \quad \text{is} \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence, L is orthogonally diagonalizable. Therefore, L is a symmetric operator, and so its matrix with respect to any orthonormal basis is symmetric. In particular, this is true of the matrix for L with respect to the standard basis.

(8) (Optional Hint: $\lambda_1 = 121$, $\lambda_2 = -121$) Answer:

$$B = \left(\tfrac{1}{11}[2,6,-9], \tfrac{1}{11}[-9,6,2], \tfrac{1}{11}[6,7,6]\right); \ \mathbf{D} = \begin{bmatrix} 121 & 0 & 0 \\ 0 & 121 & 0 \\ 0 & 0 & -121 \end{bmatrix}.$$

Other answers for B are possible since the eigenspace E_{121} is two-dimensional. Another likely answer for B is

$$\left(\frac{1}{\sqrt{85}}[-7,6,0], \frac{1}{11\sqrt{85}}[-36,-42,85], \tfrac{1}{11}[6,7,6]\right).$$

(9) $\mathbf{A}^2 = \left(\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}\right)\begin{bmatrix} 0 & 0 \\ 0 & 100 \end{bmatrix}\left(\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}\right)$.

Hence, one possible answer is

$$\mathbf{A} = \left(\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}\right)\begin{bmatrix} 0 & 0 \\ 0 & 10 \end{bmatrix}\left(\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}.$$

(10) (Optional Hint: First show that **A** and **B** orthogonally diagonalize to the same matrix.)
Answer: First, because **A** and **B** are symmetric, they are orthogonally diagonalizable by Corollary 6.23. But, since the characteristic polynomials of **A** and **B** are equal, these matrices have the same eigenvalues with the same algebraic multiplicities. Thus, there is a single diagonal matrix **D**, having these eigenvalues on its main diagonal, such that $\mathbf{D} = \mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1$ and $\mathbf{D} = \mathbf{P}_2^{-1}\mathbf{B}\mathbf{P}_2$, for some orthogonal matrices $\mathbf{P}_1$ and $\mathbf{P}_2$. Hence $\mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1 = \mathbf{P}_2^{-1}\mathbf{B}\mathbf{P}_2$, or $\mathbf{A} = \mathbf{P}_1\mathbf{P}_2^{-1}\mathbf{B}\mathbf{P}_2\mathbf{P}_1^{-1}$. Let $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2^{-1}$. Then $\mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$, where **P** is orthogonal by parts (2) and (3) of Theorem 6.6.

Answers to Test for Chapter 7 — Version A

(1) $\mathbf{a}\cdot\mathbf{b} = \overline{\mathbf{b}\cdot\mathbf{a}} = [2-2i, -1+i, 5]$

(2) $(\mathbf{P}^*\mathbf{H}\mathbf{P})^* = \mathbf{P}^*\mathbf{H}^*(\mathbf{P}^*)^*$ (by part (5) of Theorem 7.2) $= \mathbf{P}^*\mathbf{H}^*\mathbf{P}$ (by part (1) of Theorem 7.2) $= \mathbf{P}^*\mathbf{H}\mathbf{P}$ (since $\mathbf{H}$ is Hermitian). Hence $\mathbf{P}^*\mathbf{H}\mathbf{P}$ is Hermitian.

(3) (a) False
(b) False

(4) $z_1 = 1+i$, $z_2 = 1-2i$

(5) Let $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $p_{\mathbf{A}}(x) = x^2+1$, which has no real roots. Hence $\mathbf{A}$ has no real eigenvalues, and so is not diagonalizable as a real matrix. However, $p_{\mathbf{A}}(x)$ has complex roots i and $-i$. Thus, this 2×2 matrix has two complex distinct eigenvalues, and thus must be diagonalizable.

(6) (Optional Hint: $p_L(x) = x^3 - 2ix^2 - x$.)
Answer: $\lambda_1 = 0$, $\lambda_2 = i$; basis for $E_0 = \{[1+i, -i, 1]\}$, basis for $E_i = \{[i,0,1],[0,1,0]\}$.
L is diagonalizable.

(7) The statement is false. The polynomial x is in the given subset, but ix is not. Hence, the subset is not closed under scalar multiplication.

