lyryx with Open Texts

LINEAR ALGEBRA
with Applications

Open Edition

ADAPTABLE | ACCESSIBLE | AFFORDABLE

PARTIAL STUDENT SOLUTION MANUAL

VERSION 2019 – REVISION A

by W. Keith Nicholson

Champions of Access to Knowledge

OPEN TEXT

All digital forms of access to our high-quality open texts are entirely FREE! All content is reviewed for excellence and is wholly adaptable; custom editions are produced by Lyryx for those adopting Lyryx assessment. Access to the original source files is also open to anyone!

ONLINE ASSESSMENT

We have been developing superior online formative assessment for more than 15 years. Our questions are continuously adapted with the content and reviewed for quality and sound pedagogy. To enhance learning, students receive immediate personalized feedback. Student grade reports and performance statistics are also provided.

SUPPORT

Access to our in-house support team is available 7 days/week to provide prompt resolution to both student and instructor inquiries. In addition, we work one-on-one with instructors to provide a comprehensive system, customized for their course. This can include adapting the text, managing multiple sections, and more!

INSTRUCTOR SUPPLEMENTS

Additional instructor resources are also freely accessible. Product dependent, these supplements include: full sets of adaptable slides and lecture notes, solutions manuals, and multiple choice question banks with an exam building tool.

Linear Algebra with Applications

Open Edition

BE A CHAMPION OF OER!

Contribute suggestions for improvements, new content, or errata:
A new topic
A new example
An interesting new question
A new or better proof to an existing theorem
Any other suggestions to improve the material

Contact Lyryx at info@lyryx.com with your ideas.

CONTRIBUTIONS

Author
W. Keith Nicholson, University of Calgary

Lyryx Learning Team

Bruce Bauslaugh
Peter Chow
Nathan Friess
Stephanie Keyowski
Claude Laflamme
Martha Laflamme
Jennifer MacKenzie
Tamsyn Murnaghan
Bogdan Sava
Ryan Yee

LICENSE

advancing learning

Linear Algebra with Applications

Open Edition

Base Text Revision History

Current Revision: Version 2019 — Revision A

2019 A	• New Section on Singular Value Decomposition (8.6) is included. • New Example 2.3.2 and Theorem 2.2.4. Please note that this will impact the numbering of subsequent examples and theorems in the relevant sections. • Section 2.2 is renamed as *Matrix-Vector Multiplication*. • Minor revisions made throughout, including fixing typos, adding exercises, expanding explanations, and other small edits.
2018 B	• Images have been converted to LaTeX throughout. • Text has been converted to LaTeX with minor fixes throughout. Page numbers will differ from 2018A revision. Full index has been implemented.
2018 A	• Text has been released with a Creative Commons license.

Contents

1. Systems of Linear Equations

1.1 Solutions and Elementary Operations

1. b. Substitute these values of x_1, x_2, x_3 and x_4 in the equation

$$\begin{array}{rclcl} 2x_1+5x_2+9x_3+3x_4 & = & 2(2s+12t+13)+5(s)+9(-s-3t-3)+3(t) & = & -1 \\ x_1+2x_2+4x_3 & = & (2s+12t+13)+2(s)+4(-s-3t-3) & = & 1 \end{array}$$

Hence this is a solution for every value of s and t.

2. b. The equation is $2x+3y=1$. If $x=s$ then $y=\frac{1}{3}(1-2s)$ so this is one form of the general solution. Also, if $y=t$ then $x=\frac{1}{2}(1-3t)$ gives another form.

4. Given the equation $4x-2y+0z=1$, take $y=s$ and $z=t$ and solve for x: $x=\frac{1}{4}(2s+3)$. This is the general solution.

5. a. If $a=0$, no solution if $b\neq 0$, infinitely many if $b=0$.

   b. If $a\neq 0$ unique solution $x=b/a$ for all b.

7. b. The augmented matrix is $\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 1 & 1 \end{array}\right]$.

   d. The augmented matrix is $\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 1 & 2 \end{array}\right]$.

8. b. A system with this augmented matrix is

$$\begin{array}{rcrcrcr} 2x & - & y & & & = & -1 \\ -3x & + & 2y & + & z & = & 0 \\ & & y & + & z & = & 3 \end{array}$$

9. b. $\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & -1 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & -2 & -4 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & 2 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 2 \end{array}\right]$.
   Hence $x=-3$, $y=2$.

   d. $\left[\begin{array}{cc|c} 3 & 4 & 1 \\ 4 & 5 & -3 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 4 & 5 & -3 \\ 3 & 4 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1 & -4 \\ 3 & 4 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 1 & -4 \\ 0 & 1 & 13 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & -17 \\ 0 & 1 & 13 \end{array}\right]$.
   Hence $x=-17$, $y=13$.

10. b. $\left[\begin{array}{ccc|c} 2 & 1 & 1 & -1 \\ 1 & 2 & 1 & 0 \\ 3 & 0 & -2 & 5 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 1 & 1 & -1 \\ 3 & 0 & -2 & 5 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -3 & -1 & -1 \\ 0 & -6 & -5 & 5 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & -3 & 7 \end{array}\right]$
    $\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{7}{3} \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{9} \\ 0 & 1 & 0 & \frac{10}{9} \\ 0 & 0 & 1 & -\frac{7}{3} \end{array}\right]$. Hence $x=\frac{1}{9}$, $y=\frac{10}{9}$, $z=\frac{-7}{3}$.

11. b. $\left[\begin{array}{rr|r} 3 & -2 & 5 \\ -12 & 8 & 16 \end{array}\right] \to \left[\begin{array}{rr|r} 3 & -2 & 5 \\ 0 & 0 & 36 \end{array}\right]$. The last equation is $0x+0y=36$, which has no solution.

14. b. False. The system $x+y=0$, $x-y=0$ is consistent, but $x=0=y$ is the only solution.

    d. True. If the original system was consistent the final system would also be consistent because each row operation produces a system with the same set of solutions (by Theorem 1.1.1).

16. The substitution gives

$$\begin{aligned} 3(5x'-2y')+2(-7x'+3y') &= 5 \\ 7(5x'-2y')+5(-7x'+3y') &= 1 \end{aligned}$$

this simplifies to $x'=5$, $y'=1$. Hence $x=5x'-2y'=23$ and $y=-7x'+3y'=-32$.

17. As in the Hint, multiplying by $(x^2+2)(2x-1)$ gives $x^2-x+3=(ax+b)(2x-1)+c(x^2+2)$. Equating coefficients of powers of x gives equations $2a+c=1$, $-a+2b=-1$, $-b+2c=3$. Solving this linear system we find $a=-\frac{1}{9}$, $b=-\frac{5}{9}$, $c=\frac{11}{9}$.

19. If John gets \$$x$ per hour and Joe gets \$$y$ per hour, the two situations give $2x+3y=24.6$ and $3x+2y=23.9$. Solving gives $x=\$4.50$ and $y=\$5.20$.

1.2 Gaussian Elimination

1. b. No, No; no leading 1.

   d. No, Yes; not in reduced form because of the 3 and the top two 1's in the last column.

   f. No, No; the (reduced) row-echelon form would have two rows of zeros.

2. b.

$$\begin{bmatrix} 0 & -1 & 3 & 1 & 3 & 2 & 1 \\ 0 & -2 & 6 & 1 & -5 & 0 & -1 \\ 0 & 3 & -9 & 2 & 4 & 1 & -1 \\ 0 & 1 & -3 & -1 & 3 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & -3 & -1 & -3 & -2 & -1 \\ 0 & 0 & 0 & -1 & -11 & -4 & -3 \\ 0 & 0 & 0 & 5 & 13 & 7 & 2 \\ 0 & 0 & 0 & 0 & 6 & 2 & 2 \end{bmatrix}$$

$$\to \begin{bmatrix} 0 & 1 & -3 & 0 & 8 & 2 & 2 \\ 0 & 0 & 0 & 1 & 11 & 4 & 3 \\ 0 & 0 & 0 & 0 & -42 & -13 & -13 \\ 0 & 0 & 0 & 0 & 6 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & -3 & 0 & 8 & 2 & 2 \\ 0 & 0 & 0 & 1 & 11 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 & 1 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 0 & 1 & -3 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 1 & 11 & 0 & -1 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

3. b. The matrix is already in reduced row-echelon form. The nonleading variables are parameters; $x_2=r$, $x_4=s$ and $x_6=t$.
   The first equation is $x_1-2x_2+2x_4+x_6=1$, whence $x_1=1+2r-2s-t$.
   The second equation is $x_3+5x_4-3x_6=-1$, whence $x_3=-1-5s+3t$.
   The third equation is $x_5+6x_6=1$, whence $x_5=1-6t$.

   d. First carry the matrix to reduced row-echelon form.

$$\left[\begin{array}{rrrrr|r} 1 & -1 & 2 & 4 & 6 & 2 \\ 0 & 1 & 2 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 0 & 4 & 5 & 5 & 1 \\ 0 & 1 & 2 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 0 & 4 & 0 & 5 & -4 \\ 0 & 1 & 2 & 0 & -1 & -2 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The nonleading variables are parameters; $x_3 = s$, $x_5 = t$.
The first equation is $x_1 + 4x_3 + 5x_5 = -4$, whence $x_1 = -4 - 4s - 5t$.
The second equation is $x_2 + 2x_3 - x_5 = -2$, whence $x_2 = -2 - 2s + t$.
The third equation is $x_4 = 1$.

4. b. $\left[\begin{array}{rr|r} 3 & -1 & 0 \\ 2 & -3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 2 & -1 \\ 2 & -3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 2 & -1 \\ 0 & -7 & 3 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 2 & -1 \\ 0 & 1 & -\frac{3}{7} \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 0 & -\frac{1}{7} \\ 0 & 1 & -\frac{3}{7} \end{array}\right]$.
Hence $x = -\frac{1}{7}$, $y = -\frac{3}{7}$.

d. Note that the variables in the second equation are in the wrong order.
$\left[\begin{array}{rr|r} 3 & -1 & 2 \\ -6 & 2 & -4 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 3 & -1 & 2 \\ 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 \end{array}\right]$.
The nonleading variable $y = t$ is a parameter; then $x = \frac{2}{3} + \frac{1}{3}t = \frac{1}{3}(t+2)$.

f. Again the order of the variables is reversed in the second equation.
$\left[\begin{array}{rr|r} 2 & -3 & 5 \\ -2 & 3 & 2 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 2 & -3 & 5 \\ 0 & 0 & 7 \end{array}\right]$. There is no solution as the second equation is $0x + 0y = 7$.

5. b. $\left[\begin{array}{rrr|r} -2 & 3 & 3 & -9 \\ 3 & -4 & 1 & 5 \\ -5 & 7 & 2 & -14 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 3 & -4 & 1 & 5 \\ -2 & 3 & 3 & -9 \\ -5 & 7 & 2 & -14 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 4 & -4 \\ -2 & 3 & 3 & -9 \\ -5 & 7 & 2 & -14 \end{array}\right]$
$\rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 4 & -4 \\ 0 & 1 & 11 & -17 \\ 0 & 2 & 22 & -34 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 15 & -21 \\ 0 & 1 & 11 & -17 \\ 0 & 0 & 0 & 0 \end{array}\right]$.
Take $z = t$ (the nonleading variable). The equations give $x = -21 - 15t$, $y = -17 - 11t$.

d. $\left[\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 2 & 5 & -3 & 1 \\ 1 & 4 & -3 & 3 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & 1 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 7 \end{array}\right]$.
There is no solution as the third equation is $0x + 0y + 0z = 7$.

f. $\left[\begin{array}{rrr|r} 3 & -2 & 1 & -2 \\ 1 & -1 & 3 & 5 \\ -1 & 1 & 1 & -1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 3 & 5 \\ 3 & -2 & 1 & -2 \\ -1 & 1 & 1 & -1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 3 & 5 \\ 0 & 1 & -8 & -17 \\ 0 & 0 & 4 & 4 \end{array}\right]$
$\rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -5 & -12 \\ 0 & 1 & -8 & -17 \\ 0 & 0 & 1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -9 \\ 0 & 0 & 1 & 1 \end{array}\right]$. Hence $x = -7$, $y = -9$, $z = 1$.

h. $\left[\begin{array}{rrr|r} 1 & 2 & -4 & 10 \\ 2 & -1 & 2 & 5 \\ 1 & 1 & -2 & 7 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -4 & 10 \\ 0 & -5 & 10 & -15 \\ 0 & -1 & 2 & -3 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -4 & 10 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 4 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$.
Hence $z = t$, $x = 4$, $y = 3 + 2t$.

6. b. Label the rows of the augmented matrix as R_1, R_2 and R_3, and begin the gaussian algorithm on the augmented matrix keeping track of the row operations:

$$\left[\begin{array}{rrr|r} 1 & 2 & -3 & -3 \\ 1 & 3 & -5 & 5 \\ 1 & -2 & 5 & -35 \end{array}\right] \begin{array}{c} R_1 \\ R_2 \\ R_3 \end{array} \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & -5 & 5 \\ 0 & 1 & -2 & 8 \\ 0 & -4 & 8 & -32 \end{array}\right] \begin{array}{c} R_2 \\ R_2 - R_1 \\ R_3 - R_1 \end{array}$$

At this point observe that $R_3 - R_1 = -4(R_2 - R_1)$, that is $R_3 = 5R_1 - 4R_2$. This means that equation 3 is 5 times equation 1 minus 4 times equation 2, as is readily verified. (The solution is $x_1 = t - 11$, $x_2 = 2t + 8$ and $x_3 = t$.)

7. b. $\left[\begin{array}{rrrr|r} 1 & -1 & 1 & -1 & 0 \\ -1 & 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 2 & -2 & 2 & 0 \\ 0 & 2 & 0 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{array}\right]$
$\rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$. Hence $x_4 = t$; $x_1 = 0$, $x_2 = -t$, $x_3 = 0$.

d. $\left[\begin{array}{cccc|c} 1 & 1 & 2 & -1 & 4 \\ 0 & 3 & -1 & 4 & 2 \\ 1 & 2 & -3 & 5 & 0 \\ 1 & 1 & -5 & 6 & -3 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & -1 & 4 \\ 0 & 3 & -1 & 4 & 2 \\ 0 & 1 & -5 & 6 & -4 \\ 0 & 0 & -7 & 7 & -7 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 7 & -7 & 8 \\ 0 & 0 & 14 & -14 & 14 \\ 0 & 1 & -5 & 6 & -4 \\ 0 & 0 & -7 & 7 & -7 \end{array}\right]$

$\rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 7 & -7 & 8 \\ 0 & 1 & -5 & 6 & -4 \\ 0 & 0 & 14 & -14 & 14 \\ 0 & 0 & -7 & 7 & -7 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & -5 & 6 & -4 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$

Hence $x_4 = t$; $x_1 = 1$, $x_2 = 1 - t$, $x_3 = 1 + t$.

8. b. $\left[\begin{array}{cc|c} 1 & b & -1 \\ a & 2 & 5 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & b & -1 \\ 0 & 2-ab & 5+a \end{array}\right].$

Case 1 If $ab \neq 2$, it continues $\rightarrow \left[\begin{array}{cc|c} 1 & b & -1 \\ 0 & 1 & \frac{5+a}{2-ab} \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & \frac{-2-5b}{2-ab} \\ 0 & 1 & \frac{5+a}{2-ab} \end{array}\right].$

The unique solution is $x = \frac{-2-5b}{2-ab}$, $y = \frac{5+a}{2-ab}$.

Case 2 If $ab = 2$, it is $\left[\begin{array}{cc|c} 1 & b & -1 \\ 0 & 0 & 5+a \end{array}\right]$. Hence there is no solution if $a \neq -5$. If $a = -5$, then $b = \frac{-2}{5}$ and the matrix is $\left[\begin{array}{cc|c} 1 & \frac{-2}{5} & -1 \\ 0 & 0 & 0 \end{array}\right]$. Then $y = t$, $x = -1 + \frac{2}{5}t$.

d. $\left[\begin{array}{cc|c} a & 1 & 1 \\ 2 & 1 & b \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{b}{2} \\ a & 1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{b}{2} \\ 0 & 1-\frac{a}{2} & 1-\frac{ab}{2} \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{b}{2} \\ 0 & 2-a & 2-ab \end{array}\right].$

Case 1 If $a \neq 2$ it continues: $\rightarrow \left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{b}{2} \\ 0 & 1 & \frac{2-ab}{2-a} \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & \frac{b-1}{2-a} \\ 0 & 1 & \frac{2-ab}{2-a} \end{array}\right].$

The unique solution: $x = \frac{b-1}{2-a}$, $y = \frac{2-ab}{2-a}$.

Case 2 If $a = 2$ the matrix is $\left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{b}{2} \\ 0 & 0 & 2(1-b) \end{array}\right]$. Hence there is no solution if $b \neq 1$.

If $b = 1$ the matrix is $\left[\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{array}\right]$, so $y = t$, $x = \frac{1}{2} - \frac{1}{2}t = \frac{1}{2}(1-t)$.

9. b. $\left[\begin{array}{ccc|c} 2 & 1 & -1 & a \\ 0 & 2 & 3 & b \\ 1 & 0 & -1 & c \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & c \\ 0 & 2 & 3 & b \\ 2 & 1 & -1 & a \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & c \\ 0 & 2 & 3 & b \\ 0 & 1 & 1 & a-2c \end{array}\right] \rightarrow$

$\left[\begin{array}{ccc|c} 1 & 0 & -1 & c \\ 0 & 1 & 1 & a-2c \\ 0 & 2 & 3 & b \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & c \\ 0 & 1 & 1 & a-2c \\ 0 & 0 & 1 & b-2a+4c \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & b-2a+5c \\ 0 & 1 & 0 & 3a-b-6c \\ 0 & 0 & 1 & b-2a+4c \end{array}\right].$

Hence, for any values of a, b and c there is a unique solution $x = -2a+b+5c$, $y = 3a-b-6c$, and $z = -2a+b+4c$.

d. $\left[\begin{array}{ccc|c} 1 & a & 0 & 0 \\ 0 & 1 & b & 0 \\ c & 0 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & a & 0 & 0 \\ 0 & 1 & b & 0 \\ 0 & -ac & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -ab & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1+abc & 0 \end{array}\right].$

Case 1 If $abc \neq -1$, it continues: $\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -ab & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$

Hence we have the unique solution $x = 0$, $y = 0$, $z = 0$.

Case 2 If $abc = -1$, the matrix is $\left[\begin{array}{ccc|c} 1 & 0 & -ab & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$, so $z = t$, $x = abt$, $y = -bt$.

Note: It is impossible that there is no solution here: $x = y = z = 0$ always works.

f. $\left[\begin{array}{ccc|c} 1 & a & -1 & 1 \\ -1 & a-2 & 1 & -1 \\ 2 & 2 & a-2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & a & -1 & 1 \\ 0 & 2(a-1) & 0 & 0 \\ 0 & 2(a-1) & a & -1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & a & -1 & 1 \\ 0 & a-1 & 0 & 0 \\ 0 & 0 & a & -1 \end{array}\right].$

Case 1 If $a = 1$ the matrix is $\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$,

so $y = t$, $x = -t$, $z = -1$.

Case 2 If $a = 0$ the last equation is $0x + 0y + 0z = -1$, so there is no solution.

Case 3 If $a \neq 1$ and $a \neq 0$, there is a unique solution:

$$\left[\begin{array}{ccc|c} 1 & a & -1 & 1 \\ 0 & a-1 & 0 & 0 \\ 0 & 0 & a & -1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & a & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{a} \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1-\frac{1}{a} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{a} \end{array}\right].$$

Hence $x = 1 - \frac{1}{a}$, $y = 0$, $z = -\frac{1}{a}$.

10. b. $\left[\begin{array}{cccc} 2 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cccc} 1 & \frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$; rank is 1.

d. It is in row-echelon form; rank is 3.

f. $\left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$; rank is 1.

11. b. $\left[\begin{array}{ccc} -2 & 3 & 3 \\ 3 & -4 & 1 \\ -5 & 7 & 2 \end{array}\right] \to \left[\begin{array}{ccc} 1 & -1 & 4 \\ 3 & -4 & 1 \\ -5 & 7 & 2 \end{array}\right] \to \left[\begin{array}{ccc} 1 & -1 & 4 \\ 0 & -1 & -11 \\ 0 & 2 & 22 \end{array}\right] \to \left[\begin{array}{ccc} 1 & -1 & 4 \\ 0 & 1 & 11 \\ 0 & 0 & 0 \end{array}\right]$; rank is 2.

d. $\left[\begin{array}{cccc} 3 & -2 & 1 & -2 \\ 1 & -1 & 3 & 5 \\ -1 & 1 & 1 & -1 \end{array}\right] \to \left[\begin{array}{cccc} 1 & -1 & 3 & 5 \\ 3 & -2 & 1 & -2 \\ -1 & 1 & 1 & -1 \end{array}\right] \to \left[\begin{array}{cccc} 1 & -1 & 3 & 5 \\ 0 & 1 & -8 & -17 \\ 0 & 0 & 4 & 4 \end{array}\right] \to \left[\begin{array}{cccc} 1 & -1 & 3 & 5 \\ 0 & 1 & -8 & -17 \\ 0 & 0 & 1 & 1 \end{array}\right]$;
rank is 3.

f. $\left[\begin{array}{cccc} 1 & 1 & 2 & a^2 \\ 1 & 1-a & 2 & 0 \\ 2 & 2-a & 6-a & 4 \end{array}\right] \to \left[\begin{array}{cccc} 1 & 1 & 2 & a^2 \\ 0 & -a & 0 & -a^2 \\ 0 & -a & 2-a & 4-2a^2 \end{array}\right] \to \left[\begin{array}{cccc} 1 & 1 & 2 & a^2 \\ 0 & a & 0 & a^2 \\ 0 & 0 & 2-a & 4-a^2 \end{array}\right].$

If $a = 0$ we get $\left[\begin{array}{cccc} 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 4 \end{array}\right] \to \left[\begin{array}{cccc} 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$; rank $= 2$.

If $a = 2$ we get $\left[\begin{array}{cccc} 1 & 1 & 2 & 4 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cccc} 1 & 1 & 2 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$; rank $= 2$.

If $a \neq 0$, $a \neq 2$, we get $\left[\begin{array}{cccc} 1 & 1 & 2 & a^2 \\ 0 & a & 0 & a^2 \\ 0 & 0 & 2-a & 4-a^2 \end{array}\right] \to \left[\begin{array}{cccc} 1 & 1 & 2 & a^2 \\ 0 & 1 & 0 & a \\ 0 & 0 & 1 & 2+a \end{array}\right]$; rank $= 3$.

12. b. False. $A = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$.

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

f. False. The system $2x - y = 0$, $-4x + 2y = 0$ is consistent, but the system $2x - y = 1$, $-4x + 2y = 1$ is not consistent.

h. True. A has 3 rows so there can be at most 3 leading 1's. Hence the rank of A is at most 3.

14. b. We begin the row reduction $\left[\begin{array}{ccc} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{array}\right] \to \left[\begin{array}{ccc} 1 & a & b+c \\ 0 & b-a & a-b \\ 0 & c-a & a-c \end{array}\right]$. Now one of $b - a$ and $c - a$ is nonzero (by hypothesis) so that row provides the second leading 1 (its row becomes $\left[\begin{array}{ccc} 0 & 1 & -1 \end{array}\right]$). Hence further row operations give

$$\to \left[\begin{array}{ccc} 1 & a & b+c \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc} 1 & 0 & b+c+a \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right]$$

which has the given form.

16. b. Substituting the coordinates of the three points in the equation gives

$$\begin{array}{ll} 1+1+a+b+c=0 & a+b+c=-2 \\ 25+9+5a-3b+c=0 & 5a-3b+c=-34 \\ 9+9-3a-3b+c=0 & 3a+3b-c=18 \end{array}$$

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & -2 \\ 5 & -3 & 1 & -34 \\ 3 & 3 & -1 & 18 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 1 & 1 & -2 \\ 0 & -8 & -4 & -24 \\ 0 & 0 & -4 & 24 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 1 & 1 & -2 \\ 0 & 1 & \frac{1}{2} & 3 \\ 0 & 0 & 1 & -6 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{2} & -5 \\ 0 & 1 & \frac{1}{2} & 3 \\ 0 & 0 & 1 & -6 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & -6 \end{array}\right].$$

Hence $a = -2$, $b = 6$, $c = -6$, so the equation is $x^2 + y^2 - 2x + 6y - 6 = 0$.

18. Let a, b and c denote the fractions of the student population in Clubs A, B and C respectively. The new students in Club A arrived as follows: $\frac{4}{10}$ of those in Club A stayed; $\frac{2}{10}$ of those in Club B go to A, and $\frac{2}{10}$ of those in C go to A. Hence

$$a = \tfrac{4}{10}a + \tfrac{2}{10}b + \tfrac{2}{10}c$$

Similarly, looking at students in Club B and C.

$$b = \tfrac{1}{10}a + \tfrac{7}{10}b + \tfrac{2}{10}c$$

$$c = \tfrac{5}{10}a + \tfrac{1}{10}b + \tfrac{6}{10}c$$

Hence

$$\begin{aligned} -6a + 2b + 2c &= 0 \\ a - 3b + 2c &= 0 \\ 5a + b - 4c &= 0 \end{aligned}$$

$$\left[\begin{array}{rrr|r} -6 & 2 & 2 & 0 \\ 1 & -3 & 2 & 0 \\ 5 & 1 & -4 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & -3 & 2 & 0 \\ 0 & -16 & 14 & 0 \\ 0 & 16 & -14 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & -3 & 2 & 0 \\ 0 & 1 & -\frac{7}{8} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -\frac{5}{8} & 0 \\ 0 & 1 & -\frac{7}{8} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus the solution is $a = \frac{5}{8}t$, $b = \frac{7}{8}t$, $c = t$. However $a + b + c = 1$ (because every student belongs to exactly one club) which gives $t = \frac{2}{5}$. Hence $a = \frac{5}{20}$, $b = \frac{7}{20}$, $c = \frac{8}{20}$.

1.3 Homogeneous Equations

1. b. False. $A = \left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$.

   d. False. $A = \left[\begin{array}{rrr|r} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$.

   f. False. $A = \left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right]$.

   h. False. $A = \left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$.

2. b. $\left[\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 1 & 3 & 6 & 0 \\ 2 & 3 & a & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & -1 & a-2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 0 & -9 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & a+3 & 0 \end{array}\right]$.

   Hence there is a nontrivial solution when $a = -3$: $x = 9t$, $y = -5t$, $z = t$.

   d. $\left[\begin{array}{rrr|r} a & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & a & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ a & 1 & 1 & 0 \\ 1 & 1 & a & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 1-a & 1+a & 0 \\ 0 & 0 & a+1 & 0 \end{array}\right]$.

   Hence if $a \neq 1$ and $a \neq -1$, there is a unique, trivial solution. The other cases are as follows:

$$a=1:\left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right]\rightarrow\left[\begin{array}{rrr|r} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]; x=-t, y=t, z=0.$$

$$a=-1:\left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]\rightarrow\left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]; x=t, y=0, z=t.$$

3. b. Not a linear combination. If $a\mathbf{x}+b\mathbf{y}+c\mathbf{z}=\mathbf{v}$ then comparing entries gives equations $2a+b+c=4$, $a+c=3$ and $-a+b-2c=-4$. Now carry the coefficient matrix to reduced form:

$$\left[\begin{array}{rrr|r} 2 & 1 & 1 & 4 \\ 1 & 0 & 1 & 3 \\ -1 & 1 & -2 & -4 \end{array}\right]\rightarrow\left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Hence there is no solution.

d. Here, if $a\mathbf{a}+b\mathbf{y}+c\mathbf{z}=\mathbf{v}$ then comparing entries gives equations $2a+b+c=3$, $a+c=0$ and $-a+b-2c=3$. Carrying the coefficient matrix to reduced form gives

$$\left[\begin{array}{rrr|r} 2 & 1 & 1 & 3 \\ 1 & 0 & 1 & 0 \\ -1 & 1 & -2 & 3 \end{array}\right]\rightarrow\left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

so the general solution is $a=-t$, $b=3+t$ and $c=t$. Taking $t=-1$ gives the linear combination $\mathbf{v}=\mathbf{a}+2\mathbf{y}-\mathbf{z}$.

4. b. We must determine if x, y and z exist such that $\mathbf{y}=x\mathbf{a}_1+y\mathbf{a}_2+z\mathbf{a}_3$. Equating entries here gives equations $-x+3y+z=-1$, $3x+y+z=9$, $2y+z=2$ and $x+z=6$. Carrying the coefficient matrix to reduced form gives

$$\left[\begin{array}{rrrr} -1 & 3 & 1 & -1 \\ 3 & 1 & 1 & 9 \\ 0 & 2 & 1 & 2 \\ 1 & 0 & 1 & 6 \end{array}\right]\rightarrow\left[\begin{array}{rrrr} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

so the unique solution is $x=2$, $y=-1$ and $z=4$. Hence $\mathbf{y}=2\mathbf{a}_1-\mathbf{a}_2+4\mathbf{a}_3$.

5. b. Carry the augmented matrix to reduced form:

$$\left[\begin{array}{rrrrrr} 1 & 2 & -1 & 1 & 1 & 0 \\ -1 & -2 & 2 & 0 & 1 & 0 \\ -1 & -2 & 3 & 1 & 3 & 0 \end{array}\right]\rightarrow\left[\begin{array}{rrrrrr} 1 & 2 & 0 & 2 & 3 & 0 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence the general solution is $x_1=-2r-2s-3t$, $x_2=r$, $x_3=-s-2t$, $x_4=s$ and $x_5=t$. In matrix form, the general solution $\mathbf{x}=\left[\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array}\right]^T$ takes the form

$$\mathbf{x}=\left[\begin{array}{c} -2r-2s-3t \\ r \\ -s-2t \\ s \\ t \end{array}\right]=r\left[\begin{array}{r} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{array}\right]+s\left[\begin{array}{r} -2 \\ 0 \\ -1 \\ 1 \\ 0 \end{array}\right]+t\left[\begin{array}{r} -3 \\ 0 \\ -2 \\ 0 \\ 1 \end{array}\right]$$

Hence $\mathbf{x}$ is a linear combination of the basic solutions.

d. Carry the augmented matrix to reduced form:

$$\left[\begin{array}{rrrrrr} 1 & 1 & -2 & -2 & 2 & 0 \\ 2 & 2 & -4 & -4 & 1 & 0 \\ 1 & -1 & 2 & 4 & 1 & 0 \\ -2 & -4 & 8 & 10 & 1 & 0 \end{array}\right]\rightarrow\left[\begin{array}{rrrrrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -2 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence the general solution $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}^T$ is

$$\mathbf{x} = \begin{bmatrix} -t \\ 2s+3t \\ s \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 3 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Hence $\mathbf{x}$ is a linear combination of the basic solutions.

6. b. The system $\begin{array}{rcr} x+y & = & 1 \\ 2x+2y & = & 2 \\ -x-y & = & -1 \end{array}$ has nontrivial solutions with fewer variables than equations.

7. b. There are $n-r=6-1=5$ parameters by Theorem 1.2.2.

   d. The row-echelon form has four rows and, as it has a row of zeros, has at most 3 leading 1's. Hence $\operatorname{rank} A = r = 1$, 2 or 3 ($r \neq 0$, because A has nonzero entries). Thus there are $n-r=6-r=5$, 4 or 3 parameters.

9. b. Insisting that the graph of $ax+by+cz+d=0$ (the plane) contains the three points leads to three linear equations in the four variables a, b, c and d. There is a nontrivial solution by Theorem 1.3.1.

11. Since the system is consistent there are $n-r$ parameters by Theorem 1.2.2. The system has nontrivial solutions if and only if there is at least one parameter, that is if and only if $n>r$.

1.4 An Application to Network Flows

1. b. There are five flow equations, one for each junction:

$$\begin{array}{ccccccccccccccc} f_1 & - & f_2 & & & & & & & & & & & = & 25 \\ f_1 & & & + & f_3 & & & + & f_5 & & & & & = & 50 \\ & & f_2 & & & + & f_4 & & & & & + & f_7 & = & 60 \\ & & & - & f_3 & + & f_4 & & & + & f_6 & & & = & 75 \\ & & & & & & & & f_5 & + & f_6 & - & f_7 & = & 40 \end{array}$$

$$\left[\begin{array}{rrrrrrr|r} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 25 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 50 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 60 \\ 0 & 0 & -1 & 1 & 0 & 1 & 0 & 75 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 40 \end{array}\right] \to \left[\begin{array}{rrrrrrr|r} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 25 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 25 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 60 \\ 0 & 0 & -1 & 1 & 0 & 1 & 0 & 75 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 40 \end{array}\right]$$

$$\to \left[\begin{array}{rrrrrrr|r} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 50 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 25 \\ 0 & 0 & -1 & 1 & -1 & 0 & 1 & 35 \\ 0 & 0 & -1 & 1 & 0 & 1 & 0 & 75 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 40 \end{array}\right] \to \left[\begin{array}{rrrrrrr|r} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 85 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 60 \\ 0 & 0 & 1 & -1 & 1 & 0 & -1 & -35 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 40 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 40 \end{array}\right]$$

$$\to \left[\begin{array}{rrrrrrr|r} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 85 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 60 \\ 0 & 0 & 1 & -1 & 0 & -1 & 0 & -75 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 40 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

If we use f_4, f_6 , and f_7 as parameters, the solution is

$$f_1 = 85 - f_4 - f_7$$

$$
\begin{aligned}
f_2 &= 60 - f_4 - f_7 \\
f_3 &= -75 + f_4 + f_6 \\
f_5 &= 40 - f_6 + f_7
\end{aligned}
$$

2. b. The solution to (a) gives $f_1 = 55 - f_4$, $f_2 = 20 - f_4 + f_5$, $f_3 = 15 - f_5$. Closing canal BC means $f_3 = 0$, so $f_5 = 15$. Hence $f_2 = 35 - f_4$, so $f_2 \leq 30$ means $f_4 \geq 5$. Similarly $f_1 = 55 - f_4$ so $f_1 \leq 30$ implies $f_4 \geq 25$. Hence the range on f_4 is $25 \leq f_4 \leq 30$.

3. b. The road CD.

1.5 An Application to Electrical Networks

2. The junction and circuit rules give:

$$
\begin{array}{lccccccc}
\text{Left junction} & I_1 & - & I_2 & + & I_3 & = & 0 \\
\text{Right junction} & I_1 & - & I_2 & + & I_3 & = & 0 \\
\text{Top circuit} & 5I_1 & + & 10I_2 & & & = & 5 \\
\text{Lower circuit} & & & 10I_2 & + & 5I_3 & = & 10
\end{array}
$$

$$
\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 5 & 10 & 0 & 5 \\ 0 & 10 & 5 & 10 \end{array}\right] \to
\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 15 & -5 & 5 \\ 0 & 10 & 5 & 10 \end{array}\right] \to
\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \\ 0 & 2 & 1 & 2 \end{array}\right] \to
\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & -1 \\ 0 & 2 & 1 & 2 \end{array}\right]
$$

$$
\to \left[\begin{array}{rrr|r} 1 & 0 & -1 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 5 & 4 \end{array}\right] \to
\left[\begin{array}{rrr|r} 1 & 0 & -1 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 1 & \frac{4}{5} \end{array}\right] \to
\left[\begin{array}{rrr|r} 1 & 0 & 0 & -\frac{1}{5} \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & 1 & \frac{4}{5} \end{array}\right].
$$

Hence $I_1 = -\frac{1}{5}$, $I_2 = \frac{3}{5}$ and $I_3 = \frac{4}{5}$.

4. The equations are:

$$
\begin{array}{ll}
\text{Lower left junction} & I_1 - I_5 - I_6 = 0 \\
\text{Top junction} & I_2 - I_4 + I_6 = 0 \\
\text{Middle junction} & I_2 + I_3 - I_5 = 0 \\
\text{Lower right junction} & I_1 - I_3 - I_4 = 0
\end{array}
$$

Observe that the last of these follows from the others (so may be omitted).

$$
\begin{array}{lr}
\text{Left circuit} & 10I_5 - 10I_6 = 10 \\
\text{Right circuit} & -10I_3 + 10I_4 = 10 \\
\text{Lower circuit} & 10I_3 + 10I_5 = 20
\end{array}
$$

$$
\left[\begin{array}{rrrrrr|r}
1 & 0 & 0 & 0 & -1 & -1 & 0 \\
0 & 1 & 0 & -1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 10 & -1 & 10 \\
0 & 0 & -10 & 10 & 0 & 0 & 10 \\
0 & 0 & 10 & 0 & 10 & 0 & 20
\end{array}\right] \to
\left[\begin{array}{rrrrrr|r}
1 & 0 & 0 & 0 & -1 & -1 & 0 \\
0 & 1 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & -1 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 2
\end{array}\right]
$$

$$
\to \left[\begin{array}{rrrrrr|r}
1 & 0 & 0 & 0 & -1 & -1 & 0 \\
0 & 1 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 2 & -1 & -1 & 1 \\
0 & 0 & 0 & -1 & 2 & 1 & 2
\end{array}\right] \to
\left[\begin{array}{rrrrrr|r}
1 & 0 & 0 & 0 & -1 & -1 & 0 \\
0 & 1 & 0 & 0 & -2 & 0 & -2 \\
0 & 0 & 1 & 0 & 1 & 0 & 2 \\
0 & 0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 0 & 3 & 1 & 5 \\
0 & 0 & 0 & 1 & -2 & -1 & -2
\end{array}\right]
$$

$$\rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & -2 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 3 & 1 & 5 \end{array}\right] \rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 1 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 & 2 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & \frac{3}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 & \frac{3}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} \end{array}\right]. \text{ Hence } I_1 = 2,\ I_2 = 1,\ I_3 = \tfrac{1}{2},\ I_4 = \tfrac{3}{2},\ I_5 = \tfrac{3}{2},\ I_6 = \tfrac{1}{2}.$$

1.6 An Application to Chemical Reactions

2. Suppose $xNH_3 + yCuO \rightarrow zN_2 + wCu + vH_2O$ where x, y, z, w and v are positive integers. Equating the number of each type of atom on each side gives

$$\begin{array}{ll} N : x = 2z & Cu : y = w \\ H : 3x = 2v & O : y = v \end{array}$$

Taking $v = t$ these give $y = t$, $w = t$, $x = \frac{2}{3}t$ and $z = \frac{1}{2}x = \frac{1}{3}t$. The smallest value of t such that there are all integers is $t = 3$, so $x = 2$, $y = 3$, $z = 1$ and $v = 3$. Hence the balanced reaction is

$$2NH_3 + 3CuO \rightarrow N_2 + 3Cu + 3H_2O$$

4. $15Pb(N_3)_2 + 44Cr(MnO_4)_2 \rightarrow 22Cr_2O_3 + 88MnO_2 + 5Pb_3O_4 + 90NO$

Supplementary Exercises: Chapter 1

1. b. No. If the corresponding planes are parallel and distinct, there is no solution. Otherwise they either coincide or have a whole common line of solutions.

2. b. $\left[\begin{array}{cccc|c} 1 & 4 & -1 & 1 & 2 \\ 3 & 2 & 1 & 2 & 5 \\ 1 & -6 & 3 & 0 & 1 \\ 1 & 14 & -5 & 2 & 3 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 4 & -1 & 1 & 2 \\ 0 & -10 & 4 & -1 & -1 \\ 0 & -10 & 4 & -1 & -1 \\ 0 & 10 & -4 & 1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & \frac{6}{10} & \frac{6}{10} & \frac{16}{10} \\ 0 & 1 & -\frac{4}{10} & \frac{1}{10} & \frac{1}{10} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$

   Hence $x_3 = s$, $x_4 = t$ are parameters, and the equations give $x_1 = \frac{1}{10}(16 - 6s - 6t)$ and $x_2 = \frac{1}{10}(1 + 4s - t)$.

3. b. $\left[\begin{array}{ccc|c} 1 & 1 & 3 & a \\ a & 1 & 5 & 4 \\ 1 & a & 4 & a \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 3 & a \\ 0 & 1-a & 5-3a & 4-a^2 \\ 0 & a-1 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 3 & a \\ 0 & 1-a & 5-3a & 4-a^2 \\ 0 & 0 & 3(2-a) & 4-a^2 \end{array}\right].$

   If $a = 1$ the matrix is $\left[\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 3 & 3 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$, so there is no solution.

If $a = 2$ the matrix is $\left[\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$, so $x = 2 - 2t$, $y = -t$, $z = t$.

If $a \neq 1$ and $a \neq 2$ there is a unique solution.

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & a \\ 0 & 1-a & 5-3a & 4-a^2 \\ 0 & 0 & 3(2-a) & 4-a^2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 3 & a \\ 0 & 1 & \frac{3a-5}{a-1} & \frac{a^2-4}{a-1} \\ 0 & 0 & 1 & \frac{a+2}{3} \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & \frac{2}{a-1} & \frac{-a+4}{a-1} \\ 0 & 1 & \frac{3a-5}{a-1} & \frac{a^2-4}{a-1} \\ 0 & 0 & 1 & \frac{a+2}{3} \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{-5a+8}{3(a-1)} \\ 0 & 1 & 0 & \frac{-a-2}{3(a-1)} \\ 0 & 0 & 0 & \frac{a+2}{3} \end{array}\right].$$

Hence $x = \frac{8-5a}{3(a-1)}$, $y = \frac{-a-2}{3(a-1)}$, $z = \frac{a+2}{3}$.

4. If R_1 and R_2 denote the two rows, then the following indicate how they can be interchanged using row operations of the other two types:

$$\left[\begin{array}{c} R_1 \\ R_2 \end{array}\right] \to \left[\begin{array}{c} R_1 + R_2 \\ R_2 \end{array}\right] \to \left[\begin{array}{c} R_1 + R_2 \\ -R_1 \end{array}\right] \to \left[\begin{array}{c} R_2 \\ -R_1 \end{array}\right] \to \left[\begin{array}{c} R_2 \\ R_1 \end{array}\right]$$

Note that only one row operation of Type II was used — a multiplication by -1.

6. Substitute $x = 3$, $y = -1$ and $z = 2$ into the given equations. The result is

$$\begin{array}{r} 3 - a + 2c = 0 \\ 3b - c - 6 = 1 \\ 3a - 2 + 2b = 5 \end{array} \quad \text{that is} \quad \begin{array}{rcrcrcr} a & & & - & 2c & = & 3 \\ & & 3b & - & c & = & 9 \\ 3a & + & 2b & & & = & 7 \end{array}$$

This system of linear equations for a, b and c has unique solution:

$$\left[\begin{array}{rrr|r} 1 & 0 & -2 & 3 \\ 0 & 3 & -1 & 7 \\ 3 & 2 & 0 & 7 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -2 & 3 \\ 0 & 3 & -1 & 7 \\ 0 & 2 & 6 & -2 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -2 & 3 \\ 0 & 1 & -7 & 9 \\ 0 & 2 & 6 & -2 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -2 & 3 \\ 0 & 1 & -7 & 9 \\ 0 & 0 & 20 & -20 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right].$$

Hence $a = 1$, $b = 2$, $c = -1$.

8. $\left[\begin{array}{rrr|r} 1 & 1 & 1 & 5 \\ 2 & -1 & -1 & 1 \\ -3 & 2 & 2 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & 1 & 5 \\ 0 & -3 & -3 & -9 \\ 0 & 5 & 5 & 15 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & 1 & 5 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$

Hence the solution is $x = 2$, $y = 3 - t$, $z = t$. Taking $t = 3 - i$ gives $x = 2$, $y = i$, $z = 3 - i$, as required. If the real system has a unique solution, the solution is real because all the calculations in the gaussian algorithm yield real numbers (all entries in the augmented matrix are real).

2. Matrix Algebra

2.1 Matrix Addition, Scalar Multiplication, and Transposition

1. b. Equating entries gives four linear equations: $a-b=2$, $b-c=2$, $c-d=-6$, $d-a=2$. The solution is $a=-2+t$, $b=-4+t$, $c=-6+t$, $d=t$.

   d. Equating coefficients gives: $a=b$, $b=c$, $c=d$, $d=a$. The solution is $a=b=c=d=t$, t arbitrary.

2. b. $3\begin{bmatrix} 3 \\ -1 \end{bmatrix} - 5\begin{bmatrix} 6 \\ 2 \end{bmatrix} + 7\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 9 \\ -3 \end{bmatrix} - \begin{bmatrix} 30 \\ 10 \end{bmatrix} + \begin{bmatrix} 7 \\ -7 \end{bmatrix} = \begin{bmatrix} 9-30+7 \\ -3-10-7 \end{bmatrix} = \begin{bmatrix} -14 \\ -20 \end{bmatrix}$

   d. $\begin{bmatrix} 3 & -1 & 2 \end{bmatrix} - 2\begin{bmatrix} 9 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 3 & 11 & -6 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 18 & 6 & 8 \end{bmatrix} + \begin{bmatrix} 3 & 11 & -6 \end{bmatrix}$
   $= \begin{bmatrix} 3-18+3 & -1-6+11 & 2-8-6 \end{bmatrix} = \begin{bmatrix} -12 & 4 & -12 \end{bmatrix}$

   f. $\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -4 \\ -2 & 4 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 4 \\ 2 & -4 & 0 \end{bmatrix}$

   h. $3\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}^T - 2\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} = 3\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} - 2\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ 3 & 0 \end{bmatrix} - \begin{bmatrix} 2 & -2 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -1 & -6 \end{bmatrix}$

3. b. $5C - 5\begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 15 & -5 \\ 10 & 0 \end{bmatrix}$

   d. $B+D$ is not defined as B is 2×3 while D is 3×2.

   f. $(A+C)^T = \begin{bmatrix} 2+3 & 1-1 \\ 0+2 & -1+0 \end{bmatrix}^T = \begin{bmatrix} 5 & 0 \\ 2 & -1 \end{bmatrix}^T = \begin{bmatrix} 5 & 2 \\ 0 & -1 \end{bmatrix}$

   h. $A-D$ is not defined as A is 2×2 while D is 3×2.

4. b. Given $3A + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5A - 2\begin{bmatrix} 3 \\ 0 \end{bmatrix}$, subtract $3A$ from both sides to get $\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2A - 2\begin{bmatrix} 3 \\ 0 \end{bmatrix}$. Now add $2\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ to both sides: $2A = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \end{bmatrix}$. Finally, multiply both sides by $\frac{1}{2}$: $A = \frac{1}{2}\begin{bmatrix} 8 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}$.

5. b. Given $2A - B = 5(A+2B)$, add B to both sides to get

$$2A = 5(A+2B) + B = 5A + 10B + B = 5A + 11B$$

   Now subtract $5A$ from both sides: $-3A = 11B$. Multiply by $-\frac{1}{3}$ to get $A = -\frac{11}{3}B$.

6. b. Given $\begin{array}{l} 4X+3Y=A \\ 5X+4Y=B \end{array}$, subtract the first from the second to get $X+Y=B-A$. Now subtract 3 times this equation from the first equation: $X=A-3(B-A)=4A-3B$. Then $X+Y=B-A$ gives $Y=(B-A)-X=(B-A)-(4A-3B)=4B-5A$.
Note that this also follows from the Gaussian Algorithm (with matrix constants):

$$\left[\begin{array}{cc|c} 4 & 3 & A \\ 5 & 4 & B \end{array}\right] \to \left[\begin{array}{cc|c} 5 & 4 & B \\ 4 & 3 & A \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & B-A \\ 4 & 3 & A \end{array}\right]$$
$$\to \left[\begin{array}{cc|c} 1 & 1 & B-A \\ 0 & -1 & 5A-4B \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 4A-3B \\ 0 & 1 & 4B-5A \end{array}\right]$$

7. b. Given $2X-5Y=\left[\begin{array}{cc} 1 & 2 \end{array}\right]$ let $Y=T$ where T is an arbitrary 1×2 matrix. Then $2X=5T+\left[\begin{array}{cc} 1 & 2 \end{array}\right]$ so $X=\frac{5}{2}T+\frac{1}{2}\left[\begin{array}{cc} 1 & 2 \end{array}\right]$, $Y=T$. If $T=\left[\begin{array}{cc} s & t \end{array}\right]$, this gives $X=\left[\begin{array}{cc} \frac{5}{2}s+\frac{1}{2} & \frac{5}{2}t+1 \end{array}\right]$, $Y=\left[\begin{array}{cc} s & t \end{array}\right]$, where s and t are arbitrary.

8. b. $5[3(A-B+2C)-2(3C-B)-A]+2[3(3A-B+C)+2(B-2A)-2C]$
$=5[3A-3B+6C-6C+2B-A]+2[9A-3B+3C+2B-4A-2C]$
$=5[2A-B]+2[5A-B+C]$
$=10A-5B+10A-2B+2C$
$=20A-7B+2C$

9. b. Write $A=\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$. We want p, q, r and s such that

$$\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]=p\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]+q\left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right]+r\left[\begin{array}{cc} 1 & 0 \\ 1 & 0 \end{array}\right]+s\left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]=\left[\begin{array}{cc} p+q+r & q+s \\ r+s & p \end{array}\right]$$

Equating components give four linear equations in p, q, r and s:

$$\begin{array}{ccccccccc} p & + & q & + & r & & & = & a \\ & & q & & & + & s & = & b \\ & & & & r & + & s & = & c \\ p & & & & & & & = & d \end{array}$$

The solution is $p=d$, $q=\frac{1}{2}(a+b-c-d)$, $r=\frac{1}{2}(a-b+c-d)$, $s=\frac{1}{2}(-a+b+c+d)$.

11. b.
$$\begin{array}{rcll} A+A' & = & 0 & \\ -A+(A+A') & = & -A+0 & \text{(add } -A \text{ to both sides)} \\ (-A+A)+A' & = & -A+0 & \text{(associative law} \\ 0+A' & = & -A+0 & \text{(definition of } -A\text{)} \\ A' & = & -A & \text{(property of 0)} \end{array}$$

13. b. If $A=\left[\begin{array}{cccc} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{array}\right]$ and $B=\left[\begin{array}{cccc} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & b_n \end{array}\right]$,
then $A-B=\left[\begin{array}{cccc} a_1-b_1 & 0 & \cdots & 0 \\ 0 & a_2-b_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n-b_n \end{array}\right]$ so $A-B$ is also diagonal.

14. b. $\left[\begin{array}{cc} s & t \\ st & 1 \end{array}\right]$ is symmetric if and only if $t=st$; that is $t(s-1)=0$; that is $s=1$ or $t=0$.

d. This matrix is symmetric if and only if $2s = s$, $3 = t$, $3 = s + t$; that is $s = 0$ and $t = 3$.

15. b. $\begin{bmatrix} 8 & 0 \\ 3 & 1 \end{bmatrix} = \left(3A^T + 2\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right)^T = (3A^T)^T + \left(2\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right)^T = 3A + \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$.

Hence $3A = \begin{bmatrix} 8 & 0 \\ 3 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 3 & -3 \end{bmatrix}$, so $A = \frac{1}{3}\begin{bmatrix} 6 & 0 \\ 3 & -3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$.

d. $4A - 9\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = (2A^T)^T - \left(5\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\right)^T = 2A - 5\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$.

Hence $2A = \begin{bmatrix} 9 & 9 \\ -9 & 0 \end{bmatrix} - \begin{bmatrix} 5 & -5 \\ 0 & 10 \end{bmatrix} = \begin{bmatrix} 4 & 14 \\ -9 & -10 \end{bmatrix}$.

Finally $A = \frac{1}{2}\begin{bmatrix} 4 & 14 \\ -9 & 10 \end{bmatrix} = \begin{bmatrix} 2 & 7 \\ -\frac{9}{2} & -5 \end{bmatrix}$.

16. b. We have $A^T = A$ as A is symmetric. Using Theorem 2.1.2: $(kA)^T = kA^T = kA$; so kA is symmetric.

19. b. False. Take $B = -A$ for any $A \neq 0$.

d. True. The entries on the main diagonal do not change when a matrix is transposed.

f. True. Assume that A and B are symmetric, that is $A^T = A$ and $B^T = B$. Then Theorem 2.1.2 gives

$$(kA + mB)^T = (kA)^T + (mB)^T = kA^T + mB^T = kA + mB$$

for any scalars k and m. This shows that the matrix $kA + mB$ is symmetric.

20. c. If $A = S + W$ as in (b), then $A^T = S^T + W^T = S - W$. Hence $A + A^T = 2S$ and $A - A^T = 2W$, so $S = \frac{1}{2}(A + A^T)$ and $W = \frac{1}{2}(A - A^T)$.

22. b. If $A = [a_{ij}]$ then $(kp)A = [(kp)a_{ij}] = [k(pa_{ij})] = k[pa_{ij}] = k(pA)$.

2.2 Matrix-Vector Multiplication

1. b.
$$\begin{array}{rcrcrcrcr} x_1 & - & 3x_2 & - & 3x_3 & + & 3x_4 & = & 5 \\ & & 8x_2 & & & + & 2x_4 & = & 1 \\ x_1 & + & 2x_2 & + & 2x_3 & + & & = & 2 \\ & & x_2 & + & 2x_3 & - & 5x_4 & = & 0 \end{array}$$

2. b. $x_1\begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix} + x_2\begin{bmatrix} -2 \\ 0 \\ -2 \\ -4 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 1 \\ 7 \\ 9 \end{bmatrix} + x_4\begin{bmatrix} 1 \\ -2 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ 8 \\ 12 \end{bmatrix}$

3. b. By Definition 2.4:

$A\mathbf{x} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -4 \end{bmatrix} + x_3\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ -4x_2 + 5x_3 \end{bmatrix}$

By Theorem 2.2.5: $A\mathbf{x} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \cdot x_1 + 2 \cdot x_2 + 3 \cdot x_3 \\ 0 \cdot x_1 + (-4) \cdot x_2 + 5 \cdot x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ -4x_2 + 5x_3 \end{bmatrix}$

d. By Definition 2.4:

$$A\mathbf{x} = \begin{bmatrix} 3 & -4 & 1 & 6 \\ 0 & 2 & 1 & 5 \\ -8 & 7 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 \begin{bmatrix} 3 \\ 0 \\ -8 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 2 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + x_4 \begin{bmatrix} 6 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 3x_1 - 4x_2 + x_3 + 6x_4 \\ 2x_2 + x_3 + 5x_4 \\ -8x_1 + 7x_2 - 3x_3 \end{bmatrix}$$

By Theorem 2.2.5:

$$A\mathbf{x} = \begin{bmatrix} 3 & -4 & 1 & 6 \\ 0 & 2 & 1 & 5 \\ -8 & 7 & -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \cdot x_1 + (-4) \cdot x_2 + 1 \cdot x_3 + 6 \cdot x_4 \\ 0 \cdot x_1 + 2 \cdot x_2 + 1 \cdot x_3 + 5 \cdot x_4 \\ (-8) \cdot x_1 + 7 \cdot x_2 + (-3) \cdot x_3 + 0 \cdot x_4 \end{bmatrix} = \begin{bmatrix} 3x_1 - 4x_2 + x_3 + 6x_4 \\ 2x_2 + x_3 + 5x_4 \\ -8x_1 + 7x_2 - 3x_3 \end{bmatrix}$$

5. b. $\left[\begin{array}{rrr|r} 1 & -1 & -4 & -4 \\ 1 & 2 & 5 & 2 \\ 1 & 1 & 2 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & -1 & -4 & -4 \\ 0 & 3 & 9 & 6 \\ 0 & 2 & 6 & 4 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -1 & -2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$. Hence $x = t - 2$, $y = 2 - 3t$, $z = t$; that is $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2+t \\ 2-3t \\ t \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$.

Observe that $\begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}$ is a solution to the given system of equations, and $\begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$ is a solution to the associated homogeneous system.

d. $\left[\begin{array}{rrrr|r} 2 & 1 & -1 & -1 & -1 \\ 3 & 1 & 1 & -2 & -2 \\ -1 & -1 & 2 & 1 & 2 \\ -2 & -1 & 0 & 2 & 3 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 1 & -2 & 1 & -2 \\ 0 & -1 & 3 & 1 & 3 \\ 0 & -2 & 7 & 1 & 4 \\ 0 & 1 & -4 & 0 & -1 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -3 & -1 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & -1 & 1 & 2 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -4 & -9 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$.

Hence $x_1 = 3 - t$, $x_2 = 4t - 9$, $x_3 = t - 2$, $x_4 = t$,

so $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3-t \\ -9+4t \\ -2+t \\ t \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 4 \\ 1 \\ 1 \end{bmatrix}$.

Here $\begin{bmatrix} 3 \\ -9 \\ -2 \\ 0 \end{bmatrix}$ is a solution to the given equations, and $\begin{bmatrix} -1 \\ 4 \\ 1 \\ 1 \end{bmatrix}$ is a solution to the associated homogeneous equations.

6. To say that $\mathbf{x}_0$ and $\mathbf{x}_1$ are solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$ of linear equations means simply that $A\mathbf{x}_0 = \mathbf{0}$ and $A\mathbf{x}_1 = \mathbf{0}$. If $s\mathbf{x}_0 + t\mathbf{x}_1$ is any linear combination of $\mathbf{x}_0$ and $\mathbf{x}_1$, we compute:

$$A(s\mathbf{x}_0 + t\mathbf{x}_1) = A(s\mathbf{x}_0) + A(t\mathbf{x}_1) = s(A\mathbf{x}_0) + t(A\mathbf{x}_1) = s\mathbf{0} + t\mathbf{0} = \mathbf{0}$$

using Theorem 2.2.2. This shows that $s\mathbf{x}_0 + t\mathbf{x}_1$ is *also* a solution to $A\mathbf{x} = \mathbf{0}$.

8. b. The reduction of the augmented matrix is

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & 2 & 3 & -4 \\ 3 & 6 & -2 & -3 & -11 & 11 \\ -2 & 4 & -1 & 1 & -8 & 7 \\ -1 & 2 & 0 & 3 & -5 & 3 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 5 & -3 \\ 0 & 0 & 1 & 0 & -2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ so } \mathbf{x} = \begin{bmatrix} -3+2s-5t \\ s \\ -1+2t \\ 0 \\ t \end{bmatrix}$$

is the general solution. Hence $\mathbf{x} = \begin{bmatrix} -3 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \left(s \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -5 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right)$ is the desired expression.

10. b. False. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ has a zero entry, but $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ has no zero row.

d. True. The linear combination $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ equals $A\mathbf{x}$ where, by Theorem 2.2.1, $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ is the matrix with these vectors $\mathbf{a}_i$ as its columns.

f. False. If $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & 0 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ then $A\mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, and this is not a linear combination of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ because it is not a scalar multiple of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

h. False. If $A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$, there is a solution $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ for $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. But there is no solution for $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Indeed, if $\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ then $x - y + z = 1$ and $-x + y - z = 0$. This is impossible.

11. b. If $\begin{bmatrix} x \\ y \end{bmatrix}$ is reflected in the line $y = x$ the result is $\begin{bmatrix} y \\ x \end{bmatrix}$; see the diagram for Example 2.4.12. In other words, $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. So T has matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

d. If $\begin{bmatrix} x \\ y \end{bmatrix}$ is rotated *clockwise* through $\frac{\pi}{2}$ the result is $\begin{bmatrix} y \\ -x \end{bmatrix}$; see Example 2.2.14. Hence $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ so T has matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

13. b. The reflection of $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in the yz-plane keeps y and z the same and negates x. Hence

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}, \text{ so the matrix is } \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

16. Write $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ where $\mathbf{a}_i$ is column i of A for each i. If $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$ where the x_i are scalars, then $A\mathbf{x} = \mathbf{b}$ by Theorem 2.2.1 where $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$; that is $\mathbf{x}$ is a solution to the system $A\mathbf{x} = \mathbf{b}$.

18. b. We are given that $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions to $A\mathbf{x} = \mathbf{0}$; that is $A\mathbf{x}_1 = \mathbf{0}$ and $A\mathbf{x}_2 = \mathbf{0}$. If t is any scalar then, by Theorem 2.2.2, $A(t\mathbf{x}_1) = t(A\mathbf{x}_1) = t\mathbf{0} = \mathbf{0}$. That is, $t\mathbf{x}_1$ is a solution to $A\mathbf{x} = \mathbf{0}$.

22. Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ where $\mathbf{a}_i$ is column i of A for each i, and write $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$. Then

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 & x_2 + y_2 & \cdots & x_n + y_n \end{bmatrix}^T$$

Hence we have

$$\begin{aligned} A(\mathbf{x}+\mathbf{y}) &= (x_1+y_1)\mathbf{a}_1 + (x_2+y_2)\mathbf{a}_2 + \cdots + (x_n+y_n)\mathbf{a}_n && \text{Definition 2.4} \\ &= (x_1\mathbf{a}_1 + y_1\mathbf{a}_1) + (x_2\mathbf{a}_2 + y_2\mathbf{a}_2) + \cdots + (x_n\mathbf{a}_n + y_n\mathbf{a}_n) && \text{Theorem 2.1.1} \\ &= (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n) + (y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + \cdots + y_n\mathbf{a}_n) && \text{Theorem 2.1.1} \\ &= A\mathbf{x} + A\mathbf{y} && \text{Definition 2.4} \end{aligned}$$

2.3 Matrix Multiplication

1. b. $\begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 \\ 1 & 9 & 7 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2-1-2 & 3-9+0 & 1-7+4 \\ 4+0-4 & 6+0+0 & 2+0+8 \end{bmatrix} = \begin{bmatrix} -1 & -6 & -2 \\ 0 & 6 & 10 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 3 & -3 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ -2 & 1 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 3-6+0 & 0+3-18 \end{bmatrix} = \begin{bmatrix} -3 & -15 \end{bmatrix}$

f. $\begin{bmatrix} 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -8 \end{bmatrix} = \begin{bmatrix} 2-1-24 \end{bmatrix} = \begin{bmatrix} -23 \end{bmatrix}$

h. $\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6-5 & -3+3 \\ 10-10 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

j. $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} a' & 0 & 0 \\ 0 & b' & 0 \\ 0 & 0 & c' \end{bmatrix} = \begin{bmatrix} aa'+0+0 & 0+0+0 & 0+0+0 \\ 0+0+0 & 0+bb'+0 & 0+0+0 \\ 0+0+0 & 0+0+0 & 0+0+cc' \end{bmatrix} = \begin{bmatrix} aa' & 0 & 0 \\ 0 & bb' & 0 \\ 0 & 0 & cc' \end{bmatrix}$

2. b. A^2, AB, BC and C^2 are all undefined. The other products are
$BA = \begin{bmatrix} -1 & 4 & -10 \\ 1 & 2 & 4 \end{bmatrix}$, $B^2 = \begin{bmatrix} 7 & -6 \\ -1 & 6 \end{bmatrix}$, $CB = \begin{bmatrix} -2 & 12 \\ 2 & -6 \\ 1 & 6 \end{bmatrix}$, $AC = \begin{bmatrix} 4 & 10 \\ -2 & -1 \end{bmatrix}$,
$CA = \begin{bmatrix} 2 & 4 & 8 \\ -1 & -1 & -5 \\ 1 & 4 & 2 \end{bmatrix}$.

3. b. The given matrix equation becomes $\begin{bmatrix} 2a+a_1 & 2b+b_1 \\ -a+2a_1 & -b+2b_1 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ -1 & 4 \end{bmatrix}$.
Equating coefficients gives linear equations

$$\begin{aligned} 2a+a_1 &= 7 & \qquad 2b+b_1 &= 2 \\ -a+2a_1 &= -1 & \qquad -b+2b_1 &= 4 \end{aligned}$$

The solution is: $a = 3$, $a_1 = 1$; $b = 0$, $b_1 = 2$.

4. b. $A^2 - A - 6I = \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} - \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} - \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

5. b. $A(BC) = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -9 & -16 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} -14 & -17 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} -2 & -1 & -2 \\ 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 5 & 8 \end{bmatrix} = (AB)C$

6. b. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $A \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} A$ becomes $\begin{bmatrix} b & 0 \\ d & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}$ whence $b = 0$ and $a = d$. Hence A has the form $A = \begin{bmatrix} a & 0 \\ c & a \end{bmatrix}$, as required.

7. b. If A is $m \times n$ and B is $p \times q$ then $n = p$ because AB can be formed and $q = m$ because BA can be formed. So B is $n \times m$, A is $m \times n$.

8. b. (i) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

12. b. Write $A = \begin{bmatrix} P & X \\ 0 & Q \end{bmatrix}$ where $P \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, $X = \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$, and $Q = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$. Then $PX + XQ = \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix} = 0$, so $A^2 = \begin{bmatrix} P^2 & PX+XQ \\ 0 & Q^2 \end{bmatrix} = \begin{bmatrix} P^2 & 0 \\ 0 & Q^2 \end{bmatrix}$. Then

$A^4 = \begin{bmatrix} P^2 & 0 \\ 0 & Q^2 \end{bmatrix} \begin{bmatrix} P^2 & 0 \\ 0 & Q^2 \end{bmatrix} = \begin{bmatrix} P^4 & 0 \\ 0 & Q^4 \end{bmatrix}, A^6 = A^4A^2 = \begin{bmatrix} P^6 & 0 \\ 0 & Q^6 \end{bmatrix}, \dots$; in general we claim that

$$A^{2k} = \begin{bmatrix} P^{2k} & 0 \\ 0 & Q^{2k} \end{bmatrix} \text{ for } k = 1,\ 2,\ \dots \tag{$*$}$$

This holds for $k = 1$; if it holds for some $k \geq 1$ then

$$A^{2(k+1)} = A^{2k}A^2 = \begin{bmatrix} P^{2k} & 0 \\ 0 & Q^{2k} \end{bmatrix} \begin{bmatrix} P^2 & 0 \\ 0 & Q^2 \end{bmatrix} = \begin{bmatrix} P^{2(k+1)} & 0 \\ 0 & Q^{2(k+1)} \end{bmatrix}$$

Hence $(*)$ follows by induction in k.
Next $P^2 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$, $P^3 = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$, and we claim that

$$P^m = \begin{bmatrix} 1 & -m \\ 0 & 1 \end{bmatrix} \text{ for } m = 1,\ 2,\ \dots \tag{$**$}$$

It is true for $m = 1$; if it holds for some $m \geq 1$, then

$$P^{m+1} = P^mP = \begin{bmatrix} 1 & -m \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -(m+1) \\ 0 & 1 \end{bmatrix}$$

which proves $(**)$ by induction.
As to Q, $Q^2 = I$ so $Q^{2k} = I$ for all k. Hence $(*)$ and $(**)$ gives

$$A^{2k} = \begin{bmatrix} P^{2k} & 0 \\ 0 & I \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & -2k & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \text{ for } k \geq 1$$

Finally

$$A^{2k+1} = A^{2k} \cdot A = \begin{bmatrix} P^{2k} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P & X \\ 0 & Q \end{bmatrix} = \begin{bmatrix} P^{2k+1} & P^{2k}X \\ 0 & Q \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & -(2k+1) & 2 & -1 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

13. b. $\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} I^2 + X0 & -IX + XI \\ 0I + I0 & -0X + I^2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = I_{2k}$

d. $\begin{bmatrix} I & X^T \end{bmatrix} \begin{bmatrix} -X & I \end{bmatrix}^T = \begin{bmatrix} I & X^T \end{bmatrix} \begin{bmatrix} -X^T \\ I \end{bmatrix} = -IX^T + X^TI = O_k$

f. $\begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix} \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix}$

$\begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^3 = \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} 0 & X^2 \\ X & 0 \end{bmatrix}$

$\begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^4 = \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix} \begin{bmatrix} 0 & X^2 \\ X & 0 \end{bmatrix} = \begin{bmatrix} X^2 & 0 \\ 0 & X^2 \end{bmatrix}$

Continue. We claim that $\begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^{2m} = \begin{bmatrix} X^m & 0 \\ 0 & X^m \end{bmatrix}$ for $m \geq 1$. It is true if $m = 1$ and, if it holds for some m, we have

$$\begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^{2(m+1)} = \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^{2m} \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^2 = \begin{bmatrix} X^m & 0 \\ 0 & X^m \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} X^{m+1} & 0 \\ 0 & X^{m+1} \end{bmatrix}$$

Hence the result follows by induction on m. Now

$$\begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^{2m+1} = \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix}^{2m} \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix} = \begin{bmatrix} X^m & 0 \\ 0 & X^m \end{bmatrix} \begin{bmatrix} 0 & X \\ I & 0 \end{bmatrix} = \begin{bmatrix} 0 & X^{m+1} \\ X^m & 0 \end{bmatrix}$$

for all $m \geq 1$. It also holds for $m = 0$ if we take $X^0 = I$.

14. b. If $YA = 0$ for all $1 \times m$ matrices Y, let Y_i denote row i of I_m. Then row i of $I_m A = A$ is $Y_i A = 0$. Thus each row of A is zero, so $A = 0$.

16. b. $A(B+C-D)+B(C-A+D)-(A+B)C+(A-B)D = AB+AC-AD+BC-BA+BD-AC-BC+AD-BD = AB-BA$.

    d. $(A-B)(C-A)+(C-B)(A-C)+(C-A)^2 = [(A-B)-(C-B)+(C-A)](C-A) = 0(C-A) = 0$.

18. b. We are given that $AC = CA$, so $(kA)C = k(AC) = k(CA) = C(kA)$, using Theorem 2.3.3. Hence kA commutes with C.

20. Since A and B are symmetric, we have $A^T = A$ and $B^T = B$. Then Theorem 2.3.3 gives $(AB)^T = B^T A^T = BA$. Hence $(AB)^T = AB$ if and only if $BA = AB$.

22. b. Let $A = \begin{bmatrix} a & x & y \\ x & b & z \\ y & z & c \end{bmatrix}$. Then the entries on the main diagonal of A^2 are $a^2+x^2+y^2$, $x^2+b^2+z^2$, $y^2+z^2+c^2$. These are all zero if and only if $a = x = y = b = z = c = 0$; that is if and only if $A = 0$.

24. If $AB = 0$ where $A \neq 0$, suppose $BC = I$ for some matrix C. Left multiply this equation by A to get $A = AI = A(BC) = (AB)C = 0C = 0$, a contradiction. So no such matrix C exists.

26. We have $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$, and hence $A^3 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 3 & 0 & 2 & 2 \\ 2 & 0 & 1 & 1 \\ 3 & 1 & 2 & 1 \end{bmatrix}$. Hence there are 3 paths of length 3 from v_1 to v_4 because the (4, 1)-entry of A^3 is 3. Similarly, the fact that the (3, 2)-entry of A^3 is 0 means that there are no paths of length 3 from v_2 to v_3.

27. b. False. If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = J$ then $AJ = A$, but $J \neq I$.

    d. True. Since A is symmetric, we have $A^T = A$. Hence Theorem 2.1.2 gives $(I+A)^T = I^T + A^T = I + A$. In other words, $I + A$ is symmetric.

    f. False. If $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ then $A \neq 0$ but $A^2 = 0$.

    h. True. We are assuming that A commutes with $A+B$, that is $A(A+B) = (A+B)A$. Multiplying out each side, this becomes $A^2 + AB = A^2 + BA$. Subtracting A^2 from each side gives $AB = BA$; that is A commutes with B.

    j. False. Let $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$. Then $AB = 0$ is the zero matrix so *both* columns are zero. However B has *no* zero column.

    l. False. Let $A = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$ as above. Again $AB = 0$ has *both* rows zero, but A has no row of zeros.

28. b. If $A = [a_{ij}]$ the sum of the entries in row i is $\sum_{j=1}^{n} a_{ij} = 1$. Similarly for $B = [b_{ij}]$. If $AB = C = [c_{ij}]$ then c_{ij} is the dot product of row i of A with column j of B, that is $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. Hence the sum of the entries in row i of C is

$$\sum_{j=1}^{n} c_{ij} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a_{ik} \left(\sum_{j=1}^{n} b_{kj} \right) = \sum_{k=1}^{n} a_{ik} \cdot 1 = 1$$

Easier Proof: Let X be the $n \times 1$ column matrix with every entry equal to 1. Then the entries of AX are the row sums of A, so these all equal 1 if and only if $AX = X$. But if also $BX = X$ then $(AB)X = A(BX) = AX = X$, as required.

30. b. If $A = [a_{ij}]$ then the trace of A is the sum of the entries on the main diagonal, that is $\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$. Now the matrix kA is obtained by multiplying every entry of A by k, that is $kA = [ka_{ij}]$. Hence

$$\operatorname{tr}(kA) = ka_{11} + ka_{22} + \cdots + ka_{nn} = k(a_{11} + a_{22} + \cdots + a_{nn}) = k \operatorname{tr} A$$

e. If $A = [a_{ij}]$ the transpose A^T is obtained by replacing each entry a_{ij} by the entry a_{ji} directly across the main diagonal. Hence, write $A^T = [a'_{ij}]$ where $a'_{ij} = a_{ji}$ for all i and j. Let b_i denote the (i, i)-entry of AA^T. Then b_i is the dot product of row i of A and column i of A^T, that is $b_i = \sum_{k=1}^n a_{ik}a'_{ki} = \sum_{k=1}^n a_{ik}a_{ik} = \sum_{k=1}^n a_{ik}^2$. Hence we obtain

$$\operatorname{tr}(AA^T) = \sum_{i=1}^n b_i = \sum_{i=1}^n \left(\sum_{k=1}^n a_{ik}^2 \right) = \sum_{i=1}^n \sum_{k=1}^n a_{ik}^2$$

This is what we wanted.

32. e. We have $Q = P + AP - PAP$ so, since $P^2 = P$,

$$PQ = P^2 + PAP - P^2AP = P + PAP - PAP = P$$

Hence $Q^2 = (P + AP - PAP)Q = PQ + APQ - PAPQ = P + AP - PAP = Q$.

34. b. We always have

$$(A+B)(A-B) = A^2 + BA - AB - B^2$$

If $AB = BA$, this gives $(A+B)(A-B) = A^2 - B^2$. Conversely, suppose that $(A+B)(A-B) = A^2 - B^2$. Then

$$A^2 - B^2 = A^2 + BA - AB - B^2$$

Hence $0 = BA - AB$, whence $AB = BA$.

35. b. Denote $B = \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{array} \right] = \left[\begin{array}{c} \mathbf{b}_j \end{array} \right]$ where $\mathbf{b}_j$ is column j of B. Then Definition 2.9 asserts that

$$AB = \left[\begin{array}{cccc} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{array} \right] = \left[\begin{array}{c} A\mathbf{b}_j \end{array} \right],$$

that is column j of AB is $A\mathbf{b}_j$ for each j. Note that multiplying a matrix by a scalar a is the same as multiplying each column by a. This, with Definition 2.9 and Theorem 2.2.2, gives

$$\begin{aligned} a(AB) &= a[A\mathbf{b}_j] && \text{Definition 2.9} \\ &= [a(A\mathbf{b}_j)] && \text{Scalar Multiplication} \\ &= [A(a\mathbf{b}_j)] && \text{Theorem 2.2.2} \\ &= A(aB) && \text{Definition 2.9} \end{aligned}$$

Similarly,

$$a(AB) = a[A\mathbf{b}_j] \qquad \text{Definition 2.9}$$

$$\begin{aligned} &= [a(A\mathbf{b}_j)] && \text{Scalar Multiplication} \\ &= [(aA)\mathbf{b}_j)] && \text{Theorem 2.2.2} \\ &= (aA)B && \text{Definition 2.9} \end{aligned}$$

This proves that $a(AB) = A(aB) = (aA)B$, as required.

36. See the article in the mathematics journal Communications in Algebra, Volume 25, Number 7 (1997), pages 1767 to 1782.

2.4 Matrix Inverses

2. In each case we need row operations that carry A to I; these same operations carry I to A^{-1}. In short $\left[\begin{array}{cc} A & I \end{array}\right] \to \left[\begin{array}{cc} I & A^{-1} \end{array}\right]$. This is called the matrix inversion algorithm.

b. We begin by subtracting row 2 from row 1.

$$\left[\begin{array}{rr|rr} 4 & 1 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rr|rr} 1 & -1 & 1 & -1 \\ 3 & 2 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rr|rr} 1 & -1 & 1 & -1 \\ 0 & 5 & -3 & 4 \end{array}\right]$$

$$\to \left[\begin{array}{rr|rr} 1 & -1 & 1 & -1 \\ 0 & 1 & -\frac{3}{5} & \frac{4}{5} \end{array}\right] \to \left[\begin{array}{rr|rr} 1 & 0 & \frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & -\frac{3}{5} & \frac{4}{5} \end{array}\right]. \text{ Hence the inverse is } \frac{1}{5}\left[\begin{array}{rr} 2 & -1 \\ -3 & 4 \end{array}\right].$$

d. $$\left[\begin{array}{rrr|rrr} 1 & -1 & 2 & 1 & 0 & 0 \\ -5 & 7 & -11 & 0 & 1 & 0 \\ -2 & 3 & -5 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 2 & -1 & 5 & 1 & 0 \\ 0 & 1 & -1 & 2 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & 1 & 3 & 0 & 1 \\ 0 & 1 & -1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & -2 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 2 & -1 & 3 \\ 0 & 1 & 0 & 3 & 1 & -1 \\ 0 & 0 & 1 & 1 & 1 & -2 \end{array}\right]. \text{ So } A^{-1} = \left[\begin{array}{rrr} 2 & -1 & 3 \\ 3 & 1 & -1 \\ 1 & 1 & -2 \end{array}\right].$$

f. $$\left[\begin{array}{rrr|rrr} 3 & 1 & -1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 5 & -1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & -1 & 1 & -1 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 5 & -1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & -1 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 & 3 & 0 \\ 0 & 5 & 0 & -1 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|rrr} 1 & 0 & -1 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 & 3 & 0 \\ 0 & 0 & -10 & 9 & -14 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{1}{10} & -\frac{4}{10} & \frac{1}{10} \\ 0 & 1 & 0 & -\frac{2}{10} & \frac{2}{10} & \frac{2}{10} \\ 0 & 0 & 1 & -\frac{9}{10} & \frac{14}{10} & -\frac{1}{10} \end{array}\right]. \text{ Hence } A^{-1} = \frac{1}{10}\left[\begin{array}{rrr} 1 & 4 & -1 \\ -2 & 2 & 2 \\ -9 & 14 & -1 \end{array}\right].$$

h. We begin by subtracting row 2 from twice row 1:

$$\left[\begin{array}{rrr|rrr} 3 & 1 & -1 & 1 & 0 & 0 \\ 5 & 2 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & -2 & 2 & -1 & 0 \\ 5 & 2 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & -2 & 2 & -1 & 0 \\ 0 & 2 & 10 & -10 & 6 & 0 \\ 0 & 1 & 1 & -2 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|rrr} 1 & 0 & -2 & 2 & -1 & 0 \\ 0 & 1 & 5 & -5 & 3 & 0 \\ 0 & 0 & -4 & 3 & -2 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & -2 & 2 & -1 & 0 \\ 0 & 1 & 5 & -5 & 3 & 0 \\ 0 & 0 & 1 & -\frac{3}{4} & \frac{2}{4} & -\frac{1}{4} \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & \frac{2}{4} & 0 & -\frac{2}{4} \\ 0 & 1 & 0 & -\frac{5}{4} & \frac{2}{4} & \frac{5}{4} \\ 0 & 0 & 1 & -\frac{3}{4} & \frac{2}{4} & -\frac{1}{4} \end{array}\right].$$

Hence $A^{-1} = \frac{1}{4}\left[\begin{array}{rrr} 2 & 0 & -2 \\ -5 & 2 & 5 \\ -3 & 2 & -1 \end{array}\right]$.

j. $$\left[\begin{array}{rrrr|rrrr} -1 & 4 & 5 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 1 & -2 & -2 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrrr|rrrr} 1 & -4 & -5 & -2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 2 & 3 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

$$\to \left[\begin{array}{rrrr|rrrr} 1 & -4 & -5 & -2 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 2 & 3 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrrr|rrrr} 1 & 0 & -1 & -2 & -1 & 0 & 0 & -4 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 2 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 1 & 0 & -2 & -1 & 0 & -1 & -3 \\ 0 & 0 & 1 & 2 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 1 & 0 & 0 & -1 & -2 & -1 & -3 \\ 0 & 0 & 1 & 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{array}\right].$$

Hence $A^{-1} = \left[\begin{array}{rrrr} 0 & 0 & 1 & -2 \\ -1 & -2 & -1 & -3 \\ 1 & 2 & 1 & 2 \\ 0 & -1 & 0 & 0 \end{array}\right]$.

l. $\left[\begin{array}{rrrrr|rrrrr} 1 & 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrrrr|rrrrr} 1 & 0 & 0 & 0 & 0 & 1 & -2 & 6 & -30 & 210 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & -3 & 15 & -105 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -5 & 35 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$

Hence $A^{-1} = \left[\begin{array}{rrrrr} 1 & -2 & 6 & -30 & 210 \\ 0 & 1 & -3 & 15 & -105 \\ 0 & 0 & 1 & -5 & 35 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$.

3. b. The equations are $A\mathbf{x} = \mathbf{b}$ where $A = \left[\begin{array}{rr} 2 & -3 \\ 1 & -4 \end{array}\right]$, $\mathbf{x} = \left[\begin{array}{c} x \\ y \end{array}\right]$, $\mathbf{b} = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$. We have (by the algorithm or Example 2.4.4) $A^{-1} = \frac{1}{5}\left[\begin{array}{rr} 4 & -3 \\ 1 & -2 \end{array}\right]$. Left multiply $A\mathbf{x} = \mathbf{b}$ by A^{-1} to get

$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{5}\left[\begin{array}{rr} 4 & -3 \\ 1 & -2 \end{array}\right]\left[\begin{array}{c} 0 \\ 1 \end{array}\right] = \frac{1}{5}\left[\begin{array}{r} -3 \\ -2 \end{array}\right]$$

Hence $x = -\frac{3}{5}$ and $y = -\frac{2}{5}$.

d. Here $A = \left[\begin{array}{rrr} 1 & 4 & 2 \\ 2 & 3 & 3 \\ 4 & 1 & 4 \end{array}\right]$, $\mathbf{x} = \left[\begin{array}{c} x \\ y \\ z \end{array}\right]$, $\mathbf{b} = \left[\begin{array}{r} 1 \\ -1 \\ 0 \end{array}\right]$.

By the algorithm, $A^{-1} = \frac{1}{5}\left[\begin{array}{rrr} 9 & -14 & 6 \\ 4 & -4 & 1 \\ -10 & 15 & -5 \end{array}\right]$.

$\frac{1}{5}\left[\begin{array}{rrr} 9 & -14 & 6 \\ 4 & -4 & 1 \\ -10 & 15 & -5 \end{array}\right]\left[\begin{array}{r} 1 \\ -1 \\ 0 \end{array}\right] = \left[\begin{array}{c} \frac{23}{5} \\ \frac{8}{5} \\ -5 \end{array}\right]$

The equations have the form $A\mathbf{x} = \mathbf{b}$, so left multiplying by A^{-1} gives

$$\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} = \frac{1}{5}\left[\begin{array}{rrr} 9 & -14 & 6 \\ 9 & -4 & 1 \\ -10 & 15 & -5 \end{array}\right]\left[\begin{array}{r} 1 \\ -1 \\ 0 \end{array}\right] = \frac{1}{5}\left[\begin{array}{r} 23 \\ 8 \\ -25 \end{array}\right]$$

Hence $x = \frac{23}{5}$, $y = \frac{8}{5}$, and $z = -\frac{25}{5} = -5$.

4. b. We want B such that $AB = P$ where $P = \left[\begin{array}{rrr} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{array}\right]$. Since A^{-1} exists left multiply this equation by A^{-1} to get $B = A^{-1}(AB) = A^{-1}P$. [This B will satisfy our requirements because $AB = A(A^{-1}P) = IP = P$]. Explicitly

$$B = A^{-1}P = \left[\begin{array}{rrr} 1 & -1 & 3 \\ 2 & 0 & 5 \\ -1 & 1 & 0 \end{array}\right]\left[\begin{array}{rrr} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{array}\right] = \left[\begin{array}{rrr} 4 & -2 & 1 \\ 7 & -2 & 4 \\ -1 & 2 & -1 \end{array}\right]$$

5. b. By Example 2.4.4, we have $(2A)^T = \left[\begin{array}{rr} 1 & -1 \\ 2 & 3 \end{array}\right]^{-1} = \frac{1}{5}\left[\begin{array}{rr} 3 & 1 \\ -2 & 1 \end{array}\right]$. Since $(2A)^T = 2A^T$, we get $2A^T = \frac{1}{5}\left[\begin{array}{rr} 3 & 1 \\ -2 & 1 \end{array}\right]$ so $A^T = \frac{1}{10}\left[\begin{array}{rr} 3 & 1 \\ -2 & 1 \end{array}\right]$. Finally

$$A = (A^T)^T = \frac{1}{10}\left[\begin{array}{rr} 3 & 1 \\ -2 & 1 \end{array}\right]^T = \frac{1}{10}\left[\begin{array}{rr} 3 & -2 \\ 1 & 1 \end{array}\right]$$

d. We have $(I-2A^T)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ so (because $(U^{-1})^{-1} = U$ for any invertible matrix U)

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

Thus $2A^T = I - \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$.
This gives $A^T = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$, so

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

f. Given $\left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} A\right)^{-1} = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}$, take inverses to get

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

Now $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$, so left multiply by this to obtain

$$A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\left(\frac{1}{2}\begin{bmatrix} 2 & 0 \\ -2 & 1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} 2 & 0 \\ -6 & 1 \end{bmatrix}$$

h. Given $(A^{-1}-2I)^T = -2\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, take transposes to get

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

Hence $A^{-1} = 2I - 2\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - 2\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ -2 & 2 \end{bmatrix} = 2\begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}$. Finally

$$A = (A^{-1})^{-1} = \left(2\begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}\right)^{-1} = \frac{1}{2}\begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}^{-1} = \frac{1}{2}\left(\frac{1}{-1}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right) = \frac{-1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

6. b. Have $A = (A^{-1})^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & -1 \\ -2 & 1 & -1 \end{bmatrix}$ by the algorithm.

8. b. The equations are $A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} x \\ y \end{bmatrix} = B\begin{bmatrix} x' \\ y' \end{bmatrix}$ where $A = \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -5 & 4 \\ 4 & -3 \end{bmatrix}$. Thus $B = A^{-1}$ (by Example 2.4.4) so the substitution gives $\begin{bmatrix} 7 \\ 1 \end{bmatrix} = A\begin{bmatrix} x \\ y \end{bmatrix} = AB\begin{bmatrix} x' \\ y' \end{bmatrix} = I\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$. Thus $x' = 7, y' = 1$ so $\begin{bmatrix} x \\ y \end{bmatrix} = B\begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} -31 \\ 25 \end{bmatrix}$.

9. b. False. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ are both invertible, but $A+B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ is not.

d. True. If $A^4 = 3I$ then $A(\frac{1}{3}A^3) = I = (\frac{1}{3}A^3)A$, so $A^{-1} = \frac{1}{3}A^3$.

f. False. Take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $AB = B$ and $B \neq 0$, but A is not invertible by Theorem 2.4.5 since $A\mathbf{x} = \mathbf{0}$ where $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

h. True. Since A^2 is invertible, let $(A^2)B = I$. Thus $A(AB) = I$, so AB is the inverse of A by Theorem 2.4.5.

10. b. We are given $C^{-1} = A$, so $C = (C^{-1})^{-1} = A^{-1}$. Hence $C^T = (A^{-1})^T$. This also has the form $C^T = (A^T)^{-1}$ by Theorem 2.4.4. Hence $(C^T)^{-1} = A^T$.

11. b. If a solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$ exists, it can be found by left multiplication by C: $CA\mathbf{x} = C\mathbf{b}$, $I\mathbf{x} = C\mathbf{b}$, $\mathbf{x} = C\mathbf{b}$.

(i) $\mathbf{x} = C\mathbf{b} = \left[\begin{array}{c} 3 \\ 0 \end{array}\right]$ here but $\mathbf{x} = \left[\begin{array}{c} 3 \\ 0 \end{array}\right]$ is not a solution. So no solution exists.

(ii) $\mathbf{x} = C\mathbf{b} = \left[\begin{array}{c} 2 \\ -1 \end{array}\right]$ in this case and this is indeed a solution.

15. b. $B^2 = \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] = \left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \end{array}\right]$ so $B^4 = (B^2)^2 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] = I$.

Thus $B \cdot B^3 = I = B^3 B$, so $B^{-1} = B^3 = B^2 B = \left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \end{array}\right]\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] = \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right]$.

16. We use the algorithm:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ c & 1 & c & 0 & 1 & 0 \\ 3 & c & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -c & 1 & 0 \\ 0 & c & -1 & -3 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -c & 1 & 0 \\ 0 & 0 & -1 & c^2-3 & -c & 1 \end{array}\right]$$

$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & c^2-2 & -c & 1 \\ 0 & 1 & 0 & -c & 1 & 0 \\ 0 & 0 & 1 & 3-c^2 & c & -1 \end{array}\right]$. Hence $\left[\begin{array}{ccc} 1 & 0 & 1 \\ c & 1 & c \\ 3 & c & 2 \end{array}\right]^{-1} = \left[\begin{array}{ccc} c^2-2 & -c & 1 \\ -c & 1 & 0 \\ 3-c^2 & c & -1 \end{array}\right]$ for all values of c.

18. b. Suppose column j of A consists of zeros. Then $A\mathbf{y} = \mathbf{0}$ where $\mathbf{y}$ is the column with 1 in the position j and zeros elsewhere. If A^{-1} exists, left multiply by A^{-1} to get $A^{-1}A\mathbf{y} = A^{-1}\mathbf{0}$, that is $I\mathbf{y} = \mathbf{0}$; a contradiction. So A^{-1} does not exist.

d. If each column of A sums to 0, then $\mathbf{x}A = \mathbf{0}$ where $\mathbf{x}$ is the row of 1s. If A^{-1} exists, right multiply by A^{-1} to get $\mathbf{x}AA^{-1} = \mathbf{0}A^{-1}$, that is $\mathbf{x}I = \mathbf{0}$, $\mathbf{x} = \mathbf{0}$, a contradiction. So A^{-1} does not exist.

19. b. (ii) Write $A = \left[\begin{array}{ccc} 2 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \end{array}\right]$. Observe that row 1 minus row 2 minus row 3 is zero. If $\mathbf{x} = \left[\begin{array}{ccc} 1 & -1 & -1 \end{array}\right]$, this means $\mathbf{x}A = \mathbf{0}$. If A^{-1} exists, right multiply by A^{-1} to get $\mathbf{x}AA^{-1} = \mathbf{0}A^{-1}$, $\mathbf{x}I = \mathbf{0}$, $\mathbf{x} = \mathbf{0}$, a contradiction. So A^{-1} does not exist.

20. b. If A is invertible then each power A^k is also invertible by Theorem 2.4.4. In particular, $A^k \neq 0$.

21. b. If A and B both have inverses, so also does AB (by Theorem 2.4.4). But $AB = 0$ has no inverse.

22. If $a > 1$, the x-expansion $T : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $T\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} ax \\ y \end{array}\right] = \left[\begin{array}{cc} a & 0 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right]$. We have $\left[\begin{array}{cc} a & 0 \\ 0 & 1 \end{array}\right]^{-1} = \left[\begin{array}{cc} \frac{1}{a} & 0 \\ 0 & 1 \end{array}\right]$, and this is an x-compression because $\frac{1}{a} < 1$.

24. b. The condition can be written as $A(A^3 + 2A^2 - I) = 4I$, whence $A[\frac{1}{4}(A^3 + 2A^2 - I)] = I$. By Corollary 2.4.1 of Theorem 2.4.5, A is invertible and $A^{-1} = \frac{1}{4}(A^3 + 2A^2 - I)$. Alternatively, this follows directly by verifying that also $[\frac{1}{4}(A^3 + 2A^2 - I)]A = I$.

25. b. If $B\mathbf{x} = \mathbf{0}$ then $(AB)\mathbf{x} = \mathbf{0}$ so $\mathbf{x} = \mathbf{0}$ because AB is invertible. Hence B is invertible by Theorem 2.4.5. But then $A = (AB)B^{-1}$ is invertible by Theorem 2.4.4 because both AB and B^{-1} are invertible.

26. b. As in Example 2.4.11, $-B^{-1}YA^{-1} = -(-1)^{-1}\begin{bmatrix} 1 & 3 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} -13 & 8 \end{bmatrix}$, so

$$\left[\begin{array}{cc|c} 3 & 1 & 0 \\ 5 & 2 & 0 \\ \hline 1 & 3 & -1 \end{array}\right]^{-1} = \left[\begin{array}{c|c} \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}^{-1} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline \begin{bmatrix} -13 & 8 \end{bmatrix} & (-1)^{-1} \end{array}\right] = \left[\begin{array}{cc|c} 2 & -1 & 0 \\ -5 & 3 & 0 \\ \hline -13 & 8 & -1 \end{array}\right]$$

d. As in Example 2.4.11, $-A^{-1}XB^{-1} = -\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -14 & 8 \\ 16 & -9 \end{bmatrix}$, so

$$\left[\begin{array}{cc|cc} 2 & 1 & 5 & 2 \\ 1 & 1 & -1 & 0 \\ \hline 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -2 \end{array}\right]^{-1} = \left[\begin{array}{c|c} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} & \begin{bmatrix} -14 & 8 \\ 16 & -9 \end{bmatrix} \\ \hline \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix}^{-1} \end{array}\right] = \left[\begin{array}{cc|cc} 1 & -1 & -14 & 8 \\ -1 & 2 & 16 & -9 \\ \hline 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

28. d. If $A^n = 0$ write $B = I + A + A^2 + \cdots + A^{n-1}$. Then

$$\begin{aligned} (I-A)B &= (I-A)(I+A+A^2+\cdots+A^{n-1}) \\ &= (I+A+A^2+\cdots+A^{n-1}) - A - A^2 - A^3 - \cdots - A^n \\ &= I - A^n \\ &= I \end{aligned}$$

Similarly $B(I-A) = I$, so $(I-A)^{-1} = B$.

30. b. Assume that AB and BA are both invertible. Then

$$AB(AB)^{-1} = I \text{ so } AX = I \text{ where } X = B(AB)^{-1}$$
$$(BA)^{-1}BA = I \text{ so } YA = I \text{ where } Y = (BA)^{-1}B$$

But then $X = IX = (YA)X = Y(AX) = YI = Y$, so $X = Y$ is the inverse of A.
Different Proof. The fact that AB is invertible gives $A[B(AB)^{-1}] = I$. This shows that A is invertible by the Corollary 2.4.1 to Theorem 2.4.5. Similarly B is invertible.

31. b. If $A = B$ then $A^{-1}B = A^{-1}A = I$. Conversely, if $A^{-1}B = I$ left multiply by A to get $AA^{-1}B = AI$, $IB = A$, $B = A$.

32. a. Since A commutes with C, we have $AC = CA$. Left-multiply by A^{-1} to get $C = A^{-1}CA$. Then right-multiply by A^{-1} to get $CA^{-1} = A^{-1}C$. Thus A^{-1} commutes with C too.

33. b. The condition $(AB)^2 = A^2B^2$ means $ABAB = AABB$. Left multiplication by A^{-1} gives $BAB = ABB$, and then right multiplication by B^{-1} yields $BA = AB$.

34. Assume that AB is invertible; we apply Part 2 of Theorem 2.4.5 to show that B is invertible. If $B\mathbf{x} = \mathbf{0}$ then left multiplication by A gives $AB\mathbf{x} = \mathbf{0}$. Now left multiplication by $(AB)^{-1}$ yields $\mathbf{x} = (AB)^{-1}\mathbf{0} = \mathbf{0}$. Hence B is invertible by Theorem 2.4.5. But then we have $A = (AB)B^{-1}$ so A is invertible by Theorem 2.4.4 (B^{-1} and AB are both invertible).

35. b. By the hint, $B\mathbf{x} = \mathbf{0}$ where $\mathbf{x} = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$ so B is not invertible by Theorem 2.4.5.

36. Assume that A can be left cancelled. If $A\mathbf{x} = \mathbf{0}$ then $A\mathbf{x} = A\mathbf{0}$ so $\mathbf{x} = \mathbf{0}$ by left cancellation. Thus A is invertible by Theorem 2.4.5. Conversely, if A is invertible, suppose that $AB = AC$. Then left multiplication by A^{-1} yields $A^{-1}AB = A^{-1}AC$, $IB = IC$, $B = C$.

38. b. Write $U = I_n - 2XX^T$. Then U is symmetric because

$$U^T = I_n^T - 2(XX^T)^T = I_n - 2X^{TT}X^T = I_n - 2XX^T = U$$

Moreover $U^{-1} = U$ because (since $X^TX = I_n$)

$$\begin{aligned} U^2 &= (I_n - 2XX^T)(I - 2XX^T) \\ &= I_n - 2XX^T - 2XX^T + 4XX^TXX^T \\ &= I_n - 4XX^T + 4XI_mX^T \\ &= I_n \end{aligned}$$

39. b. If $P^2 = P$ then $I - 2P$ is self-inverse because

$$(I-2P)(I-2P) = I - 2P - 2P + 4P^2 = I$$

Conversely, if $I - 2P$ is self-inverse then

$$I = (I-2P)^2 = I - 4P + 4P^2$$

Hence $4P = 4P^2$; so $P = P^2$.

41. b. If A and B are any invertible matrices (of the same size), we compute:

$$A^{-1}(A+B)B^{-1} = A^{-1}AB^{-1} + A^{-1}BB^{-1} = B^{-1} + A^{-1} = A^{-1} + B^{-1}$$

Hence $A^{-1} + B^{-1}$ is invertible by Theorem 2.4.4 because each of A^{-1}, $A+B$, and B^{-1} is invertible. Furthermore

$$(A^{-1}+B^{-1})^{-1} = [A^{-1}(A+B)B^{-1}]^{-1} = (B^{-1})^{-1}(A+B)^{-1}(A^{-1})^{-1} = B(A+B)^{-1}A$$

gives the desired formula.

2.5 Elementary Matrices

1. b. Interchange rows 1 and 3 of I, $E^{-1} = E$.

d. Add (-2) times row 1 of I to row 2. $E^{-1} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$.

f. Multiply row 3 of I by 5. $E^{-1} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{5} \end{array}\right]$.

2. b. $A \to B$ is accomplished by negating row 1, so $E = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$.

   d. $A \to B$ is accomplished by subtracting row 2 from row 1, so $E = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.

   f. $A \to B$ is accomplished by interchanging rows 1 and 2, so $E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

3. b. The possibilities for E are $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$, $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$. In each case EA has a row different from C.