(8) $B = \{[1+3i, 2+i, 1-2i], [1, 2i, 3+i]\}$

(9) As a complex linear transformation:

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

As a real linear transformation:

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

(10) (Optional Hint: The two given vectors are already orthogonal.)
Answer: $\{[1, i, 1+i], [3, -i, -1-i], [0, 2, -1+i]\}$

(11) $\mathbf{P} = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

(12) $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^* = \mathbf{B}^{-1}\mathbf{A}^{-1} = (\mathbf{AB})^{-1}$, and so $\mathbf{AB}$ is unitary. Also, $\Big||\mathbf{A}|\Big|^2 = |\mathbf{A}|\,\overline{|\mathbf{A}|} = |\mathbf{A}|\,|\mathbf{A}^*|$ (by part (3) of Theorem 7.5) $= |\mathbf{A}\mathbf{A}^*| = |\mathbf{I}_n| = 1$.

(13) Proofs for the five properties of an inner product:

(1) $\langle \mathbf{f}, \mathbf{f}\rangle = \mathbf{f}^2(0) + \mathbf{f}^2(1) + \mathbf{f}^2(2) \geq 0$, since the sum of squares is always nonnegative.

(2) Suppose $\mathbf{f}(x) = ax^2 + bx + c$. Then $\langle \mathbf{f}, \mathbf{f}\rangle = 0$
$\Longrightarrow \mathbf{f}(0) = 0,\ \mathbf{f}(1) = 0,\ \mathbf{f}(2) = 0$ (since a sum of squares of real numbers can only be zero if each number is zero)
$\Longrightarrow (a,b,c)$ satisfies

$$\begin{cases} & & & & c & = & 0 \\ a & + & b & + & c & = & 0 \\ 4a & + & 2b & + & c & = & 0 \end{cases}$$

$\Longrightarrow a = b = c = 0$ (since the coefficient matrix of the homogeneous system has determinant -2).

Conversely, it is clear that $\langle \mathbf{0}, \mathbf{0} \rangle = 0$.

(3) $\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}(0)\mathbf{g}(0) + \mathbf{f}(1)\mathbf{g}(1) + \mathbf{f}(2)\mathbf{g}(2)$
$= \mathbf{g}(0)\mathbf{f}(0) + \mathbf{g}(1)\mathbf{f}(1) + \mathbf{g}(2)\mathbf{f}(2) = \langle \mathbf{g}, \mathbf{f} \rangle$

(4) $\langle \mathbf{f} + \mathbf{g}, \mathbf{h} \rangle = (\mathbf{f}(0) + \mathbf{g}(0))\mathbf{h}(0) + (\mathbf{f}(1) + \mathbf{g}(1))\mathbf{h}(1) + (\mathbf{f}(2) + \mathbf{g}(2))\mathbf{h}(2)$
$= \mathbf{f}(0)\mathbf{h}(0) + \mathbf{g}(0)\mathbf{h}(0) + \mathbf{f}(1)\mathbf{h}(1) + \mathbf{g}(1)\mathbf{h}(1) + \mathbf{f}(2)\mathbf{h}(2) + \mathbf{g}(2)\mathbf{h}(2)$
$= \mathbf{f}(0)\mathbf{h}(0) + \mathbf{f}(1)\mathbf{h}(1) + \mathbf{f}(2)\mathbf{h}(2) + \mathbf{g}(0)\mathbf{h}(0) + \mathbf{g}(1)\mathbf{h}(1) + \mathbf{g}(2)\mathbf{h}(2)$
$= \langle \mathbf{f}, \mathbf{h} \rangle + \langle \mathbf{g}, \mathbf{h} \rangle$

(5) $\langle k\mathbf{f}, \mathbf{g} \rangle = (k\mathbf{f})(0)\mathbf{g}(0) + (k\mathbf{f})(1)\mathbf{g}(1) + (k\mathbf{f})(2)\mathbf{g}(2)$
$= k\mathbf{f}(0)\mathbf{g}(0) + k\mathbf{f}(1)\mathbf{g}(1) + k\mathbf{f}(2)\mathbf{g}(2)$
$= k(\mathbf{f}(0)\mathbf{g}(0) + \mathbf{f}(1)\mathbf{g}(1) + \mathbf{f}(2)\mathbf{g}(2))$
$= k\langle \mathbf{f}, \mathbf{g} \rangle$

Also, $\langle \mathbf{f}, \mathbf{g} \rangle = 14$ and $\|\mathbf{f}\| = \sqrt{35}$.

(14) The distance between $\mathbf{x}$ and $\mathbf{y}$ is 11.

(15) Note that $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, -\mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, -\mathbf{y} \rangle = \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$. So, if $\|\mathbf{x}\| = \|\mathbf{y}\|$, then $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 = 0$.
Conversely, $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0 \Longrightarrow \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 = 0 \Longrightarrow \|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 \Longrightarrow \|\mathbf{x}\| = \|\mathbf{y}\|$.