4. If E is Type I, EA and A differ only in the interchanged rows.
   If E is of Type II, EA and A differ only in the row multiplied by a nonzero constant.
   If E is of Type II, EA and A differ only in the row to which a multiple of a row is added.

5. b. No. The zero matrix 0 is not invertible.

6. b. $\left[\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & 0 \\ 5 & 12 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & 0 \\ 0 & 2 & -6 & -5 & 1 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & -3 & -\frac{5}{2} & \frac{1}{2} \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 0 & 7 & \frac{12}{2} & -1 \\ 0 & 1 & -3 & -\frac{5}{2} & \frac{1}{2} \end{array}\right]$ so $UA = R = \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -3 \end{bmatrix}$ where $U = \frac{1}{2}\begin{bmatrix} 12 & -2 \\ -5 & 1 \end{bmatrix}$. This matrix U is the product of the elementary matrices used at each stage:

$$\begin{array}{lll} \begin{bmatrix} 1 & 2 & 1 \\ 5 & 12 & -1 \end{bmatrix} = A & & \\ \qquad\downarrow & & \\ \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -6 \end{bmatrix} = E_1A & & \text{where } E_1 = \begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix} \\ \qquad\downarrow & & \\ \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -3 \end{bmatrix} = E_2E_1A & & \text{where } E_2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \\ \qquad\downarrow & & \\ \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -3 \end{bmatrix} = E_3E_2E_1A & & \text{where } E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \end{array}$$

   d. Just as in (b), we get $UA = R$ where R is reduced row-echelon, and

$$U = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

   is a product of elementary matrices.

7. b. $\left[\begin{array}{ccc|cc} 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{array}\right] \xrightarrow{E_1} \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 0 & 1 \\ 2 & -1 & 0 & 1 & 0 \end{array}\right] \xrightarrow{E_2} \left[\begin{array}{ccc|cc} 3 & 0 & 1 & 1 & 1 \\ 2 & -1 & 0 & 1 & 0 \end{array}\right]$. So $U = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.
   Then $E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $E_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, so $U = E_2E_1$.

8. b.

$$\begin{array}{rcll}
\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} & = & A & \\
 & \downarrow & & \\
\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} & = & E_1A & \text{where } E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
 & \downarrow & & \\
\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} & = & E_2E_1A & \text{where } E_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \\
 & \downarrow & & \\
\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} & = & E_3E_2E_1A & \text{where } E_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
 & \downarrow & & \\
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & = & E_4E_3E_2E_1A & \text{where } E_4 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}
\end{array}$$

Thus $E_4E_3E_2E_1A = I$ so

$$\begin{aligned}
A &= (E_4E_3E_2E_1)^{-1} \\
&= E_1^{-1}E_2^{-2}E_3^{-1}E_4^{-1} \\
&= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

Of course a different sequence of row operations yields a different factorization of A.

d. Analogous to (b), $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$.

10. By Theorem 2.5.3, $UA = R$ for some invertible matrix U. Hence $A = U^{-1}R$ where U^{-1} is invertible.

12. b. $\left[\begin{array}{c|c} A & I \end{array}\right] = \left[\begin{array}{cc|cc} 3 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1 & -1 \\ 2 & 1 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1 & -1 \\ 0 & -1 & -2 & 3 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -3 \end{array}\right]$

so $U = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix}$. Hence, $UA = R = I_2$ in this case so $U = A^{-1}$. Thus, $r = \operatorname{rank} A = 2$ and, taking $V = I_2$, $UAV = UA = I_2$.

d. $\left[\begin{array}{c|c} A & I \end{array}\right] = \left[\begin{array}{cccc|ccc} 1 & 1 & 0 & -1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|ccc} 1 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 4 & -3 & 1 & 0 \\ 0 & -1 & 1 & 4 & -1 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|ccc} 1 & 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & 1 & -1 & -4 & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 & 1 \end{array}\right]$.

Hence, $UA = R$ where $U = \begin{bmatrix} -2 & 1 & 0 \\ 3 & -1 & 0 \\ 2 & -1 & 1 \end{bmatrix}$ and $R = \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Note that $\operatorname{rank} A = 2$. Next,

$\left[\begin{array}{c|c} R^T & I \end{array}\right] = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 & 0 \\ 3 & -4 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -3 & 4 & 0 & 1 \end{array}\right]$ so $V^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -3 & 4 & 0 & 1 \end{bmatrix}$.

Hence, $(UAV)^T = (RV)^T = V^TR^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, so $UAV = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

16. We need a sequence of elementary operations to carry $\begin{bmatrix} U & A \end{bmatrix}$ to $\begin{bmatrix} I & U^{-1}A \end{bmatrix}$. By Lemma 2.5.1 these operations can be achieved by left multiplication by elementary matrices. Observe

$$\begin{bmatrix} I & U^{-1}A \end{bmatrix} = \begin{bmatrix} U^{-1}U & U^{-1}A \end{bmatrix} = U^{-1}\begin{bmatrix} U & A \end{bmatrix} \tag{2.1}$$

Since U^{-1} is invertible, it is a product of elementary matrices (Theorem 2.5.2), say $U^{-1} = E_1E_2\cdots E_k$ where the E_i are elementary. Hence (2.1) shows that $\left[\begin{array}{cc} I & U^{-1}A \end{array}\right] = E_1E_2\cdots E_k\left[\begin{array}{cc} U & A \end{array}\right]$, so a sequence of k row operations carries $\left[\begin{array}{cc} U & A \end{array}\right]$ to $\left[\begin{array}{cc} I & U^{-1}A \end{array}\right]$. Clearly $\left[\begin{array}{cc} I & U^{-1}A \end{array}\right]$ is in reduced row-echelon form.

17. b. $A \overset{r}{\sim} A$ because $A = IA$. If $A \overset{r}{\sim} B$, let $A = UB$, U invertible. Then $B = U^{-1}A$ so $B \overset{r}{\sim} A$. Finally if $A \overset{r}{\sim} B$ and $B \overset{r}{\sim} C$, let $A = UB$ and $B = VC$ where U and V are invertible. Hence $A = U(VC) = (UV)C$ so $A \overset{r}{\sim} C$.

19. b. The matrices row-equivalent to $A = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right]$ are the matrices UA where U is invertible. If $U = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ then $UA = \left[\begin{array}{ccc} 0 & 0 & b \\ 0 & 0 & d \end{array}\right]$ where b and d are not both zero (as U is invertible). Every such matrix arises — use $U = \left[\begin{array}{cc} a & b \\ -b & a \end{array}\right]$ — it is invertible as $a^2 + b^2 \neq 0$ (Example 2.3.5).

22. b. By Lemma 2.5.1, $B = EA$ where E is elementary, obtained from I by multiplying row i by $k \neq 0$. Hence $B^{-1} = A^{-1}E^{-1}$ where E^{-1} is elementary, obtained from I by multiplying row i by $\frac{1}{k}$. But then forming the product $A^{-1}E^{-1}$ is obtained by multiplying *column* i of A^{-1} by $\frac{1}{k}$.

2.6 Matrix Transformations

1. b. Write $\mathbf{a} = \left[\begin{array}{c} 3 \\ 2 \\ -1 \end{array}\right]$, $\mathbf{b} = \left[\begin{array}{c} 2 \\ 0 \\ 5 \end{array}\right]$ and $\mathbf{x} = \left[\begin{array}{c} 5 \\ 6 \\ -13 \end{array}\right]$. We are given $T(\mathbf{a})$ and $T(\mathbf{b})$, and are asked to find $T(\mathbf{x})$. Since T is linear it is enough (by Theorem 2.6.1) to express $\mathbf{x}$ as a linear combination of $\mathbf{a}$ and $\mathbf{b}$. If we set $\mathbf{x} = r\mathbf{a} + s\mathbf{b}$, equating entries gives equations $3r + 2s = 5$, $2r = 6$ and $-r + 5s = -13$. The (unique) solution is $r = 3$, $s = -2$, so $\mathbf{x} = 3\mathbf{a} - 2\mathbf{b}$. Since T is linear we have

$$T(\mathbf{x}) = 3T(\mathbf{a}) - 2T(\mathbf{b}) = 3\left[\begin{array}{c} 3 \\ 5 \end{array}\right] - 2\left[\begin{array}{c} -1 \\ 2 \end{array}\right] = \left[\begin{array}{c} 11 \\ 11 \end{array}\right]$$

2. b. Let $\mathbf{a} = \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array}\right]$, $\mathbf{b} = \left[\begin{array}{c} -1 \\ 1 \\ 2 \\ -4 \end{array}\right]$ and $\mathbf{x} = \left[\begin{array}{c} 5 \\ -1 \\ 2 \\ 4 \end{array}\right]$. We know $T(\mathbf{a})$ and $T(\mathbf{b})$; to find $T(\mathbf{x})$ we express $\mathbf{x}$ as a linear combination of $\mathbf{a}$ and $\mathbf{b}$, and use the assumption that T is linear. If we write $\mathbf{x} = r\mathbf{a} + s\mathbf{b}$, equate entries, and solve the linear equations, we find that $r = 2$ and $s = -3$. Hence $\mathbf{x} = 2\mathbf{a} - 3\mathbf{b}$ so, since T is linear,

$$T(\mathbf{x}) = 2T(\mathbf{a}) - 3T(\mathbf{b}) = 2\left[\begin{array}{c} 5 \\ 1 \\ -3 \end{array}\right] - 3\left[\begin{array}{c} 2 \\ 0 \\ 1 \end{array}\right] = \left[\begin{array}{c} 4 \\ 2 \\ -9 \end{array}\right]$$

3. b. In $\mathbb{R}^2$, we have $\mathbf{e}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$ and $\mathbf{e}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$. We are given that $T(\mathbf{x}) = -\mathbf{x}$ for each $\mathbf{x}$ in $\mathbb{R}^2$. In particular, $T(\mathbf{e}_1) = -\mathbf{e}$ and $T(\mathbf{e}_2) = -\mathbf{e}_2$. Since T is linear, Theorem 2.6.2 gives

$$A = \left[\begin{array}{cc} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} -\mathbf{e}_1 & -\mathbf{e}_2 \end{array}\right] = \left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \end{array}\right]$$

Of course, $T\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$, so in this case we can easily see *directly* that T has matrix $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. However, sometimes Theorem 2.6.2 is necessary.

d. Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. If these vectors are rotated counterclockwise through $\frac{\pi}{4}$, some simple trigonometry shows that $T(\mathbf{e}_1) = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$ and $T(\mathbf{e}_2) = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$. Since T is linear, the matrix A of T is $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix}$.

4. b. Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ denote the standard basis of $\mathbb{R}^3$. Since $T : \mathbb{R}^3 \to \mathbb{R}^3$ is reflection in the uz-plane, we have:
$T(\mathbf{e}_1) = -\mathbf{e}_1$ because $\mathbf{e}_1$ is *perpendicular* to the uz-plane; while
$T(\mathbf{e}_2) = \mathbf{e}_2$ and $T(\mathbf{e}_3) = \mathbf{e}_3$ because $\mathbf{e}_2$ and $\mathbf{e}_3$ are *in* the uz-plane.
So $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \end{bmatrix} = \begin{bmatrix} -\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

5. b. Since $\mathbf{y}_1$ and $\mathbf{y}_2$ are both in the image of T, we have $\mathbf{y}_1 = T(\mathbf{x}_1)$ for some $\mathbf{x}_1$ in $\mathbb{R}^n$, and $\mathbf{y}_2 = T(\mathbf{x}_2)$ for some $\mathbf{x}_2$ in $\mathbb{R}^n$. Since T is linear, we have

$$T(a\mathbf{x}_1 + b\mathbf{x}_2) = aT(\mathbf{x}_1) + bT(\mathbf{x}_2) = a\mathbf{y}_1 + b\mathbf{y}_2$$

This shows that $a\mathbf{y}_1 + b\mathbf{y}_2 = T(a\mathbf{x}_1 + b\mathbf{x}_2)$ is *also* in the image of T.

7. b. It turns out that $T2$ fails for $T : \mathbb{R}^2 \to \mathbb{R}^2$. $T2$ requires that $T(a\mathbf{x}) = aT(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^2$ and all scalars a. But if $a = 2$ and $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ then

$$T\left(2\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = T\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \text{ while } 2T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = 2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Note that $T1$ also fails for this transformation T, as you can verify.

8. b. We are given $T\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x+y \\ -x+y \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \end{bmatrix}$, so T is the matrix transformation induced by the matrix $A = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. By Theorem 2.6.4 we recognize this as the matrix of the rotation $R_{-\frac{\pi}{4}}$. Hence T is rotation through $\theta = -\frac{\pi}{4}$.

d. Here $T\begin{bmatrix} x \\ y \end{bmatrix} = -\frac{1}{10}\begin{bmatrix} 8x+6y \\ 6x-8y \end{bmatrix} = \frac{1}{10}\begin{bmatrix} -8 & -6 \\ -6 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \end{bmatrix}$, so T is the matrix transformation induced by the matrix $A = \frac{1}{10}\begin{bmatrix} -8 & -6 \\ -6 & 8 \end{bmatrix}$. Looking at Theorem 2.6.5, we see that A is the matrix of Q_{-3}. Hence $T = Q_{-3}$ is reflection in the line $y = -3x$.

10. b. Since T is linear, we have $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = T\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} + T\begin{bmatrix} x \\ 0 \\ z \end{bmatrix}$. Since T is rotation about the y axis, we have $T\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$ because $\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$ is *on* the y axis. Now observe that T is rotation of the

xz-plane through the angle θ from the x axis to the z axis. By Theorem 2.6.4 the effect of T on the xz-plane is given by

$$\begin{bmatrix} x \\ z \end{bmatrix} \to \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} x\cos\theta - z\sin\theta \\ x\sin\theta + z\cos\theta \end{bmatrix}$$

Hence $T \begin{bmatrix} x \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} x\cos\theta - z\sin\theta \\ 0 \\ x\sin\theta + z\cos\theta \end{bmatrix}$, and so

$$\begin{aligned} T \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= T \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} + T \begin{bmatrix} x \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} + \begin{bmatrix} x\cos\theta - z\sin\theta \\ 0 \\ x\sin\theta + z\cos\theta \end{bmatrix} \\ &= \begin{bmatrix} x\cos\theta - z\sin\theta \\ y \\ x\sin\theta + z\cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

Hence the matrix of T is $\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$.

12. b. Let Q_0 denote reflection in the x axis, and let R_π denote rotation through π. Then Q_0 has matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and R_π has matrix $B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Then R_π followed by Q_0 is the transformation $Q_0 \circ R_\pi$, and this has matrix $AB = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ by Theorem 2.6.3. This is the matrix of reflection in the y axis.

d. Let Q_0 denote reflection in the x axis, and let $R_{\frac{\pi}{2}}$ denote rotation through $\frac{\pi}{2}$. Then Q_0 has matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and $R_{\frac{\pi}{2}}$ has matrix $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Then Q_0 followed by $R_{\frac{\pi}{2}}$ is the transformation $R_{\frac{\pi}{2}} \circ Q_0$, and this has matrix $BA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ by Theorem 2.6.3. This is the matrix of reflection Q_1 in the line with equation $y = x$.

f. Let Q_0 denote reflection in the x axis, and let Q_1 denote reflection in the line $y = x$. Then Q_0 has matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and Q_1 has matrix $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then Q_0 followed by Q_1 is the transformation $Q_1 \circ Q_0$, and this has matrix $BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ by Theorem 2.6.3. This is the matrix of rotation $R_{\frac{\pi}{2}}$ about the origin through the angle $\frac{\pi}{2}$.

13. b. Since R has matrix A, we have $R(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. By the definition of T we have

$$T(\mathbf{x}) = aR(\mathbf{x}) = a(A\mathbf{x}) = (aA)\mathbf{x}$$

for all $\mathbf{x}$ in $\mathbb{R}^n$. This shows that the matrix of T is aA.

14. b. We use Axiom $T2$: $T(-\mathbf{x}) = T[(-1)\mathbf{x}] = (-1)T(\mathbf{x}) = -T(\mathbf{x})$.

17. b. The matrix of T is B, so $T(\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Let $B^2 = I$. Then

$$T^2(\mathbf{x}) = T[T(\mathbf{x})] = B[B\mathbf{x}] = B^2\mathbf{x} = I\mathbf{x} = \mathbf{x} = 1_{\mathbb{R}^2}(\mathbf{x}) \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n.$$

Hence $T^2 = 1_{\mathbb{R}^n}$ since they have the same effect on every column $\mathbf{x}$.
Conversely, if $T^2 = 1_{\mathbb{R}^n}$ then

$$B^2\mathbf{x} = B(B\mathbf{x}) = T(T(\mathbf{x})) = T^2(\mathbf{x}) = 1_{\mathbb{R}^2}(\mathbf{x}) = \mathbf{x} = I\mathbf{x} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^n.$$

This implies that $B^2 = I$ by Theorem 2.2.6.

18. The matrices of Q_0, Q_1, Q_{-1} and $R_{\frac{\pi}{2}}$ are $\left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right]$, $\left[\begin{array}{rr} 0 & 1 \\ 1 & 0 \end{array}\right]$, $\left[\begin{array}{rr} 0 & -1 \\ -1 & 0 \end{array}\right]$ and $\left[\begin{array}{rr} 0 & -1 \\ 1 & 0 \end{array}\right]$, respectively. We use Theorem 2.6.3 repeatedly: If S has matrix A and T has matrix B then $S \circ T$ has matrix AB.

  b. The matrix of $Q_1 \circ Q_0$ is $\left[\begin{array}{rr} 0 & 1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right] = \left[\begin{array}{rr} 0 & -1 \\ 1 & 0 \end{array}\right]$, which is the matrix of $R_{\frac{\pi}{2}}$.

  d. The matrix of $Q_0 \circ R_{\frac{\pi}{2}}$ is $\left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right]\left[\begin{array}{rr} 0 & -1 \\ 1 & 0 \end{array}\right] = \left[\begin{array}{rr} 0 & -1 \\ -1 & 0 \end{array}\right]$ which is the matrix of Q_{-1}.

19. b. We have $P_m[Q_m(\mathbf{x})] = P_m(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^2$ because $Q_m(\mathbf{x})$ lies on the line $y = mx$. This means $P_m \circ Q_m = P_m$.

20. To see that T is linear, write $\mathbf{x} = \left[\begin{array}{cccc} x_1 & x_2 & \cdots & x_n \end{array}\right]^T$ and $\mathbf{y} = \left[\begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array}\right]^T$. Then:

$$\begin{aligned}
T(\mathbf{x}+\mathbf{y}) &= T\left(\left[\begin{array}{cccc} x_1+y_1 & x_2+y_2 & \cdots & x_n+y_n \end{array}\right]^T\right) \\
&= (x_1+y_1)+(x_2+y_2)+\cdots+(x_n+y_n) \\
&= (x_1+x_2+\cdots+x_n)+(y_1+y_2+\cdots+y_n) \\
&= T(\mathbf{x})+T(\mathbf{y}) \\
T(a\mathbf{x}) &= T(\left[\begin{array}{cccc} ax_1 & ax_2 & \cdots & ax_n \end{array}\right]^T) \\
&= ax_1+ax_2+\cdots+ax_n \\
&= a(x_1+x_2+\cdots+x_n) \\
&= aT(\mathbf{x})
\end{aligned}$$

Hence T is linear, so its matrix is $A = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array}\right] = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array}\right]$ by Theorem 2.6.2.

Note that this can be seen directly because

$$T\left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right] = x_1+\cdots+x_n = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right]$$

so we see immediately that T is the matrix transformation induced by $\left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array}\right]$. Note that this *also* shows that T is linear, and so avoids the tedious verification above.

22. b. Suppose that $T : \mathbb{R}^n \to \mathbb{R}$ is linear. Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ be the standard basis of $\mathbb{R}^n$, and write $T(\mathbf{e}_j) = w_j$ for each $j = 1, 2, \dots, n$. Note that each w_j is in $\mathbb{R}$. As T is linear, Theorem 2.6.2 asserts that T has matrix $A = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array}\right] = \left[\begin{array}{cccc} w_1 & w_2 & \cdots & w_n \end{array}\right]$.

  Hence, given $\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right]$ in $\mathbb{R}^n$, we have

$$T(\mathbf{x}) = A\mathbf{x} = \left[\begin{array}{cccc} w_1 & w_2 & \cdots & w_n \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right] = w_1x_1+w_2x_2+\cdots+w_nx_n = \mathbf{w}\cdot\mathbf{x} = T_{\mathbf{w}}(\mathbf{x})$$

for all $\mathbf{x}$ in $\mathbb{R}^n$ where $\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^T$. This means that $T = T_{\mathbf{w}}$. This can also be seen without Theorem 2.6.2: We have $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$ so, since T is linear,

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= x_1w_1 + x_2w_2 + \cdots + x_nw_n \\ &= \mathbf{w}\cdot\mathbf{x} \\ &= T_{\mathbf{w}}(\mathbf{x}) \end{aligned}$$

for all $\mathbf{x}$ in $\mathbb{R}^n$. Thus $T = T_{\mathbf{w}}$.

24. b. Given linear transformations $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k$ we are to show that $(S \circ T)(a\mathbf{x}) = a(S \circ T)(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^n$ and all scalars a. The proof is a straight forward computation:

$$\begin{aligned} (S\circ T)(a\mathbf{x}) &= S[T(a\mathbf{x})] && \text{Definition of } S\circ T \\ &= S[aT(\mathbf{x})] && T \text{ is linear} \\ &= a[S[T(\mathbf{x})]] && S \text{ is linear} \\ &= a[(S\circ T)(\mathbf{x})] && \text{Definition of } S\circ T \end{aligned}$$

2.7 LU-factorization

1. b. $\begin{bmatrix} 2 & 4 & 2 \\ 1 & -1 & 3 \\ -1 & 7 & -7 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 3 \\ -1 & 7 & -7 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & -3 & 2 \\ 0 & 9 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 0 \end{bmatrix} = U.$

   Hence $A = LU$ where U is above and $L = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -3 & 0 \\ -1 & 9 & 1 \end{bmatrix}$.

   d. $\begin{bmatrix} -1 & -3 & 1 & 0 & -1 \\ 1 & 4 & 1 & 1 & 1 \\ 1 & 2 & -3 & -1 & 1 \\ 0 & -2 & -4 & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & -1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & -2 & -4 & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & -1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = U.$

   Hence $A = LU$ where U is as above and $L = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}$.

   f. $\begin{bmatrix} 2 & 2 & -2 & 4 & 2 \\ 1 & -1 & 0 & 2 & 1 \\ 3 & 1 & -2 & 6 & 3 \\ 1 & 3 & -2 & 2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 & 2 & 1 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 & 2 & 1 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = U.$

   Hence $A = LU$ where U is above and $L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$.

2. b. The reduction to row-echelon form requires two row interchanges:

$$\begin{bmatrix} 0 & -1 & 2 \\ 0 & 0 & 4 \\ -1 & 2 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & -1 & 2 \\ 0 & 0 & 4 \\ -1 & 2 & 1 \end{bmatrix} \to \begin{bmatrix} -1 & 2 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 4 \end{bmatrix} \to \cdots$$

The elementary matrices corresponding (in order) to the interchanges are
$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ and $P_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, so take $P = P_2P_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.
We apply the LU-algorithm to PA:

$$PA = \begin{bmatrix} -1 & 2 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = U$$

Hence $PA = LU$ where U is as above and $L = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

d. The reduction to row-echelon form requires two row interchanges:
$$\begin{bmatrix} -1 & -2 & 3 & 0 \\ 2 & 4 & -6 & 5 \\ 1 & 1 & -1 & 3 \\ 2 & 5 & -10 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & -1 & 2 & 3 \\ 0 & 1 & -4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & -2 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & -2 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$
The elementary matrices corresponding (in order) to the interchanges are
$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ and $P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ so $P = P_2P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.
We apply the LU-algorithm to PA:
$$PA = \begin{bmatrix} -1 & -2 & 3 & 0 \\ 1 & 1 & -1 & 3 \\ 2 & 5 & -10 & 1 \\ 2 & 4 & -6 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & 3 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$
$$\rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix} = U$$
Hence $PA = LU$ where U is as above and $L = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 2 & 1 & -2 & 0 \\ 2 & 0 & 0 & 5 \end{bmatrix}$.

3. b. Write $L = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 2 & 1 \end{bmatrix}$, $U = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. The system $L\mathbf{y} = \mathbf{b}$ is
$$\begin{array}{rcrcrcr} 2y_1 & & & & & = & -2 \\ y_1 & + & 3y_2 & & & = & -1 \\ -y_1 & + & 2y_2 & + & y_3 & = & 1 \end{array}$$
and we solve this by forward substitution: $y_1 = -1$, $y_2 = \frac{1}{3}(-1-y_1) = 0$, $y_3 = 1 + y_1 - 2y_2 = 0$. The system $U\mathbf{x} = \mathbf{y}$ is
$$\begin{array}{rcrcrcr} x_1 & + & x_2 & - & x_4 & = & -1 \\ & & x_2 & + & x_4 & = & 0 \\ & & & & 0 & = & 0 \end{array}$$
and we solve this by back substitution: $x_4 = t$, $x_3 = 5$, $x_2 = -x_4 = -t$, $x_1 = -1 + x_4 - x_2 = -1 + 2t$.

d. Analogous to (b). The solution is: $\mathbf{y} = \begin{bmatrix} 2 \\ 8 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 8-2t \\ 6-t \\ -1-t \\ t \end{bmatrix}$, t arbitrary.

5. If the rows in question are R_1 and R_2, they can be interchanged thus:
$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \rightarrow \begin{bmatrix} R_1 + R_2 \\ R_2 \end{bmatrix} \rightarrow \begin{bmatrix} R_1 + R_2 \\ -R_1 \end{bmatrix} \rightarrow \begin{bmatrix} R_1 \\ -R_1 \end{bmatrix} \rightarrow \begin{bmatrix} R_2 \\ R_1 \end{bmatrix}$$

6. b. Let $A = LU = L_1U_1$ be LU-factorizations of the invertible matrix A. Then U and U_1 have no row of zeros so (being row-echelon) are upper triangular with 1's on the main diagonal. Thus $L_1^{-1}L = U_1U^{-1}$ is both lower triangular ($L_1^{-1}L$) and upper triangular (U_1U^{-1}) and so is diagonal. But it has 1's on the diagonal (U_1 and U do) so it is I. Hence $L_1 = L$ and $U_1 = U$.

7. We proceed by induction on n where A and B are $n \times n$. It is clear if $n = 1$. In general, write $A = \begin{bmatrix} a & 0 \\ X & A_1 \end{bmatrix}$ and $B = \begin{bmatrix} b & 0 \\ Y & B_1 \end{bmatrix}$ where A_1 and B_1 are lower triangular. Then $AB = \begin{bmatrix} ab & 0 \\ Xb + A_1Y & A_1B_1 \end{bmatrix}$ by Theorem 2.2.5, and A_1B_1 is upper triangular by induction. Hence AB is upper triangular.

9. b. Let $A = LU = L_1U_1$ be two such factorizations. Then $UU_1^{-1} = L^{-1}L_1$; write this matrix as $D = UU_1^{-1} = L^{-1}L_1$. Then D is lower triangular (apply Lemma 2.7.1 to $D = L^{-1}L_1$), and D is also upper triangular (consider UU_1^{-1}). Hence D is diagonal, and so $D = I$ because L^{-1} and L_1 are unit triangular. Since $A = LU$, this completes the proof.

2.8 An Application to Input-Output Economic Models

1. b. $I - E = \begin{bmatrix} .5 & 0 & -.5 \\ -.1 & .1 & -.2 \\ -.4 & -.1 & .7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 0 & -1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$. The equilibrium price structure $\mathbf{p}$ is the solution to $(I - E)\mathbf{p} = \mathbf{0}$; the general solution is $\mathbf{p} = \begin{bmatrix} t & 3t & t \end{bmatrix}^T$.

   d. $-E = \begin{bmatrix} .5 & 0 & -.1 & -.1 \\ -.2 & .3 & 0 & -.1 \\ -.1 & -.2 & .2 & -.2 \\ -.2 & -.1 & -.1 & .4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -2 & 2 \\ 5 & 0 & -1 & -1 \\ -2 & 3 & 0 & -1 \\ -2 & -1 & -1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -2 & 2 \\ 0 & -10 & 9 & -11 \\ 0 & 7 & -4 & 3 \\ 0 & 3 & -5 & 8 \end{bmatrix}$.

   Now add 3 times row 4 to row 2 to get:

$$\begin{bmatrix} 1 & 2 & -2 & 2 \\ 0 & -1 & -6 & 13 \\ 0 & 7 & -4 & 3 \\ 0 & 3 & -5 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -14 & 28 \\ 0 & 1 & 6 & -13 \\ 0 & 0 & -46 & 94 \\ 0 & 0 & -23 & 47 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -14 & 28 \\ 0 & 1 & 6 & -13 \\ 0 & 0 & 1 & -\frac{47}{23} \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & -\frac{14}{23} \\ 0 & 1 & 0 & -\frac{17}{23} \\ 0 & 0 & 1 & -\frac{47}{23} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

   The equilibrium price structure $\mathbf{p}$ is the solution to $(I - E)\mathbf{p} = \mathbf{0}$. The solution is $\mathbf{p} = \begin{bmatrix} 14t & 17t & 47t & 23t \end{bmatrix}^T$.

2. Here the input-output matrix is $E = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ so we get

$$I - E = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

   Thus the solution to $(I - E)\mathbf{p}$ is $p_1 = p_2 = p_3 = t$. Thus all three industries produce the same output.

4. $I - E = \begin{bmatrix} 1-a & -b \\ -1+a & b \end{bmatrix} \rightarrow \begin{bmatrix} 1-a & -b \\ 0 & 0 \end{bmatrix}$ so the possible equilibrium price structures are $\mathbf{p} = \begin{bmatrix} bt \\ (1-a)t \end{bmatrix}$, t arbitrary. This is nonzero for some t unless $b = 0$ and $a = 1$, and in that case $\mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a solution. If the entries of A are positive then $\mathbf{p} = \begin{bmatrix} b \\ 1-a \end{bmatrix}$ has positive entries.

7. b. One such example is $E = \begin{bmatrix} .4 & .8 \\ .7 & .2 \end{bmatrix}$, because $(I - E)^{-1} = -\frac{5}{4}\begin{bmatrix} 8 & 8 \\ 7 & 6 \end{bmatrix}$.

8. If $E = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $I - E = \begin{bmatrix} 1-a & -b \\ -c & 1-d \end{bmatrix}$. We have $\det(I - E) = (1 - a)(1 - d) - bc = 1 - (a + d) + (ad - bc) = 1 - \text{tr } E + \det E$. If $\det(I - E) \neq 0$ then Example 2.3.5 gives $(I - E)^{-1} = \frac{1}{\det(I-E)}\begin{bmatrix} 1-d & b \\ c & 1-a \end{bmatrix}$. The entries $1 - d$, b, c, and $1 - a$ are all between 0 and 1 so $(I - E)^{-1} \geq 0$ if $\det(I - E) > 0$, that is if $\text{tr } E < 1 + \det E$.

9. b. If $\mathbf{p} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ then $\mathbf{p} > E\mathbf{p}$ so Theorem 2.8.2 applies.

d. If $\mathbf{p} = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}$ then $\mathbf{p} > E\mathbf{p}$ so Theorem 2.8.2 applies.

2.9 An Application to Markov Chains

1. b. Not regular. Every power of P has the $(1,\ 2)$- and $(3,\ 2)$-entries zero.

2. b. $I-P = \begin{bmatrix} \frac{1}{2} & -1 \\ -\frac{1}{2} & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$ so $(I-P)\mathbf{s} = \mathbf{0}$ has solutions $\mathbf{s} = \begin{bmatrix} 2t \\ t \end{bmatrix}$. The entries of $\mathbf{s}$ sum to 1 if $t = \frac{1}{3}$, so $\mathbf{s} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$ is the steady state vector. Given $\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, we get $\mathbf{s}_1 = P\mathbf{s}_0 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$, $\mathbf{s}_2 = P\mathbf{s}_1 \begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \end{bmatrix}$, $\mathbf{s}_3 = P\mathbf{s}_2 = \begin{bmatrix} \frac{5}{8} \\ \frac{3}{8} \end{bmatrix}$. So it is in state 2 after three transitions with probability $\frac{3}{8}$.

d. $I-P = \begin{bmatrix} .6 & -.1 & -.5 \\ -.2 & .4 & -.2 \\ -.4 & -.3 & .7 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1 \\ 0 & 11 & -11 \\ 0 & -11 & 11 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ so $(I-P)\mathbf{s} = \mathbf{0}$ has solution $\mathbf{s} = \begin{bmatrix} t \\ t \\ t \end{bmatrix}$. The entries sum to 1 if $t = \frac{1}{3}$ so the steady state vector is $\mathbf{s} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$. Given $\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{s}_1 = P\mathbf{s}_0 = \begin{bmatrix} .4 \\ .2 \\ .4 \end{bmatrix}$, $\mathbf{s}_2 = P\mathbf{s}_1 = \begin{bmatrix} .38 \\ .28 \\ .34 \end{bmatrix}$, $\mathbf{s}_3 = P\mathbf{s}_2 = \begin{bmatrix} .350 \\ .312 \\ .338 \end{bmatrix}$. Hence it is in state 2 after three transitions with probability .312.

f. $I-P = \begin{bmatrix} .9 & -.3 & -.3 \\ -.3 & .9 & -.6 \\ -.6 & -.6 & .9 \end{bmatrix} \to \begin{bmatrix} 1 & -3 & 2 \\ 0 & 24 & -21 \\ 0 & -24 & 21 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -\frac{5}{8} \\ 0 & 1 & -\frac{7}{8} \\ 0 & 0 & 0 \end{bmatrix}$, so $(I-P)\mathbf{s} = \mathbf{0}$ has solution $\mathbf{s} = \begin{bmatrix} 5t \\ 7t \\ 8t \end{bmatrix}$. The entries sum to 1 if $t = \frac{1}{20}$ so the steady state vector is $\mathbf{s} = \begin{bmatrix} \frac{5}{20} \\ \frac{7}{20} \\ \frac{8}{20} \end{bmatrix}$. Given $\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{s}_1 = P\mathbf{s}_0 = \begin{bmatrix} .1 \\ .3 \\ .6 \end{bmatrix}$, $\mathbf{s}_2 = P\mathbf{s}_1 = \begin{bmatrix} .28 \\ .42 \\ .30 \end{bmatrix}$, $\mathbf{s}_3 = P\mathbf{s}_1 = \begin{bmatrix} .244 \\ .306 \\ .450 \end{bmatrix}$. Hence it is in state 2 after three transitions with probability .306.

4. b. The transition matrix is $P = \begin{bmatrix} .7 & .1 & .1 \\ .1 & .8 & .3 \\ .2 & .1 & .6 \end{bmatrix}$ where the columns (and rows) represent the upper, middle and lower classes respectively and, for example, the last column asserts that, for children of lower class people, 10% become upper class, 30% become middle class and 60% remain lower class. Hence $I-P = \begin{bmatrix} .3 & -.1 & -.1 \\ -.1 & .2 & -.3 \\ -.2 & -.1 & -.4 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 3 \\ 0 & 5 & -10 \\ 0 & -5 & 10 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$. Thus the general solution to $(I-P)\mathbf{s} = \mathbf{0}$ is $\mathbf{s} = \begin{bmatrix} t \\ 2t \\ t \end{bmatrix}$, so $\mathbf{s} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$ is the steady state solution. Eventually, upper, middle and lower classes will comprise 25%, 50% and 25% of this society respectively.

6. Let States 1 and 2 be "late" and "on time" respectively. Then the transition matrix is $P = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix}$. Here column 1 describes what happens if he was late one day: the two entries sum to 1 and the top entry is twice the bottom entry by the information we are given. Column 2 is determined similarly. Now if Monday is the initial state, we are given that $\mathbf{s}_0 = \begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \end{bmatrix}$. Hence $\mathbf{s}_1 = P\mathbf{s}_0 = \begin{bmatrix} \frac{3}{8} \\ \frac{5}{8} \end{bmatrix}$ and $\mathbf{s}_2 = P\mathbf{s}_1 = \begin{bmatrix} \frac{7}{16} \\ \frac{9}{16} \end{bmatrix}$. Hence the probabilities that he is late and on time Wednesdays are $\frac{7}{16}$ and $\frac{9}{16}$ respectively.

8. Let the states be the five compartments. Since each tunnel entry is equally likely,

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{5} & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & 0 & \frac{2}{5} & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{5} & \frac{2}{4} & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & 0 \end{bmatrix}$$

a. Since he starts in compartment 1,

$$\mathbf{s}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{s}_1 = P\mathbf{s}_0 = \begin{bmatrix} 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ \frac{1}{3} \end{bmatrix}, \mathbf{s}_2 = P\mathbf{s}_1 = \begin{bmatrix} \frac{2}{5} \\ 0 \\ \frac{3}{10} \\ \frac{7}{30} \\ \frac{1}{15} \end{bmatrix}, \mathbf{s}_3 = P\mathbf{s}_2 = \begin{bmatrix} \frac{7}{75} \\ \frac{23}{120} \\ \frac{69}{200} \\ \frac{53}{300} \\ \frac{29}{150} \end{bmatrix}$$

Hence the probability that he is in compartment 1 after three moves is $\frac{7}{75}$.

b. The steady state vector $\mathbf{s}$ satisfies $(I-P)\mathbf{s} = \mathbf{0}$. As

$$(I-P) = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{5} & 0 & -\frac{1}{2} \\ -\frac{1}{3} & 1 & 0 & -\frac{1}{4} & 0 \\ -\frac{1}{3} & 0 & \frac{3}{5} & -\frac{1}{4} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{5} & \frac{1}{2} & 0 \\ -\frac{1}{3} & 0 & -\frac{1}{5} & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & -\frac{3}{2} \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -\frac{5}{2} \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

so the steady state is $\mathbf{s} = \frac{1}{16}\begin{bmatrix} 3 \\ 2 \\ 5 \\ 4 \\ 2 \end{bmatrix}$. Hence, in the long run, he spends most of his time in compartment 3 (in fact $\frac{5}{16}$ of his time).

12. a. $\begin{bmatrix} 1-p & q \\ p & 1-q \end{bmatrix} \cdot \frac{1}{p+q}\begin{bmatrix} q \\ p \end{bmatrix} = \frac{1}{p+q}\begin{bmatrix} (1-p)q+qp \\ pq+(1-q)p \end{bmatrix} = \frac{1}{p+q}\begin{bmatrix} q \\ p \end{bmatrix}$. Since the entries of $\frac{1}{p+q}\begin{bmatrix} q \\ p \end{bmatrix}$ add to 1, it is the steady state vector.

b. If $m = 1$

$$\begin{aligned} \frac{1}{p+q}\begin{bmatrix} q & q \\ p & p \end{bmatrix} + \frac{1-p-q}{p+q}\begin{bmatrix} p & -q \\ -p & q \end{bmatrix} &= \frac{1}{p+q}\begin{bmatrix} q+p-p^2-pq & q-q+pq+q^2 \\ p-p\_p^2+pq & p+q-pq-q^2 \end{bmatrix} \\ &= \frac{1}{p+q}\begin{bmatrix} (p+q)(1-p) & (p+q)q \\ (p+q)p & (p+q)(1-q) \end{bmatrix} \\ &= P \end{aligned}$$

In general, write $X = \begin{bmatrix} q & q \\ p & p \end{bmatrix}$ and $Y = \begin{bmatrix} p & -q \\ -p & q \end{bmatrix}$. Then $PX = X$ and $PY = (1-p-q)Y$. Hence if $P^m = \frac{1}{p+q}X + \frac{(1-p-q)^m}{p+q}Y$ for some $m \geq 1$, then

$$\begin{aligned} P^{m+1} = PP^m &= \tfrac{1}{p+q}PX + \tfrac{(1-p-q)^m}{p+q}PY \\ &= \tfrac{1}{p+q}X + \tfrac{(1-p-q)^m}{p+q}(1-p-q)Y \\ &= \tfrac{1}{p+q}X + \tfrac{(1-p-q)^{m+1}}{p+q}Y \end{aligned}$$

Hence the formula holds for all $m \geq 1$ by induction.

Now $0 < p < 1$ and $0 < q < 1$ imply $0 < p+q < 2$, so that $-1 < (p+q-1) < 1$. Multiplying through by -1 gives $1 > (1-p-q) > -1$, so $(1-p-q)^m$ converges to zero as m increases.

Supplementary Exercises: Chapter 2

2. b. We have $0 = p(U) = U^3 - 5U^2 + 11U - 4I$ so that $U(U^2 - 5U + 11I) = 4I = (U^2 - 5U + 11I)U$. Hence $U^{-1} = \frac{1}{4}(U^2 - 5U + 11I)$.

4. b. If $\mathbf{x}_h = \mathbf{x}_m$, then $\mathbf{y} = k(\mathbf{y} - \mathbf{z}) = \mathbf{y} + m(\mathbf{y} - \mathbf{z})$, whence $(k - m)(\mathbf{y} - \mathbf{z}) = 0$. But the matrix $\mathbf{y} - \mathbf{z} \neq 0$ (because $\mathbf{y} \neq \mathbf{z}$) so $k - m = 0$ by Example 2.1.7.

6. d. Using (c), $I_{pq}AI_{rs} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}I_{pq}I_{ij}I_{rs}$. Now (b) shows that $I_{pq}I_{ij}I_{rs} = 0$ unless $i = q$ and $j = r$, when it equals I_{ps}. Hence the double sum for $I_{pq}AI_{rs}$ has only one nonzero term — the one for which $i = q$, $j = r$. Hence $I_{pq}AI_{rs} = a_{qr}I_{ps}$.

7. b. If $n = 1$ it is clear. If $n > 1$, Exercise 6(d) gives

$$a_{qr}I_{ps} = I_{pq}AI_{rs} = I_{pq}I_{rs}A$$

because $AI_{rs} = I_{rs}A$. Hence $a_{qr} = 0$ if $q \neq r$ by Exercise 6(b). If $r = q$ then $a_{qq}I_{ps} = I_{ps}A$ is the same for each value of q. Hence $a_{11} = a_{22} = \cdots = a_{nn}$, so A is a scalar matrix.

3. Determinants and Diagonalization

3.1 The Cofactor Expansion

If A is a square matrix, we write $\det A = |A|$ for convenience.

1. b. Take 3 out of row 1, then subtract 4 times row 1 from row 2:
$$\left|\begin{array}{cc} 6 & 9 \\ 8 & 12 \end{array}\right| = 3\left|\begin{array}{cc} 2 & 3 \\ 8 & 12 \end{array}\right| = 3\left|\begin{array}{cc} 2 & 3 \\ 0 & 0 \end{array}\right| = 0$$

   d. Subtract row 2 from row 1: $\left|\begin{array}{cc} a+1 & a \\ a & a-1 \end{array}\right| = \left|\begin{array}{cc} 1 & 1 \\ a & a-1 \end{array}\right| = (a-1)-a = -1$

   f. Subtract 2 times row 2 from row 1, then expand along row 2:
$$\left|\begin{array}{rrr} 2 & 0 & -3 \\ 1 & 2 & 5 \\ 0 & 3 & 0 \end{array}\right| = \left|\begin{array}{rrr} 0 & -4 & -13 \\ 1 & 2 & 5 \\ 0 & 3 & 0 \end{array}\right| = -\left|\begin{array}{rr} -4 & -13 \\ 3 & 0 \end{array}\right| = -39$$

   h. Expand along row 1: $\left|\begin{array}{ccc} 0 & a & 0 \\ b & c & d \\ 0 & e & 0 \end{array}\right| = -a\left|\begin{array}{cc} b & d \\ 0 & 0 \end{array}\right| = -a(0) = 0$

   j. Expand along row 1:
$$\left|\begin{array}{ccc} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{array}\right| = -a\left|\begin{array}{cc} a & c \\ b & 0 \end{array}\right| + b\left|\begin{array}{cc} a & 0 \\ b & c \end{array}\right| = -a(-bc)+b(ac) = 2abc$$

   l. Subtract multiples of row 1 from rows 2, 3 and 4, then expand along column 1:
$$\left|\begin{array}{rrrr} 1 & 0 & 3 & 1 \\ 2 & 2 & 6 & 0 \\ -1 & 0 & -3 & 1 \\ 4 & 1 & 12 & 0 \end{array}\right| = \left|\begin{array}{rrrr} 1 & 0 & 3 & 1 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & -4 \end{array}\right| = \left|\begin{array}{rrr} 2 & 0 & -2 \\ 0 & 0 & 2 \\ 1 & 0 & -4 \end{array}\right| = 0$$

   n. Subtract multiples of row 4 from rows 1 and 2, then expand along column 1:
$$\left|\begin{array}{rrrr} 4 & -1 & 3 & -1 \\ 3 & 1 & 0 & 2 \\ 0 & 1 & 2 & 2 \\ 1 & 2 & -1 & 1 \end{array}\right| = \left|\begin{array}{rrrr} 0 & -9 & 7 & -5 \\ 0 & -5 & 3 & -1 \\ 0 & 1 & 2 & 2 \\ 1 & 2 & -1 & 1 \end{array}\right| = -\left|\begin{array}{rrr} -9 & 7 & -5 \\ -5 & 3 & -1 \\ 1 & 2 & 2 \end{array}\right|$$
   Again, subtract multiples of row 3 from rows 1 and 2, then expand along column 1:
$$\left|\begin{array}{rrrr} 4 & -1 & 3 & -1 \\ 3 & 1 & 0 & 2 \\ 0 & 1 & 2 & 2 \\ 1 & 2 & -1 & 1 \end{array}\right| = -\left|\begin{array}{rrr} 0 & 25 & 13 \\ 0 & 13 & 9 \\ 1 & 2 & 2 \end{array}\right| = \left|\begin{array}{rr} 25 & 13 \\ 13 & 9 \end{array}\right| = -\left|\begin{array}{rr} -1 & -5 \\ 13 & 9 \end{array}\right| = -(-9+65) = -56$$

   p. Keep expanding along row 1:
$$\left|\begin{array}{cccc} 0 & 0 & 0 & a \\ 0 & 0 & b & p \\ 0 & c & q & k \\ d & s & t & u \end{array}\right| = -a\left|\begin{array}{ccc} 0 & 0 & b \\ 0 & v & q \\ d & s & t \end{array}\right| = -a\left(b\left|\begin{array}{cc} 0 & c \\ d & s \end{array}\right|\right) = -ab(-cd) = abcd$$

5. b. $\left|\begin{array}{rrr} -1 & 3 & 1 \\ 2 & 5 & 3 \\ 1 & -2 & 1 \end{array}\right| = \left|\begin{array}{rrr} -1 & 3 & 1 \\ 0 & 11 & 5 \\ 0 & 1 & 2 \end{array}\right| = -\left|\begin{array}{rrr} -1 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 11 & 5 \end{array}\right| = -\left|\begin{array}{rrr} -1 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -17 \end{array}\right| = -(17) = -17$

d. $\begin{vmatrix} 2 & 3 & 1 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 5 & 1 & 1 \\ 1 & 1 & 2 & 5 \end{vmatrix} = -\begin{vmatrix} 1 & 1 & 2 & 5 \\ 0 & 2 & -1 & 3 \\ 0 & 5 & 1 & 1 \\ 2 & 3 & 1 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 1 & 2 & 5 \\ 0 & 2 & -1 & 3 \\ 0 & 5 & 1 & 1 \\ 0 & 1 & -3 & -9 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 & 5 \\ 0 & 1 & -3 & -9 \\ 0 & 5 & 1 & 1 \\ 0 & 2 & -1 & 3 \end{vmatrix} =$

$\begin{vmatrix} 1 & 1 & 2 & 5 \\ 0 & 1 & -3 & -9 \\ 0 & 0 & 16 & 46 \\ 0 & 0 & 5 & 21 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 & 5 \\ 0 & 1 & -3 & -9 \\ 0 & 0 & 1 & -17 \\ 0 & 0 & 0 & 106 \end{vmatrix} = 106$

6. b. Subtract row 1 from row 2: $\begin{vmatrix} a & b & c \\ a+b & 2b & c+b \\ 2 & 2 & 2 \end{vmatrix} = \begin{vmatrix} a & b & c \\ b & b & b \\ 2 & 2 & 2 \end{vmatrix} = 0$ by Theorem 3.1.2(4).

7. b. Take -2 and 3 out of rows 1 and 2, then subtract row 3 from row 2, then take 2 out of row 2:

$$\begin{vmatrix} -2a & -2b & -2c \\ 2p+x & 2q+y & 2r+z \\ 3x & 3y & 3z \end{vmatrix} = -6\begin{vmatrix} a & b & c \\ 2p+x & 2q+y & 2r+z \\ x & y & z \end{vmatrix} = -6\begin{vmatrix} a & b & c \\ 2p & 2q & 2r \\ x & y & z \end{vmatrix} = -12\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix} = 12$$

8. b. First add rows 2 and 3 to row 1:

$$\begin{vmatrix} 2a+p & 2b+q & 2c+r \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{vmatrix} = \begin{vmatrix} 3a+3p+3x & 3b+3q+3y & 3c+3r+3z \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{vmatrix} = 3\begin{vmatrix} a+p+x & b+q+y & c+r+z \\ 2p+x & 2q+y & 2r+z \\ 2x+a & 2y+b & 2z+c \end{vmatrix}$$

Now subtract row 1 from rows 2 and 3, and then add row 2 plus twice row 3 to row 1, to get

$$= 3\begin{vmatrix} a+p+x & b+q+y & c+r+z \\ p-a & q-b & r-c \\ x-p & y-q & z-r \end{vmatrix} = 3\begin{vmatrix} 3x & 3y & 3z \\ p-a & q-b & r-c \\ x-p & y-q & z-r \end{vmatrix}$$

Next take 3 out of row 1, and then add row 3 to row 2, to get

$$= 9\begin{vmatrix} x & y & z \\ p-a & q-b & r-c \\ -p & -q & -r \end{vmatrix} = 9\begin{vmatrix} x & y & z \\ -a & -b & -c \\ -p & -q & -r \end{vmatrix}$$

Now use row interchanges and common row factors to get

$$= -9\begin{vmatrix} -p & -q & -r \\ -a & -b & -c \\ x & y & z \end{vmatrix} = 9\begin{vmatrix} -a & -b & -c \\ -p & -q & -r \\ x & y & z \end{vmatrix} = 9\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$$

9. b. False. The matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ has zero determinant, but no two rows are equal.

d. False. The reduced row-echelon form of $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ is $R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, but $\det A = 2$ while $\det R = 1$.

f. False. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\det A = 1 = \det A^T$.

h. False. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ then $\det A = \det B = 1$. In fact, it is a theorem that $\det A = \det A^T$ holds for *every* square matrix A.

10. b. Partition the matrix as follows and use Theorem 3.1.5:

$$\left|\begin{array}{cc|ccc} 1 & 2 & 0 & 3 & 0 \\ -1 & 3 & 1 & 4 & 0 \\ \hline 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & -1 & 0 & 2 \\ 0 & 0 & 3 & 0 & 1 \end{array}\right| = \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix}\begin{vmatrix} 2 & 1 & 1 \\ -1 & 0 & 2 \\ 3 & 0 & 1 \end{vmatrix} = 5\left(-1\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix}\right) = -5(-7) = 35$$

11. b. Use Theorem 3.1.5 twice:

$$\left|\begin{array}{cc|c} A & 0 & 0 \\ X & B & 0 \\ \hline Y & Z & C \end{array}\right| = \det \left[\begin{array}{cc} A & 0 \\ X & B \end{array}\right] \det C = (\det A \det B) \det C = 2(-1)3 = -6$$

d. $$\left|\begin{array}{cc|c} A & X & 0 \\ 0 & B & 0 \\ \hline Y & Z & C \end{array}\right| = \det \left[\begin{array}{cc} A & X \\ 0 & B \end{array}\right] \det C = (\det A \det B) \det C = 2(-1)3 = -6$$

14. b. Follow the Hint, take out the common factor in row 1, subtract multiples of column 1 from columns 2 and 3, and expand along row 1:

$$\begin{aligned} \det \left|\begin{array}{ccc} x-1 & -3 & 1 \\ 2 & -1 & x-1 \\ -3 & x+2 & -2 \end{array}\right| &= \left|\begin{array}{ccc} x-2 & x-2 & x-2 \\ 2 & -1 & x-1 \\ -3 & x+2 & -2 \end{array}\right| = (x-2)\left|\begin{array}{ccc} 1 & 1 & 1 \\ 2 & -1 & x-1 \\ -3 & x+2 & -2 \end{array}\right| \\ &= (x-2)\left|\begin{array}{ccc} 1 & 0 & 0 \\ 2 & -3 & x-3 \\ -3 & x+5 & 1 \end{array}\right| = (x-2)\left|\begin{array}{cc} -3 & x-3 \\ x+5 & 1 \end{array}\right| \\ &= (x-2)(-x^2-2x+12) = -(x-2)(x^2+2x-12) \end{aligned}$$

15. b. If we expand along column 2, the coefficient of z is $-\left|\begin{array}{cc} 2 & -1 \\ 1 & 3 \end{array}\right| = -(6+1) = -7$. So $c = -7$.

16. b. Compute $\det A$ by adding multiples of row 1 to rows 2 and 3, and then expanding along column 1:

$$\begin{aligned} \det A = \left|\begin{array}{ccc} 1 & x & x \\ -x & -2 & x \\ -x & -x & -3 \end{array}\right| &= \left|\begin{array}{ccc} 1 & x & x \\ 0 & x^2-2 & x^2+x \\ 0 & x^2-x & x^2-3 \end{array}\right| = \left|\begin{array}{cc} x^2-2 & x^2+x \\ x^2-x & x^2-3 \end{array}\right| \\ &= (x^2-2)(x^2-3)-(x^2+x)(x^2-x) = (x^4-5x^2+6)-x^2(x^2-1) = 6-4x^2 \end{aligned}$$

Hence $\det A = 0$ means $x^2 = \frac{3}{2} = \frac{6}{4}$, so $x = \pm\frac{\sqrt{6}}{2}$.

d. Expand along column 1, and use Theorem 3.1.4:

$$\begin{aligned} \det A = \left|\begin{array}{cccc} x & y & 0 & 0 \\ 0 & x & y & 0 \\ 0 & 0 & x & y \\ y & 0 & 0 & x \end{array}\right| &= x\left|\begin{array}{ccc} x & y & 0 \\ 0 & x & y \\ 0 & 0 & x \end{array}\right| - y\left|\begin{array}{ccc} y & 0 & 0 \\ x & y & 0 \\ 0 & x & y \end{array}\right| = x \cdot x^3 - y \cdot y^3 \\ &= x^4 - y^4 = (x^2-y^2)(x^2+y^2) = (x-y)(x+y)(x^2+y^2) \end{aligned}$$

Hence $\det A = 0$ means $x = y$ or $x = -y$ ($x^2+y^2 = 0$ only if $x = y = 0$).

21. Let $\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right]$, $\mathbf{y} = \left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array}\right]$, and $A = \left[\begin{array}{ccccc} \mathbf{c}_1 & \cdots & \mathbf{x}+\mathbf{y} & \cdots & \mathbf{c}_n \end{array}\right]$ where $\mathbf{x}+\mathbf{y}$ is in column j. Expanding $\det A$ along column j we obtain

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) = \det A &= \sum_{i=1}^{n} (x_i+y_i)c_{ij}(A) \\ &= \sum_{i=1}^{n} x_i c_{ij}(A) + \sum_{i=1}^{n} y_i c_{ij}(A) \\ &= T(\mathbf{x}) + T(\mathbf{y}) \end{aligned}$$

where the determinant at the second step is expanded along column 1. Similarly, $T(a\mathbf{x}) = aT(\mathbf{x})$ for any scalar a.

24. Suppose A is $n \times n$. B can be found from A by interchanging the following pairs of columns: 1 and n, 2 and $n-1$, $\dots$. There are two cases according as n is even or odd:

Case 1. $n = 2k$. Then we interchange columns 1 and n, 2 and $n-1$, $\dots$, k and $k+1$, k interchanges in all. Thus $\det B = (-1)^k \det A$ in this case.

Case 2. $n = 2k+1$. Now we interchange columns 1 and n, 2 and $n-1$, $\dots$, k and $k+2$, leaving column k fixed. Again k interchanges are used so $\det B = (-1)^k \det A$.

Thus in both cases: $\det B = (-1)^k \det A$ where A is $n \times n$ and $n = 2k$ or $n = 2k+1$.

Remark: Observe that, in each case, k and $\frac{1}{2}n(n-1)$ are both even or both odd, so $(-1)^k = (-1)^{\frac{1}{2}n(n-1)}$. Hence, if A is $n \times n$, we have $\det B = (-1)^{\frac{1}{2}n(n-1)} \det A$.

3.2 Determinants and Matrix Inverses

1. b. The cofactor matrix is
$$\begin{bmatrix} \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} & -\begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 0 & -1 \end{vmatrix} \\ -\begin{vmatrix} -1 & 2 \\ -1 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 0 & -1 \end{vmatrix} \\ \begin{vmatrix} -1 & 2 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 1 & -3 & -3 \\ -1 & 1 & 1 \\ -2 & 6 & 4 \end{bmatrix}.$$

The adjugate is the transpose of the cofactor matrix: $\operatorname{adj} A = \begin{bmatrix} 1 & -1 & -2 \\ -3 & 1 & 6 \\ -3 & 1 & 4 \end{bmatrix}$.

d. In computing the cofactor matrix, we use the fact that $\det \left[\frac{1}{3}M\right] = \frac{1}{9} \det M$ for any 2×2 matrix M. Thus the cofactor matrix is
$$\begin{bmatrix} \frac{1}{9}\begin{vmatrix} -1 & 2 \\ 2 & -1 \end{vmatrix} & -\frac{1}{9}\begin{vmatrix} 2 & 2 \\ 2 & -1 \end{vmatrix} & \frac{1}{9}\begin{vmatrix} 2 & -1 \\ 2 & 2 \end{vmatrix} \\ -\frac{1}{9}\begin{vmatrix} 2 & 2 \\ 2 & -1 \end{vmatrix} & \frac{1}{9}\begin{vmatrix} -1 & 2 \\ 2 & -1 \end{vmatrix} & -\frac{1}{9}\begin{vmatrix} -1 & 2 \\ 2 & 2 \end{vmatrix} \\ \frac{1}{9}\begin{vmatrix} 2 & 2 \\ -1 & 2 \end{vmatrix} & -\frac{1}{9}\begin{vmatrix} -1 & 2 \\ 2 & 2 \end{vmatrix} & \frac{1}{9}\begin{vmatrix} -1 & 2 \\ 2 & -1 \end{vmatrix} \end{bmatrix} = \frac{1}{9}\begin{bmatrix} -3 & 6 & 6 \\ 6 & -3 & 6 \\ 6 & 6 & -3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

The adjugate is the transpose of the cofactor matrix: $\operatorname{adj} A = \frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$. Note that the cofactor matrix is symmetric here. Note also that the adjugate actually equals the original matrix in this case.

2. b. We compute the determinant by first adding column 3 to column 2:
$$\begin{vmatrix} 0 & c & -c \\ -1 & 2 & -1 \\ c & -c & c \end{vmatrix} = \begin{vmatrix} 0 & 0 & -c \\ -1 & 1 & -1 \\ c & 0 & c \end{vmatrix} = (-c)\begin{vmatrix} -1 & 1 \\ c & 0 \end{vmatrix} = (-c)(-c) = c^2$$
This is zero if and only if $c = 0$, so the matrix is invertible if and only if $c \neq 0$.

d. Begin by subtracting row 1 from row 3, and then subtracting column 1 from column 3:
$$\begin{vmatrix} 4 & c & 3 \\ c & 2 & c \\ 5 & c & 4 \end{vmatrix} = \begin{vmatrix} 4 & c & 3 \\ c & 2 & c \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 4 & c & -1 \\ c & 2 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 1\begin{vmatrix} c & -1 \\ 2 & 0 \end{vmatrix} = 2$$
This is nonzero for *all* values of c, so the matrix is invertible for all c.

f. Begin by subtracting c times row 1 from row 2:

$$\begin{vmatrix} 1 & c & -1 \\ c & 1 & 1 \\ 0 & 1 & c \end{vmatrix} = \begin{vmatrix} 1 & c & -1 \\ 0 & 1-c^2 & 1+c \\ 0 & 1 & c \end{vmatrix} = \begin{vmatrix} 1-c^2 & 1+c \\ 1 & c \end{vmatrix} = \begin{vmatrix} (1+c)(1-c) & 1+c \\ 1 & c \end{vmatrix}$$

Now take the common factor $(1+c)$ out of row 1:

$$\begin{vmatrix} 1 & c & -1 \\ c & 1 & 1 \\ 0 & 1 & c \end{vmatrix} = (1+c)\begin{vmatrix} 1-c & 1 \\ 1 & c \end{vmatrix} = (1+c)[c(1-c)-1] = -(1+c)(c^2-c+1) = -(c^3+1)$$

This is zero if and only if $c=-1$ (the roots of c^2-c+1 are not real). Hence the matrix is invertible if and only if $c \neq -1$.

3. b.
$$\begin{aligned} \det(B^2C^{-1}AB^{-1}C^T) &= \det B^2 \det C^{-1} \det A \det B^{-1} \det C^T \\ &= (\det B)^2 \tfrac{1}{\det C} \det A \tfrac{1}{\det B} \det C \\ &= \det B \det A \\ &= -2 \end{aligned}$$

4. b. $\det(A^{-1}B^{-1}AB) = \det A^{-1} \det B^{-1} \det A \det B = \frac{1}{\det A}\frac{1}{\det B} \det A \det B = 1.$

Note that the following proof is **wrong**:

$$\det(A^{-1}B^{-1}AB) = \det(A^{-1}AB^{-1}B) = \det(I \cdot I) = \det I = 1$$

The reason is that $A^{-1}B^{-1}AB$ may not equal $A^{-1}AB^{-1}B$ because $B^{-1}A$ need not equal AB^{-1}.

6. b. Since C is 3×3, the same is true for C^{-1}, so $\det(2C^{-1}) = 2^3 \cdot \det C^{-1} = \frac{8}{\det C}$. Now we compute $\det C$ by taking 2 and 3 out of columns 2 and 3, subtracting column 3 from column 2:

$$\det C = \begin{vmatrix} 2p & -a+u & 3u \\ 2q & -b+v & 3v \\ 2r & -c+w & 3w \end{vmatrix} = 6\begin{vmatrix} p & -a+u & u \\ q & -b+v & v \\ r & -c+w & w \end{vmatrix} = 6\begin{vmatrix} p & -a & u \\ q & -b & v \\ r & -c & w \end{vmatrix}$$

Now take -1 from column 2, interchange columns 1 and 2, and apply Theorem 3.2.3:

$$\det C = -6\begin{vmatrix} p & a & u \\ q & b & v \\ r & c & w \end{vmatrix} = 6\begin{vmatrix} a & p & u \\ b & q & v \\ c & r & w \end{vmatrix} = 6\begin{vmatrix} a & b & c \\ p & q & r \\ u & v & w \end{vmatrix} = 6\cdot 3 = 18$$

Finally $\det 2C^{-1} = \frac{8}{\det C} = \frac{8}{18} = \frac{4}{9}$.

7. b. Begin by subtracting row 2 from row 3, and then expand along column 2:

$$\begin{vmatrix} 2b & 0 & 4d \\ 1 & 2 & -2 \\ a+1 & 2 & 2(c-1) \end{vmatrix} = \begin{vmatrix} 2b & 0 & 4d \\ 1 & 2 & -2 \\ a & 0 & 2c \end{vmatrix} = 2\begin{vmatrix} 2b & 4d \\ a & 2c \end{vmatrix} = 4\begin{vmatrix} b & 2d \\ a & 2c \end{vmatrix} = 8\begin{vmatrix} b & d \\ a & c \end{vmatrix}$$

Interchange rows and use Theorem 3.2.3, to get

$$= -8\begin{vmatrix} a & c \\ b & d \end{vmatrix} = -8\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -8(-2) = 16$$

8. b. $x = \dfrac{\begin{vmatrix} 9 & 4 \\ -1 & -1 \end{vmatrix}}{\begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix}} = \frac{-5}{-11} = \frac{5}{11}$, $y = \dfrac{\begin{vmatrix} 3 & 9 \\ 2 & -1 \end{vmatrix}}{\begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix}} = \frac{-21}{-11} = \frac{21}{11}$

d. The coefficient matrix has determinant:
$\begin{vmatrix} 4 & -1 & 3 \\ 6 & 2 & -1 \\ 3 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 0 \\ 14 & 2 & 5 \\ 15 & 3 & 11 \end{vmatrix} = -(-1)\begin{vmatrix} 14 & 5 \\ 15 & 11 \end{vmatrix} = 79$
Hence Cramer's rule gives

$x = \frac{1}{79}\begin{vmatrix} 1 & -1 & 3 \\ 0 & 2 & -1 \\ -1 & 3 & 2 \end{vmatrix} = \frac{1}{79}\begin{vmatrix} 1 & -1 & 3 \\ 0 & 2 & -1 \\ 0 & 2 & 5 \end{vmatrix} = \frac{1}{79}\begin{vmatrix} 2 & -1 \\ 2 & 5 \end{vmatrix} = \frac{12}{79}$

$y = \frac{1}{79}\begin{vmatrix} 4 & 1 & 3 \\ 6 & 0 & -1 \\ 3 & -1 & 2 \end{vmatrix} = \frac{1}{79}\begin{vmatrix} 4 & 1 & 3 \\ 6 & 0 & -1 \\ 7 & 0 & 5 \end{vmatrix} = -\frac{1}{79}\begin{vmatrix} 6 & -1 \\ 7 & 5 \end{vmatrix} = -\frac{37}{79}$

$z = \frac{1}{79}\begin{vmatrix} 4 & -1 & 1 \\ 6 & 2 & 0 \\ 3 & 3 & -1 \end{vmatrix} = \frac{1}{79}\begin{vmatrix} 4 & -1 & 1 \\ 6 & 2 & 0 \\ 7 & 2 & 0 \end{vmatrix} = \frac{1}{79}\begin{vmatrix} 6 & 2 \\ 7 & 2 \end{vmatrix} = -\frac{2}{79}$

9. b. $A^{-1} = \frac{1}{\det A}\operatorname{adj} A = \frac{1}{\det A}\left[C_{ij}\right]^T$ where $\left[C_{ij}\right]$ is the cofactor matrix. Hence the $(2,\ 3)$-entry of A^{-1} is $\frac{1}{\det A}C_{32}$. Now $C_{32} = -\begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix} = -4$. Since $\det A = \begin{vmatrix} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 0 & 4 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & -5 & 4 \\ 0 & 4 & 7 \end{vmatrix} = \begin{vmatrix} -5 & 4 \\ 4 & 7 \end{vmatrix} = -51$, the $(2,\ 3)$ entry of A^{-1} is $\frac{-4}{-51} = \frac{4}{51}$.

10. b. If $A^2 = I$ then $\det A^2 = \det I = 1$, that is $(\det A)^2 = 1$. Hence $\det A = 1$ or $\det A = -1$.

d. If $PA = P$, P invertible, then $\det PA = \det P$, that is $\det P \det A = \det P$. Since $\det P \neq 0$ (as P is invertible), this gives $\det A = 1$.

f. If $A = -A^T$, A is $n \times n$, then A^T is also $n \times n$ so, using Theorem 3.1.3 and Theorem 3.2.3,

$$\det A = \det\left(-A^T\right) = \det\left[(-1)A^T\right] = (-1)^n \det A^T = (-1)^n \det A$$

If n is even this is $\det A = \det A$ and so gives no information about $\det A$. But if n is odd it reads $\det A = -\det A$, so $\det A = 0$ in this case.

15. Write $d = \det A$, and let C denote the cofactor matrix of A. Here

$$A^T = A^{-1} = \tfrac{1}{d}\operatorname{adj} A = \tfrac{1}{d}C^T$$

Take transposes to get $A = \frac{1}{d}C$, whence $C = dA$.

19. b. Write $A = \begin{bmatrix} 0 & c & -c \\ -1 & 2 & -1 \\ c & -c & c \end{bmatrix}$. Then $\det A = c^2$ (Exercise 2) and the cofactor matrix is

$$\left[C_{ij}\right] = \begin{bmatrix} \begin{vmatrix} 2 & -1 \\ -c & c \end{vmatrix} & -\begin{vmatrix} -1 & -1 \\ c & c \end{vmatrix} & \begin{vmatrix} -1 & 2 \\ c & -c \end{vmatrix} \\ -\begin{vmatrix} c & -c \\ -c & c \end{vmatrix} & \begin{vmatrix} 0 & -c \\ c & c \end{vmatrix} & -\begin{vmatrix} 0 & c \\ c & -c \end{vmatrix} \\ \begin{vmatrix} c & -c \\ 2 & -1 \end{vmatrix} & -\begin{vmatrix} 0 & -c \\ -1 & -1 \end{vmatrix} & \begin{vmatrix} 0 & c \\ -1 & 2 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} c & 0 & -c \\ 0 & c^2 & c^2 \\ c & c & c \end{bmatrix}$$

Hence $A^{-1} = \frac{1}{\det A}\operatorname{adj} A = \frac{1}{c^2}\left[C_{ij}\right]^T = \frac{1}{c^2}\begin{bmatrix} c & 0 & c \\ 0 & c^2 & c \\ -c & c^2 & c \end{bmatrix} = \frac{1}{c}\begin{bmatrix} 1 & 0 & 1 \\ 0 & c & 1 \\ -1 & c & 1 \end{bmatrix}$ for any $c \neq 0$.

d. Write $A = \begin{bmatrix} 4 & c & 3 \\ c & 2 & c \\ 5 & c & 4 \end{bmatrix}$. Then $\det A = 2$ (Exercise 2) and the cofactor matrix is

$$[C_{ij}] = \begin{bmatrix} \begin{vmatrix} 2 & c \\ c & 4 \end{vmatrix} & -\begin{vmatrix} c & c \\ 5 & 4 \end{vmatrix} & \begin{vmatrix} c & 2 \\ 5 & c \end{vmatrix} \\ -\begin{vmatrix} c & 3 \\ c & 4 \end{vmatrix} & \begin{vmatrix} 4 & 3 \\ 5 & 4 \end{vmatrix} & -\begin{vmatrix} 4 & c \\ 5 & c \end{vmatrix} \\ \begin{vmatrix} c & 3 \\ 2 & c \end{vmatrix} & -\begin{vmatrix} 4 & 3 \\ c & c \end{vmatrix} & \begin{vmatrix} 4 & c \\ c & 2 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 8-c^2 & c & c^2-10 \\ -c & 1 & c \\ c^2-6 & -c & 8-c^2 \end{bmatrix}$$

Hence $A^{-1} = \frac{1}{\det A}\,\mathrm{adj}\, A = \frac{1}{2}\left[C_{ij}\right]^T = \frac{1}{2}\begin{bmatrix} 8-c^2 & -c & c^2-6 \\ c & 1 & -c \\ c^2-10 & c & 8-c^2 \end{bmatrix}$.

f. Write $A = \begin{bmatrix} 1 & c & -1 \\ c & 1 & 1 \\ 0 & 1 & c \end{bmatrix}$. Then $\det A = -(c^3+1)$ (Exercise 2) so $\det A = 0$ means $c \neq -1$ (c is real). The cofactor matrix is

$$[C_{ij}] = \begin{bmatrix} \begin{vmatrix} 1 & 1 \\ 1 & c \end{vmatrix} & -\begin{vmatrix} c & 1 \\ 0 & c \end{vmatrix} & \begin{vmatrix} c & 1 \\ 0 & 1 \end{vmatrix} \\ -\begin{vmatrix} c & -1 \\ 1 & c \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & c \end{vmatrix} & -\begin{vmatrix} 1 & c \\ 0 & 1 \end{vmatrix} \\ \begin{vmatrix} c & -1 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ c & 1 \end{vmatrix} & \begin{vmatrix} 1 & c \\ c & 1 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} c-1 & -c^2 & c \\ -(c^2+1) & c & -1 \\ c+1 & -(1+c) & 1-c^2 \end{bmatrix}$$

Hence $A^{-1} = \frac{1}{\det A}\,\mathrm{adj}\, A = \frac{-1}{c^3+1}[C_{ij}]^T = \frac{-1}{c^3+1}\begin{bmatrix} c-1 & -(c^2+1) & c+1 \\ -c^2 & c & -(c+1) \\ c & -1 & 1-c^2 \end{bmatrix} = \frac{1}{c^3+1}\begin{bmatrix} 1-c & c^2+1 & -c-1 \\ c^2 & -c & c+1 \\ -c & 1 & c^2-1 \end{bmatrix}$, where $c \neq -1$.

20. b. True. Write $d = \det A$, so that $d \cdot A^{-1} = \mathrm{adj}\, A$. Since $\mathrm{adj}\, A = A^{-1}$ by hypothesis, this gives $dA^{-1} = A^{-1}$, that is $(d-1)A^{-1} = 0$. It follows that $d = 1$ because $A^{-1} \neq 0$ (see Example 2.1.7).

d. True. Since $AB = AC$ we get $A(B-C) = 0$. As A is invertible, this means $B = C$. More precisely, left multiply by A^{-1} to get $A^{-1}A(B-C) = A^{-1}0 = 0$; that is $I(B-C) = 0$; that is $B - C = 0$, so $B = C$.

f. False. If $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ then $\mathrm{adj}\, A = 0$. However $A \neq 0$.

h. False. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ then $\mathrm{adj}\, A = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$, and this has no row of zeros.

j. False. If $A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ then $\det(I + A) = -1$ but $1 + \det A = 1$.

l. False. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ then $\det A = 1$, but $\mathrm{adj}\, A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \neq A$.

22. b. If $p(x) = r_0 + r_1x + r_2x^2$, the conditions give linear equations for r_0, r_1 and r_2:

$$\begin{array}{ccccccccc} r_0 & & & & & = & p(0) & = & 5 \\ r_0 & + & r_1 & + & r_2 & = & p(1) & = & 3 \\ r_0 & + & 2r_1 & + & 4r_2 & = & p(2) & = & 5 \end{array}$$

The solution is $r_0 = 5$, $r_1 = -4$, $r_2 = 2$, so $p(x) = 5 - 4x + 2x^2$.

23. b. If $p(x) = r_0 + r_1x + r_2x^2 + r_3x^3$, the conditions give linear equations for r_0, r_1, r_2 and r_3:

$$\begin{array}{rcrcrcrcrcr} r_0 & & & & & & & = & p(0) & = & 1 \\ r_0 & + & r_1 & + & r_2 & + & r_3 & = & p(1) & = & 1 \\ r_0 & - & r_1 & + & r_2 & - & r_3 & = & p(-1) & = & 2 \\ r_0 & - & 2r_1 & + & 4r_2 & - & 8r_3 & = & p(-2) & = & -3 \end{array}$$

The solution is $r_0 = 1$, $r_1 = \frac{-5}{3}$, $r_2 = \frac{1}{2}$, $r_3 = \frac{7}{6}$, so $p(x) = 1 - \frac{5}{3}x + \frac{1}{2}x^2 + \frac{7}{6}x^3$.

24. b. If $p(x) = r_0 + r_1x + r_2x^2 + r_3x^3$, the conditions give linear equations for r_0, r_1, r_2 and r_3:

$$\begin{array}{rcrcrcrcrcr} r_0 & & & & & & & = & p(0) & = & 1 \\ r_0 & + & r_1 & + & r_2 & + & r_3 & = & p(1) & = & 1.49 \\ r_0 & + & 2r_1 & + & 4r_2 & + & 8r_3 & = & p(2) & = & -0.42 \\ r_0 & + & 3r_1 & + & 9r_2 & + & 27r_3 & = & p(3) & = & -11.33 \end{array}$$

The solution is $r_0 = 1$, $r_1 = -0.51$, $r_2 = 2.1$, $r_3 = -1.1$, so

$$p(x) = 1 - 0.51x + 2.1x^2 - 1.1x^3$$

The estimate for the value of y corresponding to $x = 1.5$ is

$$y = p(1.5) = 1 - 0.51(1.5) + 2.1(1.5)^2 - 1.1(1.5)^3 = 1.25$$

to two decimals.

26. b. Let A be an upper triangular, invertible, $n \times n$ matrix. We use induction on n. If $n = 1$ it is clear (every 1×1 matrix is upper triangular). If $n > 1$ write $A = \left[\begin{array}{cc} a & X \\ 0 & B \end{array}\right]$ and $A^{-1} = \left[\begin{array}{cc} b & Y \\ Z & C \end{array}\right]$ in block form. Then

$$\left[\begin{array}{cc} 1 & 0 \\ 0 & I \end{array}\right] = AA^{-1} = \left[\begin{array}{cc} ab + XZ & aY + XC \\ BZ & BC \end{array}\right]$$

So $BC = I$, $BZ = 0$. Thus $C = B^{-1}$ is upper triangular by induction (B is upper triangular because A is) and $BZ = 0$ gives $Z = 0$ because B is invertible. Hence $A^{-1} = \left[\begin{array}{cc} b & Y \\ 0 & C \end{array}\right]$ is upper triangular.

28. Write $d = \det A$. Then $\frac{1}{d} = \det(A^{-1}) = \det \left[\begin{array}{rrr} 3 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 1 & -1 \end{array}\right] = -21$. Hence $d = \frac{-1}{21}$. By Theorem 3.2.4, we have $A \cdot \mathrm{adj}\, A = dI$, so $\mathrm{adj}\, A = A^{-1}(dI) = dA^{-1} = \frac{-1}{21} \left[\begin{array}{rrr} 3 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 1 & -1 \end{array}\right]$.

34. b. Write $d = \det A$ so $\det A^{-1} = \frac{1}{d}$. Now the adjugate for A^{-1} gives

$$A^{-1}(\mathrm{adj}\, A^{-1}) = \tfrac{1}{d}I$$

Take inverses to get $(\mathrm{adj}\, A^{-1})^{-1}A = dI$. But $dI = (\mathrm{adj}\, A)A$ by the adjugate formula for A. Hence

$$(\mathrm{adj}\, A^{-1})^{-1}A = (\mathrm{adj}\, A)A$$

Since A is invertible, we get $\left[\mathrm{adj}\, A^{-1}\right]^{-1} = \mathrm{adj}\, A$, and the result follows by taking inverses again.

d. The adjugate formula gives

$$AB \operatorname{adj}(AB) = \det AB \cdot I = \det A \cdot \det B \cdot I$$

On the other hand

$$\begin{aligned} AB \operatorname{adj} B \cdot \operatorname{adj} A &= A[(\det B)I] \operatorname{adj} A \\ &= A \cdot \operatorname{adj} A \cdot (\det B)I \\ &= (\det A)I \cdot (\det B)I \\ &= \det A \det B \cdot I \end{aligned}$$

Thus $AB \operatorname{adj}(AB) = AB \cdot \operatorname{adj} B \cdot \operatorname{adj} A$, and the result follows because AB is invertible.

3.3 Diagonalization and Eigenvalues

1. b. $c_A(x) = \begin{vmatrix} x-2 & 4 \\ 1 & x+1 \end{vmatrix} = x^2 - x - 6 = (x-3)(x+2)$; hence the eigenvalues are $\lambda_1 = 3$, and $\lambda_2 = -2$. Take these values for x in the matrix $xI - A$ for $c_A(x)$:
$\lambda_1 = 3$: $\begin{bmatrix} 1 & 4 \\ 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 4 \\ 0 & 0 \end{bmatrix}$; $\mathbf{x}_1 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$.
$\lambda_2 = -2$: $\begin{bmatrix} -4 & 4 \\ 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$; $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
So $P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix}$ has $P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$.

d. To compute $c_A(x)$ we first add row 1 to row 3:

$$\begin{aligned} c_A(x) &= \begin{vmatrix} x-1 & -1 & 3 \\ -2 & x & -6 \\ -1 & 1 & x-5 \end{vmatrix} = \begin{vmatrix} x-1 & -1 & 3 \\ -2 & x & -6 \\ x-2 & 0 & x-2 \end{vmatrix} = \begin{vmatrix} x-1 & -1 & -x+4 \\ -2 & x & -4 \\ x-2 & 0 & 0 \end{vmatrix} \\ &= (x-2)\begin{vmatrix} -1 & -x+4 \\ x & -4 \end{vmatrix} = (x-2)[x^2-4x+4] = (x-2)^3 \end{aligned}$$

So the eigenvalue is $\lambda_1 = 2$ of multiplicity 3. Taking $x = 3$ in the matrix $xI - A$ for $c_A(x)$:

$$\begin{bmatrix} 1 & -1 & 3 \\ -2 & 2 & 6 \\ -1 & 1 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} s-3t \\ s \\ t \end{bmatrix}; \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}; \mathbf{x}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

Hence there are not $n = 3$ basic eigenvectors, so A is *not* diagonalizable.

f. Here $c_A(x) = \begin{vmatrix} x & -1 & 0 \\ -3 & x & -1 \\ -2 & 0 & x \end{vmatrix} = x^3 - 3x - 2$. Note that -1 is a root of $c_A(x)$ so $x+1$ is a factor. Long division gives $c_A(x) = (x+1)(x^2 - x - 2)$. But $x^2 - x - 2 = (x+1)(x-2)$, so $c_A(x) = (x+1)^2(x-2)$. Hence, the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 2$. Substitute $\lambda_1 = -1$ in the matrix $xI - c_A(x)$ gives

$$\begin{bmatrix} -1 & -1 & 0 \\ -3 & -1 & -1 \\ -2 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

so the solution involves only 1 parameter. As the multiplicity of λ_1 is 2, A is not diagonalizable by Theorem 3.3.5. Note that this matrix and the matrix in Example 3.3.9 have the same characteristic polynomial but the matrix in Example 3.3.9 is diagonalizable, while this one is not.

h. $c_A(x) = \begin{vmatrix} x-2 & -1 & -1 \\ 0 & x-1 & 0 \\ -1 & 1 & x-2 \end{vmatrix} = (x-1)\begin{vmatrix} x-2 & -1 \\ -1 & x-2 \end{vmatrix} = (x-1)^2(x-3)$. Hence the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 3$. Take these values for x in the matrix $xI - A$ for $c_A(x)$:
$\lambda_1 = 1$: $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ -1 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $\mathbf{x}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$
$\lambda_2 = 3$: $\begin{bmatrix} 1 & -1 & -1 \\ 0 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$
Since $n = 3$ and there are only two basic eigenvectors, A is *not* diagonalizable.

2. b. As in Exercise 1, we find $\lambda_1 = 2$ and $\lambda_2 = -1$; with corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, so $P = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ satisfies $P^{-1}AP = D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$. Next compute

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = P_0^{-1}\mathbf{v} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 7 \\ -5 \end{bmatrix}$$

Hence $b_1 = \frac{7}{3}$ so, as λ_1 is dominant, $\mathbf{x}_k \cong b_1\lambda_1^k\mathbf{x}_1 = \frac{7}{3}2^k\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

d. Here $\lambda_1 = 3$, $\lambda_2 = -2$ and $\lambda_3 = 1$; $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}$ and $\mathbf{x}_3 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, and $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 1 & -3 & 3 \end{bmatrix}$. Now $P^{-1} = \frac{1}{6}\begin{bmatrix} 3 & 6 & 3 \\ 2 & -2 & -2 \\ 1 & -4 & -1 \end{bmatrix}$, so $P_0^{-1}\mathbf{v}_0 = \frac{1}{6}\begin{bmatrix} 9 \\ 2 \\ 1 \end{bmatrix}$ and hence $b_1 = \frac{3}{2}$. Hence $\mathbf{v}_k \cong \frac{3}{2}3^k\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

4. If λ is an eigenvalue for A, let $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq \mathbf{0}$. Then

$$A_1\mathbf{x} = (A - \alpha I)\mathbf{x} = A\mathbf{x} - \alpha\mathbf{x} = \lambda\mathbf{x} - \alpha\mathbf{x} = (\lambda - \alpha)\mathbf{x}$$

So $\lambda - \alpha$ is an eigenvalue of $A_1 = A - \alpha I$ (with the same eigenvector). Conversely, if $\lambda - \alpha$ is an eigenvalue of A_1, then $A_1\mathbf{y} = (\lambda - \alpha)\mathbf{y}$ for some $\mathbf{y} \neq 5$. Thus, $(A - \alpha I)\mathbf{y} = (\lambda - \alpha)\mathbf{y}$, whence $A\mathbf{y} - \alpha\mathbf{y} = \lambda\mathbf{y} - \alpha\mathbf{y}$. Thus $A\mathbf{y} = \lambda\mathbf{y}$ so λ is an eigenvalue of A.

8. b. Direct computation gives $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. Since $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^n = \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix}$, the hint gives $A^n = P\begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix}P^{-1} = \begin{bmatrix} 9-8\cdot 2^n & 12(1-2^n) \\ 6(2^n-1) & 9\cdot 2^n - 8 \end{bmatrix}$.

9. b. $A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$. We have $c_A(x) = x(x-2)$ so A has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 2$ with basic eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Since $\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ is invertible, it is a diagonalizing matrix for A. On the other hand, $D + A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable by Example 3.3.10.

11. b. Since A is diagonalizable, let $P^{-1}AP = D$ be diagonal. Then $P^{-1}(kA)P = k(P^{-1}AP) = dD$ is also diagonal, so kA is diagonalizable too.

d. Again let $P^{-1}AP = D$ be diagonal. The matrix $Q = U^{-1}P$ is invertible and

$$Q^{-1}(U^{-1}AU)Q = P^{-1}U(U^{-1}AU)U^{-1}P = P^{-1}AP = D \text{ is diagonal.}$$

This shows that $U^{-1}AU$ is diagonalizable with diagonalizing matrix $Q = U^{-1}P$.

12. $\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{cc} 2 & 1 \\ 0 & -1 \end{array}\right] + \left[\begin{array}{cc} -1 & 0 \\ 0 & 2 \end{array}\right]$ and both $\left[\begin{array}{cc} 2 & 1 \\ 0 & -1 \end{array}\right]$ and $\left[\begin{array}{cc} -1 & 0 \\ 0 & 2 \end{array}\right]$ are diagonalizable. However, $\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]$ is not diagonalizable by Example 3.3.10.

14. If A is $n \times n$, let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues, all either 0 or 1. Since A is diagonalizable (by hypothesis), we have $P^{-1}AP = D$ where $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ is the diagonal matrix with $\lambda_1, \dots, \lambda_n$ down the main diagonal. Since each $\lambda_i = 0, 1$ it follows that $\lambda_i^2 = \lambda_i$ for each i. Thus $D^2 = \text{diag}(\lambda_1^2, \dots, \lambda_n^2) = \text{diag}(\lambda_1, \dots, \lambda_n) = D$. Since $P^{-1}AP = D$, we have $A = PDP^{-1}$. Hence

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD^2P^{-1} = PDP^{-1} = A$$

18. b. Since $r \neq 0$ and A is $n \times n$, we have

$$c_{rA}(x) = \det[xI - rA] = \det\left[r\left(\tfrac{x}{r}I - A\right)\right] = r^n \det\left[\tfrac{x}{r}I - A\right]$$

As $c_A(x) = \det[xI - A]$, this shows that $c_{rA}(x) = r^n c_A\left(\frac{x}{r}\right)$.

20. b. If μ is an eigenvalue of A^{-1} then $A^{-1}\mathbf{x} = \mu\mathbf{x}$ for some column $\mathbf{x} \neq \mathbf{0}$. Note that $\mu \neq 0$ because A^{-1} is invertible and $\mathbf{x} \neq \mathbf{0}$. Left multiplication by A gives $\mathbf{x} = \mu A\mathbf{x}$, whence $A\mathbf{x} = \frac{1}{\mu}\mathbf{x}$. Thus, $\frac{1}{\mu}$ is an eigenvalue of A; call it $\lambda = \frac{1}{\mu}$. Hence, $\mu = \frac{1}{\lambda}$ as required. Conversely, if λ is any eigenvalue of A then $\lambda \neq 0$ by (a) and we claim that $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}. We have $A\mathbf{x} = \lambda\mathbf{x}$ for some column $\mathbf{x} \neq \mathbf{0}$. Multiply on the left by A^{-1} to get $\mathbf{x} = \lambda A^{-1}\mathbf{x}$; whence $A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$. Thus $\frac{1}{\lambda}$ is indeed an eigenvalue of A^{-1}.

21. b. We have $A\mathbf{x} = \lambda\mathbf{x}$ for some column $\mathbf{x} \neq \mathbf{0}$. Hence, $A^2\mathbf{x} = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$, $A^3 = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$, so

$$(A^3 - 2A + 3I)\mathbf{x} = A^3\mathbf{x} - 2A\mathbf{x} + 3\mathbf{x} = \lambda^3\mathbf{x} - 2\lambda\mathbf{x} + 3\mathbf{x} = (\lambda^3 - 2\lambda + 3)\mathbf{x}$$

23. b. If λ is an eigenvalue of A, let $A\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \neq \mathbf{0}$. Then $A^2\mathbf{x} = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$, $A^3\mathbf{x} = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$, ... We claim that $A^k\mathbf{x} = \lambda^k\mathbf{x}$ holds for every $k \geq 1$. We have already checked this for $k = 1$. If it holds for some $k \geq 1$, then $A^k\mathbf{x} = \lambda^k\mathbf{x}$, so

$$A^{k+1}\mathbf{x} = A(A^k\mathbf{x}) = A(\lambda^k\mathbf{x}) = \lambda^k A\mathbf{x} = \lambda^k(\lambda\mathbf{x}) = \lambda^{k+1}\mathbf{x}$$

Hence, it also holds for $k + 1$, and so $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for all $k \geq 1$ by induction. In particular, if $A^m = 0$, $m \geq 1$, then $\lambda^m\mathbf{x} = A^m\mathbf{x} = 0\mathbf{x} = \mathbf{0}$. As $\mathbf{x} \neq \mathbf{0}$, this implies that $\lambda^m = 0$, so $\lambda = 0$.

24. a. Let A be diagonalizable with $A^m = I$. If λ is any eigenvalue of A, say $A\mathbf{x} = \lambda\mathbf{x}$ for some column $\mathbf{x} \neq \mathbf{0}$, then (see the solution to 23(b) above) $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for all $k \geq 1$. Taking $k = m$ we have $\mathbf{x} = A^m\mathbf{x} = \lambda^m\mathbf{x}$, whence $\lambda^m = 1$. Thus λ is a complex m^{th} root of unity and so lies on the unit circle by Theorem A.3. But we are assuming that λ is a *real* number so $\lambda = \pm 1$, so $\lambda^2 = 1$. Also A is diagonalizable, say $P^{-1}AP = D = \text{diag}(\lambda_1, \dots, \lambda_n)$ where the λ_i are the eigenvalues of A. Hence $D^2 = \text{diag}(\lambda_1^2, \dots, \lambda_n^2) = I$ because $\lambda_i^2 = 1$ for each i. Finally, since $A = PDP^{-1}$ we obtain $A^2 = PD^2P^{-1} = PIP^{-1} = I$.

27. a. If A is diagonalizable and has only one eigenvalue λ, then the diagonalization algorithm asserts that $P^{-1}AP = \lambda I$. But then $A = P(\lambda I)P^{-1} = \lambda I$, as required.

    b. Here the characteristic polynomial is $c_A(x) = (x-1)^2$, so the only eigenvalue is $\lambda = 1$. Hence A is not diagonalizable by (a).

31. b. The matrix in Example 3.3.1 is $\left[\begin{array}{cc} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{array}\right]$. In this case $A = \left[\begin{array}{cc} \frac{1}{4} & \frac{1}{4} \\ 3 & 0 \end{array}\right]$ so

$$c_A(x) = \det \left[\begin{array}{cc} x-\frac{1}{4} & -\frac{1}{4} \\ -3 & x \end{array}\right] = x^2 - \tfrac{1}{4}x - \tfrac{3}{4} = (x-1)(x+\tfrac{3}{4})$$

    Hence the dominant eigenvalue is $\lambda = 1$, and the population stabilizes.

    d. In this case $A = \left[\begin{array}{cc} \frac{3}{5} & \frac{1}{5} \\ 3 & 0 \end{array}\right]$ so $c_A(x) = \det \left[\begin{array}{cc} x-\frac{3}{5} & -\frac{1}{5} \\ -3 & x \end{array}\right] = x^2 - \frac{3}{5}x - \frac{3}{5}$. By the quadratic formula, the roots are $\frac{1}{10}[3 \pm \sqrt{69}]$, so the dominant eigenvalue is $\frac{1}{10}[3+\sqrt{69}] \approx 1.13 > 1$, so the population diverges.

34. Here the matrix A in Example 3.3.1 is $A = \left[\begin{array}{cc} \alpha & \frac{2}{5} \\ 2 & 0 \end{array}\right]$ where α is the adult reproduction rate. Hence $c_A(x) = \det \left[\begin{array}{cc} x-\alpha & -\frac{2}{5} \\ -2 & x \end{array}\right] = x^2 - \alpha x - \frac{4}{5}$, and the roots are $\frac{1}{2}\left[\alpha \pm \sqrt{\alpha^2 + \frac{16}{5}}\right]$. Thus the dominant eigenvalue is $\lambda_1 = \frac{1}{2}\left[\alpha + \sqrt{\alpha^2 + \frac{16}{5}}\right]$, and this equals 1 if and only if $\alpha = \frac{1}{5}$. So the population stabilizes if $\alpha = \frac{1}{5}$. In fact it is easy to see that the population becomes extinct ($\lambda_1 < 1$) if and only if $\alpha < \frac{1}{5}$, and the population diverges ($\lambda_1 > 1$) if and only if $\alpha > \frac{1}{5}$.

3.4 An Application to Linear Recurrences

1. b. In this case $x_{k+2} = 2x_k - x_{k+1}$, so $\mathbf{v}_{k+1} = \left[\begin{array}{c} x_{k+1} \\ 2x_k - x_{k+1} \end{array}\right] = \left[\begin{array}{cc} 0 & 1 \\ 2 & -1 \end{array}\right]\left[\begin{array}{c} x_k \\ x_{k+1} \end{array}\right] = A\mathbf{v}_k$. Diagonalizing A gives $P = \left[\begin{array}{cc} 1 & 1 \\ 1 & -2 \end{array}\right]$ and $D = \left[\begin{array}{cc} 1 & 0 \\ 0 & -2 \end{array}\right]$. Hence

$$\mathbf{b} = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] = P_0^{-1}\mathbf{v}_0 = \tfrac{1}{3}\left[\begin{array}{cc} 2 & 1 \\ 1 & -1 \end{array}\right]\left[\begin{array}{c} 1 \\ 2 \end{array}\right] = \left[\begin{array}{c} \frac{4}{3} \\ -\frac{1}{3} \end{array}\right]$$

    Thus $\left[\begin{array}{c} x_k \\ x_{k+1} \end{array}\right] = \frac{4}{3}1^k\left[\begin{array}{c} 1 \\ 1 \end{array}\right] - \frac{1}{3}(-2)^k\left[\begin{array}{c} 1 \\ -2 \end{array}\right]$ for each k. Comparing top entries gives

$$x_k = \tfrac{4}{3} - \tfrac{1}{3}(-2)^k = \tfrac{1}{3}[4-(-2)^k] = \tfrac{(-2)^k}{3}\left[1 - 4\left(\tfrac{1}{-2}\right)^k\right] \approx -\tfrac{(-2)^k}{3} \text{ for large } k.$$

    Here -2 is the dominant eigenvalue, so $x_k = \frac{1}{3}(-2)^k[\frac{4}{(-2)^k} - 1] \approx -\frac{1}{3}(-2)^k$ if k is large.

    d. Here $x_{k+2} = 6x_k - x_{k+1}$, so $\mathbf{v}_{k+1} = \left[\begin{array}{c} x_{k+1} \\ 6x_k - x_{k+1} \end{array}\right] = \left[\begin{array}{cc} 0 & 1 \\ 6 & -1 \end{array}\right]\left[\begin{array}{c} x_k \\ x_{k+1} \end{array}\right] = A\mathbf{v}_k$. Diagonalizing A gives $P = \left[\begin{array}{cc} 1 & 1 \\ 2 & -3 \end{array}\right]$ and $D = \left[\begin{array}{cc} 2 & 0 \\ 0 & -3 \end{array}\right]$. Hence

$$\mathbf{b} = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] = P_0^{-1}\mathbf{v} = \tfrac{1}{5}\left[\begin{array}{cc} 3 & 1 \\ 2 & -1 \end{array}\right]\left[\begin{array}{c} 1 \\ 1 \end{array}\right] = \left[\begin{array}{c} \frac{4}{5} \\ \frac{1}{5} \end{array}\right]$$

Now $\begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix} = \frac{4}{5}2^k \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \frac{1}{5}(-3)^k \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, and so, looking at the top entries we get

$$x_k = \frac{4}{5}2^k + \frac{1}{5}(-3)^k = \frac{1}{5}[2^{k+2} + (-3)^k]$$

Here $x_k = \frac{(-3)^k}{5}\left[1 + r\left(\frac{2}{-3}\right)^k\right] \approx \frac{(-3)^k}{5}$ for large k, so -3 is dominant.

2. b. Let $\mathbf{v}_k = \begin{bmatrix} x_k \\ x_{k+1} \\ x_{k+2} \end{bmatrix}$. Then $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix}$, and diagonalization gives $P = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 1 & 4 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = P_0^{-1}\mathbf{v}_0 = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$, giving the general formula

$$\mathbf{v}_k = \frac{1}{2}(-1)1^k \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + (0)(-2)^k \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} + \frac{1}{2}1^k \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Thus equating first entries give

$$x_k = \frac{1}{2}(-1)^k + \frac{1}{2}1^k = \frac{1}{2}[(-1)^k + 1]$$

Note that the sequence x_k here is 0, 1, 0, 1, 0, 1, ... which does not converge to any fixed value for large k.

3. b. If a bus is parked at one end of the row, the remaining spaces can be filled in x_k ways to fill it in; if a truck is at the end, there are x_{k+2} ways; and if a car is at the end, there are x_{k+3} ways. Since one (and only one) of these three possibilities must occur, we have $x_{k+4} = x_k + x_{k+2} + x_{k+3}$ must hold for all $k \geq 1$. Since $x_1 = 1$, $x_2 = 2$ (cc or t), $x_3 = 3$ (ccc, ct or tc) and $x_4 = 6$ (cccc, cct, ctc, tcc, tt, b), we get successively, $x_5 = 10$, $x_6 = 18$, $x_7 = 31$, $x_8 = 55$, $x_9 = 96$, $x_{10} = 169$.