(16) $\{x^2, x, 3 - 5x^2\}$

(17) $\mathbf{w}_1 = [-8, -5, 16]$, $\mathbf{w}_2 = [8, 6, -19]$

Answers to Test for Chapter 7 — Version B

(1) $\mathbf{A}^*\mathbf{B} = \begin{bmatrix} -1+14i & 33-3i \\ -5-5i & 5+2i \end{bmatrix}$

(2) $(i\mathbf{H})^* = \bar{i}\mathbf{H}^*$ (by part (3) of Theorem 7.2) $= (-i)\mathbf{H}^* = (-i)\mathbf{H}$ (since $\mathbf{H}$ is Hermitian) $= -(i\mathbf{H})$. Hence, $i\mathbf{H}$ is skew-Hermitian.

(3) (a) False

(b) True

(4) The determinant of the matrix of coefficients is 1, and so the inverse matrix is $\begin{bmatrix} 6+5i & -2-14i \\ -2+i & 4+i \end{bmatrix}$. Using this gives $z_1 = 2+i$ and $z_2 = -i$.

(5) Let $\mathbf{A}$ be an $n \times n$ complex matrix. By the Fundamental Theorem of Algebra, $p_{\mathbf{A}}(x)$ factors into n linear factors. Hence, $p_{\mathbf{A}}(x) = (x-\lambda_1)^{k_1} \cdots (x-\lambda_m)^{k_m}$, where $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of $\mathbf{A}$, and $k_1, \ldots, k_m$ are their corresponding algebraic multiplicities. But, since the degree of $p_{\mathbf{A}}(x)$ is n, $\sum_{i=1}^m k_i = n$.

(6) $\lambda = i$; basis for $E_i = \{[1, 2i]\}$. L is not diagonalizable.

(7) The set of Hermitian $n \times n$ matrices is not closed under scalar multiplication, and so is not a subspace of $\mathcal{M}_{nn}^{\mathbb{C}}$. To see this, note that $i\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}$. However, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is Hermitian and $\begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}$ is not. Since the set of Hermitian $n \times n$ matrices is not a vector space, it does not have a dimension.

(8) $B = \{[1, 0, i], [0, 1, 1-i]\}$

(9) Note that $L\left(i\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = L\left(\begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}^* = \begin{bmatrix} -i & 0 \\ 0 & 0 \end{bmatrix}$, but $iL\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = i\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^* = i\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}$.

(10) $\{[1+i, 1-i, 2], [3, i, -1+i], [0, 2, -1-i]\}$

(11) Now, $\mathbf{Z}^* = \begin{bmatrix} 1-i & 1+i \\ -1-i & 1-i \end{bmatrix}$ and $\mathbf{Z}\mathbf{Z}^* = \begin{bmatrix} 4 & 4i \\ -4i & 4 \end{bmatrix} = \mathbf{Z}^*\mathbf{Z}$. Hence, $\mathbf{Z}$ is normal. Thus, by Theorem 7.9, $\mathbf{Z}$ is unitarily diagonalizable.

(12) Now, $\mathbf{A}$ unitary implies $\mathbf{A}\mathbf{A}^* = \mathbf{I}_n$. Suppose $\mathbf{A}^2 = \mathbf{I}_n$. Hence $\mathbf{A}\mathbf{A}^* = \mathbf{A}\mathbf{A}$. Multiplying both sides by $\mathbf{A}^{-1}\,(= \mathbf{A}^*)$ on the left yields $\mathbf{A}^* = \mathbf{A}$, and $\mathbf{A}$ is Hermitian. Conversely, if $\mathbf{A}$ is Hermitian, then $\mathbf{A} = \mathbf{A}^*$. Hence $\mathbf{A}\mathbf{A}^* = \mathbf{I}_n$ yields $\mathbf{A}^2 = \mathbf{I}_n$.

(13) Proofs for the five properties of an inner product:

(1) $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{A}\mathbf{x} \cdot \mathbf{A}\mathbf{x} \geq 0$, since $\cdot$ is an inner product.

(2) $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \implies \mathbf{A}\mathbf{x} \cdot \mathbf{A}\mathbf{x} = 0 \implies \mathbf{A}\mathbf{x} = \mathbf{0}$ (since $\cdot$ is an inner product)
$\implies \mathbf{x} = \mathbf{0}$ (since $\mathbf{A}$ is nonsingular)

Conversely, $\langle \mathbf{0}, \mathbf{0} \rangle = \mathbf{A}\mathbf{0} \cdot \mathbf{A}\mathbf{0} = \mathbf{0} \cdot \mathbf{0} = 0$.