5. Let x_k denote the number of ways to form words of k letters. A word of $k+2$ letters must end in either a or b. The number of words that end in b is x_{k+1} — just add a b to a $(k+1)$-letter word. But the number ending in a is x_k since the second-last letter must be a b (no adjacent a's) so we simply add ba to any k-letter word. This gives the recurrence $x_{k+2} = x_{k+1} + x_k$ which is the same as in Example 3.4.2, but with different initial conditions: $x_0 = 1$ (since the "empty" word is the only one formed with no letters) and $x_1 = 2$. The eigenvalues, eigenvectors, and diagonalization remain the same, and so

$$\mathbf{v}_k = b_1\lambda_1^k \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} + b_2\lambda_2^k \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}$$

where $\lambda_1 = \frac{1}{2}(1+\sqrt{5})$ and $\lambda_2 = \frac{1}{2}(1-\sqrt{5})$. Comparing top entries gives

$$x_k = b_1\lambda_1^k + b_2\lambda_2^k$$

By Theorem 2.4.1, the constants b_1 and b_2 come from $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = P_0^{-1}\mathbf{v}_0$. However, we vary the method and use the initial conditions to determine the values of b_1 and b_2 directly. More precisely, $x_0 = 1$ means $1 = b_1 + b_2$ while $x_1 = 2$ means $2 = b_1\lambda_1 + b_2\lambda_2$. These equations have unique solution $b_1 = \frac{\sqrt{5}-3}{2\sqrt{5}}$ and $b_2 = \frac{\sqrt{5}-3}{2\sqrt{5}}$. It follows that

$$x_k = \frac{1}{2\sqrt{5}}\left[(3+\sqrt{5})\left(\frac{1+\sqrt{5}}{2}\right)^k + (-3+\sqrt{5})\left(\frac{1-\sqrt{5}}{2}\right)^k\right] \text{ for each } k \geq 0$$

7. In a stack of $k+2$ chips, if the last chip is gold then (to avoid having two gold chips together) the second last chip must be either red or blue. This can happen in $2x_k$ ways. But there are x_{k+1} ways that the last chip is red (or blue) so there are $2x_{k+1}$ ways these possibilities can occur. Hence $x_{k+2} = 2x_k + 2x_{k+1}$. The matrix is $A = \left[\begin{array}{cc} 0 & 1 \\ 2 & 2 \end{array}\right]$ with eigenvalues $\lambda_1 = 1+\sqrt{3}$ and $\lambda_2 = 1-\sqrt{3}$ and corresponding eigenvectors $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ \lambda_1 \end{array}\right]$ and $\mathbf{x}_2 = \left[\begin{array}{c} 1 \\ \lambda_2 \end{array}\right]$. Given the initial conditions $x_0 = 1$ and $x_1 = 3$, we get

$$\left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] = P_0^{-1}\mathbf{v}_0 = \frac{1}{\sqrt{3}}\left[\begin{array}{cc} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{array}\right]\left[\begin{array}{c} 1 \\ 3 \end{array}\right] = \frac{1}{-2\sqrt{3}}\left[\begin{array}{c} -2-\sqrt{3} \\ 2-\sqrt{3} \end{array}\right] = \frac{1}{2\sqrt{3}}\left[\begin{array}{c} 2+\sqrt{3} \\ -2+\sqrt{3} \end{array}\right]$$

Since Theorem 2.4.1 gives

$$\mathbf{v}_k = b_1\lambda_1^k\left[\begin{array}{c} 1 \\ \lambda_1 \end{array}\right] + b_2\lambda_2^k\left[\begin{array}{c} 1 \\ \lambda_2 \end{array}\right]$$

comparing top entries gives

$$x_k = b_1\lambda_1^k + b_2\lambda_2^k = \frac{1}{2\sqrt{3}}\left[(2+\sqrt{3})(1+\sqrt{3})^k + (-2+\sqrt{3})(1-\sqrt{3})^k\right]$$

9. Let y_k be the yield for year k. Then the yield for year $k+2$ is $y_{k+2} = \frac{y_k + y_{k+1}}{2} = \frac{1}{2}y_k + \frac{1}{2}y_{k+1}$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -\frac{1}{2}$, with corresponding eigenvectors $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ 1 \end{array}\right]$ and $\mathbf{x}_2 = \left[\begin{array}{c} -2 \\ 1 \end{array}\right]$. Given that $k=0$ for the year 1990, we have the initial conditions $y_0 = 10$ and $y_1 = 12$. Thus

$$\left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] = P_0^{-1}\mathbf{v}_0 = \frac{1}{3}\left[\begin{array}{cc} 1 & 2 \\ -1 & 1 \end{array}\right]\left[\begin{array}{c} 10 \\ 12 \end{array}\right] = \frac{1}{3}\left[\begin{array}{c} 34 \\ 2 \end{array}\right]$$

Since

$$\mathbf{v}_k = \frac{34}{3}(1)^k\left[\begin{array}{c} 1 \\ 1 \end{array}\right] + \frac{2}{3}\left(-\frac{1}{2}\right)^k\left[\begin{array}{c} -2 \\ 1 \end{array}\right]$$

then

$$y_k = \frac{34}{3}(1)^k + \frac{2}{3}(-2)\left(-\frac{1}{2}\right)^k = \frac{34}{3} - \frac{4}{3}\left(-\frac{1}{2}\right)^k$$

For large k, $y_k \approx \frac{34}{3}$ so the long term yield is $11\frac{1}{3}$ million tons of wheat.

11. b. We have $A = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & b & c \end{array}\right]$ so $c_A(x) = x^3 - (a + bx + cx^2)$. If λ is any eigenvalue of A, and we write $\mathbf{x} = \left[\begin{array}{c} 1 \\ \lambda \\ \lambda^2 \end{array}\right]$, we have

$$A\mathbf{x} = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & b & c \end{array}\right]\left[\begin{array}{c} 1 \\ \lambda \\ \lambda^2 \end{array}\right] = \left[\begin{array}{c} \lambda \\ \lambda^2 \\ a + b\lambda + c\lambda^2 \end{array}\right] = \left[\begin{array}{c} \lambda \\ \lambda^2 \\ \lambda^3 \end{array}\right] = \lambda\mathbf{x}$$

because $c_A(\lambda) = 0$. Hence $\mathbf{x}$ is a λ-eigenvector.

12. b. We have $p = \frac{5}{6}$ from (a), so $y_k = x_k + \frac{5}{6}$ satisfies $y_{k+2} = y_{k+1} + 6y_k$ with $y_0 = y_1 = \frac{11}{6}$. Here $A = \left[\begin{array}{cc} 0 & 1 \\ 6 & 1 \end{array}\right]$ with eigenvalues 3 and -2, and diagonalizing matrix $P = \left[\begin{array}{cc} 1 & -1 \\ 3 & 2 \end{array}\right]$. This gives $y_k = \frac{11}{30}\left[3^{k+1} - (-2)^{k+1}\right]$, so $x_k = \frac{11}{30}\left[3^{k+1} - (-2)^{k+1}\right] - \frac{5}{6}$.

13. a. If p_k is a solution of (3.12) and q_k is a solution of (3.13) then

$$\begin{aligned} q_{k+2} &= aq_{k+1} + bq_k \\ p_{k+2} &= ap_{k+1} + bp_k + c(k) \end{aligned}$$

for all k. Adding these equations we obtain

$$p_{k+2} + q_{k+2} = a(p_{k+1} + q_{k+1}) + b(p_k + q_k) + c(k)$$

that is $p_k + q_k$ is also a solution of (3.12).

b. If r_k is any solution of (3.12) then $r_{k+2} = ar_{k+1} + br_k + c(k)$. Define $q_k = r_k - p_k$ for each k. Then it suffices to show that q_k is a solution of (3.13). But

$$q_{k+2} = r_{k+2} - p_{k+2} = (ar_{k+1} + br_k + c(k)) - (ap_{k+1} + bp_k + c(k)) = aq_{k+1} + bq_k$$

which is what we wanted.

3.5 An Application to Systems of Differential Equations

1. b. The matrix of the system is $A = \begin{bmatrix} -1 & 5 \\ 1 & 3 \end{bmatrix}$ so $c_A(x) = \begin{vmatrix} x+1 & -5 \\ -1 & x-3 \end{vmatrix} = (x-4)(x+2)$.
$\lambda_1 = 4$: $\begin{bmatrix} 5 & -5 \\ -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
$\lambda_2 = -2$: $\begin{bmatrix} -1 & -5 \\ -1 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 5 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$.
Thus $P^{-1}AP = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}$ where $P = \begin{bmatrix} 1 & 5 \\ 1 & -1 \end{bmatrix}$. The general solution is

$$\mathbf{f} = c_1\mathbf{x}_1 e^{\lambda_1 x} + c_2\mathbf{x}_2 e^{\lambda_2 x} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4x} + c_2 \begin{bmatrix} 5 \\ -1 \end{bmatrix} e^{-2x}$$

Hence, $f_1(x) = c_1e^{4x} + 5c_2e^{-2x}$, $f_2(x) = c_1e^{4x} - c_2e^{-2x}$. The boundary condition is $f_1(0) = 1$, $f_2(0) = -1$; that is

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{f}(0) = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

Thus $c_1 + 5c_2 = 1$, $c_1 - c_2 = -1$; the solution is $c_1 = -\frac{2}{3}$, $c_2 = \frac{1}{3}$, so the specific solution is

$$f_1(x) = \tfrac{1}{3}(5e^{-2x} - 2e^{4x}),\ f_2(x) = -\tfrac{1}{3}(2e^{4x} + e^{-2x})$$

d. Now $A = \begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & -2 \\ 3 & 1 & 1 \end{bmatrix}$. To evaluate $c_A(x)$, first subtract row 1 from row 3:

$$\begin{aligned} c_A(x) &= \begin{vmatrix} x-2 & -1 & -2 \\ -2 & x-2 & 2 \\ -3 & -1 & x-1 \end{vmatrix} = \begin{vmatrix} x-2 & -1 & -2 \\ -2 & x-2 & 2 \\ -x-1 & 0 & x+1 \end{vmatrix} = \begin{vmatrix} x-4 & -1 & -2 \\ 0 & x-2 & 2 \\ 0 & 0 & x+1 \end{vmatrix} \\ &= (x+1)(x-2)(x-4) \end{aligned}$$

$\lambda_1 = -1$: $\begin{bmatrix} -3 & -1 & -2 \\ -2 & -3 & 2 \\ -3 & -1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -6 \\ 2 & 3 & -2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{8}{7} \\ 0 & 1 & -\frac{10}{7} \\ 0 & 0 & 0 \end{bmatrix}$; $\mathbf{x}_1 = \begin{bmatrix} -8 \\ 10 \\ 7 \end{bmatrix}$

$\lambda_2 = 2$: $\begin{bmatrix} 0 & -1 & -2 \\ -2 & 0 & 2 \\ -3 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$; $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

$\lambda_3 = 4$: $\begin{bmatrix} 2 & -1 & -2 \\ -2 & 2 & 2 \\ -3 & -1 & 3 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & -2 \\ 0 & 2 & 0 \\ -3 & -1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Thus $P^{-1}AP = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ where $P = \begin{bmatrix} -8 & 1 & 1 \\ 10 & -2 & 0 \\ 7 & 1 & 1 \end{bmatrix}$. The general solution is

$$\mathbf{f} = c_1\mathbf{x}_1e^{-x} + c_2\mathbf{x}_2e^{2x} + c_3\mathbf{x}_3e^{4x} = c_1 \begin{bmatrix} -8 \\ 10 \\ 7 \end{bmatrix} e^{-x} + c_2 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{2x} + c_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} e^{4x}$$

That is

$$\begin{aligned} f_1(x) &= -8c_1e^{-x} + c_2e^{2x} + c_3e^{4x} \\ f_2(x) &= 10c_1e^{-x} - 2c_2e^{2x} \\ f_3(x) &= 7c_1e^{-x} + c_2e^{2x} + c_3e^{4x} \end{aligned}$$

If we insist on the boundary conditions $f_1(0) = f_2(0) = f_3(0) = 1$, we get

$$\begin{array}{rcrcrcl} -8c_1 & + & c_2 & + & c_3 & = & 1 \\ 10c_1 & - & 2c_2 & & & = & 1 \\ 2c_1 & + & c_2 & + & c_3 & = & 1 \end{array}$$

The coefficient matrix is P is invertible, so the solution is unique: $c_1 = 0$, $c_2 = -\frac{1}{2}$, $c_3 = \frac{3}{2}$. Hence

$$\begin{aligned} f_1(x) &= \tfrac{1}{2}(3e^{4x} - e^{2x}) \\ f_2(x) &= e^{2x} \\ f_3(x) &= \tfrac{1}{2}(3e^{4x} - e^{2x}) \end{aligned}$$

Note that $f_1(x) = f_3(x)$ happens to hold.

3. b. Have $m'(t) = km(t)$, so $m(t) = ce^{kt}$ by Theorem 3.5.1. Then the requirement that $m(0) = 10$ gives $c = 10$. Also we ask that $m(3) = 8$, whence $10e^{3k} = 8$, $e^{3k} = \frac{4}{5}$. Hence $(e^k)^3 = \frac{4}{5}$, so $(e^k) = (\frac{4}{5})^{1/3}$. Thus $m(t) = 10(\frac{4}{5})^{t/3}$. Now, we want the half-life t_0 satisfying $m(t_0) = \frac{1}{2}m(0)$, that is $10(\frac{4}{5})^{t_0/3} = 5$ so $t_0 = \frac{3\ln(1/2)}{\ln(4/5)} = 9.32$ hours.

5. a. Assume that a $\mathbf{g}' = A\mathbf{g}$ where A is $n \times n$. Put $\mathbf{f} = \mathbf{g} - A^{-1}\mathbf{b}$ where b is a column of constant functions. Then $\mathbf{f}' = \mathbf{g}' = A\mathbf{g} = A(\mathbf{f} + A^{-1}\mathbf{b}) = A\mathbf{f} + \mathbf{b}$, as required.

6. b. Assume that $f_1' = a_1f_1 + f_2$ and $f_2' = a_2f_1$. Differentiating gives $f_1'' = a_1f_1' + f_2' = a_1f_1' + a_2f_1$. This shows that f_1 satisfies $(*)$.

3.6 Proof of the Cofactor Expansion Theorem

2. Consider the rows $R_p, R_{p+1}, \dots, R_{q-1}, R_q$. Using adjacent interchanges we have

$$\begin{bmatrix} R_p \\ R_{p+1} \\ \vdots \\ R_{q-1} \\ R_q \end{bmatrix} \xrightarrow[\text{interchanges}]{q-p} \begin{bmatrix} R_{p+1} \\ \vdots \\ R_{q-1} \\ R_q \\ R_p \end{bmatrix} \xrightarrow[\text{interchanges}]{q-p-1} \begin{bmatrix} R_q \\ R_{p+1} \\ \vdots \\ R_{q-1} \\ R_p \end{bmatrix}$$

Hence $2(q-p)-1$ interchanges are used in all.

Supplementary Exercises: Chapter 3

2. b. Proceed by induction on n where A is $n \times n$. If $n = 1$, $A^T = A$. In general, induction and (a) give

$$\det[A_{ij}] = \det\left[\left(A_{ij}\right)^T\right] = \det[(A^T)_{ij}]$$

Write $A^T = [a'_{ij}]$ where $a'_{ij} = a_{ji}$, and expand $\det(A^T)$ along column 1:

$$\det(A^T) = \sum_{j=1}^{n} a'_{j1}(-1)^{j+1}\det[(A^T)_{j1}] = \sum_{j=1}^{n} a_{1j}(-1)^{1+j}\det[A_{1j}] = \det A$$

where the last equality is the expansion of $\det A$ along row 1.

4. Vector Geometry

4.1 Vectors and Lines

1. b. $\left\| \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right\| = \sqrt{1^2+(-1)^2+2^2} = \sqrt{6}$

   d. $\left\| \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\| = \sqrt{(-1)^2+0^2+2^2} = \sqrt{5}$

   f. $\left\| -3 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\| = |-3| \sqrt{1^2+1^2+2^2} = 3\sqrt{6}$

2. b. A vector $\mathbf{u}$ in the direction of $\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$ must have the form $\mathbf{u} = t \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$ for a scalar $t > 0$. Since $\mathbf{u}$ is a unit vector, we want $\|\mathbf{u}\| = 1$; that is $1 = |t| \sqrt{(-2)^2+(-1)^2+2^2} = 3t$, which gives $t = \frac{1}{3}$. Hence $\mathbf{u} = \frac{1}{3} \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$.

4. b. Write $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$. The distance between $\mathbf{u}$ and $\mathbf{v}$ is the length of their difference: $\|\mathbf{u}-\mathbf{v}\| = \left\| \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\| = \sqrt{0^2+(-1)^2+1^2} = \sqrt{2}$.

   d. As in (b), the distance is $\left\| \begin{bmatrix} 4 \\ 0 \\ -2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} \right\| = \sqrt{1^2+(-2)^2+(-2)^2} = 3$.

6. b. In the diagram, let E and F be the midpoints of sides BC and AC respectively. Then $\overrightarrow{FC} = \frac{1}{2}\overrightarrow{AC}$ and $\overrightarrow{CE} = \frac{1}{2}\overrightarrow{CB}$. Hence

$$\overrightarrow{FE} = \overrightarrow{FC} + \overrightarrow{CE} = \tfrac{1}{2}\overrightarrow{AC} + \tfrac{1}{2}\overrightarrow{CB} = \tfrac{1}{2}(\overrightarrow{AC} + \overrightarrow{CB}) = \tfrac{1}{2}\overrightarrow{AB}$$

7. Two nonzero vectors are parallel if and only if one is a scalar multiple of the other.

   b. Yes, they are parallel: $\mathbf{u} = (-3)\mathbf{v}$.

   d. Yes, they are parallel: $\mathbf{v} = (-4)\mathbf{u}$.

8. b. $\overrightarrow{QR} = \mathbf{p}$ because $OPQR$ is a parallelogram (where O is the origin).

   d. $\overrightarrow{RO} = -(\mathbf{p}+\mathbf{q})$ because $\overrightarrow{OR} = \mathbf{p}+\mathbf{q}$.

9. b. $\overrightarrow{PQ} = \begin{bmatrix} 1 \\ -1 \\ 6 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix}$, so $\left\| \overrightarrow{PQ} \right\| = \sqrt{(-1)^2 + (-1)^2 + 5^2} = \sqrt{27} = 3\sqrt{3}$.

d. Here $P = Q$ are equal points, so $\overrightarrow{PQ} = \mathbf{0}$. Hence $\left\| \overrightarrow{PQ} \right\| = 0$.

f. $\overrightarrow{PQ} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} - \begin{bmatrix} 3 \\ -1 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$. Hence $\left\| \overrightarrow{PQ} \right\| = |2| \sqrt{(-1)^2 + 1^2 + (-1)^2} = 2\sqrt{3}$.

10. b. Given $Q(x,\ y,\ z)$ let $\mathbf{q} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $\mathbf{p} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$ be the vectors of Q and P. Then $\overrightarrow{PQ} = \mathbf{q} - \mathbf{p}$. Let $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

(i) If $\overrightarrow{PQ} = \mathbf{v}$ then $\mathbf{q} - \mathbf{p} = \mathbf{v}$, so $\mathbf{q} = \mathbf{p} + \mathbf{v} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$. Thus $Q = Q(5,\ -1,\ 2)$.

(ii) If $\overrightarrow{PQ} = -\mathbf{v}$ then $\mathbf{q} - \mathbf{p} = -\mathbf{v}$, so $\mathbf{q} = \mathbf{p} - \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}$. Thus $Q = Q(1,\ 1,\ -4)$.

11. b. If $2(3\mathbf{v} - \mathbf{x}) = 5\mathbf{w} + \mathbf{u} - 3\mathbf{x}$ then $6\mathbf{v} - 2\mathbf{x} = 5\mathbf{w} + \mathbf{u} - 3\mathbf{x}$, so

$$\mathbf{x} = 5\mathbf{w} + \mathbf{u} - 6\mathbf{v} = \begin{bmatrix} -5 \\ 5 \\ 25 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 24 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} -26 \\ 4 \\ 19 \end{bmatrix}$$

12. b. We have $a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \begin{bmatrix} a \\ a \\ 2a \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ 2b \end{bmatrix} + \begin{bmatrix} c \\ 0 \\ -c \end{bmatrix} = \begin{bmatrix} a+c \\ a+b \\ 2a+2b-c \end{bmatrix}$. Hence setting $a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$ gives equations

$$\begin{array}{rcrcrcr} a & & & + & c & = & 1 \\ a & + & b & & & = & 3 \\ 2a & + & 2b & - & c & = & 0 \end{array}$$

The solution is $a = -5$, $b = 8$, $c = 6$.

13. b. Suppose $\begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} = a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \begin{bmatrix} 3a+4b+c \\ -a+c \\ b+c \end{bmatrix}$. Equating coefficients gives linear equations for a, b, c:

$$\begin{array}{rcrcrcr} 3a & + & 4b & + & c & = & 5 \\ -a & & & + & c & = & 6 \\ & & b & + & c & = & -1 \end{array}$$

This system has no solution, so no such a, b, c exist.

14. b. Write $P = P(x,\ y,\ z)$ and let $\mathbf{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ and $\mathbf{p}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ be the vectors of P, P_1 and P_2 respectively. Then

$$\mathbf{p} - \mathbf{p}_2 = \overrightarrow{P_2P} = \mathbf{p}_2 + \tfrac{1}{4}(\overrightarrow{P_2P_1}) = \mathbf{p}_2 + \tfrac{1}{4}(\mathbf{p}_1 - \mathbf{p}_2) = \tfrac{1}{4}\mathbf{p}_1 + \tfrac{3}{4}\mathbf{p}_2$$

Since $\mathbf{p}_1$ and $\mathbf{p}_2$ are known, this gives

$$\mathbf{p} = \frac{1}{4}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} + \frac{3}{4}\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 5 \\ -5 \\ -2 \end{bmatrix}$$

Hence $P = P\left(\frac{5}{4},\ -\frac{5}{4},\ -\frac{1}{2}\right)$.

17. b. Let $\mathbf{p} = \overrightarrow{OP}$ and $\mathbf{q} = \overrightarrow{OQ}$ denote the vectors of the points P and Q respectively. Then $\mathbf{q} - \mathbf{p} = \overrightarrow{PQ} = \begin{bmatrix} -1 \\ 4 \\ 7 \end{bmatrix}$ and $\mathbf{p} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$, so $\mathbf{q} = (\mathbf{q}-\mathbf{p}) + \mathbf{p} = \begin{bmatrix} -1 \\ 4 \\ 7 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 3 \end{bmatrix}$.
Hence $Q = Q(0,\ 7,\ 3)$.

18. b. We have $\|\mathbf{u}\|^2 = 20$, so the given equation is $3\mathbf{u} + 7\mathbf{v} = 20(2\mathbf{x} + \mathbf{v})$. Solving for $\mathbf{x}$ gives

$$40\mathbf{x} = 3\mathbf{u} - 13\mathbf{v} = \begin{bmatrix} 6 \\ 0 \\ -12 \end{bmatrix} - \begin{bmatrix} 26 \\ 13 \\ -26 \end{bmatrix} = \begin{bmatrix} -20 \\ -13 \\ 14 \end{bmatrix}$$

Hence $\mathbf{x} = \frac{1}{40}\begin{bmatrix} -20 \\ -13 \\ 14 \end{bmatrix}$.

20. b. Let S denote the fourth point. We have $\overrightarrow{RS} = \overrightarrow{PQ}$, so

$$\overrightarrow{OS} = \overrightarrow{OR} + \overrightarrow{RS} = \overrightarrow{OR} + \overrightarrow{PQ} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$$

Hence $S = S(-1,\ 3,\ 2)$.

21. b. True. If $\|\mathbf{v} - \mathbf{w}\| = 0$ then $\mathbf{v} - \mathbf{w} = \mathbf{0}$ by Theorem 4.1.1, so $\mathbf{v} = \mathbf{w}$.
d. False. $\|\mathbf{v}\| = \|-\mathbf{v}\|$ for all $\mathbf{v}$ but $\mathbf{v} = -\mathbf{v}$ only holds if $\mathbf{v} = \mathbf{0}$.
f. False. If $t < 0$ they have opposite directions.
h. False. By Theorem 4.1.1, $\|-5\mathbf{v}\| = |-5|\,\|\mathbf{v}\| = 5\,\|\mathbf{v}\|$ so it fails if $\mathbf{v} \neq \mathbf{0}$.
j. False. If $\mathbf{w} = -\mathbf{v}$ where $\mathbf{v} \neq \mathbf{0}$, then $\|\mathbf{v} + \mathbf{w}\| = \mathbf{0}$ but $\|\mathbf{v}\| + \|\mathbf{w}\| = 2\,\|\mathbf{v}\| \neq \mathbf{0}$.

22. b. One direction vector is $\mathbf{d} = \overrightarrow{QP} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$. Let $\mathbf{p}_0 = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}$ be the vector of P. Then the vector equation of the line is

$$\mathbf{p} = \mathbf{p}_0 + t\mathbf{d} = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} + t\begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix} \text{ when } \mathbf{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

is the vector of an arbitrary point on the line. Equating coefficients gives the parametric equations of the line

$$\begin{aligned} x &= 3 + 2t \\ y &= -1 - t \\ z &= 4 + 5t \end{aligned}$$

d. Now $\mathbf{p}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ because $P_1(1,\ 1,\ 1)$ is on the line, and take $\mathbf{d} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ because the line is to be parallel to $\mathbf{d}$. Hence the vector equation is $\mathbf{p} = \mathbf{p}_0 + t\mathbf{d} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Taking $\mathbf{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, the scalar equations are $\begin{array}{l} x = 1+t \\ y = 1+t \\ z = 1+t \end{array}$.

f. The line with parametric equations $\begin{array}{l} x = 2-t \\ y = 1 \\ z = t \end{array}$ has direction vector $\mathbf{d} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ — the components are the coefficients of t. Since our line is parallel to this one, $\mathbf{d}$ will do as direction vector. We are given the vector $\mathbf{p}_0 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ of a point on the line, so the vector equation is

$$\mathbf{p} = \mathbf{p}_0 + t\mathbf{d} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

The scalar equations are

$$\begin{aligned} x &= 2-t \\ y &= -1 \\ z &= 1+t \end{aligned}$$

23. b. $P(2,\,3,\,-3)$ lies on the line $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4-t \\ 3 \\ 1-2t \end{bmatrix}$ since it corresponds to $t = 2$. Similarly $Q(-1,\,3,\,-9)$ corresponds to $t = 5$, so Q lies on the line too.

24. b. If $P = P(x,\ y,\ z)$ is a point on both lines then

$$\begin{array}{l} x = 1-t \\ y = 2+2t \\ z = -1+3t \end{array} \quad \text{for some } t \text{ because } P \text{ lies on the first line.}$$

$$\begin{array}{l} x = 2s \\ y = 1+s \\ z = 3 \end{array} \quad \text{for some } s \text{ because } P \text{ lies on the second line.}$$

If we eliminate x, y, and z we get three equations for s and t:

$$\begin{aligned} 1-t &= 2s \\ 2+2t &= 1+s \\ -1+3t &= 3 \end{aligned}$$

The last two equations require $t = \frac{4}{3}$ and $s = \frac{11}{3}$, but these values do *not* satisfy the first equation. Hence no such s and t exist, so the lines do *not* intersect.

d. If $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is the vector of a point on both lines, then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 5 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \text{ for some } t \text{ (first line)}$$
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -7 \\ 12 \end{bmatrix} + s \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix} \text{ for some } s \text{ (second line).}$$

Eliminating $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ gives $\begin{bmatrix} 4 \\ -1 \\ 5 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -7 \\ 12 \end{bmatrix} + s \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}$. Equating coefficients gives three equations for s and t:

$$\begin{aligned} 4+t &= 2 \\ -1 &= -7-2s \\ 5+t &= 12+3s \end{aligned}$$

This has a (unique) solution $t=-2$, $s=-3$ so the lines *do* intersect. The point of intersection has vector

$$\begin{bmatrix} 4 \\ -1 \\ 5 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 5 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

(equivalently $\begin{bmatrix} 2 \\ -7 \\ 12 \end{bmatrix} + s \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ -7 \\ 12 \end{bmatrix} - 3 \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$).

29. Let $\mathbf{a} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ be the vectors of A and B. Then $\mathbf{d} = \mathbf{b} - \mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ is a direction vector for the line through A and B, so the vector $\mathbf{c}$ of C is given by $\mathbf{c} = \mathbf{a} + t\mathbf{d}$ for some t. Then

$$\left\| \overrightarrow{AC} \right\| = \|\mathbf{c}-\mathbf{a}\| = \|t\mathbf{d}\| = |t|\ \|\mathbf{d}\| \quad \text{and} \quad \left\| \overrightarrow{BC} \right\| = \|\mathbf{c}-\mathbf{b}\| = \|(t-1)\mathbf{d}\| = |t-1|\,\|\mathbf{d}\|$$

Hence $\left\| \overrightarrow{AC} \right\| = 2 \left\| \overrightarrow{BC} \right\|$ means $|t| = 2\,|t-1|$, so $t^2 = 4(t-1)^2$, whence $0 = 3t^2 - 8t + 4 = (t-2)(3t-2)$. Thus $t=2$ or $t=\frac{2}{3}$. Since $\mathbf{c} = \mathbf{a} + t\mathbf{d}$, this means $\mathbf{c} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ or $\mathbf{c} = \begin{bmatrix} \frac{5}{3} \\ -\frac{1}{3} \\ \frac{4}{3} \end{bmatrix}$.

31. b. If there are $2n$ points, then P_k and P_{n+k} are opposite ends of a diameter of the circle for each $k = 1, 2, \dots$. Hence $\overrightarrow{CP}_k = -\overrightarrow{CP}_{n+k}$ so these terms cancel in the sum $\overrightarrow{CP}_1 + \overrightarrow{CP}_2 + \cdots + \overrightarrow{CP}_{2n}$. Thus all terms cancel and the sum is $\mathbf{0}$.

33. We have $2\overrightarrow{EA} = \overrightarrow{DA}$ because E is the midpoint of side AD, and $2\overrightarrow{AF} = \overrightarrow{FC}$ because F is $\frac{1}{3}$ the way from A to C. Finally $\overrightarrow{DA} = \overrightarrow{CB}$ because $ABCD$ is a parallelogram. Thus

$$2\overrightarrow{EF} = 2(\overrightarrow{EA} + \overrightarrow{AF}) = 2\overrightarrow{EA} + 2\overrightarrow{AF} = \overrightarrow{DA} + \overrightarrow{FC} = \overrightarrow{CB} + \overrightarrow{FC} = \overrightarrow{FB}$$

Hence $\overrightarrow{EF} = \frac{1}{2}\overrightarrow{FB}$ so F is in the line segment EB, $\frac{1}{3}$ the way from E to B. Hence F is the trisection point of both AC and EB.

4.2 Projections and Planes

1. b. $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{u} = 1^2 + 2^2 + (-1)^2 = 6$

   d. $\mathbf{u} \cdot \mathbf{v} = 3 \cdot 6 + (-1)(-7) + 5(-5) = 18 + 7 - 25 = 0$

   f. $\mathbf{v} = \mathbf{0}$ so $\mathbf{u} \cdot \mathbf{v} = a \cdot 0 + b \cdot 0 + c \cdot 0 = 0$

2. b. $\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{-18-2+0}{\sqrt{10}\sqrt{40}} = \frac{-20}{20} = -1$. Hence $\theta = \pi$.

   d. $\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{6+6-3}{\sqrt{6}(3\sqrt{6})} = \frac{1}{2}$. Hence $\theta = \frac{\pi}{3}$.

   f. $\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{0-21-4}{\sqrt{25}\sqrt{100}} = -\frac{1}{2}$. Hence $\theta = \frac{2\pi}{3}$.

3. b. Writing $\mathbf{u} = \left[\begin{array}{r} 2 \\ -1 \\ 1 \end{array}\right]$ and $\mathbf{v} = \left[\begin{array}{c} 1 \\ x \\ 2 \end{array}\right]$, the requirement is

$$\frac{1}{2} = \cos\frac{\pi}{3} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{2-x+2}{\sqrt{6}\sqrt{x^2+5}}$$

   Hence $6(x^2+5) = 4(4-x)^2$, whence $x^2 + 16x - 17 = 0$. The roots are $x = -17$ and $x = 1$.

4. b. The conditions are $\mathbf{u}_1 \cdot \mathbf{v} = 0$ and $\mathbf{u}_2 \cdot \mathbf{v} = 0$, yielding equations

$$\begin{array}{rcrcrcr} 3x & - & y & + & 2z & = & 0 \\ 2x & & & + & z & = & 0 \end{array}$$

   The solutions are $x = -t$, $y = t$, $z = 2t$, so $\mathbf{v} = t\left[\begin{array}{r} -1 \\ 1 \\ 2 \end{array}\right]$.

   d. The conditions are $\mathbf{u}_1 \cdot \mathbf{v} = 0$ and $\mathbf{u}_2 \cdot \mathbf{v} = 0$, yielding equations

$$\begin{aligned} 2x - y + 3z &= 0 \\ 0 &= 0 \end{aligned}$$

   The solutions are $x = s$, $y = 2s + 3t$, $z = t$, so $\mathbf{v} = s\left[\begin{array}{c} 1 \\ 2 \\ 0 \end{array}\right] + t\left[\begin{array}{c} 0 \\ 3 \\ 1 \end{array}\right]$.

6. b. $\left\|\overrightarrow{PQ}\right\|^2 = \left\|\left[\begin{array}{r} 3 \\ -2 \\ 4 \end{array}\right]\right\|^2 = 9 + 4 + 16 = 29$

   $\left\|\overrightarrow{QR}\right\|^2 = \left\|\left[\begin{array}{c} 2 \\ 7 \\ 2 \end{array}\right]\right\|^2 = 4 + 49 + 4 = 57$

   $\left\|\overrightarrow{PR}\right\|^2 = \left\|\left[\begin{array}{c} 5 \\ 5 \\ 6 \end{array}\right]\right\|^2 = 25 + 25 + 36 = 86$

   Hence $\left\|\overrightarrow{PR}\right\| = \left\|\overrightarrow{PQ}\right\|^2 + \left\|\overrightarrow{QR}\right\|^2$. Note that this *implies* that the triangle is right angled, that PR is the hypotenuse, and hence that the angle at Q is a right angle. Of course, we can confirm this latter fact by computing $\overrightarrow{PQ} \cdot \overrightarrow{QR} = 6 - 14 + 8 = 0$.

8. b. We have $\overrightarrow{AB} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ and $\overrightarrow{AC} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ so the angle α at A is given by

$$\cos\alpha = \frac{\overrightarrow{AB}\cdot\overrightarrow{AC}}{\left\|\overrightarrow{AB}\right\|\left\|\overrightarrow{AC}\right\|} = \frac{2+2-1}{\sqrt{6}\sqrt{6}} = \frac{1}{2}$$

Hence $\alpha = \frac{\pi}{3}$ or $60°$. Next $\overrightarrow{BA} = \begin{bmatrix} -2 \\ -1 \\ -1 \end{bmatrix}$ and $\overrightarrow{BC} = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$ so the angle β at B is given by

$$\cos\beta = \frac{\overrightarrow{BA}\cdot\overrightarrow{BC}}{\left\|\overrightarrow{BA}\right\|\left\|\overrightarrow{BC}\right\|} = \frac{2-1+2}{\sqrt{6}\sqrt{6}} = \frac{1}{2}$$

Hence $\beta = \frac{\pi}{3}$. Since the angles in any triangle add to π, the angle γ at C is $\pi - \frac{\pi}{3} - \frac{\pi}{3} = \frac{\pi}{3}$. However, $\overrightarrow{CA} = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$ and $\overrightarrow{CB} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, this can also be seen directly from

$$\cos\gamma = \frac{\overrightarrow{CA}\cdot\overrightarrow{CB}}{\left\|\overrightarrow{CA}\right\|\left\|\overrightarrow{CB}\right\|} = \frac{-1+2+2}{\sqrt{6}\sqrt{6}} = \frac{1}{2}$$

10. b. $\text{proj}_{\mathbf{v}}\,\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{12-2+1}{16+1+1}\begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} = \frac{11}{18}\begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$

d. $\text{proj}_{\mathbf{v}}\,\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{-18-8-2}{36+16+4}\begin{bmatrix} -6 \\ 4 \\ 2 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} -6 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}$

11. b. Take $\mathbf{u}_1 = \text{proj}_{\mathbf{v}}\,\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{-6+1+0}{4+1+16}\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} = \frac{-5}{21}\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}$. Then $\mathbf{u}_2 = \mathbf{u} - \mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \frac{5}{21}\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} = \frac{1}{21}\begin{bmatrix} 53 \\ 26 \\ 20 \end{bmatrix}$. As a check, verify that $\mathbf{u}_2\cdot\mathbf{v} = 0$, that is $\mathbf{u}_2$ is orthogonal to $\mathbf{v}$.

d. Take $\mathbf{u}_1 = \text{proj}_{\mathbf{v}}\,\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{-18-8-1}{36+16+1}\begin{bmatrix} -6 \\ 4 \\ -1 \end{bmatrix} = \frac{27}{53}\begin{bmatrix} 6 \\ -4 \\ 1 \end{bmatrix}$. Then $\mathbf{u}_2$ is given by $\mathbf{u}_2 = \mathbf{u} - \mathbf{u}_1 = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} - \frac{27}{53}\begin{bmatrix} 6 \\ -4 \\ 1 \end{bmatrix} = \frac{1}{53}\begin{bmatrix} -3 \\ 2 \\ 26 \end{bmatrix}$. As a check, verify that $\mathbf{u}_2\cdot\mathbf{v} = \mathbf{0}$, that is $\mathbf{u}_2$ is orthogonal to $\mathbf{v}$.

12. b. Write $\mathbf{p}_0 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{d} = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ and write $\mathbf{u} = \overrightarrow{P_0P} = \mathbf{p} - \mathbf{p}_0 = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$. Write $\mathbf{u}_1 = \overrightarrow{P_0Q}$ and compute it as $\mathbf{u}_1 = \text{proj}_{\mathbf{d}}\,\overrightarrow{P_0P} = \frac{0-1+16}{9+1+16}\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} = \frac{15}{26}\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$. Then the distance from P to the line is $\left\|\overrightarrow{QP}\right\| = \|\mathbf{u} - \mathbf{u}_1\| = \left\|\frac{1}{26}\begin{bmatrix} -45 \\ -41 \\ 44 \end{bmatrix}\right\| = \frac{1}{26}\sqrt{5642}$. To compute Q let $\mathbf{q}$ be its vector. Then

$$\mathbf{q} = \mathbf{p}_0 + \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{15}{26}\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} = \frac{1}{26}\begin{bmatrix} 71 \\ 15 \\ 34 \end{bmatrix}$$

Hence $Q = Q(\frac{71}{26}, \frac{15}{26}, \frac{34}{26})$.

13. b. $\mathbf{u}\times\mathbf{v} = \det\begin{bmatrix} \mathbf{i} & 3 & -6 \\ \mathbf{j} & -1 & 2 \\ \mathbf{k} & 0 & 0 \end{bmatrix} = 0\mathbf{i} - 0\mathbf{j} + 0\mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$

d. $\mathbf{u}\times\mathbf{v} = \det\begin{bmatrix} \mathbf{i} & 2 & 1 \\ \mathbf{j} & 0 & 4 \\ \mathbf{k} & -1 & 7 \end{bmatrix} = 4\mathbf{i} - 15\mathbf{j} + 8\mathbf{k} = \begin{bmatrix} 4 \\ -15 \\ 8 \end{bmatrix}$

14. b. A normal is $\mathbf{n} = \overrightarrow{AB}\times\overrightarrow{AC} = \begin{bmatrix} -1 \\ 1 \\ -5 \end{bmatrix}\times\begin{bmatrix} 3 \\ 8 \\ -17 \end{bmatrix} = \det\begin{pmatrix} \mathbf{i} & -1 & 3 \\ \mathbf{j} & 1 & 8 \\ \mathbf{k} & -5 & -17 \end{pmatrix} = \begin{bmatrix} 23 \\ -32 \\ -11 \end{bmatrix}$.

Since the plane passes through $B(0,\ 0,\ 1)$ the equation is

$$23(x-0) - 32(y-0) - 11(z-1) = 0, \text{ that is } -23x + 32y + 11z = 11$$

d. The plane with equation $2x - y + z = 3$ has normal $\mathbf{n} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$. Since our plane is parallel to this one, $\mathbf{n}$ will serve as normal. The point $P(3,\ 0,\ -1)$ lies on our plane, the equation is $2(x-3) - (y-0) + (z-(-1) = 0$, that is $2x - y + z = 5$.

f. The plane contains $P(2,\ 1,\ 0)$ and $P_0(3,\ -1,\ 2)$, so the vector $\mathbf{u} = \overrightarrow{PP_0} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$ is parallel to the plane. Also the direction vector $\mathbf{d} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ of the line is parallel to the plane. Hence $\mathbf{n} = \mathbf{u}\times\mathbf{d} = \det\begin{bmatrix} \mathbf{i} & 1 & 1 \\ \mathbf{j} & -2 & 0 \\ \mathbf{k} & 2 & -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}$ is perpendicular to the plane and so serves as a normal. As $P(2,\ 1,\ 0)$ is in the plane, the equation is

$$2(x-2) + 3(y-1) + 2(z-0) = 0, \text{ that is } 2x + 3y + 2z = 7$$

h. The two direction vectors $\mathbf{d}_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ and $\mathbf{d}_2 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ are parallel to the plane, so $\mathbf{n} = \mathbf{d}_1\times\mathbf{d}_2 = \det\begin{bmatrix} \mathbf{i} & 1 & 2 \\ \mathbf{j} & -1 & 1 \\ \mathbf{k} & 3 & -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \\ 3 \end{bmatrix}$ will serve as normal. The plane contains $P(3,\ 1,\ 0)$ so the equation is

$$-2(x-3) + 7(y-1) + 3(z-0) = 0, \text{ that is } -2x + 7y + 3z = 1$$

Note that this plane contains the line $\begin{bmatrix} x \\ a \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ by construction; it contains the *other* line because it contains $P(0,\ -2,\ 5)$ and is parallel to $\mathbf{d}_2$. This implies that the lines intersect (both are in the same plane). In fact the point of intersection is $P(4,\ 0,\ 3)$ [$t = 1$ on the first line and $t = 2$ on the second line].

j. The set of all points $R(x,\ y,\ z)$ equidistant from both $P(0,\ 1,\ -1)$ and $Q(2,\ -1,\ -3)$ is determined as follows: The condition is $\left\|\overrightarrow{PR}\right\| = \left\|\overrightarrow{QR}\right\|$, that is $\left\|\overrightarrow{PR}\right\|^2 = \left\|\overrightarrow{QR}\right\|^2$, that is

$$x^2 + (y-1)^2 + (z+1)^2 = (x-2)^2 + (y+1)^2 + (z+3)^2$$

This simplifies to $x^2+y^2+z^2-2y+2z+2 = x^2+y^2+z^2-4x+2y+6z+14$; that is $4x-4y-4z = 12$; that is $x-y-z = 3$.

15. b. The normal $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ to the given plane will serve as direction vector for the line. Since the line passes through $P(2,\ -1,\ 3)$, the vector equation is $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + t\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

d. The given lines have direction vectors $\mathbf{d}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$ and $\mathbf{d}_2 = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$, so

$$\mathbf{d} = \mathbf{d}_1 \times \mathbf{d}_2 = \det \begin{bmatrix} \mathbf{i} & 1 & 1 \\ \mathbf{j} & 1 & 2 \\ \mathbf{k} & -2 & -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ is perpendicular to both lines.}$$

Hence $\mathbf{d}$ is a direction vector for the line we seek. As $P(1,\ 1,\ -1)$ is on the line, the vector equation is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

f. Each point on the given line has the form $Q(2+t,\ 1+t,\ t)$ for some t. So $\overrightarrow{PQ} = \begin{bmatrix} 1+t \\ t \\ t-2 \end{bmatrix}$. This is perpendicular to the given line if $\overrightarrow{PQ} \cdot \mathbf{d} = 0$ (where $\mathbf{d} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is the direction vector of the given line). This condition is $(1+t)+t+(t-2) = 0$, that is $t = \frac{1}{3}$. Hence the line we want has direction vector $\begin{bmatrix} \frac{4}{3} \\ \frac{1}{3} \\ -\frac{5}{3} \end{bmatrix}$. For convenience we use $\mathbf{d} = \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$. As the line we want passes through $P(1,\ 1,\ 2)$, the vector equation is $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + t \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$. [Note that $Q\left(\frac{7}{3},\ \frac{4}{3},\ \frac{1}{3}\right)$ is the point of intersection of the two lines.]

16. b. Choose a point P_0 in the plane, say $P_0(0,\ 6,\ 0)$, and write $\mathbf{u} = \overrightarrow{P_0P} = \begin{bmatrix} 3 \\ -5 \\ -1 \end{bmatrix}$. Now write $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ for the normal to the plane. Compute

$$\mathbf{u}_1 = \operatorname{proj}_{\mathbf{n}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \mathbf{n} = \frac{2}{6} \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

The distance from P to the plane is $\|\mathbf{u}_1\| = \frac{1}{3}\sqrt{6}$.
Since $\mathbf{p}_0 = \begin{bmatrix} 0 \\ 6 \\ 0 \end{bmatrix}$ and $\mathbf{q}$ are the vectors of P_0 and Q, we get

$$\mathbf{q} = \mathbf{p}_0 + (\mathbf{u} - \mathbf{u}_1) = \begin{bmatrix} 0 \\ 6 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -5 \\ -1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 7 \\ 2 \\ -2 \end{bmatrix}$$

Hence $Q = Q\left(\frac{7}{3},\ \frac{2}{3},\ \frac{-2}{3}\right)$.

17. b. A normal to the plane is given by

$$\mathbf{n} = \overrightarrow{PQ} \times \overrightarrow{PR} = \begin{bmatrix} -2 \\ 2 \\ -4 \end{bmatrix} \times \begin{bmatrix} -3 \\ -1 \\ -3 \end{bmatrix} = \det \begin{bmatrix} \mathbf{i} & -2 & -3 \\ \mathbf{j} & 2 & -1 \\ \mathbf{k} & -4 & -3 \end{bmatrix} = \begin{bmatrix} -10 \\ 6 \\ 8 \end{bmatrix}$$

Thus, as $P(4,\ 0,\ 5)$ is in the plane, the equation is

$$-10(x-4) + 6(y-0) + 8(z-5) = 0; \text{ that is } 5x - 3y - 4z = 0.$$

The plane contains the origin $P(0,\ 0,\ 0)$.

19. b. The coordinates of points of intersection satisfy both equations:

$$\begin{aligned} 3x+y-2z &= 1 \\ x+y+z &= 5 \end{aligned}$$

Solve

$$\left[\begin{array}{rrr|r} 3 & 1 & -2 & 1 \\ 1 & 1 & 1 & 5 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & 1 & 5 \\ 0 & -2 & -5 & -14 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & -\frac{3}{2} & -2 \\ 0 & 1 & \frac{5}{2} & 7 \end{array}\right]$$

Take $z = 2t$, to eliminate fractions, whence $x = -2+3t$ and $y = 7-5t$. Thus

$$\left[\begin{array}{c} x \\ y \\ z \end{array}\right] = \left[\begin{array}{c} -2+3t \\ 7-5t \\ 2t \end{array}\right] = \left[\begin{array}{r} -2 \\ 7 \\ 0 \end{array}\right] = +t\left[\begin{array}{r} 3 \\ -5 \\ 2 \end{array}\right]$$

is the line of intersection.

20. b. If $P(x,\ y,\ z)$ is an intersection point, then $x = 1+2t$, $y = -2+5t$, $z = 3-t$ since P is on the line. Substitution in the equation of the plane gives $2(1+2t)-(-2+5t)-(3-t) = 5$, that is $1 = 5$. Thus there is no such t, so the line does not intersect the plane.

d. If $P(x,\ y,\ z)$ is an intersection point, then $x = 1+2t$, $y = -2+5t$ and $z = 3-t$ since P is on the line. Substitution in the equation of the plane gives $-1(1+2t)-4(-2+5t)-3(3-t) = 6$, whence $t = \frac{-8}{19}$. Thus $\left[\begin{array}{c} x \\ y \\ z \end{array}\right] = \left[\begin{array}{c} \frac{3}{19} \\ -\frac{78}{19} \\ \frac{65}{19} \end{array}\right]$ so $P\left(\frac{3}{19},\ -\frac{78}{19},\ \frac{65}{19}\right)$ is the point of intersection.

21. b. The line has direction vector $\mathbf{d} = \left[\begin{array}{c} 3 \\ 0 \\ 2 \end{array}\right]$ which is a normal to all such planes. If $P_0(x_0,\ y_0,\ z_0)$ is any point, the plane $3(x-x_0) = 0(y-y_0)+2(z-z_0) = 0$ is perpendicular to the line. This can be written $3x+2z = 3x_0+2z_0$, so $3x+2z = d$, d arbitrary.

d. If the normal is $\mathbf{n} = \left[\begin{array}{c} a \\ b \\ c \end{array}\right] \neq \mathbf{0}$, the plane is $a(x-3)+b(y-2)+c(z+4) = 0$, where a, b and c are not all zero.

f. The vector $\mathbf{u} = \overrightarrow{PQ} = \left[\begin{array}{r} -1 \\ 1 \\ -1 \end{array}\right]$ is parallel to these planes so the normal $\mathbf{n} = \left[\begin{array}{c} a \\ b \\ c \end{array}\right]$ is orthogonal to $\mathbf{u}$. Thus $0 = \mathbf{u} \cdot \mathbf{n} = -a+b-c$. Hence $c = b-a$ and $\mathbf{n} = \left[\begin{array}{c} a \\ b \\ b-a \end{array}\right]$. The plane passes through $Q(1,\ 0,\ 0)$ so the equation is $a(x-1)+b(y-0)+(b-a)(z-0) = 0$, that is $ax+by+(b-a)z = a$. Here a and b are not both zero (as $\mathbf{n} \neq \mathbf{0}$). As a check, observe that this plane contains $P(2,\ -1,\ 1)$ and $Q(1,\ 0,\ 0)$.

h. Such a plane contains $P_0(3,\ 0,\ 2)$ and its normal $\mathbf{n} = \left[\begin{array}{c} a \\ b \\ c \end{array}\right]$ must be orthogonal to the direction vector $\mathbf{d} = \left[\begin{array}{r} 1 \\ -2 \\ -1 \end{array}\right]$ of the line. Thus $0 = \mathbf{d} \cdot \mathbf{n} = a-2b-c$, whence $c = a-2b$ and $\mathbf{n} = \left[\begin{array}{c} a \\ b \\ a-2b \end{array}\right]$ (where a and b are not both zero as $\mathbf{n} \neq \mathbf{0}$). Thus the equation is

$$a(x-3)+b(y-0)+(a-2b)(z-2) = 0, \text{ that is } ax+by+(a-2b)z = 5a-4b$$

where a and b are not both zero. As a check, observe that the plane contains every point $P(3+t,\ -2t,\ 2-t)$ on the line.

23. b. Choose $P_1(3,\ 0,\ 2)$ on the first line. The distance in question is the distance from P_1 to the second line. Choose $P_2(-1,\ 2,\ 2)$ on the second line and let $\mathbf{u} = \overrightarrow{P_2P_1} = \left[\begin{array}{r} 4 \\ -2 \\ 0 \end{array}\right]$. If $\mathbf{d} = \left[\begin{array}{r} 3 \\ 1 \\ 0 \end{array}\right]$ is the direction vector for the line, compute

$$\mathbf{u}_1 = \operatorname{proj}_{\mathbf{d}} \mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{d}}{\|\mathbf{d}\|^2}\mathbf{d} = \frac{10}{10}\left[\begin{array}{rrr} 3 & 1 & 0 \end{array}\right]^T = \left[\begin{array}{rrr} 3 & 1 & 0 \end{array}\right]^T$$

then the required distance is $\|\mathbf{u}-\mathbf{u}_1\| = \left\|\left[\begin{array}{rrr} 3 & 1 & 0 \end{array}\right]^T\right\| = \sqrt{10}$.

24. b. The cross product $\mathbf{n} = \left[\begin{array}{r} 1 \\ 1 \\ 1 \end{array}\right] \times \left[\begin{array}{r} 3 \\ 1 \\ 0 \end{array}\right] = \left[\begin{array}{r} -1 \\ 3 \\ -2 \end{array}\right]$ of the two direction vectors serves as a normal to the plane. Given $P_1(1,\ -1,\ 0)$ and $P_2(-2,\ -1,\ 3)$ on the lines, let $\mathbf{u} = \overrightarrow{P_1P_2} = \left[\begin{array}{r} 1 \\ 0 \\ 3 \end{array}\right]$. Compute

$$\mathbf{u}_1 = \operatorname{proj}_{\mathbf{n}} \mathbf{u} = \frac{-7}{14}\left[\begin{array}{r} -1 \\ 3 \\ -2 \end{array}\right] = \frac{1}{2}\left[\begin{array}{r} 1 \\ -3 \\ 2 \end{array}\right]$$

The required distance is $\|\mathbf{u}_1\| = \frac{1}{2}\sqrt{1+9+4} = \frac{1}{2}\sqrt{14}$.
Now let $A = A(1+s,\ -1+s,\ s)$ and $B = B(2+3t,\ -1+t,\ 3)$ be the points on the two lines that are closest together. Then $\overrightarrow{AB} = \left[\begin{array}{c} 1+3t-s \\ t-s \\ 3-s \end{array}\right]$ is orthogonal to both direction vectors $\mathbf{d}_1 = \left[\begin{array}{r} 1 \\ 1 \\ 1 \end{array}\right]$ and $\mathbf{d}_2 = \left[\begin{array}{r} 3 \\ 1 \\ 0 \end{array}\right]$. By Theorem 4.2.3 this means $\mathbf{d}_1 \cdot \overrightarrow{AB} = 0 = \mathbf{d}_2 \cdot \overrightarrow{AB}$, giving equations $4t-3s=-4$, $10t-4s=-3$. The solution is $t=\frac{1}{2}$, $s=2$, so the points are $A = A(3,\ 1,\ 2)$ and $B = B\left(\frac{7}{2},\ -\frac{1}{2},\ 3\right)$.

d. Analogous to (b). The distance is $\frac{\sqrt{6}}{6}$, and the points are $A(\frac{19}{3},\ 2,\ \frac{1}{3})$ and $B = B\left(\frac{37}{6},\ \frac{13}{6},\ 0\right)$.

26. b. Position the cube with one vertex at the origin and sides along the positive axes. Assume each side has length a and consider the diagonal with direction $\mathbf{d} = \left[\begin{array}{c} a \\ a \\ a \end{array}\right]$. The face diagonals that do not meet $\mathbf{d}$ are: $\pm\left[\begin{array}{r} a \\ -a \\ 0 \end{array}\right]$, $\pm\left[\begin{array}{r} a \\ 0 \\ -a \end{array}\right]$ and $\pm\left[\begin{array}{r} 0 \\ a \\ -a \end{array}\right]$, and all are orthogonal to d (the dot product is 0).

28. Position the solid with one vertex at the origin and sides, of lengths a, b, c, along the positive x, y and z axes respectively. The diagonals are $\pm\left[\begin{array}{r} a \\ b \\ c \end{array}\right]$, $\pm\left[\begin{array}{r} -a \\ b \\ c \end{array}\right]$, $\pm\left[\begin{array}{r} a \\ -b \\ c \end{array}\right]$ and $\pm\left[\begin{array}{r} a \\ b \\ -c \end{array}\right]$. The possible dot products are $\pm(-a^2+b^2+c^2)$, $\pm(a^2-b^2+c^2)$, $\pm(a^2+b^2-c^2)$ and one of these is zero if and only if the sum of two of a^2, b^2 and c^2 equals the third.

34. b. The sum of the squares of the lengths of the diagonals equals the sum of the squares of the lengths of the four sides.

38. b. The angle θ between $\mathbf{u}$ and $\mathbf{u}+\mathbf{v}+\mathbf{w}$ is given by

$$\cos\theta = \frac{\mathbf{u}\cdot(\mathbf{u}+\mathbf{v}+\mathbf{w})}{\|\mathbf{u}\|\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|} = \frac{\mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{u}\|\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|} = \frac{\|\mathbf{u}\|^2+0+0}{\|\mathbf{u}\|\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|} = \frac{\|\mathbf{u}\|}{\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|}$$

Similarly the angles φ, ψ between $\mathbf{v}$ and $\mathbf{w}$ and $\mathbf{u}+\mathbf{v}+\mathbf{w}$ are given by

$$\cos\varphi = \frac{\|\mathbf{v}\|}{\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|} \quad \text{and} \quad \cos\psi = \frac{\|\mathbf{w}\|}{\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|}$$

Since $\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\|$ we get $\cos\theta = \cos\varphi = \cos\psi$, whence $\theta = \varphi = \psi$.

NOTE: $\|\mathbf{u}+\mathbf{v}+\mathbf{w}\| = \sqrt{\|\mathbf{u}\|^2+\|\mathbf{v}\|^2+\|\mathbf{w}\|^2} = \|\mathbf{u}\|\sqrt{3}$ by part (a), so $\cos\theta = \cos\varphi = \cos\psi = \frac{1}{\sqrt{3}}$. Thus, in fact $\theta = \varphi = \psi = .955$ radians, (54.7°).

39. b. If $P_1(x, y)$ is on the line then $ax+by+c=0$. Hence $\mathbf{u} = \overrightarrow{P_1P_0} = \begin{bmatrix} x_0-x_1 \\ y_0-y_1 \end{bmatrix}$ so the distance is

$$\|\operatorname{proj}_{\mathbf{n}}\mathbf{u}\| = \left\|\frac{\mathbf{u}\cdot\mathbf{n}}{\|\mathbf{n}\|^2}\mathbf{n}\right\| = \frac{|\mathbf{u}\cdot\mathbf{n}|}{\|\mathbf{n}\|} = \frac{|a(x_0-x)+b(y_0-y)|}{\sqrt{a^2+b^2}} = \frac{|ax_0+by_0+c|}{\sqrt{a^2+b^2}}$$

41. b. This follows from (a) because $\|\mathbf{v}\|^2 = a^2+b^2+c^2$.

44. d. Take $x_1 = z_2 = x$, $y_1 = x_2 = y$ and $z_1 = y_2 = z$ in (c).

4.3 More on the Cross Product

3. b. One vector orthogonal to $\mathbf{u}$ and $\mathbf{v}$ is $\mathbf{u}\times\mathbf{v} = \det\begin{bmatrix} \mathbf{i} & 1 & 3 \\ \mathbf{j} & 2 & 1 \\ \mathbf{k} & -1 & 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \\ -5 \end{bmatrix}$. We have $\|\mathbf{u}\times\mathbf{v}\| = 5\left\|\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}\right\| = 5\sqrt{3}$. Hence the unit vectors parallel to $\mathbf{u}\times\mathbf{v}$ are $\pm\frac{1}{5\sqrt{3}}\begin{bmatrix} 5 \\ -5 \\ -5 \end{bmatrix} = \pm\frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$.

4. b. The area of the triangle is $\frac{1}{2}$ the area of the parallelogram $ABCD$. By Theorem 4.3.4,

$$\text{Area of triangle} = \tfrac{1}{2}\left\|\overrightarrow{AB}\times\overrightarrow{AC}\right\| = \tfrac{1}{2}\left\|\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}\times\begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix}\right\| = \tfrac{1}{2}\left\|\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\right\| = 0$$

Hence $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are parallel.

d. Analogous to (b). Area $= \sqrt{5}$.

5. b. We have $\mathbf{u}\times\mathbf{v} = \begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix}$ so $\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) = -7$. The volume is $|\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})| = |-7| = 7$ by Theorem 4.3.5.

6. b. The line through P_0 perpendicular to the plane has direction vector $\mathbf{n}$, and so has vector equation $\mathbf{p} = \mathbf{p}_0 + t\mathbf{n}$ where $\mathbf{p} = \begin{bmatrix} x & y & z \end{bmatrix}^T$. If $P(x, y, z)$ also lies in the plane, then $\mathbf{n}\cdot\mathbf{p} = ax+by+cz = d$. Using $\mathbf{p} = \mathbf{p}_0 + t\mathbf{n}$ we find

$$d = \mathbf{n}\cdot\mathbf{p} = \mathbf{n}\cdot\mathbf{p}_0 + t(\mathbf{n}\cdot\mathbf{n}) = \mathbf{n}\cdot\mathbf{p}_0 + t\|\mathbf{n}\|^2$$

Hence $t = \frac{d-\mathbf{n}\cdot\mathbf{p}_0}{\|\mathbf{n}\|^2}$, so $\mathbf{p} = \mathbf{p}_0 + \left(\frac{d-\mathbf{n}\cdot\mathbf{p}_0}{\|\mathbf{n}\|^2}\right)\mathbf{n}$. Finally, the distance from P_0 to the plane is

$$\left\|\overrightarrow{PP_0}\right\| = \|\mathbf{p}-\mathbf{p}_0\| = \left\|\left(\frac{d-\mathbf{n}\cdot\mathbf{p}_0}{\|\mathbf{n}\|^2}\right)\mathbf{n}\right\| = \frac{|d-\mathbf{n}\cdot\mathbf{p}_0|}{\|\mathbf{n}\|}$$

10. The points A, B and C are all on one line if and only if the parallelogram they determine has area zero. Since this area is $\|\overrightarrow{AB} \times \overrightarrow{AC}\|$, this happens if and only if $\overrightarrow{AB} \times \overrightarrow{AC} = \mathbf{0}$.

12. If $\mathbf{u}$ and $\mathbf{v}$ are perpendicular, Theorem 4.3.4 shows that $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\|$. Moreover, if $\mathbf{w}$ is perpendicular to both $\mathbf{u}$ and $\mathbf{v}$, it is parallel to $\mathbf{u} \times \mathbf{v}$ so $\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \pm \|\mathbf{w}\| \|\mathbf{u} \times \mathbf{v}\|$ because the angle between them is either 0 or π. Finally, the rectangular parallepiped has volume

$$|\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})| = \|\mathbf{w}\| \|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{w}\| (\|\mathbf{u}\| \|\mathbf{v}\|)$$

using Theorem 4.3.5.

15. b. If $\mathbf{u} = \left[\begin{array}{c} x \\ y \\ z \end{array}\right]$, $\mathbf{v} = \left[\begin{array}{c} p \\ q \\ r \end{array}\right]$ and $\mathbf{w} = \left[\begin{array}{c} l \\ m \\ n \end{array}\right]$ then, by the row version of Exercise 3.1.19 Section 3.1, we get

$$\begin{aligned} \mathbf{u} \times (\mathbf{v} + \mathbf{w}) &= \det \left[\begin{array}{ccc} \mathbf{i} & x & l+p \\ \mathbf{j} & y & m+q \\ \mathbf{k} & z & n+r \end{array}\right] \\ &= \det \left[\begin{array}{ccc} \mathbf{i} & x & p \\ \mathbf{j} & y & q \\ \mathbf{k} & z & r \end{array}\right] + \det \left[\begin{array}{ccc} \mathbf{i} & x & l \\ \mathbf{j} & y & m \\ \mathbf{k} & z & n \end{array}\right] = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w} \end{aligned}$$

16. b. Let $\mathbf{v} = \left[\begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array}\right]$, $\mathbf{w} = \left[\begin{array}{c} w_1 \\ w_2 \\ w_3 \end{array}\right]$ and $\mathbf{u} = \left[\begin{array}{c} u_1 \\ u_2 \\ u_3 \end{array}\right]$. Compute

$$\begin{aligned} \mathbf{v} \cdot [(\mathbf{u} \times \mathbf{v}) + (\mathbf{v} \times \mathbf{w}) + (\mathbf{w} \times \mathbf{u})] &= \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) + \mathbf{v} \cdot (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) \\ &= 0 + 0 + \det \left[\begin{array}{ccc} v_1 & w_1 & u_1 \\ v_2 & w_2 & u_2 \\ v_3 & w_3 & u_3 \end{array}\right] \end{aligned}$$

by Theorem 4.3.1. Similarly

$$\mathbf{w} \cdot [[(\mathbf{u} \times \mathbf{v}) + (\mathbf{v} \times \mathbf{w}) + (\mathbf{w} \times \mathbf{u})]] = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \det \left[\begin{array}{ccc} w_1 & u_1 & v_1 \\ w_2 & u_2 & v_2 \\ w_3 & u_3 & v_3 \end{array}\right]$$

These determinants are equal because each can be obtained from the other by two column interchanges. The result follows because $(\mathbf{v} - \mathbf{w}) \cdot \mathbf{x} = \mathbf{v} \cdot \mathbf{x} - \mathbf{w} \cdot \mathbf{x}$ for any vector $\mathbf{x}$.

22. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are vectors of points in the planes (so $\mathbf{v}_1 \cdot \mathbf{n} = d_1$ and $\mathbf{v}_2 \cdot \mathbf{n} = d_2$), the distance is the length of the projection of $\mathbf{v}_2 - \mathbf{v}_1$ along $\mathbf{n}$; that is

$$\|\operatorname{proj}_{\mathbf{n}} (\mathbf{v}_2 - \mathbf{v}_1)\| = \left\| \left(\frac{(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}}{\|\mathbf{n}\|^2} \right) \mathbf{n} \right\| = \frac{|(\mathbf{v}_2 - \mathbf{v}_1) \cdot \mathbf{n}|}{\|\mathbf{n}\|} = \frac{|d_2 - d_1|}{\|\mathbf{n}\|}$$

4.4 Linear Operators on $\mathbb{R}^3$

1. b. By inspection, $A = \frac{1}{2} \left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]$; by the formulas preceding Theorem 4.4.2, this is the matrix of projection on $y = -x$.

d. By inspection, $A = \frac{1}{5}\begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}$; by the formulas precedinging Theorem 4.4.2, this is the matrix of reflection in $y = 2x$.

f. By inspection, $A = \frac{1}{2}\begin{bmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix}$; by Theorem 2.6.4 this is the matrix of rotation through $\frac{\pi}{3}$.

2. b. For any slope m, projection on the line $y = mx$ has matrix $\frac{1}{1+m^2}\begin{bmatrix} 1 & m \\ m & m^2 \end{bmatrix}$ (see the discussion preceding Theorem 4.4.2). Hence the projections on the lines $y = x$ and $y = -x$ have matrices $\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, respectively, so the first followed by the second has matrix (note the order)

$$\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

It follows that projection on $y = x$ followed by projection on $y = -x$ is the zero transformation. Note that this conclusion can also be reached geometrically. Given any vector **v**, its projection **p** on the line $y = x$ points along that line. But the line $y = -x$ is *perpendicular* to the line $y = x$, so the projection of **p** along $y = -x$ will be the zero vector. Since **v** was arbitrary, this shows again that projection on $y = x$ followed by projection on $y = -x$ is the zero transformation.

3. b. By Theorem 4.4.3: $\frac{1}{21}\begin{bmatrix} 17 & 2 & -8 \\ 2 & 20 & 4 \\ -8 & 4 & 5 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} = \frac{1}{21}\begin{bmatrix} 26 \\ 8 \\ -11 \end{bmatrix}$

d. By Theorem 4.4.3: $\frac{1}{30}\begin{bmatrix} 22 & -4 & 20 \\ -4 & 28 & 10 \\ 20 & 10 & -20 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} = \frac{1}{15}\begin{bmatrix} -32 \\ -1 \\ 35 \end{bmatrix}$

f. By Theorem 4.4.2: $\frac{1}{25}\begin{bmatrix} 9 & 0 & 12 \\ 0 & 0 & 0 \\ 12 & 0 & 16 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = \frac{1}{25}\begin{bmatrix} 93 \\ 0 \\ 124 \end{bmatrix}$

h. By Theorem 4.4.2: $\frac{1}{11}\begin{bmatrix} -9 & 2 & -6 \\ 2 & -9 & -6 \\ -6 & -6 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ -5 \\ 0 \end{bmatrix} = \frac{1}{11}\begin{bmatrix} -28 \\ 49 \\ 18 \end{bmatrix}$

4. b. This is Example 4.4.1 with $\theta = \frac{\pi}{6}$. Since $\cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$ and $\sin\frac{\pi}{6} = \frac{1}{2}$, the matrix is $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 & 0 \\ 1 & \sqrt{3} & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Hence the rotation of $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ is $\frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 & 0 \\ 1 & \sqrt{3} & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{3} \\ 1 \\ 6 \end{bmatrix}$.

6. Denote the rotation by $R_{L,\ \theta}$. Here the rotation takes place about the y-axis, so $R_{L,\ \theta}(\mathbf{j}) = \mathbf{j}$. In the xz-plane the effect of $R_{L,\ \theta}$ is to rotate counterclockwise through θ, and this has matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ Theorem 2.6.4. So, in the xz-plane, $R_{L,\ \theta}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ and $R_{L,\ \theta}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$. Hence $R_{L,\ \theta}(\mathbf{i}) = \begin{bmatrix} \cos\theta \\ 0 \\ \sin\theta \end{bmatrix}$ and $R_{L,\ \theta}(\mathbf{k}) = \begin{bmatrix} -\sin\theta \\ 0 \\ \cos\theta \end{bmatrix}$. Finally, the matrix of $R_{L,\ \theta}$ is $\begin{bmatrix} R_{L,\ \theta}(\mathbf{i}) & R_{L,\ \theta}(\mathbf{j}) & R_{L,\ \theta}(\mathbf{k}) \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$.

9. a. Write $\mathbf{v} = \begin{bmatrix} x & y \end{bmatrix}^T$.

Then $P_L(\mathbf{v}) = \text{proj}_{\mathbf{d}}\ \mathbf{v} = \left(\frac{\mathbf{v}\cdot\mathbf{d}}{\|\mathbf{d}\|^2}\right)\mathbf{d} = \left(\frac{ax+by}{a^2+b^2}\right)\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{a^2+b^2}\begin{bmatrix} a^2x+aby \\ abx+b^2y \end{bmatrix} = \frac{1}{a^2+b^2}\begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$.

Hence the matrix of P_L is $\frac{1}{a^2+b^2}\left[\begin{array}{cc} a^2 & ab \\ ab & b^2 \end{array}\right]$. Note that if the line L has slope m this retrieves the formula $\frac{1}{1+m^2}\left[\begin{array}{cc} 1 & m \\ m & m^2 \end{array}\right]$ preceding Theorem 4.4.2. However the present matrix works for vertical lines, where $\mathbf{d} = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$.

4.5 An Application to Computer Graphics

1. b. Translate to the origin, rotate and then translate back. As in Example 4.5.1, we compute

$$\left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{array}\right]\left[\begin{array}{ccc} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{array}\right]\left[\begin{array}{ccc} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right]\left[\begin{array}{ccccc} 0 & 6 & 5 & 1 & 3 \\ 0 & 0 & 3 & 3 & 9 \\ 1 & 1 & 1 & 1 & 1 \end{array}\right]$$

$$= \frac{1}{2}\left[\begin{array}{ccccc} \sqrt{2}+2 & 7\sqrt{2}+2 & 3\sqrt{2}+2 & -\sqrt{2}+2 & -5\sqrt{2}+2 \\ -3\sqrt{2}+4 & 3\sqrt{2}+4 & 5\sqrt{2}+4 & \sqrt{2}+4 & 9\sqrt{2}+4 \\ 2 & 2 & 2 & 2 & 2 \end{array}\right]$$

5. b. The line has a point $\mathbf{w} = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$, so we translate by $-\mathbf{w}$, then reflect in $y = 2x$, and then translate back by $\mathbf{w}$. The line $y = 2x$ has matrix $\frac{1}{5}\left[\begin{array}{cc} -3 & 4 \\ 4 & 3 \end{array}\right]$. Thus the matrix (for homogeneous coordinates) is

$$\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]\frac{1}{5}\left[\begin{array}{ccc} -3 & 4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 5 \end{array}\right]\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right] = \frac{1}{5}\left[\begin{array}{ccc} -3 & 4 & -4 \\ 4 & 3 & 2 \\ 0 & 0 & 5 \end{array}\right]$$

Hence for $\mathbf{w} = \left[\begin{array}{c} 1 \\ 4 \\ 1 \end{array}\right]$ we get $\frac{1}{5}\left[\begin{array}{ccc} -3 & 4 & -4 \\ 4 & 3 & 2 \\ 0 & 0 & 5 \end{array}\right]\left[\begin{array}{c} 1 \\ 4 \\ 1 \end{array}\right] = \frac{1}{5}\left[\begin{array}{c} 9 \\ 18 \\ 5 \end{array}\right]$. Hence the point is $P\left(\frac{9}{5}, \frac{18}{5}\right)$.

Supplementary Exercises: Chapter 4

4. Let $\mathbf{p}$ and $\mathbf{w}$ be the velocities of the airplane and the wind. Then $\|\mathbf{p}\| = 100$ knots and $\|\mathbf{w}\| = 75$ knots and the resulting actual velocity of the airplane is $\mathbf{v} = \mathbf{w} + \mathbf{p}$. Since $\mathbf{w}$ and $\mathbf{p}$ are orthogonal. Pythagoras' theorem gives $\|\mathbf{v}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{p}\|^2 = 75^2 + 100^2 = 25^2(3^2+4^2) = 25^2 \cdot 5^2$. Hence $\|\mathbf{v}\| = 25 \cdot 5 = 125$ knots. The angle θ satisfies $\cos\theta = \frac{\|\mathbf{w}\|}{\|\mathbf{v}\|} = \frac{75}{125} = 0.6$ so $\theta = 0.93$ radians or $53°$.

6. Let $\mathbf{v} = \left[\begin{array}{cc} x & y \end{array}\right]^T$ denote the velocity of the boat in the water. If $\mathbf{c}$ is the current velocity then $\mathbf{c} = (0, -5)$ because it flows south at 5 knots. We want to choose $\mathbf{v}$ so that the resulting actual velocity $\mathbf{w}$ of the boat has easterly direction. Thus $\mathbf{w} = \left[\begin{array}{c} z \\ 0 \end{array}\right]$ for some z. Now $\mathbf{w} = \mathbf{v} + \mathbf{c}$ so $\left[\begin{array}{c} z \\ 0 \end{array}\right] = \left[\begin{array}{c} x \\ y \end{array}\right] + \left[\begin{array}{c} 0 \\ -5 \end{array}\right] = \left[\begin{array}{c} x \\ y-5 \end{array}\right]$. Hence $z = x$ and $y = 5$. Finally, $13 = \|\mathbf{v}\| = \sqrt{x^2+y^2} = \sqrt{x^2+25}$ gives $x^2 = 144$, $x = \pm 12$. But $x > 0$ as w heads *east*, so $x = 12$. Thus he steers $\mathbf{v} = \left[\begin{array}{cc} 12 & 5 \end{array}\right]^T$, and the resulting actual speed is $\|\mathbf{w}\| = z = 12$ knots.

5. The Vector Space $\mathbb{R}^n$

5.1 Subspaces and Spanning

1. b. Yes. In fact, $U = \text{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$ so Theorem 5.1.1 applies.

   d. No. $\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$ is in U but $2\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$ is not in U.

   f. No. $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$ is in U but $(-1)\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is not in U.

2. b. No. If $\mathbf{x} = a\mathbf{y} + b\mathbf{z}$ equating first and third components gives $1 = 2a + b$, $15 = -3b$; whence $a = 3$, $b = -5$. This does not satisfy the second component which requires that $2 = -a - b$.

   d. Yes. $\mathbf{x} = 3\mathbf{y} + 4\mathbf{z}$.

3. b. No. Write these vectors as $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and $\mathbf{a}_4$, and let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix}$ be the matrix with these vectors as columns. Then $\det A = 0$, so A is not invertible. By Theorem 2.4.5, this means that the system $A\mathbf{x} = \mathbf{b}$ has *no* solution for some column $\mathbf{b}$. But this says that $\mathbf{b}$ is *not* a linear combination of the $\mathbf{a}_i$ by Definition 2.5. That is, the $\mathbf{a}_i$ do not span $\mathbb{R}^4$.

   For a more direct proof, $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is not a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and $\mathbf{a}_4$.

10. Since $a_i\mathbf{x}_i$ is in $\text{span}\{\mathbf{x}_i\}$ for each i, Theorem 5.1.1 shows that $\text{span}\{a_i\mathbf{x}_i\} \subseteq \text{span}\{\mathbf{x}_i\}$. Since $\mathbf{x}_i = a_i^{-1}(a_i\mathbf{x}_i)$ is in $\text{span}\{a_i\mathbf{x}_i\}$, we get $\text{span}\{\mathbf{x}_i\} \subseteq \text{span}\{a_i\mathbf{x}_i\}$, again by Theorem 5.1.1.

12. We have $U = \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ so, if $\mathbf{y}$ is in U, write $\mathbf{y} = t_1\mathbf{x}_1 + \cdots + t_k\mathbf{x}_k$ where the t_i are in $\mathbb{R}$. Then $A\mathbf{y} = t_1A\mathbf{x}_1 + \cdots + t_1A\mathbf{x}_k = t_1\mathbf{0} + \cdots + t_k\mathbf{0} = \mathbf{0}$.

15. b. $\mathbf{x} = (\mathbf{x} + \mathbf{y}) - \mathbf{y}$ is in U because $\mathbf{x} + \mathbf{y}$ and $-\mathbf{y} = (-1)\mathbf{y}$ are both in U and U is a subspace.

16. b. True. If we take $r = 1$ we see that $\mathbf{x} = 1\mathbf{x}$ is in U.

    d. True. We have $\text{span}\{\mathbf{y}, \mathbf{z}\} \subseteq \text{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ by Theorem 5.1.1 because both $\mathbf{y}$ and $\mathbf{z}$ are in $\text{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$. In other words, $U \subseteq \text{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$. For the other inclusion, it is clear that $\mathbf{y}$ and $\mathbf{z}$ are both in $U = \text{span}\{\mathbf{y}, \mathbf{z}\}$, and we are given that $\mathbf{x}$ is in U. Hence $\text{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\} \subseteq U$ by Theorem 5.1.1.

    f. False. Every vector in $\text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}\right\}$ has second component zero.

20. If U is a subspace then S2 and S3 certainly hold. Conversely, suppose that S2 and S3 hold. It is here that we need the condition that U is nonempty. Because we can then choose some $\mathbf{x}$ in U, and so $\mathbf{0} = 0\mathbf{x}$ is in U by S3. So U is a subspace.

22. b. First, $\mathbf{0}$ is in $U + W$ because $\mathbf{0} = \mathbf{0} + \mathbf{0}$ (and $\mathbf{0}$ is in both U and W). Now suppose that P and Q are both in $U + W$, say $\mathbf{p} = \mathbf{x}_1 + \mathbf{y}_1$ and $\mathbf{q} = \mathbf{x}_2 + \mathbf{y}_2$ where $\mathbf{x}_1$ and $\mathbf{x}_2$ are in U, and $\mathbf{y}_1$ and $\mathbf{y}_2$ are in W. Hence

$$\mathbf{p} + \mathbf{q} = (\mathbf{x}_1 + \mathbf{y}_1) + (\mathbf{x}_2 + \mathbf{y}_2) = (\mathbf{x}_1 + \mathbf{x}_2) + (\mathbf{y}_1 + \mathbf{y}_2)$$

so $\mathbf{p} + \mathbf{q}$ is in $U + W$ because $\mathbf{x}_1 + \mathbf{x}_2$ is in U (both $\mathbf{x}_1$ and $\mathbf{x}_2$ are in U), and $\mathbf{y}_1 + \mathbf{y}_2$ is in W. Similarly

$$aP = a(\mathbf{x}_1 + \mathbf{y}_1) = a\mathbf{x}_1 + a\mathbf{y}_1$$

is in $\mathbf{p} + \mathbf{q}$ because $a\mathbf{x}_1$ is in $\mathbf{p}$ and $a\mathbf{y}_1$ is in Q. Hence $U + W$ is a subspace.

5.2 Independence and Dimension

1. b. Yes. The matrix with these vectors as columns has determinant $-2 \neq 0$, so Theorem 5.2.3 applies.

   d. No. $(1, 1, 0, 0) - (1, 0, 1, 0) + (0, 0, 1, 1) - (0, 1, 0, 1) = (0, 0, 0, 0)$ is a nontrivial linear combination that vanishes.

2. b. Yes. If $a(\mathbf{x} + \mathbf{y}) + b(\mathbf{y} + \mathbf{z}) + c(\mathbf{z} + \mathbf{x}) = \mathbf{0}$ then $(a+c)\mathbf{x} + (a+b)\mathbf{y} + (b+c)\mathbf{z} = \mathbf{0}$. Since we are assuming that $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is independent, this means $a + c = 0$, $a + b = 0$, $b + c = 0$. The only solution is $a = b = c = 0$.

   d. No. $(\mathbf{x} + \mathbf{y}) - (\mathbf{y} + \mathbf{z}) + (\mathbf{z} + \mathbf{w}) - (\mathbf{w} + \mathbf{x}) = \mathbf{0}$ is a nontrivial linear combination that vanishes.

3. b. Write $\mathbf{x}_1 = (2, 1, 0, -1)$, $\mathbf{x}_2 = (-1, 1, 1, 1)$, $\mathbf{x}_3 = (2, 7, 4, 1)$, and write $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. Observe that $\mathbf{x}_3 = 3\mathbf{x}_1 + 4\mathbf{x}_2$ so $U = \{\mathbf{x}_1, \mathbf{x}_2\}$. This is a basis because $\{\mathbf{x}_1, \mathbf{x}_2\}$ is independent, so the dimension is 2.

   d. Write $\mathbf{x}_1 = (-2, 0, 3, 1)$, $\mathbf{x}_2 = (1, 2, -1, 0)$, $\mathbf{x}_3 = (-2, 8, 5, 3)$, $\mathbf{x}_4 = (-1, 2, 2, 1)$ and write $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$. Then $\mathbf{x}_3 = 3\mathbf{x}_1 + 4\mathbf{x}_2$ and $\mathbf{x}_4 = \mathbf{x}_1 + \mathbf{x}_2$ so the space is $\text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$. As this is independent, it is a basis so the dimension is 2.

4. b. $(a+b, a-b, b, a) = a(1, 1, 0, 1) + b(1, -1, 1, 0)$ so $U = \text{span}\{(1, 1, 0, 1), (1, -1, 1, 0)\}$. This is a basis so $\dim U = 2$.

   d. $(a - b, b + c, a, b + c) = a(1, 0, 1, 0) + b(-1, 1, 0, 1) + c(0, 1, 0, 1)$. Hence $U = \text{span}\{(1, 0, 1, 0), (-1, 1, 0, 1), (0, 1, 0, 1)\}$. This is a basis so $\dim U = 3$.

   f. If $a + b = c + d$ then $a = -b + c + d$. Hence $U = \{(-b + c + d, b, c, d) \mid b, c, d \text{ in } \mathbb{R}\}$ so $U = \text{span}\{(-1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1)\}$. This is a basis so $\dim U = 3$.

5. b. Let $a(\mathbf{x}+\mathbf{w})+b(\mathbf{y}+\mathbf{w})+c(\mathbf{z}+\mathbf{w})+d\mathbf{w}=\mathbf{0}$, that is $a\mathbf{x}+b\mathbf{y}+c\mathbf{z}+(a+b+c+d)\mathbf{w}=\mathbf{0}$. As $\{\mathbf{x},\ \mathbf{y},\ \mathbf{z},\ \mathbf{w}\}$ is independent, this implies that $a=0$, $b=0$, $c=0$ and $a+b+c+d=0$. Hence $d=0$ too, proving that $\{\mathbf{x}+\mathbf{w},\ \mathbf{y}+\mathbf{w},\ \mathbf{z}+\mathbf{w},\ \mathbf{w}\}$ is independent. It is a basis by Theorem 5.2.7 because $\dim \mathbb{R}^4=4$.

6. b. Yes. They are independent (the matrix with them as columns has determinant -2) and so are a basis of $\mathbb{R}^3$ by Theorem 5.2.7 (since $\dim \mathbb{R}^3=3$).

   d. Yes. They are independent (the matrix with them as columns has determinant -6) and so are a basis of $\mathbb{R}^3$ by Theorem 5.2.7 (since $\dim \mathbb{R}^3=3$).

   f. No. The determinant of the matrix with these vectors as its columns is zero, so they are not independent (by Theorem 5.2.3). Hence they are not a basis of $\mathbb{R}^4$ because $\dim \mathbb{R}^4=4$.

7. b. True. If $s\mathbf{y}+t\mathbf{z}=\mathbf{0}$ then $0\mathbf{x}+s\mathbf{y}+t\mathbf{z}=\mathbf{0}$, so $s=t=0$ by the independence of $\{\mathbf{x},\ \mathbf{y},\ \mathbf{z}\}$.

   d. False. If $\mathbf{x}\neq\mathbf{0}$ let $k=2$, $\mathbf{x}_1=\mathbf{x}$ and $\mathbf{x}_2=-\mathbf{x}$. Then each $\mathbf{x}_i\neq\mathbf{0}$ but $\{\mathbf{x}_1,\ \mathbf{x}_2\}$ is not independent.

   f. False. If $\mathbf{y}=-\mathbf{x}$ and $\mathbf{z}=\mathbf{0}$ then $1\mathbf{x}+1\mathbf{y}+1\mathbf{z}=\mathbf{0}$, but $\{\mathbf{x},\ \mathbf{y},\ \mathbf{z}\}$ is certainly not independent.

   h. True. The $\mathbf{x}_i$ are not independent so, by definition, some nontrivial linear combination vanishes.

10. If $r\mathbf{x}_2+s\mathbf{x}_3+t\mathbf{x}_5=\mathbf{0}$ then $0\mathbf{x}_1+r\mathbf{x}_2+s\mathbf{x}_3+0\mathbf{x}_4+t\mathbf{x}_5+0\mathbf{x}_6=\mathbf{0}$. Since the larger set is independent, this implies $r=s=t=0$.

12. If $t_1\mathbf{x}_1+t_2(\mathbf{x}_1+\mathbf{x}_2)+\cdots+t_k(\mathbf{x}_1+\mathbf{x}_2+\cdots+\mathbf{x}_k)=\mathbf{0}$ then, collecting terms in $\mathbf{x}_1,\ \mathbf{x}_2,\ \dots,$

$$(t_1+t_2+\cdots+t_k)\mathbf{x}_1+(t_2+\cdots+t_k)\mathbf{x}_2+\cdots+(t_{k-1}+t_k)\mathbf{x}_{k-1}+t_k\mathbf{x}_k=\mathbf{0}$$

Since $\{\mathbf{x}_1,\ \mathbf{x}_2,\ \dots,\ \mathbf{x}_k\}$ is independent we get

$$\begin{aligned} t_1+t_2+\cdots+t_k &= 0 \\ t_2+\cdots+t_k &= 0 \\ &\ \vdots \\ t_{k-1}+t_k &= 0 \\ t_k &= 0 \end{aligned}$$

The solution (from the bottom up) is $t_k=0$, $t_{k-1}=0$, $\dots$, $t_2=0$, $t_1=0$.

16. b. We show that A^T is invertible. Suppose $A^T\mathbf{x}=\mathbf{0x}$ in $\mathbb{R}^2$. By Theorem 2.4.5, we must show that $\mathbf{x}=\mathbf{0}$. If $\mathbf{x}=\left[\begin{smallmatrix} s \\ t \end{smallmatrix}\right]$ then $A^T\mathbf{x}=\mathbf{0}$ gives $as+ct=0$, $bs+dt=0$. But then $s(a\mathbf{x}+b\mathbf{y})+t(c\mathbf{x}+d\mathbf{y})=(sa+tc)\mathbf{x}+(sb+td)\mathbf{y}=\mathbf{0}$. Hence $s=t=0$ because $\{a\mathbf{x}+b\mathbf{y},\ c\mathbf{x}+d\mathbf{y}\}$ is independent.

17. b. Note first that each $V^{-1}\mathbf{x}_i$ is in null (AV) because $(AV)(V^{-1}\mathbf{x}_i)=A\mathbf{x}_i=\mathbf{0}$. If $t_1V^{-1}\mathbf{x}_1+\cdots+t_kV^{-1}\mathbf{x}_k=\mathbf{0}$ then $V^{-1}(t_1\mathbf{x}_1+\cdots+t_k\mathbf{x}_k)=\mathbf{0}$ so $t_1\mathbf{x}_1+\cdots+t_k\mathbf{x}_k=\mathbf{0}$ (by multiplication by V). Thus $t_1=\cdots=t_k=0$ because $\{\mathbf{x}_1,\ \dots,\ \mathbf{x}_k\}$ is independent. So $\left\{V^{-1}\mathbf{x}_1,\ \dots,\ V^{-1}\mathbf{x}_k\right\}$ is independent. To see that it spans null (AV), let $\mathbf{y}$ be in null (AV), so that $AV\mathbf{y}=\mathbf{0}$. Then $V\mathbf{y}$ is in null A so $V\mathbf{y}=s_1\mathbf{x}_1+\cdots+s_n\mathbf{x}_n$ because $\{\mathbf{x}_1,\ \dots,\ \mathbf{x}_n\}$ spans null A. Hence $\mathbf{y}=s_1V^{-1}\mathbf{x}_1++s_kV^{-1}\mathbf{x}_k$, as required.

20. We have $\{\mathbf{0}\}\subseteq U\subseteq W$ where $\dim\{\mathbf{0}\}=0$ and $\dim W=1$. Hence $\dim U$ is an integer between 0 and 1 (by Theorem 5.2.8), so $\dim U=0$ or $\dim U=1$. If $\dim U=0$ then $U=\{\mathbf{0}\}$ by Theorem 5.2.8 (because $\{\mathbf{0}\}\subseteq U$ and both spaces have dimension 0); if $\dim U=1$ then $U=W$ again by Theorem 5.2.8 (because $U\subseteq W$ and both spaces have dimension 1).

5.3 Orthogonality

1. b. $\left\{\frac{1}{\sqrt{3}}(1,\ 1,\ 1),\ \frac{1}{\sqrt{42}}(4,\ 1,\ -5),\ \frac{1}{\sqrt{14}}(3,\ -3,\ 1)\right\}$ where in each case we divide by the norm of the vector.

3. b. Write $\mathbf{e}_1 = (1,\ 0,\ -1)$, $\mathbf{e}_2 = (1,\ 4,\ 1)$, $\mathbf{e}_3 = (2,\ -1,\ 2)$. Then

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = 1+0-1 = 0,\ \mathbf{e}_1 \cdot \mathbf{e}_3 = 2+0-2 = 0,\ \mathbf{e}_2 \cdot \mathbf{e}_3 = 2-4+2 = 0$$

so $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3\}$ is orthogonal and hence a basis of $\mathbb{R}^3$. If $\mathbf{x} = (a,\ b,\ c)$, Theorem 5.3.6 gives

$$\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \frac{\mathbf{x}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 + \frac{\mathbf{x}\cdot\mathbf{e}_3}{\|\mathbf{e}_3\|^2}\mathbf{e}_3 = \frac{a-c}{2}\mathbf{e}_1 + \frac{a+4b+c}{18}\mathbf{e}_2 + \frac{2a-b+2c}{9}\mathbf{e}_3$$

d. Write $\mathbf{e}_1 = (1,\ 1,\ 1)$, $\mathbf{e}_2 = (1,\ -1,\ 0)$, $\mathbf{e}_3 = (1,\ 1,\ -2)$. Then

$$\mathbf{e}_1 \cdot \mathbf{e}_2 = 1-1+0 = 0,\ \mathbf{e}_1 \cdot \mathbf{e}_3 = 1+1-2 = 0, \text{ and } \mathbf{e}_2 \cdot \mathbf{e}_3 = 1-1+0 = 0$$

Hence $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3\}$ is orthogonal and hence is a basis of $\mathbb{R}^3$. If $\mathbf{x} = (a,\ b,\ c)$, Theorem 5.3.6 gives

$$\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \frac{\mathbf{x}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 + \frac{\mathbf{x}\cdot\mathbf{e}_3}{\|\mathbf{e}_3\|^2}\mathbf{e}_3 = \frac{a+b+c}{3}\mathbf{e}_1 + \frac{a-b}{2}\mathbf{e}_2 + \frac{a+b-2c}{6}\mathbf{e}_3$$

4. b. If $\mathbf{e}_1 = (2,\ -1,\ 0,\ 3)$ and $\mathbf{e}_2 = (2,\ 1,\ -2,\ -1)$ then $\{\mathbf{e}_1,\ \mathbf{e}_2\}$ is orthogonal because $\mathbf{e}_1 \cdot \mathbf{e}_2 = 4-1+0-3 = 0$. Hence $\{\mathbf{e}_1,\ \mathbf{e}_2\}$ is an orthogonal basis of the space U it spans. If $\mathbf{x} = (14,\ 1,\ -8,\ 5)$ *is* in U, Theorem 5.3.6 gives

$$\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \frac{\mathbf{x}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 = \frac{42}{14}\mathbf{e}_1 + \frac{40}{10}\mathbf{e}_2 = 3\mathbf{e}_1 + 4\mathbf{e}_2$$

We check that these are indeed equal. [We shall see in Section 8.1 that in any case, $\mathbf{x} - \left(\frac{\mathbf{x}\cdot\mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \frac{\mathbf{x}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2\right)$ is orthogonal to every vector in U.]

5. b. The condition that $(a,\ b,\ c,\ d)$ is orthogonal to each of the other three vectors gives the following equations for a, b, c, and d.

$$\begin{array}{rcrcrcrcl} a & & & - & c & + & d & = & 0 \\ 2a & + & b & + & c & - & d & = & 0 \\ a & - & 3b & + & c & & & = & 0 \end{array}$$

Solving we get:

$$\left[\begin{array}{rrrr|r} 1 & 0 & -1 & 1 & 0 \\ 2 & 1 & 1 & -1 & 0 \\ 1 & -3 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 3 & -3 & 0 \\ 1 & -3 & 2 & -1 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 3 & -3 & 0 \\ 0 & 0 & 11 & -10 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & \frac{1}{11} & 0 \\ 0 & 1 & 0 & -\frac{3}{11} & 0 \\ 0 & 0 & 1 & -\frac{10}{11} & 0 \end{array}\right]$$

The solution is $(a,\ b,\ c,\ d) = t(-1,\ 3,\ 10,\ 11)$, t in $\mathbb{R}$.

6. b.
$$\begin{aligned} \|2\mathbf{x}+7\mathbf{y}\|^2 &= (2\mathbf{x}+7\mathbf{y})\cdot(2\mathbf{x}+7\mathbf{y}) \\ &= 4(\mathbf{x}\cdot\mathbf{x})+14(\mathbf{x}\cdot\mathbf{y})+14(\mathbf{y}\cdot\mathbf{x})+49(\mathbf{y}\cdot\mathbf{y}) \\ &= 4\|\mathbf{x}\|^2+28(\mathbf{x}\cdot\mathbf{y})+49\|\mathbf{y}\|^2 \\ &= 36-56+49 \\ &= 29 \end{aligned}$$

d.
$$\begin{aligned} (\mathbf{x}-2\mathbf{y})\cdot(3\mathbf{x}+5\mathbf{y}) &= 3(\mathbf{x}\cdot\mathbf{x})+5(\mathbf{x}\cdot\mathbf{y})-6(\mathbf{y}\cdot\mathbf{x})-10(\mathbf{y}\cdot\mathbf{y}) \\ &= 3\|\mathbf{x}\|^2-(\mathbf{x}\cdot\mathbf{y})-10\|\mathbf{y}\|^2 \\ &= 27+2-10 \\ &= 19 \end{aligned}$$

7. b. False. For example, if $\mathbf{x}=(1,\ 0)$ and $\mathbf{y}=(0,\ 1)$ in $\mathbb{R}^2$, then $\{\mathbf{x},\ \mathbf{y}\}$ is orthogonal but $\mathbf{x}+\mathbf{y}=(1,\ 1)$ is not orthogonal to $\mathbf{x}$.

d. True. Let $\mathbf{x}$ and $\mathbf{y}$ be distinct vectors in the larger set. Then either both are $\mathbf{x}_i$'s, both are $\mathbf{y}_i$'s, or one is an $\mathbf{x}_i$ and one is a $\mathbf{y}_i$. In the first two cases $\mathbf{x}\cdot\mathbf{y}=\mathbf{0}$ because $\{\mathbf{x}_i\}$ and $\{\mathbf{y}_j\}$ are orthogonal sets; in the last case $\mathbf{x}\cdot\mathbf{y}=\mathbf{0}$ by the given condition.

f. True. Every pair of *distinct* vectors in $\{\mathbf{x}\}$ are orthogonal (there are *no* such pairs). As $\mathbf{x}\neq\mathbf{0}$, this shows that $\{\mathbf{x}\}$ is an orthogonal set.

9. Row i of A^T is $\mathbf{c}_i^T$ so the $(i,\ j)$ entry of A^TA is $\mathbf{c}_i^T\mathbf{c}_j=\mathbf{c}_i\cdot\mathbf{c}_j$. This is 0 if $i\neq j$, and 1 if $i=j$. That is $A^TA=I$.

11. b. Take $\mathbf{x}=(1,\ 1,\ 1)$ and $\mathbf{y}=(r_1,\ r_2,\ r_3)$. Then $|\mathbf{x}\cdot\mathbf{y}|\leq\|\mathbf{x}\|\,\|\mathbf{y}\|$ by Theorem 5.3.2; that is $|r_1+r_2+r_3|\leq\sqrt{3}\sqrt{r_1^2+r_2^2+r_3^2}$. Squaring both sides gives

$$r_1^2+r_2^2+r_3^2+2(r_1r_2+r_1r_3+r_2r_3)\leq 3(r_1^2+r_2^2+r_3^2)$$

Simplifying we obtain $r_1r_2+r_1r_3+r_2r_3\leq r_1^2+r_2^2+r_3^2$, as required.