(3) $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{A}\mathbf{x} \cdot \mathbf{A}\mathbf{y} = \mathbf{A}\mathbf{y} \cdot \mathbf{A}\mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle$

(4) $\langle \mathbf{x}+\mathbf{y}, \mathbf{z}\rangle = \mathbf{A}(\mathbf{x}+\mathbf{y}) \cdot \mathbf{Az} = (\mathbf{Ax}+\mathbf{Ay}) \cdot \mathbf{Az} = \mathbf{Ax} \cdot \mathbf{Az} + \mathbf{Ay} \cdot \mathbf{Az} = \langle \mathbf{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle$

(5) $\langle k\mathbf{x}, \mathbf{y}\rangle = \mathbf{A}(k\mathbf{x}) \cdot \mathbf{Ay} = (k\mathbf{Ax}) \cdot (\mathbf{Ay}) = k((\mathbf{Ax}) \cdot (\mathbf{Ay})) = k\langle \mathbf{x}, \mathbf{y}\rangle$

Also, $\langle \mathbf{x}, \mathbf{y}\rangle = 35$ and $\|\mathbf{x}\| = 7$.

(14) The distance between $\mathbf{w}$ and $\mathbf{z}$ is 8.

(15) $\frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) = \frac{1}{4}(\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle)$
$= \frac{1}{4}(\langle \mathbf{x}+\mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{x}+\mathbf{y}, \mathbf{y}\rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{y}\rangle)$
$= \frac{1}{4}(\langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle - \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{y}, \mathbf{y}\rangle)$
$= \frac{1}{4}(4\langle \mathbf{x}, \mathbf{y}\rangle) = \langle \mathbf{x}, \mathbf{y}\rangle$.

(16) $\{x^2, 17x - 9x^2, \ 2 - 3x + x^2\}$

(17) $\mathbf{w}_1 = \frac{25}{4}x^2$, $\mathbf{w}_2 = x + 3 - \frac{25}{4}x^2$

Answers to Test for Chapter 7 — Version C

(1) $\mathbf{A}^T\mathbf{B}^* = \begin{bmatrix} 14-5i & 1+4i \\ 8-7i & 8-3i \end{bmatrix}$

(2) $\mathbf{HJ} = (\mathbf{HJ})^* = \mathbf{J}^*\mathbf{H}^* = \mathbf{JH}$, because $\mathbf{H}$, $\mathbf{J}$, and $\mathbf{HJ}$ are all Hermitian.

(3) (a) True

(b) True

(4) $z_1 = 4-3i$, $z_2 = -1+i$

(5) Let $\mathbf{A}$ be an $n\times n$ complex matrix that is not diagonalizable. By the Fundamental Theorem of Algebra, $p_{\mathbf{A}}(x)$ factors into n linear factors. Hence, $p_{\mathbf{A}}(x) = (x-\lambda_1)^{k_1}\cdots(x-\lambda_m)^{k_m}$, where $\lambda_1,\ldots,\lambda_m$ are the eigenvalues of $\mathbf{A}$, and $k_1,\ldots,k_m$ are their corresponding algebraic multiplicities. Now, since the degree of $p_{\mathbf{A}}(x)$ is n, $\sum_{i=1}^m k_i = n$. Hence, if each geometric multiplicity actually equaled each algebraic multiplicity for each eigenvalue, then the sum of the geometric multiplicities would also be n, implying that $\mathbf{A}$ is diagonalizable. However, since $\mathbf{A}$ is not diagonalizable, some geometric multiplicity must be less than its corresponding algebraic multiplicity.

(6) $\lambda_1 = 2$, $\lambda_2 = 2i$; basis for $E_2 = \{[1,1]\}$, basis for $E_{2i} = \{[1,i]\}$. L is diagonalizable.

(7) We need to find normal matrices $\mathbf{Z}$ and $\mathbf{W}$ such that $(\mathbf{Z}+\mathbf{W})$ is not normal. Note that $\mathbf{Z} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and $\mathbf{W} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$ are normal matrices ($\mathbf{Z}$ is Hermitian and $\mathbf{W}$ is skew-Hermitian). Let $\mathbf{Y} = \mathbf{Z}+\mathbf{W} = \begin{bmatrix} 1 & 1+i \\ 1+i & 0 \end{bmatrix}$. Then $\mathbf{YY}^* = \begin{bmatrix} 1 & 1+i \\ 1+i & 0 \end{bmatrix}\begin{bmatrix} 1 & 1-i \\ 1-i & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1-i \\ 1+i & 2 \end{bmatrix}$, while $\mathbf{Y}^*\mathbf{Y} = \begin{bmatrix} 1 & 1-i \\ 1-i & 0 \end{bmatrix}\begin{bmatrix} 1 & 1+i \\ 1+i & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1+i \\ 1-i & 2 \end{bmatrix}$, and so $\mathbf{Y}$ is not normal.