12. b. Observe first that

$$(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})=\|\mathbf{x}\|^2-\|\mathbf{y}\|^2 \qquad (*)$$

holds for all vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

If $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ are orthogonal then $(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})=0$, so $\|\mathbf{x}\|^2=\|\mathbf{y}\|^2$ by $(*)$. Taking positive square roots gives $\|\mathbf{x}\|=\|\mathbf{y}\|$.

Conversely, if $\|\mathbf{x}\|=\|\mathbf{y}\|$ then certainly $\|\mathbf{x}\|^2=\|\mathbf{y}\|^2$, so $(*)$ gives $(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})=0$. This means that $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ are orthogonal.

15. If λ is an eigenvalue of A^TA, let $(A^TA)\mathbf{x}=\lambda\mathbf{x}$ where $\mathbf{x}\neq\mathbf{0}$ in $\mathbb{R}^n$. Then:

$$\|A\mathbf{x}\|^2=(A\mathbf{x})^T(A\mathbf{x})=(\mathbf{x}^TA^T)A\mathbf{x}=\mathbf{x}^T(A^TA\mathbf{x})=\mathbf{x}^T(\lambda\mathbf{x})=\lambda\|\mathbf{x}\|^2$$

Since $\|\mathbf{x}\|\neq 0$ (because $\mathbf{x}\neq\mathbf{0}$), this gives $\lambda=\frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2}\geq 0$.

5.4 Rank of a Matrix

1. b. $\begin{bmatrix} 2 & -1 & 1 \\ -2 & 1 & 1 \\ 4 & -2 & 3 \\ -6 & 3 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

   Hence, $\operatorname{rank} A = 2$ and $\left\{ \begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}^T, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \right\}$ is a basis of $\operatorname{row} A$. Thus $\left\{ \begin{bmatrix} 2 & -1 & 1 \end{bmatrix}^T, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \right\}$ is also a basis of $\operatorname{row} A$. Since the leading 1's are in columns 1 and 3, columns 1 and 3 of A are a basis of $\operatorname{col} A$.

   d. $\begin{bmatrix} 1 & 2 & -1 & 3 \\ -3 & -6 & 3 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

   Hence, $\operatorname{rank} A = 2$ and $\left\{ \begin{bmatrix} 1 & 2 & -1 & 3 \end{bmatrix}^T, \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T \right\}$ is a basis of $\operatorname{row} A$. Since the leading 1's are in columns 1 and 4, columns 1 and 4 of A are a basis of $\operatorname{col} A$.

2. b. Apply the gaussian algorithm to the matrix with these vectors as rows:

$$\begin{bmatrix} 1 & -1 & 2 & 5 & 1 \\ 3 & 1 & 4 & 2 & 7 \\ 1 & 1 & 0 & 0 & 0 \\ 5 & 1 & 6 & 7 & 8 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 2 & 5 & 1 \\ 0 & -2 & 4 & 2 & 7 \\ 0 & -4 & 6 & 7 & 8 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -\frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{3}{2} & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

   Hence, $\left\{ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T, \begin{bmatrix} 0 & 2 & -2 & -5 & -1 \end{bmatrix}^T, \begin{bmatrix} 0 & 0 & 2 & -3 & 6 \end{bmatrix}^T \right\}$ is a basis of U (where we have cleared fractions using scalar multiples).

   d. Write these columns as the rows of the following matrix:

$$\begin{bmatrix} 1 & 5 & -6 \\ 2 & 6 & -8 \\ 3 & 7 & -10 \\ 4 & 8 & 12 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -6 \\ 0 & -4 & 4 \\ 0 & -8 & 8 \\ 0 & -12 & 36 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -6 \\ 0 & 1 & -1 \\ 0 & 0 & 24 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -6 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

   Hence, $\left\{ \begin{bmatrix} 1 \\ 5 \\ -6 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of U.

3. b. No. If the 3 columns were independent, the rank would be 3.
   No. If the 4 rows were independent, the rank would be 4, a contradiction here as the rank cannot exceed the number of columns.

   d. No. Suppose that A is $m \times n$. If the rows are independent then $\operatorname{rank} A = \dim(\operatorname{row} A) = m$ (the number of rows). Similarly if the columns are independent then $\operatorname{rank} A = n$ (the number of columns). So if *both* the rows and columns are independent then $m = \operatorname{rank} A = n$, that is A is square.

   f. No. Then $\dim(\operatorname{null} A) = n - r = 4 - 2 = 2$, contrary to $\operatorname{null}(A) = \mathbb{R}\mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$.

4. Let $\mathbf{c}_j$ denote column j of A. If $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T \in \mathbb{R}^n$ then $A\mathbf{x} = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n$ by Definition 2.5. Hence

$$\operatorname{col} A = \operatorname{span}\{\mathbf{c}_1, \ldots, \mathbf{c}_n\} = \{x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n \mid x_j \in \mathbb{R}\} = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

7. b. The null space of A is the set of columns $\mathbf{x}$ such that $A\mathbf{x}=\mathbf{0}$. Applying gaussian elimination to the augmented matrix gives:

$$\left[\begin{array}{rrrrr|r} 3 & 5 & 5 & 2 & 0 & 0 \\ 1 & 0 & 2 & 2 & 1 & 0 \\ 1 & 1 & 1 & -2 & -2 & 0 \\ -2 & 0 & -4 & -4 & -2 & 0 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 0 & 2 & 2 & 1 & 0 \\ 0 & 5 & -1 & -4 & -3 & 0 \\ 0 & 1 & -1 & -4 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{rrrrr|r} 1 & 0 & 2 & 2 & 1 & 0 \\ 0 & 1 & -1 & -4 & -3 & 0 \\ 0 & 0 & 4 & 16 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 0 & 0 & -6 & -5 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence, the set of solutions is $\operatorname{null} A = \left\{ \left[\begin{array}{c} 6s+5t \\ 0 \\ -4s-3t \\ s \\ t \end{array}\right] \mid s,\ t \text{ in } \mathbb{R} \right\} = \operatorname{span} B$ where

$B = \left\{ \left[\begin{array}{r} 6 \\ 0 \\ -4 \\ 1 \\ 0 \end{array}\right], \left[\begin{array}{r} 5 \\ 0 \\ -3 \\ 0 \\ 1 \end{array}\right] \right\}$. Since B is independent, it is the required basis of $\operatorname{null} A$. We have $r = \operatorname{rank} A = 3$ by the above reduction, so $n-r = 5-3 = 2$. This is the dimension of $\operatorname{null} A$, as Theorem 5.4.3 asserts.

8. b. Since A is $m \times n$, $\dim(\operatorname{null} A) = n - \operatorname{rank} A$. To compute $\operatorname{rank} A$, let $R = \left[\begin{array}{cccc} r_1 & r_2 & \cdots & r_n \end{array}\right]$. Then $A = CR = \left[\begin{array}{cccc} r_1C & r_2C & \cdots & r_nC \end{array}\right]$ by block multiplication, so

$$\operatorname{col} A = \operatorname{span}\{r_1C,\ r_2C,\ \dots,\ r_nC\} = \operatorname{span}\{C\}$$

because some $r_i \neq 0$. Hence $\operatorname{rank} A = 1$, so $\dim(\operatorname{null} A) = n - \operatorname{rank} A = n-1$.

9. b. Let $A = \left[\begin{array}{ccc} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{array}\right]$ where $\mathbf{c}_j$ is the j^{th} column of A; we must show that $\{\mathbf{c}_1,\ \dots,\ \mathbf{c}_n\}$ is independent. Suppose that $x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = 0$, x_i in $\mathbb{R}$. If we write $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \cdots & x_n \end{array}\right]^T$, this reads $A\mathbf{x} = \mathbf{0}$ by Definition 2.5. But then $\mathbf{x}$ is in $\operatorname{null} A$, and $\operatorname{null} A = 0$ by hypothesis. So $\mathbf{x} = \mathbf{0}$, that is each $x_i = 0$. This shows that $\{\mathbf{c}_1,\ \dots,\ \mathbf{c}_n\}$ is independent.

10. b. If $A^2 = 0$ then $A(A\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$ in $\mathbb{R}^n$, that is $\{A\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^n\} \subseteq \operatorname{null} A$. But $\operatorname{col} A = \{A\mathbf{x} \mid \mathbf{x}$ in $\mathbb{R}^n\}$, so this shows that $\operatorname{col} A \subseteq \operatorname{null} A$. If we write $r = \operatorname{rank} A$, taking dimensions gives $r = \dim(\operatorname{col} A) \leq \dim(\operatorname{null} A) = n-r$ by Theorem 5.4.3. It follows that $2r \leq n$; that is $r \leq \frac{n}{2}$.

12. We have $\operatorname{rank}(A) = \dim[\operatorname{col}(A)]$ and $\operatorname{rank}(A^T) = \dim[\operatorname{row}(A^T)]$. Let $\{\mathbf{c}_1,\ \mathbf{c}_2,\ \dots,\ \mathbf{c}_k\}$ be a basis of col(A); it suffices to show that $\{\mathbf{c}_1^T,\ \mathbf{c}_2^T,\ \dots,\ \mathbf{c}_k^T\}$ is a basis of $\operatorname{row}(A^T)$. But if $t_1\mathbf{c}_1^T + t_2\mathbf{c}_2^T + \cdots + t_k\mathbf{c}_k^T = \mathbf{0}$, t_j in $\mathbb{R}$, then (taking transposes) $t_1\mathbf{c}_1 + t_2\mathbf{c}_2 + \cdots + t_k\mathbf{c}_k = \mathbf{0}$ so each $t_j = 0$. Hence $\{\mathbf{c}_1^T,\ \mathbf{c}_2^T,\ \dots,\ \mathbf{c}_k^T\}$ is independent. Given $\mathbf{v}$ in $\operatorname{row}(A^T)$ then $\mathbf{v}^T$ is in $\operatorname{col}(A)$, say $\mathbf{v}^T = s_1\mathbf{c}_1 + s_2\mathbf{c}_2 + \cdots + s_k\mathbf{c}_k$, s_j in $\mathbb{R}$. Hence $\mathbf{v} = s_1\mathbf{c}_1^T + s_2\mathbf{c}_2^T + \cdots + s_k\mathbf{c}_k^T$ so $\{\mathbf{c}_1^T,\ \mathbf{c}_2^T,\ \dots,\ \mathbf{c}_k^T\}$ spans $\operatorname{row}(A^T)$, as required.

15. b. Let $\{\mathbf{c}_1,\ \dots,\ \mathbf{c}_r\}$ be a basis of $\operatorname{col} A$ where $r = \operatorname{rank} A$. Since $A\mathbf{x} = \mathbf{b}$ has no solution, $\mathbf{b}$ is not in $\operatorname{col} A = \operatorname{span}\{\mathbf{c}_1,\ \cdots,\ \mathbf{c}_r\}$ by Exercise 12. It follows that $\{\mathbf{c}_1,\ \dots,\ \mathbf{c}_r,\ \mathbf{b}\}$ is independent [If $a_1\mathbf{c}_1 + \cdots + a_r\mathbf{c}_r + a\mathbf{b} = \mathbf{0}$ then $a = 0$ (since $\mathbf{b}$ is not in $\operatorname{col} A$), whence each $a_i = 0$ by the independence of the $\mathbf{c}_i$]. Hence, it suffices to show that $\operatorname{col}[A,\ B] = \operatorname{span}\{\mathbf{c}_1,\ \cdots,\ \mathbf{c}_r,\ \mathbf{b}\}$. It is clear that $\mathbf{b}$ is in $\operatorname{col}[A,\ \mathbf{b}]$, and each $\mathbf{c}_j$ is in $\operatorname{col}[A,\ \mathbf{b}]$ because it is a linear combination of columns of A (and so those of $[A,\ \mathbf{b}]$). Hence

$$\operatorname{span}\{\mathbf{c}_1,\ \dots,\ \mathbf{c}_r,\ \mathbf{b}\} \subseteq \operatorname{col}[A,\ \mathbf{b}]$$

On the other hand, each column $\mathbf{x}$ in col $[A, \mathbf{b}]$ is a linear combination of $\mathbf{b}$ and the columns of A. Since these columns are themselves linear combinations of the $\mathbf{c}_j$, so $\mathbf{x}$ is a linear combination of $\mathbf{b}$ and the $\mathbf{c}_j$. That is, $\mathbf{x}$ is in span $\{\mathbf{c}_1, \dots, \mathbf{c}_r, \mathbf{b}\}$.

5.5 Similarity and Diagonalization

1. b. $\det A = -5$, $\det B = -1$ (so A and B are not similar). However, $\operatorname{tr} A = 2 = \operatorname{tr} B$, and $\operatorname{rank} A = 2 = \operatorname{rank} B$ (both are invertible).

   d. $\operatorname{tr} A = 5$, $\operatorname{tr} B = 4$ (so A and B are not similar). However, $\det A = 7 = \det B$, so $\operatorname{rank} A = 2 = \operatorname{rank} B$ (both are invertible).

   f. $\operatorname{tr} A = -5 = \operatorname{tr} B$; $\det A = 0 = \det B$; however $\operatorname{rank} A = 2$, $\operatorname{rank} B = 1$ (so A and B are not similar).

3. b. We have $A \sim B$, say $B = P^{-1}AP$. Hence $B^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}(P^{-1})^{-1}$, so $A^{-1} \sim B^{-1}$ because P^{-1} is invertible.

4. b. $c_A(x) = \begin{vmatrix} x-3 & 0 & -6 \\ 0 & x+3 & 0 \\ -5 & 0 & x-2 \end{vmatrix} = (x+3)(x^2-5x-24) = (x+3)^2(x-8)$. So the eigenvalues are $\lambda_1 = -3$, $\lambda_2 = 8$. To find the associated eigenvectors:
   $\lambda_1 = -3$: $\begin{bmatrix} -6 & 0 & -6 \\ 0 & 0 & 0 \\ -5 & 0 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; basic eigenvectors $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.
   $\lambda_2 = 8$: $\begin{bmatrix} 5 & 0 & -6 \\ 0 & 11 & 0 \\ -5 & 0 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -\frac{6}{5} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; basic eigenvector $\begin{bmatrix} 6 \\ 0 \\ 5 \end{bmatrix}$.
   Since $\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ 5 \end{bmatrix} \right\}$ is a basis of eigenvectors, A is diagonalizable and
   $P = \begin{bmatrix} -1 & 0 & 6 \\ 0 & 1 & 0 \\ 1 & 0 & 5 \end{bmatrix}$ will satisfy $P^{-1}AP = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 8 \end{bmatrix}$.

   d. $c_A(x) = \begin{vmatrix} x-4 & 0 & 0 \\ 0 & x-2 & -2 \\ 2 & -3 & x-1 \end{vmatrix} = (x-4)^2(x+1)$. For $\lambda = 4$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & -2 \\ 2 & -3 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$;
   $E_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. Hence A is not diagonalizable by Theorem 5.5.6 because the dimension of $E_4(A) = 1$ while the eigenvalue 4 has multiplicity 2.

8. b. If $B = P^{-1}AP$ and $A^k = 0$, then $B^k = (P^{-1}AP)^k = P^{-1}A^kP = P^{-1}0P = 0$.

9. b. Let the diagonal entries of A all equal λ. If A is diagonalizable then $P^{-1}AP = \lambda I$ by Theorem 5.5.3 for some invertible matrix P. Hence $A = P(\lambda I)P^{-1} = \lambda(PIP^{-1}) = \lambda I$.

10. b. Let $P^{-1}AP = D = \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Since A and D are similar matrices, they have the same trace by Theorem 5.5.1. That is

$$\operatorname{tr} A = \operatorname{tr}(P^{-1}AP) = \operatorname{tr} D = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

12. b. $T_P(A)T_P(B) = (P^{-1}AP)(P^{-1}BP) = P^{-1}AIBP = P^{-1}ABP = T_P(AB)$

13. b. Assume that A is diagonalizable, say $A \sim D$ where D is diagonal, say $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of A. But A and A^T have the same eigenvalues (Example 3.3.5) so $A^T \sim D$ also. Hence $A \sim D \sim A^T$, so $A \sim A^T$ as required.

17. b. We use Theorem 5.5.7. The characteristic polynomial of B is computed by first adding rows 2 and 3 to row 1. For convenience, write $s = a+b+c$, $k = a^2+b^2+c^2-(ab+ac+bc)$.

$$\begin{aligned} c_B(x) &= \begin{vmatrix} x-c & -a & -b \\ -a & x-b & -c \\ -b & -c & x-a \end{vmatrix} = \begin{vmatrix} x-s & x-s & x-s \\ -a & x-b & -c \\ -b & -c & x-a \end{vmatrix} = \begin{vmatrix} x-s & 0 & 0 \\ -a & x+(a-b) & a-c \\ -b & b-c & x-(a-b) \end{vmatrix} \\ &= (x-s)\left[x^2-(a-b)^2-(a-c)(b-c)\right] \\ &= (x-s)(x^2-k) \end{aligned}$$

Hence, the eigenvalues of B are s, $\sqrt{k}$ and $-\sqrt{k}$. These must be real by Theorem 5.5.7, so $k \geq 0$. Thus $a^2+b^2+c^2 \geq ab+ac+bc$.

20. b. To compute $c_A(x) = \det(xI-A)$, add x times column 2 to column 1, and expand along row 1:

$$c_A(x) = \begin{vmatrix} x & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & x & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & x & -1 \\ -r_0 & -r_1 & -r_2 & -r_3 & \cdots & -r_{k-2} & x-r_{k-1} \end{vmatrix}$$

$$= \begin{vmatrix} 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ x^2 & x & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & x & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & 0 & & x & -1 \\ -r_0-r_1x & -r_1 & -r_2 & -r_3 & & -r_{k-2} & x-r_{k-1} \end{vmatrix}$$

Now expand along row 1 to get

$$c_A(x) = \begin{vmatrix} x^2 & -1 & & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & \cdots & x & -1 \\ -r_0-r_1x & -r_2 & -r_3 & \cdots & -r_{k-2} & x-r_{k-1} \end{vmatrix}$$

This matrix has the same form as $xI-A$, so repeat this procedure. It leads to the given expression for $\det(xI-A)$.

5.6 Best Approximation and Least Squares

1. b. Here $A = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 3 & -1 \\ 2 & -1 & 1 \\ 3 & -3 & 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 6 \\ 1 \\ 0 \\ 8 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Hence, $A^TA = \begin{bmatrix} 26 & -2 & 12 \\ -2 & 20 & -12 \\ 12 & -12 & 12 \end{bmatrix}$.

This is invertible and the inverse is

$$(A^TA)^{-1} = \frac{1}{144}\begin{bmatrix} 96 & -120 & -216 \\ -120 & 168 & 288 \\ -216 & 288 & 516 \end{bmatrix} = \frac{1}{36}\begin{bmatrix} 24 & -30 & -54 \\ -30 & 42 & 72 \\ -54 & 72 & 129 \end{bmatrix}$$

Here the (unique) best approximation is

$$\mathbf{z} = (A^TA)^{-1}A^T\mathbf{b} = \frac{1}{36}\begin{bmatrix} 24 & -30 & -54 \\ -30 & 42 & 72 \\ -54 & 72 & 129 \end{bmatrix}\begin{bmatrix} 44 \\ -15 \\ 29 \end{bmatrix} = \frac{1}{36}\begin{bmatrix} -60 \\ 138 \\ 285 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} -20 \\ 46 \\ 95 \end{bmatrix}$$

Of course this can be found more efficiently using gaussian elimination on the normal equations for $\mathbf{z}$.

2. b. Here $M^TM = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 21 \\ 21 & 133 \end{bmatrix}$, $M^T\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 7 & 8 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 42 \end{bmatrix}$. We solve the normal equation $(M^TM)A = M^T\mathbf{y}$ by inverting M^TM:

$$A = (M^TM)^{-1}M^T\mathbf{y} = \frac{1}{91}\begin{bmatrix} 133 & -21 \\ -21 & 4 \end{bmatrix}\begin{bmatrix} 10 \\ 42 \end{bmatrix} = \frac{1}{91}\cdot\begin{bmatrix} 448 \\ -42 \end{bmatrix} = \frac{1}{13}\begin{bmatrix} 64 \\ -6 \end{bmatrix}$$

Hence the best fitting line has equation $y = \frac{64}{13} - \frac{6}{13}x$.

d. Analogous to (b). The best fitting line is $y = -\frac{4}{10} - \frac{17}{10}x$.

3. b. Now

$$M^TM = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -2 & 0 & 3 & 4 \\ 4 & 0 & 9 & 16 \end{bmatrix}\begin{bmatrix} 1 & -2 & 4 \\ 1 & 0 & 0 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 29 \\ 5 & 29 & 83 \\ 29 & 83 & 353 \end{bmatrix}$$

$$M^T\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -2 & 0 & 3 & 4 \\ 4 & 0 & 9 & 16 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 16 \\ 70 \end{bmatrix}$$

We use $(MM^T)^{-1}$ to solve the normal equations even though it is more efficient to solve them by gaussian elimination.

$$A = (M^TM)^{-1}(M^T\mathbf{y}) = \frac{1}{4248}\begin{bmatrix} 3348 & 642 & -426 \\ 642 & 571 & -187 \\ -426 & -187 & 91 \end{bmatrix}\begin{bmatrix} 6 \\ 16 \\ 70 \end{bmatrix} = \frac{1}{4248}\begin{bmatrix} 540 \\ -102 \\ 822 \end{bmatrix} = \begin{bmatrix} .127 \\ -.024 \\ .194 \end{bmatrix}$$

Hence the best fitting quadratic has equation $y = .127 - .024x + .194x^2$.

4. b. In the notation of Theorem 5.6.3: $\mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ 5 \\ 10 \end{bmatrix}$, $M = \begin{bmatrix} 0 & 0^2 & 2^0 \\ 1 & 1^2 & 2^1 \\ 2 & 2^2 & 2^2 \\ 3 & 3^2 & 2^3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 4 & 4 \\ 3 & 9 & 8 \end{bmatrix}$.

Hence, $M^TM = \begin{bmatrix} 14 & 36 & 34 \\ 36 & 98 & 90 \\ 34 & 90 & 85 \end{bmatrix}$, and $(M^TM)^{-1} = \frac{1}{92}\begin{bmatrix} 230 & 0 & -92 \\ 0 & 34 & -36 \\ -92 & -36 & 76 \end{bmatrix} = \frac{1}{46}\begin{bmatrix} 115 & 0 & -46 \\ 0 & 17 & -18 \\ -46 & -18 & 38 \end{bmatrix}$.
Thus, the (unique) solution to the normal equation is

$$\mathbf{z} = (M^TM)^{-1}M^T\mathbf{y} = \frac{1}{46}\begin{bmatrix} 115 & 0 & -46 \\ 0 & 17 & -18 \\ -46 & -18 & 38 \end{bmatrix}\begin{bmatrix} 41 \\ 111 \\ 103 \end{bmatrix} = \frac{1}{46}\begin{bmatrix} -23 \\ 33 \\ 30 \end{bmatrix}$$

The best fitting function is thus $\frac{1}{46}[-23x + 33x^2 + 30(2)^x]$.

5. b. Here $\mathbf{y}=\begin{bmatrix}\frac{1}{2}\\1\\5\\9\end{bmatrix}$, $M=\begin{bmatrix}1 & (-1)^2 & \sin\left(-\frac{\pi}{2}\right)\\1 & 0^2 & \sin(0)\\1 & 2^2 & \sin(\pi)\\1 & 3^2 & \sin\left(\frac{3\pi}{2}\right)\end{bmatrix}=\begin{bmatrix}1 & 1 & -1\\1 & 0 & 0\\1 & 4 & 0\\1 & 9 & -1\end{bmatrix}$. Hence

$$M^TM=\begin{bmatrix}4 & 14 & 0\\14 & 98 & -10\\0 & -10 & 2\end{bmatrix} \text{ and } (M^TM)^{-1}=\frac{1}{2}\begin{bmatrix}-24 & 7 & 35\\7 & -2 & -10\\35 & -10 & -49\end{bmatrix}$$

Thus, the (unique) solution to the normal equations is

$$\mathbf{z}=(M^TM)^{-1}M^T\mathbf{y}=\frac{1}{40}\begin{bmatrix}24 & -2 & 14\\-2 & 12 & 3\\14 & 3 & 49\end{bmatrix}\begin{bmatrix}\frac{31}{2}\\\frac{203}{2}\\-\frac{19}{2}\end{bmatrix}=\frac{1}{20}\begin{bmatrix}18\\21\\28\end{bmatrix}$$

Hence, the best fitting functions

$$\frac{1}{20}[18+21x^2+28\sin\left(\frac{\pi x}{2}\right)]$$

7. To fit $s=a+bx$ where $x=t^2$, we have

$$M^TM=\begin{bmatrix}1 & 1 & 1\\1 & 4 & 9\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 4\\1 & 9\end{bmatrix}=\begin{bmatrix}3 & 14\\14 & 98\end{bmatrix}$$
$$M^T\mathbf{y}=\begin{bmatrix}1 & 1 & 1\\1 & 4 & 9\end{bmatrix}\begin{bmatrix}95\\80\\56\end{bmatrix}=\begin{bmatrix}231\\919\end{bmatrix}$$

Hence $A=(M^TM)^{-1}M^T\mathbf{y}=\frac{1}{98}\begin{bmatrix}98 & -14\\-14 & 3\end{bmatrix}\begin{bmatrix}231\\919\end{bmatrix}=\frac{1}{98}\begin{bmatrix}9772\\-477\end{bmatrix}=\begin{bmatrix}99.71\\-4.87\end{bmatrix}$ to two decimal places. Hence the best fitting equation is

$$y=99.71-4.87x=99.71-4.87t^2$$

Hence the estimate for g comes from $-\frac{1}{2}g=-4.87$, $g=9.74$ (the true value of g is 9.81). Now fit $s=a+bt+ct^2$. In this case

$$M^TM=\begin{bmatrix}1 & 1 & 1\\1 & 2 & 3\\1 & 4 & 9\end{bmatrix}\begin{bmatrix}1 & 1 & 1\\1 & 2 & 4\\1 & 3 & 9\end{bmatrix}=\begin{bmatrix}3 & 6 & 14\\6 & 14 & 36\\14 & 36 & 98\end{bmatrix}$$
$$M^T\mathbf{y}=\begin{bmatrix}1 & 1 & 1\\1 & 2 & 3\\1 & 4 & 9\end{bmatrix}\begin{bmatrix}95\\80\\56\end{bmatrix}=\begin{bmatrix}231\\423\\919\end{bmatrix}$$

Hence

$$A=(M^TM)^{-1}(M^T\mathbf{y})=\frac{1}{4}\begin{bmatrix}76 & -84 & 20\\-84 & 98 & -24\\20 & -24 & 6\end{bmatrix}\begin{bmatrix}231\\423\\919\end{bmatrix}=\frac{1}{4}\begin{bmatrix}404\\-6\\-18\end{bmatrix}=\begin{bmatrix}101\\-\frac{3}{2}\\-\frac{9}{2}\end{bmatrix}$$

so the best quadratic is $y=101-\frac{3}{2}t-\frac{9}{2}t^2$. This gives $-\frac{9}{2}=-\frac{1}{2}g$ so the estimate for g is $g=9$ in this case.

9. We want r_0, r_1, r_2, and r_3 to satisfy

$$r_0+50r_1+18r_2+10r_3=28$$
$$r_0+40r_1+20r_2+16r_3=30$$
$$r_0+35r_1+14r_2+10r_3=21$$

$$r_0 + 40r_1 + 12r_2 + 12r_3 = 23$$
$$r_0 + 30r_1 + 16r_2 + 14r_3 = 23$$

We settle for a best approximation. Here

$$A = \begin{bmatrix} 1 & 50 & 18 & 10 \\ 1 & 40 & 20 & 16 \\ 1 & 35 & 14 & 10 \\ 1 & 40 & 12 & 12 \\ 1 & 30 & 16 & 14 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 28 \\ 30 \\ 21 \\ 23 \\ 23 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 5 & 195 & 80 & 62 \\ 195 & 7825 & 3150 & 2390 \\ 80 & 3150 & 1320 & 1008 \\ 62 & 2390 & 1008 & 796 \end{bmatrix}$$
$$(A^T A)^{-1} = \frac{1}{50160} \begin{bmatrix} 1035720 & -16032 & 10080 & -45300 \\ -16032 & 416 & -632 & 800 \\ 10080 & -632 & 2600 & -2180 \\ -45300 & 800 & -2180 & 3950 \end{bmatrix}$$

So the best approximation

$$\mathbf{z} = (A^T A)^{-1}(A^T \mathbf{b}) = \frac{1}{50160} \begin{bmatrix} 1035720 & -16032 & 10080 & -45300 \\ -16032 & 416 & -632 & 800 \\ 10080 & -632 & 2600 & -2180 \\ -45300 & 800 & -2180 & 3950 \end{bmatrix} \begin{bmatrix} 125 \\ 4925 \\ 2042 \\ 1568 \end{bmatrix} = \begin{bmatrix} -5.19 \\ 0.34 \\ 0.51 \\ 0.71 \end{bmatrix}$$

The best fitting function is

$$y = -5.19 + 0.34x_1 + 0.51x_2 + 0.71x_3$$

10. b. $f(x) = a_0$ here so the sum of squares is

$$\begin{aligned} s &= (y_1 - a_0)^2 + (y_2 - a_0)^2 + \cdots + (y_n - a_0)^2 \\ &= \sum_{i=1}^{n} (y_i - a_0)^2 \\ &= \sum_{i=1}^{n} (a_0^2 - 2a_0 y_i + y_i^2) \\ &= n a_0^2 - \left(2 \sum y_i\right) a_0 + \left(\sum y_i^2\right) \end{aligned}$$

— a quadratic in a_0. Completing the square gives

$$s = n\left[a_0 - \tfrac{1}{n}\sum y_i\right]^2 - \left[\sum y_i^2 - \tfrac{1}{n}\left(\sum y_i\right)^2\right]$$

This is minimal when $a_0 = \frac{1}{n}\sum y_i$.

13. b. It suffices to show that the columns of $M = \begin{bmatrix} 1 & e^{x_1} \\ \vdots & \vdots \\ 1 & e^{x_n} \end{bmatrix}$ are independent. If $r_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + r_1 \begin{bmatrix} e^{x_1} \\ \vdots \\ e^{x_2} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, then $r_0 + r_1 e^{x_i} = 0$ for each i. Thus, $r_1(e^{x_i} - e^{x_j}) = 0$ for all i and j, so $r_1 = 0$ because two x_i are distinct. Then $r_0 = r_1 e^{x_1} = 0$ too.

5.7 An Application to Correlation and Variance

2. Let $\mathbf{x} = \left[\begin{array}{cccc} x_1 & x_2 & \cdots & x_{10} \end{array} \right] = \left[\begin{array}{ccccc} 12 & 16 & 13 & \cdots & 14 \end{array} \right]$ denote the number of years of education. Then $\overline{x} = \frac{1}{10}\sum x_i = 15.3$, and $s_x^2 = \frac{1}{n-1}\sum(x_i - \overline{x})^2 = 9.12$ (so $s_x = 3.02$).

   Let $\mathbf{y} = \left[\begin{array}{cccc} y_1 & y_2 & \cdots & y_{10} \end{array} \right] = \left[\begin{array}{ccccc} 31 & 48 & 35 & \cdots & 35 \end{array} \right]$ denote the number of dollars (in thousands) of yearly income. Then $\overline{y} = \frac{1}{10}\sum t_i = 40.3$, and $s_y^2 = \frac{1}{n-1}\sum(y_i - \overline{y})^2 = 114.23$ (so $s_y = 10.69$).

   The correlation is $r = \frac{\mathbf{x}\cdot\mathbf{y} - 10\overline{x}\overline{y}}{9 s_x s_y} = 0.599$.

4. b. We have $z_i = a + bx_i$ for each i, so

$$\overline{z} = \frac{1}{n}\sum(a + bx_i) = \frac{1}{n}\left(na + b\sum x_i\right) = a + b\left(\frac{1}{n}\sum x_i\right) = a + b\overline{x}$$

   Hence

$$s_z^2 = \frac{1}{n-1}\sum(z_i - \overline{z})^2 = \frac{1}{n-1}\sum[(a + bx_i) - (a + b\overline{x})]^2 = \frac{1}{n-1}\sum b^2(x_i - \overline{x})^2 = b^2 s_x^2$$

   The result follows because $\sqrt{b^2} = |b|$.

Supplementary Exercises: Chapter 5

1. b. False. If $r = 0$ then $r\mathbf{x}$ is in U for *any* $\mathbf{x}$.

   d. True. If $\mathbf{x}$ is in U then $-\mathbf{x} = (-1)\mathbf{x}$ is also in U by axiom S3 in Section 5.1.

   f. True. If $r\mathbf{x} + s\mathbf{y} = \mathbf{0}$ then $r\mathbf{x} + s\mathbf{y} + 0\mathbf{z} = \mathbf{0}$ so $r = s = 0$ because $\{\mathbf{x},\ \mathbf{y},\ \mathbf{z}\}$ is independent.

   h. False. Take $n = 2$, $\mathbf{x}_1 = \left[\begin{array}{c} 1 \\ 1 \end{array} \right]$ and $\mathbf{x}_2 = \left[\begin{array}{c} -1 \\ -1 \end{array} \right]$. Then both $\mathbf{x}_1$ and $\mathbf{x}_2$ are nonzero, but $\{\mathbf{x}_1,\ \mathbf{x}_2\}$ is not independent.

   j. False. If $a = b = c = 0$ then $a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = \mathbf{0}$ for *any* $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$.

   l. True. If $t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_n\mathbf{x}_n = \mathbf{0}$ implies that each $t_i = 0$, then $\{\mathbf{x}_1,\ \mathbf{x}_2,\ \dots,\ \mathbf{x}_n\}$ is independent, contrary to assumption.

   n. False. $\left\{ \left[\begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \end{array} \right], \left[\begin{array}{c} -1 \\ 0 \\ 0 \\ 0 \end{array} \right], \left[\begin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array} \right], \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \end{array} \right] \right\}$ is not independent.

   p. False. $\{\mathbf{x},\ \mathbf{x} + \mathbf{y},\ \mathbf{y}\}$ is never independent because $1\mathbf{x} + (-1)(\mathbf{x} + \mathbf{y}) + 1\mathbf{y} = \mathbf{0}$ is a nontrivial vanishing linear combination.

   r. False. Every basis of $\mathbb{R}^3$ must contain exactly 3 vectors (by Theorem 5.2.5). Of course a nonempty subset of a basis will be independent, but it will not span $\mathbb{R}^3$ if it contains fewer than 3 vectors.

6. Vector Spaces

6.1 Examples and Basic Properties

1. b. No: S5 fails $1(x,\ y,\ z) = (1x,\ 0,\ 1z) = (x,\ 0,\ z) \neq (x,\ y,\ z)$ for all $(x,\ y,\ z)$ in V. Note that the other nine axioms do hold.

   d. No: S4 and S5 fail: S5 fails because $1(x,\ y,\ z) = (2x,\ 2y,\ 2z) \neq (x,\ y,\ z)$; and S4 fails because $a[b(x,\ y,\ z)] = a(2bx,\ 2by,\ 2bz) = (4abx,\ 4aby,\ 4abz) \neq (2abx,\ 2aby,\ 2abz) = ab(x,\ y,\ z)$. Note that the eight other axioms hold.

2. b. No: A1 fails — for example $(x^3 + x + 1) + (-x^3 + x + 1) = 2x + 2$ is not in the set.

   d. No: A1 and S1 both fail. For example $x + x^2$ and $2x$ are not in the set. Hence none of the other axioms make sense.

   f. Yes. First verify A1 and S1. Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ are in V, so $a + c = b + d$ and $x + z = y + w$. Then $A + B = \begin{bmatrix} a+x & b+y \\ c+z & d+w \end{bmatrix}$ is in V because

   $$(a+x) + (c+z) = (a+c) + (x+z) = (b+d) + (y+w) = (b+y) + (d+w)$$

   Also $rA = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}$ is in V for all r in $\mathbb{R}$ because $ra + rc = r(a+c) = r(b+d) = rb + rd$.
   A2, A3, S2, S3, S4, S5. These hold for matrices in general.
   A4. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is in V and so serves as the zero of V.
   A5. Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $a + c = b + d$, then $-A = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$ is also in V because $-a - c = -(a+c) = -(b+d) = -b - d$. So $-A$ is the negative of A in V.

   h. Yes. The vector space axioms are the basic laws of arithmetic.

   j. No. S4 and S5 fail. For S4, $a(b(x,\ y)) = a(bx,\ -by) = (abx,\ aby)$, and this need not equal $ab(x,\ y) = (abx,\ -aby)$; as to S5, $1(x,\ y) = (x,\ -y) \neq (x,\ y)$ if $y \neq 0$.
   Note that the other axioms do hold here:
   A1, A2, A3, A4 and A5 hold because they hold in $\mathbb{R}^2$.
   S1 is clear; S2 and S3 hold because they hold in $\mathbb{R}^2$.

   l. No. S3 fails: Given $f : \mathbb{R} \to \mathbb{R}$ and $a,\ b$ in $\mathbb{R}$, we have

   $$\begin{aligned} [(a+b)f](x) &= f((a+b)x) = f(ax+bx) \\ (af+bf)(x) &= (af)(x) + (bf)(x) = f(ax) + f(bx) \end{aligned}$$

   These need not be equal: for example, if f is the function defined by $f(x) = x^2$;
   Then $f(ax+bx) = (ax+bx)^2$ need not equal $(ax)^2 + (bx)^2 = f(ax) + f(bx)$.

Note that the other axioms hold. A1-A4 hold by Example 6.1.7 as we are using pointwise addition.

$$\begin{array}{llll} \text{S2.} & a(f+g)(x) & = (f+g)(ax) & \text{definition of scalar multiplication in } V \\ & & = f(ax)+g(ax) & \text{definition of pointwise addition} \\ & & = (af)(x)+(ag)(x) & \text{definition of scalar multiplication in } V \\ & & = (af+ag)(x) & \text{definition of pointwise addition} \end{array}$$

As this is true for all x, $a(f+g) = af+ag$.

S4. $[a(bf)](x) = (bf)(ax) = f[b(ax)] = f[(ba)x] = [(ba)f](x) = [abf](x)$ for all x, so $a(bf) = (ab)f$.

S5. $(1f)(x) = f(1x) = f(x)$ for all x, so $1f = f$.

n. No. S4, S5 fail: $a*(b*X) = a*(bX^T) = a(bX^T)^T = abX^{TT} = abX$, while $(ab)*X = abX^T$. These need not be equal. Similarly: $1*X = 1X^T = X^T$ need not equal X.

Note that the other axioms do hold:

A1-A5. These hold for matrix addition generally.

S1. $a*X = aX^T$ is in V.

S2. $a*(X+Y) = a(X+Y)^T = a(X^T+Y^T) = aX^T+aY^T = a*X+a*Y$.

S3 $(a+b)*X = (a+b)X^T = aX^T+bX^T = a*X+b*X$.

4. A1. $(x,\ y)+(x_1,\ y_1) = (x+x_1,\ y+y_1+1)$ is in V for all $(x,\ y)$ and $(x_1,\ y_1)$ in V.

A2. $(x,\ y)+(x_1,\ y_1) = (x+x_1,\ y+y_1+1) = (x_1+x,\ y_1+y+1) = (x_1,\ y_1)+(x_1,\ y)$.

$$\begin{array}{lrl} \text{A3.} & (x,\ y)+((x_1,\ y_1)+(x_2,\ y_2)) & = (x,\ y)+(x_1+x_2,\ y_1+y_2+1) \\ & & = (x+(x_1+x_2),\ y+(y_1+y_2+1)+1) \\ & & = (x+x_1+x_2,\ y+y_1+y_2+2) \\ & ((x,\ y)+(x_1,\ y_1))+(x_2,\ y_2) & = (x+x_1,\ y+y_1+1)+(x_2,\ y_2) \\ & & = ((x+x_1)+x_2,\ (y+y_1+1)+y_2+1) \\ & & = (x+x_1+x_2,\ y+y_1+y_2+2) \end{array}$$

These are equal for all $(x,\ y)$, $(x_1,\ y_1)$ and $(x_2,\ y_2)$ in V.

A4. $(x,\ y)+(0,\ -1) = (x+0,\ y+(-1)+1) = (x,\ y)$ for all $(x,\ y)$, so $(0,\ -1)$ is the zero of V.

A5. $(x,\ y)+(-x,\ -y-2) = (x+(-x),\ y+(-y-2)+1) = (0,\ -1)$ is the zero of V (from A4) so the negative of $(x,\ y)$ is $(-x,\ -y-2)$.

S1. $a(x,\ y) = (ax,\ ay+a-1)$ is in V for all $(x,\ y)$ in V and a in $\mathbb{R}$.

$$\begin{array}{lrl} \text{S2.} & a[(x,\ y)+(x_1,\ y_1)] & = a(x+x_1,\ y+y_1+1) = (a(x+x_1),\ a(y+y_1+1)+a-1) \\ & & \qquad = (ax+ax_1,\ ay+ay_1+2a-1) \\ & a(x,\ y)+a(x_1,\ y_1) & = (ax,\ ay+a-1)+(ax_1,\ ay_1-a-1) \\ & & = ((ax+ax_1),\ (ay+a-1)+(ay_1+a-1)+1) \\ & & = (ax+ax_1,\ ay+ay_1+2a-1) \end{array}$$

These are equal.

S4. $a[b(x,\ y)] = a(bx,\ by+b-1) = (a(bx),\ a(by+b-1)+a-1) = (abx,\ aby+ab-1) = (ab)(x,\ y)$.

S5. $1(x,\ y) = (1x,\ 1y+1-1) = (x,\ y)$ for all $(x,\ y)$ in V.

5. b. Subtract the first equation from the second to get $\mathbf{x}-3\mathbf{y} = \mathbf{v}-\mathbf{u}$, whence $\mathbf{x} = 3\mathbf{y}+\mathbf{v}-\mathbf{u}$. Substitute in the first equation to get

$$\begin{aligned} 3(3\mathbf{y}+\mathbf{v}-\mathbf{u})-2\mathbf{y} &= \mathbf{u} \\ 7\mathbf{y} &= 4\mathbf{u}-3\mathbf{v} \end{aligned}$$

$$\mathbf{y} = \tfrac{4}{7}\mathbf{u} - \tfrac{3}{7}\mathbf{v}$$

Substitute this in the first equation to get $\mathbf{x} = \frac{5}{7}\mathbf{u} - \frac{2}{7}\mathbf{v}$.

It is worth noting that these equations can also be solved by gaussian elimination using **u** and **v** as the constants.

6. b. $a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0}$ becomes $\left[\begin{array}{cc} a & 0 \\ 0 & a \end{array}\right] + \left[\begin{array}{cc} 0 & b \\ b & 0 \end{array}\right] + \left[\begin{array}{cc} c & c \\ c & -c \end{array}\right] = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right]$.
Equating corresponding entries gives equations for a and b.

$$a + c = 0,\ b + c = 0,\ b + c = 0,\ a - c = 0$$

The only solution is $a = b = c = 0$.

d. $a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0}$ means $a \sin x + b \cos x + c1 = 0$ for all choices of x. If $x = 0,\ \frac{\pi}{2},\ \pi$, we get, respectively, equations $b + c = 0$, $a + c = 0$, and $-b + c = 0$. The only solution is $a = b = c = 0$.

7. b. $4(3\mathbf{u} - \mathbf{v} + \mathbf{w}) - 2[(3\mathbf{u} - 2\mathbf{v}) - 3(\mathbf{v} - \mathbf{w})] + 6(\mathbf{w} - \mathbf{u} - \mathbf{v})$
$= (12\mathbf{u} - 4\mathbf{v} + 4\mathbf{w}) - 2[3\mathbf{u} - 2\mathbf{v} - 3\mathbf{v} + 3\mathbf{w}] + (6\mathbf{w} - 6\mathbf{u} - 6\mathbf{v})$
$= (12\mathbf{u} - 4\mathbf{v} + 4\mathbf{w}) - (6\mathbf{u} - 10\mathbf{v} + 6\mathbf{w}) + (6\mathbf{w} - 6\mathbf{u} - 6\mathbf{v})$
$= 4\mathbf{w}$

10. Suppose that a vector **z** has the property that $\mathbf{z} + \mathbf{v} = \mathbf{v}$ for all **v** in V. Since $\mathbf{0} + \mathbf{v} = \mathbf{v}$ also holds for all **v**, we obtain $\mathbf{z} + \mathbf{v} = \mathbf{0} + \mathbf{v}$, so $\mathbf{z} = \mathbf{0}$ by cancellation.

12. b. $(-a)\mathbf{v} + a\mathbf{v} = (-a + a)\mathbf{v} = 0\mathbf{v} = \mathbf{0}$. Since also $-(a\mathbf{v}) + a\mathbf{v} = \mathbf{0}$ we get $(-a)\mathbf{v} + a\mathbf{v} = -(a\mathbf{v}) + a\mathbf{v}$. Thus $(-a)\mathbf{v} = -(a\mathbf{v})$ by cancellation.
Alternatively: $(-a)\mathbf{v} = [(-1)a]\mathbf{v} = (-1)(a\mathbf{v}) = -a\mathbf{v}$ using part 4 of Theorem 6.1.3.

13. b. We proceed by induction on n (see Appendix A). The case $n = 1$ is clear. If the equation holds for some $n \geq 1$, we have

$$\begin{array}{rcll} (a_1 + a_2 + \cdots + a_n + a_{n+1})\mathbf{v} & = & [(a_1 + a_2 + \cdots + a_n) + a_{n+1}]\mathbf{v} & \\ & = & (a_1 + a_2 + \cdots + a_n)\mathbf{v} + a_{n+1}\mathbf{v} & \text{by S3} \\ & = & (a_1\mathbf{v} + a_2\mathbf{v} + \cdots + a_n\mathbf{v}) + a_{n+1}\mathbf{v} & \text{by induction} \\ & = & a_1\mathbf{v} + a_2\mathbf{v} + \cdots + a_n\mathbf{v} + a_{n+1}\mathbf{v} & \end{array}$$

Hence it holds for $n + 1$, and the induction is complete.

15. c. Since $a \neq 0$, a^{-1} exists in $\mathbb{R}$. Hence $a\mathbf{v} = a\mathbf{w}$ gives $a^{-1}a\mathbf{v} = a^{-1}a\mathbf{w}$; that is $1\mathbf{v} = 1\mathbf{w}$, that is $\mathbf{v} = \mathbf{w}$.
Alternatively: $a\mathbf{v} = a\mathbf{w}$ gives $a\mathbf{v} - a\mathbf{w} = \mathbf{0}$, so $a(\mathbf{v} - \mathbf{w}) = \mathbf{0}$. As $a \neq 0$, it follows that $\mathbf{v} - \mathbf{w} = \mathbf{0}$ by Theorem 6.1.3, that is $\mathbf{v} = \mathbf{w}$.

6.2 Subspaces and Spanning Sets

1. b. Yes. U is a subset of $\mathbf{P}_3$ because $xg(x)$ has degree one more than the degree of $g(x)$. Clearly $0 = x \cdot 0$ is in U. Given $\mathbf{u} = xg(x)$ and $\mathbf{v} = xh(x)$ in U (where $g(x)$ and $b(x)$ are in $\mathbf{P}_2$) we have

$$\mathbf{u}+\mathbf{v} = x(g(x)+h(x)) \text{ is in } U \text{ because } g(x)+h(x) \text{ is in } \mathbf{P}_2$$
$$k\mathbf{u} = x(kg(x)) \text{ is in } U \text{ for all } k \text{ in } \mathbb{R} \text{ because } kg(x) \text{ is in } \mathbf{P}_2$$

d. Yes. As in (b), U is a subset of $\mathbf{P}_3$. Clearly $0 = x \cdot 0 + (1-x) \cdot 0$ is in U. If $\mathbf{u} = xg(x) + (1-x)h(x)$ and $\mathbf{v} = xg_1(x) + (1-x)h_1(x)$ are in U then

$$\begin{aligned} \mathbf{u}+\mathbf{v} &= x[g(x)+g_1(x)] + (1-x)\,[h(x)+h_1(x)] \\ k\mathbf{u} &= x[kg(x)] + (1-x)[kh(x)] \end{aligned}$$

both lie in U because $g(x)+g_1(x)$ and $h(x)+h_1(x)$ are in $\mathbf{P}_2$.

f. No. U is not closed under addition (for example $\mathbf{u} = 1+x^3$ and $\mathbf{v} = x - x^3$ are in U but $\mathbf{u}+\mathbf{v} = 1+x$ is not in U). Also, the zero polynomial is not in U.

2. b. Yes. Clearly $\mathbf{0} = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right]$ is in U. If $\mathbf{u} = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ and $\mathbf{u}_1 = \left[\begin{array}{cc} a_1 & b_1 \\ c_1 & d_1 \end{array}\right]$ are in U then $\mathbf{u}+\mathbf{u}_1 = \left[\begin{array}{cc} a+a_1 & b+b_1 \\ c+c_1 & d+d_1 \end{array}\right]$ is in U because

$$\begin{aligned} (a+a_1)+(b+b_1) &= (a+b)+(a_1+b_1) \\ &= (c+d)+(c_1+d1) \\ &= (c+c_1)+(d+d_1) \end{aligned}$$

$k\mathbf{u} = \left[\begin{array}{cc} ka & kb \\ kc & kd \end{array}\right]$ is in U because $ka+kb = k(a+b) = k(c+d) = kc+kd$.

d. Yes. Here 0 is in U as $0B = 0$. If A and A_1 are in U then $AB = 0$ and $A_1B = 0$, so $(A+A_1)B = AB + A_1B = 0+0 = 0$ and $(kA)B = k(AB) = k0 = 0$ for all k in $\mathbb{R}$. This shows that $A+A_1$ and kA are also in U.

f. No. U is not closed under addition. In fact, $A = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right]$ and $A_1 = \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$ are both in U, but $A + A_1 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$ is *not* in U.