(8) One possibility: $B = \{[1,3+i,i,1-i],[1+i,3+4i,-1+i,2],[1,0,0,0],[0,0,1,0]\}$

(9) Let j be such that w_j, the jth entry of $\mathbf{w}$, does not equal zero. Then, $L(i\mathbf{e}_j) = w_j(\bar{i}) = w_1(-i)$. However, $iL(\mathbf{e}_j) = i(w_j)(1)$. Setting these equal gives $-iw_j = iw_j$, which implies $0 = 2iw_j$, and thus $w_j = 0$. But this contradicts the assumption that $w_j \neq 0$.

(10) (Optional Hint: The last vector in the given basis is already orthogonal to the first two.)
Answer: $\{[2-i,1+2i,1],[5i,6,-1+2i],[-1,0,2+i]\}$

(11) Because the rows of $\mathbf{Z}$ form an orthonormal basis for $\mathbb{C}^n$, by part (1) of Theorem 7.7, $\mathbf{Z}$ is a unitary matrix. Hence, by part (2) of Theorem 7.7, the columns of $\mathbf{Z}$ form an orthonormal basis for $\mathbb{C}^n$.

(12) (a) Note that $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A} = \begin{bmatrix} 1825 & 900i \\ -900i & 1300 \end{bmatrix}$.

(b) (Optional Hint: $p_{\mathbf{A}}(x) = x^2 - 25ix + 1250$, or, $\lambda_1 = 25i$ and $\lambda_2 = -50i$.)
Answer: $\mathbf{P} = \frac{1}{5}\begin{bmatrix} 3 & 4 \\ 4i & -3i \end{bmatrix}$

(13) Proofs for the five properties of an inner product:
Property 1: $\langle \mathbf{x}, \mathbf{x} \rangle = 2x_1x_1 - x_1x_2 - x_2x_1 + x_2x_2 = x_1^2 + x_1^2 - 2x_1x_2 + x_2^2 = x_1^2 + (x_1 - x_2)^2 \geq 0$.
Property 2: $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ precisely when $x_1 = 0$ and $x_1 - x_2 = 0$, which is when $x_1 = x_2 = 0$.
Property 3: $\langle \mathbf{y}, \mathbf{x} \rangle = 2y_1x_1 - y_1x_2 - y_2x_1 + y_2x_2 = 2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2 = \langle \mathbf{x}, \mathbf{y} \rangle$
Property 4: Let $\mathbf{z} = [z_1, z_2]$. Then, $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle$
$= 2(x_1 + y_1)z_1 - (x_1 + y_1)z_2 - (x_2 + y_2)z_1 + (x_2 + y_2)z_2$
$= 2x_1z_1 + 2y_1z_1 - x_1z_2 - y_1z_2 - x_2z_1 - y_2z_1 + x_2z_2 + y_2z_1$
$= 2x_1z_1 - x_1z_2 - x_2z_1 + x_2z_2 + 2y_1z_1 - y_1z_2 - y_2z_1 + y_2z_1$
$= \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
Property 5: $\langle k\mathbf{x}, \mathbf{y} \rangle = 2(kx_1)y_1 - (kx_1)y_2 - (kx_2)y_1 + (kx_2)y_2$
$= k(2x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2) = k\langle \mathbf{x}, \mathbf{y} \rangle$
Also, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ and $\|\mathbf{x}\| = \sqrt{5}$.

(14) Distance between $\mathbf{f}$ and $\mathbf{g}$ is $\sqrt{\pi}$.

(15) $\mathbf{x} \perp \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle = 0 \iff 2\langle \mathbf{x}, \mathbf{y} \rangle = 0 \iff \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$
$\iff \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \iff \langle \mathbf{x}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$
$\iff \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \iff \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$

(16) (Optional Hint: $\langle [-1, 1, 1], [3, -4, -1] \rangle = 0$) Answer: $\{[-1, 1, 1], [3, -4, -1], [-1, 2, 0]\}$

(17) $\mathbf{w}_1 = 2x - \frac{1}{3}$, $\mathbf{w}_2 = x^2 - 2x + \frac{1}{3}$