3. b. No. U is not closed under addition. For example if f and g are defined by $f(x) = x+1$ and $g(x) = x^2+1$, then f and g are in U but $f+g$ is not in U because $(f+g)(0) = f(0)+g(0) = 1+1 = 2$.

d. No. U is not closed under scalar multiplication. For example, if f is defined by $f(x) = x$, then f is in U but $(-1)f$ is not in U (for example $[(-1)f](\frac{1}{2}) = -\frac{1}{2}$ so is not in U).

f. Yes. 0 is in U because $0(x+y) = 0 = 0+0 = 0(x)+0(y)$ for all x and y in $[0,\ 1]$. If f and g are in U then, for all k in $\mathbb{R}$:

$$\begin{aligned} (f+g)(x+y) &= f(x+y)+g(x+y) \\ &= (f(x)+f(y))+(g(x)+g(y)) \end{aligned}$$

$$\begin{aligned} &= (f(x)+g(x)) + (f(y)+g(y)) \\ &= (f+g)(x) + (f+g)(y) \end{aligned}$$

$$\begin{aligned} (kf)(x+y) = k[f(x+y)] = k[f(x)+f(y)] &= k[f(x)] + k[f(y)] \\ &= (kf)(x) + (kf)(y) \end{aligned}$$

Hence $f+g$ and kf are in U.

5. b. Suppose $\mathbf{x} = \left[\begin{array}{c} x_1 \\ \vdots \\ x_n \end{array}\right] \neq 0$, say $x_k \neq 0$. Given $\mathbf{y} = \left[\begin{array}{c} y_1 \\ \vdots \\ y_n \end{array}\right]$ let A be the $m \times n$ matrix with k^{th} column $x_k^{-1}\mathbf{y}$ and the other columns zero. Then $\mathbf{y} = A\mathbf{x}$ by matrix multiplication, so $\mathbf{y}$ is in U. Since $\mathbf{y}$ was an arbitrary column in $\mathbb{R}^n$, this shows that $U = \mathbb{R}^m$.

6. b. We want r, s and t such that $2x^2 - 3x + 1 = r(x+1) + s(x^2+x) + t(x^2+2)$. Equating coefficients of x^2, x and 1 gives $s+t=2$, $r+s=-3$, $r+2t=1$. The unique solution is $r=-3$, $s=0$, $t=2$.

   d. As in (b), $x = \frac{2}{3}(x+1) + \frac{1}{3}(x^2+x) - \frac{1}{3}(x^2+2)$.

7. b. If $\mathbf{v} = s\mathbf{u} + t\mathbf{w}$ then $x = s(x^2+1) + t(x+2)$. Equating coefficients gives $0 = s$, $1 = t$ and $0 = s + 2t$. Since there is *no* solution to these equations, $\mathbf{v}$ does *not* lie in span $\{\mathbf{u},\ \mathbf{w}\}$.

   d. If $\mathbf{v} = s\mathbf{u} + t\mathbf{w}$, then $\left[\begin{array}{rr} 1 & -4 \\ 5 & 3 \end{array}\right] = s\left[\begin{array}{rr} 1 & -1 \\ 2 & 1 \end{array}\right] + t\left[\begin{array}{rr} 2 & 1 \\ 1 & 0 \end{array}\right]$. Equating corresponding entries gives $s+2t=1$, $-s+t=-4$, $2s+t=5$ and $s=3$. These equations have the unique solution $t=-1$, $s=3$, so $\mathbf{v}$ is in span $\{\mathbf{u},\ \mathbf{w}\}$; in fact $\mathbf{v} = 3\mathbf{u} - \mathbf{w}$.

8. b. Yes. The trigonometry identity $1 = \sin^2 x + \cos^2 x$ for all x means that 1 is in span $\{\sin^2 x,\ \cos^2 x\}$.

   d. Suppose $1 + x^2 = s\sin^2 x + t\cos^2 x$ for some s and t. This must hold for all x. Taking $x = 0$ gives $1 = t$; taking $x = \pi$ gives $1 + \pi^2 = -t$. Thus $2 + \pi^2 = 0$, a contradiction. So no such s and t exist, that is $1 + x^2$ is *not* in span $\{\sin^2 x,\ \cos^2 x\}$.

9. b. Write $U = $ span $\{1+2x^2,\ 3x,\ 1+x\}$, then successively

$$\begin{aligned} x &= \tfrac{1}{3}(3x) \text{ is in } U \\ 1 &= (1+x) - x \text{ is in } U \\ x^2 &= \tfrac{1}{2}[(1+2x^2) - 1] \text{ is in } U \end{aligned}$$

Since $\mathbf{P}_2 = $ span $\{1,\ x,\ x^2\}$, this shows that $\mathbf{P}_2 \subseteq U$. Clearly $U \subseteq \mathbf{P}_2$, so $U = \mathbf{P}_2$.

11. b. The vectors $\mathbf{u} - \mathbf{v} = 1\mathbf{u} + (-1)\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, and $\mathbf{w}$ are all in span $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ so span $\{\mathbf{u} - \mathbf{v},\ \mathbf{u} + \mathbf{w},\ \mathbf{w}\} \subseteq$ span $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ by Theorem 6.2.2. The other inclusion also follows from Theorem 6.2.2 because

$$\begin{aligned} \mathbf{u} &= (\mathbf{u} + \mathbf{w}) - \mathbf{w} \\ \mathbf{v} &= -(\mathbf{u} - \mathbf{v}) + (\mathbf{u} + \mathbf{w}) - \mathbf{w} \\ \mathbf{w} &= \mathbf{w} \end{aligned}$$

show that $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are all in span $\{\mathbf{u} - \mathbf{v},\ \mathbf{u} + \mathbf{v},\ \mathbf{w}\}$.

14. No. For example $(1,\ 1,\ 0)$ is not even *in* span $\{(1,\ 2,\ 0),\ (1,\ 1,\ 1)\}$. Indeed $(1,\ 1,\ 0) = s(1,\ 2,\ 0) + t(1,\ 1,\ 1)$ requires that $s+t=1$, $2s+t=1$, $t=0$, and this has no solution.

18. Write $W = \text{span}\ \{\mathbf{u},\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$. Since $\mathbf{u}$ is in V we have $W \subseteq V$. But the fact that $a_1 \neq 0$ means

$$\mathbf{v}_1 = \tfrac{1}{a_1}\mathbf{u} - \tfrac{a_2}{a_1}\mathbf{v}_2 - \cdots - \tfrac{a_n}{a_1}\mathbf{v}_n$$

so $\mathbf{v}_1$ is in W. Since $\mathbf{v}_2,\ \dots,\ \mathbf{v}_n$ are all in W, this shows that $V = \text{span}\ \{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\} \subseteq W$. Hence $V = W$.

21. b. If $\mathbf{u}$ and $\mathbf{u}+\mathbf{v}$ are in U then $\mathbf{v} = (\mathbf{u}+\mathbf{v}) - \mathbf{u} = (\mathbf{u}+\mathbf{v}) + (-1)\mathbf{u}$ is in U because U is closed under addition and scalar multiplication.

22. If U is a subspace then, $\mathbf{u}_1 + a\mathbf{u}_2$ is in U for any $\mathbf{u}_i$ in U and a in $\mathbb{R}$ by the subspace test. Conversely, assume that this condition holds for U. Then, in the subspace test, conditions (2) and (3) hold for U (because $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v}$ in V), so it remains to show that $\mathbf{0}$ is in U. This is where we use the assumption that U is nonempty because, if $\mathbf{u}$ is any vector in U then $\mathbf{u} + (-1)\mathbf{u}$ is in U by assumption, that is $\mathbf{0} \in U$.

6.3 Linear Independence and Dimension

1. b. Independent. If $rx^2 + s(x+1) + t(1-x-x^2) = 0$ then, equating coefficients of x^2, x and 1, we get $r-t=0$, $s-t=0$, $s+t=0$. The only solution is $r=s=t=0$.

   d. Independent. If $r\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right] + s\left[\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right] + t\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right] + u\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, then $r+t+u=0$, $r+s+u=0$, $r+s+t=0$, $s+t+u=0$. The only solution is $r=s=t=u=0$.

2. b. Dependent. $3(x^2-x+3) - 2(2x^2+x+5) + (x^2+5x+1) = 0$

   d. Dependent. $2\left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] + 0\left[\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

   f. Dependent. $\frac{5}{x^2+x-6} + \frac{1}{x^2-5x+6} - \frac{6}{x^2-9} = 0$.

3. b. Dependent. $1 - \sin^2 x - \cos^2 x = 0$ for all x.

4. b. If $r(2,\ x,\ 1) + s(1,\ 0,\ 1) + t(0,\ 1,\ 3) = (0,\ 0,\ 0)$ then, equating components:

$$\begin{array}{rcrcrcl} 2r & + & s & & & = & 0 \\ xr & & & + & t & = & 0 \\ r & + & s & + & 3t & = & 0 \end{array}$$

   Gaussian elimination gives

$$\left[\begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ x & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 2 & 1 & 0 & 0 \\ x & 0 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & -x & 1-3x & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 1+3x & 0 \end{array}\right]$$

This has only the trivial solution $r = s = t = 0$ if and only if $x \neq \frac{-1}{3}$. Alternatively, the coefficient matrix has determinant

$$\det \begin{bmatrix} 2 & 1 & 0 \\ x & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} = \det \begin{bmatrix} 2 & 1 & 0 \\ x & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix} = -\det \begin{bmatrix} x & 1 \\ -1 & 3 \end{bmatrix} = -(1+3x)$$

This is nonzero if and only if $x \neq -\frac{1}{3}$.

5. b. **Independence**: If $r(-1,\ 1,\ 1) + s(1,\ -1,\ 1) + t(1,\ 1,\ -1) = (0,\ 0,\ 0)$ then $-r+s+t = 0$, $r-s+t = 0$, $r+s-t = 0$. The only solution is $r = s = t = 0$.

   Spanning: Write $U = \text{span}\ \{(-1,\ 1,\ 1),\ (1,\ -1,\ 1),\ (1,\ 1,\ -1)\}$.
   Then $(1,\ 0,\ 0) = \frac{1}{2}[(1,\ 1,\ -1) + (1,\ -1,\ 1)]$ is in U; similarly $(0,\ 1,\ 0)$ and $(0,\ 0,\ 1)$ are in U. As $\mathbb{R}^3 = \text{span}\ \{(1,\ 0,\ 0),\ (0,\ 1,\ 0),\ (0,\ 0,\ 1)\}$, we have $\mathbb{R}^3 \subseteq U$. Clearly $U \subseteq \mathbb{R}^3$, so we have $\mathbb{R}^3 = U$.

   d. **Independence**: If $r(1+x) + s(x+x^2) + t(x^2+x^3) + ux^3 = 0$ then

$$r + (r+s)x + (s+t)x^2 + (t+u)x^3 = 0$$

   so $r = 0$, $r+s = 0$, $s+t = 0$, $t+u = 0$. The only solution is $r = s = t = u = 0$.

   Spanning: Write $U = \text{span}\ \left\{1+x,\ x+x^2,\ x^2+x^3,\ x^3\right\}$. Then x^3 is in U; whence $x^2 = (x^2 + x^3) - x^3$ is in U; whence $x = (x+x^2) - x^2$ is in U; whence $1 = (1+x) - x$ is in U. Hence $\mathbf{P}_3 = \text{span}\ \left\{1,\ x,\ x^2,\ x^3\right\}$ is contained in U. As $U \subseteq \mathbf{P}_3$, we have $U = \mathbf{P}_3$.

6. b. Write $U = \left\{a + b(x+x^2) \mid a,\ b \text{ in } \mathbb{R}\right\} = \text{span}\, B$ where $B = \left\{1,\ x+x^2\right\}$. But B is independent because $s + t(x+x^2) = 0$ implies $s = t = 0$. Hence B is a basis of U, so $\dim U = 2$.

   d. Write $U = \{p(x) \mid p(x) = p(-x)\}$. As $U \subseteq \mathbf{P}_2$, write $p(x) = a + bx + cx^2$ be any member of U. The condition $p(x) = p(-x)$ becomes $a + bx + cx^2 = a - bx + cx^2$, so $b = 0$. Thus $U = \left\{a + bx^2 \mid a,\ b \text{ in } \mathbb{R}\right\} = \text{span}\ \left\{1,\ x^2\right\}$. As $\left\{1,\ x^2\right\}$ is independent ($s + tx^2 = 0$ implies $s = 0 = t$), it is a basis of U, so $\dim U = 2$.

7. b. Write $U = \left\{A \mid A \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} A\right\}$. If $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$, A is in U if and only if $\begin{bmatrix} x & y \\ z & w \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix}$, that is $\begin{bmatrix} x-y & x \\ z-w & z \end{bmatrix} = \begin{bmatrix} x+z & y+w \\ -x & -y \end{bmatrix}$.
   This holds if and only if $x = y+w$ and $z = -y$, that is

$$A = \begin{bmatrix} y+w & y \\ -y & w \end{bmatrix} = y\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} + w\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

   Hence $U = \text{span}\, B$ where $B = \left\{\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix},\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$. But B is independent here because $s\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} + t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ means $s+t = 0$, $s = 0$, $-s = 0$, $t = 0$, so $s = t = 0$. Thus B is a basis of U, so $\dim U = 2$.

   d. Write $U = \left\{A \mid A \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} A\right\}$. If $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ then A is in U if and only if $\begin{bmatrix} x & y \\ z & w \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix}$; that is $\begin{bmatrix} x-y & x \\ z-w & z \end{bmatrix} = \begin{bmatrix} z & w \\ z-x & w-y \end{bmatrix}$.
   This holds if and only if $z = x-y$ and $x = w$; that is

$$A = \begin{bmatrix} x & y \\ x-y & x \end{bmatrix} = x\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + y\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Thus $U = \text{span } B$ where $B = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$. But B is independent because $s \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + t \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ implies $s = t = 0$. Hence B is a basis of U, so $\dim U = 2$.

8. b. If $X = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ the condition $AX = X$ is $\begin{bmatrix} x+z & y+w \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ and this holds if and only if $z = w = 0$. Hence $X = \begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix} = x \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. So $U = \text{span } B$ where $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$. As B is independent, it is a basis of U, so $\dim U = 2$.

10. b. If the common column sum is m, V has the form
$V = \left\{ \begin{bmatrix} a & q & r \\ b & p & s \\ m-a-b & m-p-q & m-r-s \end{bmatrix} \mid a,\ b,\ p,\ q,\ r,\ s,\ m \text{ in } \mathbb{R} \right\} = \text{span } B$ where
$B = \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix} \right\}$.
The set B is independent (a linear combination using coefficients a, b, p, q, r, s, and m yields the matrix in V, and this is 0 if and only if $a = b = p = q = r = s = m = 0$). Hence B is a basis of B, so $\dim V = 7$.

11. b. A general polynomial in $\mathbf{P}_3$ has the form $p(x) = a + bx + cx^2 + dx^3$, so

$$\begin{aligned} V &= \left\{ (x^2 - x)(a + bx + cx^2 + dx^3) \mid a,\ b,\ c,\ d \text{ in } \mathbb{R} \right\} \\ &= \left\{ a(x^2 - x) + bx(x^2 - x) + cx^2(x^2 - x) + dx^3(x^2 - x) \mid a,\ b,\ c,\ d \text{ in } \mathbb{R} \right\} \\ &= \text{span } B \end{aligned}$$

where $B = \left\{ (x^2 - x),\ x(x^2 - x),\ x^2(x^2 - x),\ x^3(x^2 - x) \right\}$. We claim that B is independent. For if $a(x^2 - x) + bx(x^2 - x) + cx^2(x^2 - x) + dx^3(x^2 - x) = 0$ then $(a + bx + cx^2 + dx^3)(x^2 - x) = 0$, whence $a + bx + cx^2 + dx^3 = 0$ by the hint in (a). Thus $a = b = c = d = 0$. [This also follows by comparing coefficients.] Thus B is a basis of V, so $\dim V = 4$.

12. b. No. If $\mathbf{P}_3 = \text{span } \{f_1(x),\ f_2(x),\ f_3(x),\ f_4(x)\}$ where $f_i(0) = 0$ for each i, then each polynomial $p(x)$ in $\mathbf{P}_3$ is a linear combination

$$p(x) = a_1 f_1(x) + a_2 f_2(x) + a_3 f_3(x) + a_4 f_4(x)$$

when the a_i are in $\mathbb{R}$. But then

$$p(0) = a_1 f_1(0) + a_2 f_2(0) + a_3 f_3(0) + a_4 f_4(0) = 0$$

for every $p(x)$ in $\mathbf{P}_3$. This is not the case, so no such basis of $\mathbf{P}_3$ can exist. [Indeed, no such *spanning set* of $\mathbf{P}_3$ can exist.]

d. No. $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ is a basis of invertible matrices.
Independent: $r \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + s \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + t \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + u \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ gives
$r + s + t = 0$, $s + u = 0$, $t + u = 0$, $r + s + t + u = 0$. The only solution is $r = s = t = u = 0$.

Spanning: $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is in span B

$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is in span B

$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ is in span B

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is in span B

Hence $\mathbf{M}_{22} = \text{span}\left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\} \subseteq \text{span } B$. Clearly span $B \subseteq \mathbf{M}_{22}$.

f. Yes. Indeed, $0\mathbf{u}+0\mathbf{v}+0\mathbf{w}=\mathbf{0}$ for *any* $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, independent or not!

h. Yes. If $s\mathbf{u}+t(\mathbf{u}+\mathbf{v})=\mathbf{0}$ then $(s+t)\mathbf{u}+t\mathbf{v}=\mathbf{0}$, so $s+t=0$ and $t=0$ (because $\{\mathbf{u},\ \mathbf{v}\}$ is independent). Thus $s=t=0$.

j. Yes. If $s\mathbf{u}+t\mathbf{v}=\mathbf{0}$ then $s\mathbf{u}+t\mathbf{v}+0\mathbf{w}=\mathbf{0}$, so $s=t=0$ (because $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent). This shows that $\{\mathbf{u},\ \mathbf{v}\}$ is independent.

l. Yes. Since $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent, the vector $\mathbf{u}+\mathbf{v}+\mathbf{w}$ is not zero. Hence $\{\mathbf{u}+\mathbf{v}+\mathbf{w}\}$ is independent (see Example 5.2.5).

n. Yes. If I is a set of independent vectors, then $|I| \leq n$ by the fundamental theorem because V contains a spanning set of n vectors (any basis).

15. If a linear combination of the vectors in the subset vanishes, it is a linear combination of the vectors in the larger set (take the coefficients outside the subset to be zero). Since it still vanishes, all the coefficients are zero because the larger set is independent.

19. We have $s\mathbf{u}'+t\mathbf{v}=s(a\mathbf{u}+b\mathbf{v})+t(c\mathbf{u}+d\mathbf{v})=(sa+tc)\mathbf{u}+(sb+td)\mathbf{v}$. Since $\{\mathbf{u},\ \mathbf{v}\}$ is independent, we have

$$\begin{aligned} s\mathbf{u}'+t\mathbf{v}'=\mathbf{0} \quad &\text{if and only if} \quad sa+tc=0 \text{ and } sb+td=0 \\ &\text{if and only if} \quad \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

Hence $\{\mathbf{u}',\ \mathbf{v}'\}$ is independent if and only if $\begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ implies $\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

By Theorem 2.4.5, this is equivalent to A being invertible.

23. b. **Independent**: If $r(\mathbf{u}+\mathbf{v})+s(\mathbf{v}+\mathbf{w})+t(\mathbf{w}+\mathbf{u})=\mathbf{0}$ then $(r+t)\mathbf{u}+(r+s)\mathbf{v}+(s+t)\mathbf{w}+0\mathbf{z}=\mathbf{0}$. Thus $r+t=0$, $r+s=0$, $s+t=0$ (because $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w},\ \mathbf{z}\}$ is independent). Hence $r=s=t=0$.

d. **Dependent**: $(\mathbf{u}+\mathbf{v})-(\mathbf{v}+\mathbf{w})+(\mathbf{w}+\mathbf{z})-(\mathbf{z}+\mathbf{u})=\mathbf{0}$ is a nontrivial linear combination that vanishes.

26. If $rz+sz^2=0$, $r,\ s$ in $\mathbb{R}$, then $z(r+sz)=0$. If z is not real then $z \neq 0$ so $r+sz=0$. Thus $s=0$ (otherwise $z=\frac{-r}{s}$ is real), whence $r=0$. Conversely, if z is real then $rz+sz^2=0$ when $r=z$, $s=-1$, so $\{z,\ z^2\}$ is not independent.

29. b. If U is not invertible, let $U\mathbf{x}=\mathbf{0}$ where $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$ (Theorem 2.4.5). We claim that no set $\{A_1U,\ A_2U,\ \dots\}$ can span $\mathbf{M}_{mn}$ (let alone be a basis). For if it did, we could write any matrix B in $\mathbf{M}_{mn}$ as a linear combination

$$B = a_1A_1U + a_2A_2U + \cdots$$

Then $B\mathbf{x} = a_1AU\mathbf{x} + a_2A_2U\mathbf{x} + \cdots = \mathbf{0} + \mathbf{0} + \cdots = \mathbf{0}$, a contradiction. In fact, if entry k of $\mathbf{x}$ is nonzero, then $B\mathbf{x} \neq \mathbf{0}$ where all entries of B are zero except column k, which consists of 1's.

33. b. Suppose $U \cap W = 0$. If $s\mathbf{u} + t\mathbf{w} = \mathbf{0}$ with $\mathbf{u}$ and $\mathbf{w}$ nonzero in U and W, then $s\mathbf{u} = -t\mathbf{w}$ is in $U \cap W = \{\mathbf{0}\}$. Hence $s\mathbf{u} = \mathbf{0} = t\mathbf{w}$. So $s = 0 = t$ (as $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{w} \neq \mathbf{0}$). Thus $\{\mathbf{u}, \mathbf{v}\}$ is independent. Conversely, assume that the condition holds. If $\mathbf{v} \neq \mathbf{0}$ lies in $U \cap W$, then $\{\mathbf{v}, -\mathbf{v}\}$ is independent by the hypothesis, a contradiction because $1\mathbf{v} + 1(-\mathbf{v}) = \mathbf{0}$.

36. b. If $p(x) = a_0 + a_1x + \cdots + a_nx^n$ is in O_n, then $p(-x) = -p(x)$, so

$$a_0 - a_1x + a_2x^2 + a_3x^3 + a_4x^4 - \cdots = -a_0 - a_1x - a_2x^2 - a_3x^3 - a_4x^4 - \cdots$$

Hence $a_0 = a_2 = a_4 = \cdots = 0$ and $p(x) = a_1x + a_3x^3 + a_5x^5 + \cdots$. Thus $O_n = \text{span}\left\{x, x^3, x^5, \dots\right\}$ is spanned by the odd powers of x in $\mathbf{P}_n$. The set $B = \left\{x, x^3, x^5, \dots\right\}$ is independent (because $\left\{1, x, x^2, x^3, x^4, \dots\right\}$ is independent) so it is a basis of O_n. If n is even, $B = \left\{x, x^3, x^5, \dots, x^{n-1}\right\}$ has $\frac{n}{2}$ members, so $\dim O_n = \frac{n}{2}$. If n is odd, $B = \left\{x, x^3, x^5, \dots, x^n\right\}$ has $\frac{n+1}{2}$ members, so $\dim O_n = \frac{n+1}{2}$.

6.4 Finite Dimensional Spaces

1. b. $B = \{(1, 0, 0), (0, 1, 0), (0, 1, 1)\}$ is independent as $r(1, 0, 0) + s(0, 1, 0) + t(0, 1, 1) = (0, 0, 0)$ implies $r = 0$, $s + t = 0$, $t = 0$, whence $r = s = t = 0$. Hence B is a basis by Theorem 6.4.3 because $\dim \mathbb{R}^3 = 3$.

   d. $B = \left\{1, x, x^2 - x + 1\right\}$ is independent because $r1 + sx + t(x^2 - x - 1) = 0$ implies $r - t = 0$, $s - t = 0$, and $t = 0$; whence $r = s = t = 0$. Hence B is a basis by Theorem 6.4.3 because $\dim \mathbf{P}_2 = 3$.

2. b. As $\dim \mathbf{P}_2 = 3$, any independent set of three vectors is a basis by Theorem 6.4.3. But we have $-(x^2 + 3) + 2(x + 2) + (x^2 - 2x - 1) = 0$, $\left\{x^2 + 3, x + 2, x^2 - 2x - 1\right\}$, so is dependent. However any other subset of three vectors from $\left\{x^2 + 3, x + 2, x^2 - 2x - 1, x^2 + x\right\}$ is independent (verify).

3. b. $B = \{(0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 1, 1), (1, 1, 1, 1)\}$ spans $\mathbb{R}^4$ because

$$\begin{array}{ll} (1, 0, 0, 0) = (1, 1, 1, 1) - (0, 1, 0, 0) - (0, 0, 1, 1) & \text{is in span } B \\ (0, 0, 0, 1) = (0, 0, 1, 1) - (0, 0, 1, 0) & \text{is in span } B \end{array}$$

and, of course, $(0, 1, 0, 0)$ and $(0, 0, 1, 0)$ are in span B. Hence B is a basis of $\mathbb{R}^4$ by Theorem 6.4.3 because $\dim \mathbb{R}^4 = 4$.

   d. $B = \left\{1, x^2 + x, x^2 + 1, x^3\right\}$ spans $\mathbf{P}_3$ because $x^2 = (x^2 + 1) - 1$ and $x = (x^2 + x) - x^2$ are in span B (together with 1 and x^3). So B is a basis of $\mathbf{P}_3$ by Theorem 6.4.3 because $\dim \mathbf{P}_3 = 4$.

4. b. Let $z = a + bi$; a, b in $\mathbb{R}$. Then $b \neq 0$ as z is not real and $a \neq 0$ as z is not pure imaginary. Since $\dim \mathbb{C} = 2$, it suffices (by Theorem 6.4.3) to show that $\{z, \overline{z}\}$ is independent. If $rz + s\overline{z} = 0$ then $0 = r(a + bi) + s(a - bi) = (r + s)a + (r - s)bi$. Hence $(r + s)a = 0 = (r - s)b$ so (because $a \neq 0 \neq b$) $r + s = 0 = r - s$. Thus $r = s = 0$.

5. b. The four polynomials in S have distinct degrees. Use Example 6.3.4.

6. b. $\{4,\ 4x,\ 4x^2,\ 4x^3\}$ is such a basis. There is no basis of $\mathbf{P}_3$ consisting of polynomials have the property that their coefficients sum to zero. For if it did then *every* polynomial in $\mathbf{P}_3$ would have this property (since sums and scalar multiples of such polynomials have the same property).

7. b. Not a basis because $(2\mathbf{u}+\mathbf{v}+3\mathbf{w})-(3\mathbf{u}+\mathbf{v}-\mathbf{w})+(\mathbf{u}-4\mathbf{w})=\mathbf{0}$.

   d. Not a basis because $2\mathbf{u}-(\mathbf{u}+\mathbf{w})-(\mathbf{u}-\mathbf{w})+0(\mathbf{v}+\mathbf{w})=\mathbf{0}$.

8. b. Yes, four vectors can span $\mathbb{R}^3$ — say any basis together with any other vector.
   No, four vectors in $\mathbb{R}^3$ cannot be independent by the fundamental theorem (Theorem 6.3.2) because $\mathbb{R}^3$ is spanned by 3 vectors ($\dim \mathbb{R}^3 = 3$).

10. We have $\det A = 0$ if and only if A is not invertible. This holds if and only if the rows of A are dependent by Theorem 5.2.3. This in turn holds if and only if some row is a linear combination of the rest by the dependent lemma (Lemma 6.4.3).

11. b. No. Take $X = \{(0,\ 1),\ (1,\ 0)\}$ and $D = \{(0,\ 1),\ (1,\ 0),\ (1,\ 1)\}$. Then D is dependent, but its subset X is independent.

    d. Yes. This is follows from Exercise 15 Section 6.3 (solution above).

15. Let $\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m\}$, $m \leq k$, be a basis of U so $\dim U = m$. If $\mathbf{v} \in U$ then $W = U$ by Theorem 6.2.2, so certainly $\dim W = \dim U$. On the other hand, if $\mathbf{v} \notin U$ then $\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m,\ \mathbf{v}\}$ is independent by the independent lemma (Lemma 6.4.1). Since $W = \operatorname{span}\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m,\ \mathbf{v}\}$, again by Theorem 6.2.2, it is a basis of W and so $\dim W = 1 + \dim U$.

18. b. The two-dimensional subspaces of $\mathbb{R}^3$ are the planes through the origin, and the one-dimensional subspaces are the lines through the origin. Hence part (a) asserts that if U and W are distinct planes through the origin, then $U \cap W$ is a line through the origin.

23. b. Let $\mathbf{v}_n$ denote the sequence with 1 in the n^{th} coordinate and zeros elsewhere. Thus $\mathbf{v}_0 = (1,\ 0,\ 0,\ \dots)$, $\mathbf{v}_1 = (0,\ 1,\ 0,\ \dots)$ etc. Then $a_0\mathbf{v}_0 + a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = (a_0,\ a_1,\ \dots,\ a_n,\ 0,\ 0,\ \dots)$ so $a_0\mathbf{v}_0 + a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = 0$ implies $a_0 = a_1 = \cdots = a_n = 0$. Thus $\{\mathbf{v}_0,\ \mathbf{v}_1,\ \dots,\ \mathbf{v}_n\}$ is an independent set of $n+1$ vectors. Since n is arbitrary, $\dim V$ cannot be finite by the fundamental theorem.

25. b. Observe that $\mathbb{R}\mathbf{u} = \{s\mathbf{u} \mid s \text{ in } \mathbb{R}\}$. Hence $\mathbb{R}\mathbf{u} + \mathbb{R}\mathbf{v} = \{s\mathbf{u} + t\mathbf{v} \mid s \text{ in } \mathbb{R},\ t \text{ in } \mathbb{R}\}$ is the set of all linear combinations of $\mathbf{u}$ and $\mathbf{v}$. But this is the definition of $\operatorname{span}\{\mathbf{u},\ \mathbf{v}\}$.

6.5 An Application to Polynomials

2. b. $f^{(0)}(x) = f(x) = x^3 + x + 1$, so $f^{(1)}(x) = 3x^2 + 1$, $f^{(2)}(x) = 6x$, $f^{(3)}(x) = 6$. Hence, Taylor's theorem gives

$$\begin{aligned} f(x) &= f^{(0)}(1) + f^{(1)}(1)(x-1) + \tfrac{f^{(2)}(1)}{2!}(x-1)^2 + \tfrac{f^{(3)}(1)}{3!}(x-1)^3 \\ &= 3 + 4(x-1) + 3(x-1)^2 + (x-1)^3 \end{aligned}$$

d. $f^{(0)}(x) = f(x) = x^3 - 3x^2 + 3x$, $f^{(1)}(x) = 3x^2 - 6x + 3$, $f^{(2)}(x) = 6x - 6$, $f^{(3)}(x) = 6$. Hence, Taylor's theorem gives

$$\begin{aligned} f(x) &= f^{(0)}(1) + f^{(1)}(1)(x-1) + \tfrac{f^{(2)}(1)}{2!}(x-1)^2 + \tfrac{f^{(3)}(1)}{3!}(x-1)^3 \\ &= 1 + 0(x-1) + \tfrac{0}{2!}(x-1)^2 + 1(x-1)^3 \\ &= 1 + (x-1)^3 \end{aligned}$$

6. b. The three polynomials are $x^2 - 3x + 2 = (x-1)(x-2)$, $x^2 - 4x + 3 = (x-1)(x-3)$ and $x^2 - 5x + 6 = (x-2)(x-3)$, so use $a_0 = 3$, $a_1 = 2$, $a_2 = 1$, in Theorem 6.5.2.

7. b. The Lagrange polynomials for $a_0 = 1$, $a_1 = 2$, $a_2 = 3$, are

$$\begin{aligned} \delta_0(x) &= \tfrac{(x-2)(x-3)}{(1-2)(1-3)} = \tfrac{1}{2}(x-2)(x-3) \\ \delta_1(x) &= \tfrac{(x-1)(x-3)}{(2-1)(2-3)} = -(x-1)(x-3) \\ \delta_2(x) &= \tfrac{(x-1)(x-2)}{(3-1)(3-2)} = \tfrac{1}{2}(x-1)(x-2) \end{aligned}$$

Given $f(x) = x^2 + x + 1$:

$$\begin{aligned} f(x) &= f(1)\delta_0(x) + f(2)\delta_1(x) + f(3)\delta_2(x) \\ &= \tfrac{3}{2}(x-2)(x-3) - 7(x-1)(x-3) + \tfrac{13}{2}(x-1)(x-2) \end{aligned}$$

10. b. If $r(x-a)^2 + s(x-a)(x-b) + t(x-b)^2 = 0$, then taking $x = a$ gives $t(a-b)^2 = 0$, so $t = 0$ because $a \neq b$; and taking $x = b$ gives $r(b-a)^2 = 0$, so $r = 0$. Thus, we are left with $s(x-a)(x-b) = 0$. If x is any number except a, b, this implies $s = 0$. Thus $B = \left\{(x-a)^2,\ (x-a)(x-b),\ (x-b)^2\right\}$ is independent in $\mathbf{P}_2$; since $\dim \mathbf{P}_2 = 3$, B is a basis.

11. b. Have $U_n = \{f(x) \text{ in } \mathbf{P}_n \mid f(a) = 0 = f(b)\}$. Let $\{p_1(x), \dots, p_{n-1}(x)\}$ be a basis of $\mathbf{P}_{n-2}$; it suffices to show that

$$B = \{(x-a)(x-b)p_1(x),\ \dots,\ (x-a)(x-b)p_{n-1}(x)\}$$

is a basis of U_n. Clearly $B \subseteq U_n$.

Independent: Let $s_1(x-a)(x-b)p_1(x) + \cdots + s_{n-1}(x-a)(x-b)p_{n-1}(x) = 0$. Then $(x-a)(x-b)[s_1p_1(x) + \cdots + s_{n-1}p_{n-1}(x)] = 0$, so (by the hint) $s_1p_1(x) + \cdots + s_{n-1}p_{n-1}(x) = 0$. Thus $s_1 = s_2 = \cdots = s_{n-1} = 0$.

Spanning: Given $f(x)$ in $\mathbf{P}_n$ with $f(a)=0$, we have $f(x)=(x-a)g(x)$ for some polynomial $g(x)$ in $\mathbf{P}_{n-1}$ by the factor theorem. But $0=f(b)=(b-a)g(b)$ so (as $b\neq a$) $g(b)=0$. Then $g(x)=(x-b)h(x)$ with $h(x)=r_1p_1(x)+\cdots+r_{n-1}p_{n-1}(x)$, r_i in $\mathbb{R}$, whence

$$\begin{aligned}f(x)&=(x-a)g(x)\\&=(x-a)(x-b)g(x)\\&=(x-a)(x-b)[r_1p_1(x)+\cdots+r_{n-1}p_{n-1}(x)]\\&=r_1(x-a)(x-b)p_1(x)+\cdots+r_{n-1}(x-a)(x-b)p_{n-1}(x)\end{aligned}$$

6.6 An Application to Differential Equations

1. b. By Theorem 6.6.1, $f(x)=ce^{-x}$ for some constant c. We have $1=f(1)=ce^{-1}$, so $c=e$. Thus $f(x)=e^{1-x}$.

   d. The characteristic polynomial is $x^2+x-6=(x-2)(x+3)$. Hence $f(x)=ce^{2x}+de^{-3x}$ for some c, d. We have $0=f(0)=c+d$ and $1=f(1)=ce^2+de^{-3}$. Hence, $d=-c$ and $c=\frac{1}{e^2-e^{-3}}$ so $f(x)=\frac{e^{2x}-e^{-3x}}{e^2-e^{-3}}$.

   f. The characteristic polynomial is $x^2-4x+4=(x-2)^2$. Hence, $f(x)=ce^{2x}+dxe^{2x}=(c+dx)e^{2x}$ for some c, d. We have $2=f(0)=c$ and $0=f(-1)=(c-d)e^{-2}$. Thus $c=d=2$ and $f(x)=2(1+x)e^{2x}$.

   h. The characteristic polynomial is $x^2-a^2=(x-a)(x+a)$, so (as $a\neq -a$) $f(x)=ce^{ax}+de^{-ax}$ for some c, d. We have $1=f(0)=c+d$ and $0=f(1)=ce^a+de^{-a}$. Thus $d=1-c$ and $c=\frac{1}{1-e^{2a}}$ whence

   $$f(x)=c^{ax}+(1-c)e^{-ax}=\frac{e^{ax}-e^{a(2-x)}}{1-e^{2a}}$$

   j. The characteristic polynomial is x^2+4x+5. The roots are $\lambda=-2\pm i$, so

   $$f(x)=e^{-2x}(c\sin x+d\cos x)\text{ for some real }c\text{ and }d.$$

   We have $0=f(0)=d$ and $1=f\left(\frac{\pi}{2}\right)=e^{-\pi}(c)$. Hence $f(x)=e^{\pi-2x}\sin x$.

4. b. If $f(x)=g(x)+2$ then $f'+f=2$ becomes $g'+g=0$, whence $g(x)=ce^{-x}$ for some c. Thus $f(x)=ce^{-x}+2$ for some constant c.

5. b. If $f(x)=-\frac{x^3}{3}$ then $f'(x)=-x^2$ and $f''(x)=-2x$, so

   $$f''(x)+f'(x)-6f(x)=-2x-x^2+2x^3$$

   Hence, $f(x)=\frac{-x^3}{3}$ is a particular solution. Now, if $h=h(x)$ is any solution, write $g(x)=h(x)-f(x)=h(x)+\frac{x^3}{3}$. Then

   $$g''+g'-6g=(h'+h'-6h)-(f''+f'-6f)=0$$

   So, to find g, the characteristic polynomial is $x^2+x-6=(x-2)(x+3)$. Hence we have $g(x)=ce^{-3x}+de^{2x}$, where c and d are constants, so

   $$h(x)=ce^{-3x}+de^{2x}-\frac{x^3}{3}$$

6. b. The general solution is $m(t) = 10(\frac{4}{5})^{1/3}$. Hence $10(\frac{4}{5})^{t/3} = 5$ so $t = \frac{3\ln(1/2)}{\ln(4/5)} = 9.32$ hours.

7. b. If $m = m(t)$ is the mass at time t, then the rate $m'(t)$ of decay is proportional to $m(t)$, that is $m'(t) = km(t)$ for some k. Thus, $m' - km = 0$ so $m = ce^{kt}$ for some constant c. Since $m(0) = 10$, we obtain $c = 10$, whence $m(t) = 10e^{kt}$. Also, $8 = m(3) = 10e^{3k}$ so $e^{3k} = \frac{4}{5}$, $e^k = \left(\frac{4}{5}\right)^{1/3}$, $m(t) = 10(e^k)^t = 10\left(\frac{4}{5}\right)^{t/3}$.

9. In Example 6.6.4, we found that the period of oscillation is $\frac{2\pi}{\sqrt{k}}$. Hence $\frac{2\pi}{\sqrt{k}} = 30$ so we obtain $k = \left(\frac{\pi}{15}\right)^2 = 0.044$.

Supplementary Exercises: Chapter 6

2. b. Suppose $\{Ax_1, \dots, Ax_n\}$ is a basis of $\mathbb{R}^n$. To show that A is invertible, we show that $YA = 0$ implies $Y = 0$. (This shows A^T is invertible by Theorem 2.4.4, so A is invertible). So assume that $YA = 0$. Let $\mathbf{c}_1, \dots, \mathbf{c}_m$ denote the columns of I_m, so $I_m = \left[\begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_m \end{array}\right]$. Then $Y = YI_m = Y\left[\begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_m \end{array}\right] = \left[\begin{array}{cccc} Y\mathbf{c}_1 & Y\mathbf{c}_2 & \cdots & Y\mathbf{c}_m \end{array}\right]$, so it suffices to show that $Y\mathbf{c}_j = \mathbf{0}$ for each j. But $\mathbf{c}_j$ is in $\mathbb{R}^n$ so our hypothesis shows that $\mathbf{c}_j = r_1A\mathbf{v}_1 + \cdots + r_nA\mathbf{v}_n$ for some r_j in $\mathbb{R}$. Hence,
$$\mathbf{c}_j = A(r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n)$$
so $Y\mathbf{c}_j = YA(r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n) = \mathbf{0}$, as required.

4. Assume that A is $m \times n$. If $\mathbf{x}$ is in null A, then $A\mathbf{x} = \mathbf{0}$ so $(A^TA)\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$. Thus $\mathbf{x}$ is in null A^TA, so null $A \subseteq$ null A^TA. Conversely, let $\mathbf{x}$ be in null A^TA; that is $A^TA\mathbf{x} = \mathbf{0}$. Write
$$A\mathbf{x} = \mathbf{y} = \left[\begin{array}{c} y_1 \\ \vdots \\ y_m \end{array}\right]$$
Then $y_1^2 + y_2^2 + \cdots + y_m^2 = \mathbf{y}^T\mathbf{y} = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$. Since the y_i are real numbers, this implies that $y_1 = y_2 = \cdots = y_m = 0$; that is $\mathbf{y} = \mathbf{0}$, that is $A\mathbf{x} = \mathbf{0}$, that is $\mathbf{x}$ is in null A.

7. Linear Transformations

7.1 Examples and Elementary Properties

1. b. $T(X) = XA$ where $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ and $X = (x,\ y,\ z)$ is thought of as a row matrix. Hence, matrix algebra gives $T(X+Y) = A(X+Y) = AX + AY = T(X) + T(Y)$ and $T(r\lambda) = A(rX) = rA(X) = rT(X)$.

   d. $T(A+B) = P(A+B)Q = PAQ + PBQ = T(A) + T(B)$; $T(rA) = P(rA)Q = rPAQ = rT(A)$.

   f. Here $T[p(x)] = p(0)$ for all polynomials $p(x)$ in $\mathbf{P}_n$. Thus

$$\begin{aligned} T[(p+q)(x)] &= T[p(x)+q(x)] = p(0)+q(0) = T[p(x)] + T[q(x)] \\ T[rp(x)] &= rp(0) = r[Tp(x)] \end{aligned}$$

   h. Here $\mathbf{z}$ is fixed in $\mathbb{R}^n$ and $T(\mathbf{x}) = \mathbf{x}\cdot\mathbf{z}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. We use Theorem 5.3.1:

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= (\mathbf{x}+\mathbf{y})\cdot\mathbf{z} = \mathbf{x}\cdot\mathbf{z} + \mathbf{y}\cdot\mathbf{z} = T(\mathbf{x}) + T(\mathbf{y}) \\ T(r\mathbf{x}) &= (r\mathbf{x})\cdot\mathbf{z} = r(\mathbf{x}\cdot\mathbf{z}) = rT(\mathbf{x}) \end{aligned}$$

   j. If $\mathbf{v} = (r_1 \cdots r_n)$ and $\mathbf{w} = (s_1 \cdots s_n)$ then, $\mathbf{v}+\mathbf{w} = (r_1+s_1 \cdots r_n+s_n)$. Hence:

$$\begin{aligned} T(\mathbf{v}+\mathbf{w}) &= (r_1+s_1)\mathbf{e}_1 + \cdots + (r_n+s_n)\mathbf{e}_n \\ &= (r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n) + (s_1\mathbf{e}_1 + \cdots + s_n\mathbf{e}_n) = T(\mathbf{v}) + T(\mathbf{w}) \end{aligned}$$

   Similarly, for a in $\mathbb{R}$, we have $a\mathbf{v} = (ar_1 \cdots ar_n)$ so

$$T(a\mathbf{v}) = (ar_1)\mathbf{e}_1 + \cdots + (ar_n)\mathbf{e}_n = a(r_1\mathbf{e}_1 + \cdots + r_n\mathbf{e}_n) = aT(\mathbf{v})$$

2. b. Let $A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$, then $A+B = \begin{bmatrix} 2 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$. Thus, $T(A) = \text{rank } A = 2$, $T(B) = \text{rank } B = 2$ and $T(A+B) = \text{rank } (A+B) = 1$. Thus $T(A+B) \neq T(A) + T(B)$.

   d. Here $T(\mathbf{v}) = \mathbf{v}+\mathbf{u}$, $T(\mathbf{w}) = \mathbf{w}+\mathbf{u}$, and $T(\mathbf{v}+\mathbf{w}) = \mathbf{v}+\mathbf{w}+\mathbf{u}$. Thus if $T(\mathbf{v}+\mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$ then $\mathbf{v}+\mathbf{w}+\mathbf{u} = (\mathbf{v}+\mathbf{u}) + (\mathbf{w}+\mathbf{u})$, so $\mathbf{u} = 2\mathbf{u}$, $\mathbf{u} = \mathbf{0}$. This is contrary to assumption. Alternatively, $T(\mathbf{0}) = \mathbf{0}+\mathbf{u} \neq \mathbf{0}$, so T cannot be linear by Theorem 7.1.1.

3. b. Because T is linear, $T(3\mathbf{v}_1 + 2\mathbf{v}_2) = 3T(\mathbf{v}_1) + 2T(\mathbf{v}_2) = 3(2) + 2(-3) = 0$.

d. Since we know the action of T on $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, it suffices to express $\begin{bmatrix} 1 \\ -7 \end{bmatrix}$ as a linear combination of these vectors.

$$\begin{bmatrix} 1 \\ -7 \end{bmatrix} = r\begin{bmatrix} 1 \\ -1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Comparing components gives $1 = r+s$ and $-7 = -r+s$. The solution is $r = 4$, $s = -3$, so

$$T\begin{bmatrix} 1 \\ -7 \end{bmatrix} = T\left(4\begin{bmatrix} 1 \\ -1 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 4T\begin{bmatrix} 1 \\ -1 \end{bmatrix} - 3T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 0 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

f. We know $T(1)$, $T(x+2)$ and $T(x^2+x)$, so we express $2-x+3x^2$ as a linear combination of these vectors:

$$2-x+3x^2 = r\cdot 1 + s(x+2) + t(x^2+x)$$

Equating coefficients gives $2 = r+2s$, $-1 = s+t$ and $3 = t$. The solution is $r = 10$, $s = -4$ and $t = 3$, so

$$\begin{aligned} T(2-x+3x^2) &= T[r\cdot 1 + s(x+2) + t(x^2+x)] \\ &= rT(1) + sT(x+2) + tT(x^2+x) \\ &= 5r + s + 0 \\ &= 46 \end{aligned}$$

In fact, we can find the action of T on any vector $a+bx+cx^2$ in the same way. Observe that

$$a+bx+cx^2 = (a-2b+2c)\cdot 1 + (b-c)(x+2) + c(x^2+x)$$

for any a, b and c, so

$$\begin{aligned} T(a+bx+cx^2) &= (a-2b+2c)T(1) + (b-c)T(x+2) + cT(x^2+x) \\ &= (a-2b+2c)\cdot 5 + (b-c)\cdot 1 + c\cdot 0 \\ &= 5a-9b+9c \end{aligned}$$

This retrieves the above result when $a = 2$, $b = -1$ and $c = 3$.

4. b. Since $B = \{(2, -1), (1, 1)\}$ is a basis of $\mathbb{R}^2$, *any* vector (x, y) in $\mathbb{R}^2$ is a linear combination $(x, y) = r(2, -1) + s(1, 1)$. Indeed, equating components gives $x = 2r+s$ and $y = -r+s$ so $r = \frac{1}{3}(x-y)$, $s = \frac{1}{3}(x+2y)$. Hence,

$$\begin{aligned} T(x, y) &= T[r(2, -1) + s(1, 1)] \\ &= rT(2, -1) + sT(1, 1) \\ &= \tfrac{1}{3}(x-y)(1, -1, 1) + \tfrac{1}{3}(x+2y)(0, 1, 0) \\ &= \left(\tfrac{1}{3}(x-y), y, \tfrac{1}{3}(x-y)\right) \\ &= \tfrac{1}{3}(x-y, 3y, x-y) \end{aligned}$$

In particular, $T(\mathbf{v}) = T(-1, 2) = \frac{1}{3}(-3, 6, -3) = (-1, 2, -1)$.

This works in general. Observe that $(x, y) = \frac{x-y}{3}(2, -1) + \frac{x+2y}{3}(1, 1)$ for any x and y, so since T is linear,

$$T(x, y) = \tfrac{x-y}{3}T(2, -1) + \tfrac{x+2y}{3}T(1, 1)$$

for any choice of $T(2, -1)$ and $T(1, 1)$.

d. Since $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis of $\mathbf{M}_{22}$, every vector $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a linear combination

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = r\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + s\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + t\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + u\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Indeed, equating components and solving for r, s, t and u gives $r = a - c + b$, $s = b$, $t = c - b$, $u = d$. Thus,

$$\begin{aligned} T\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= rT\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + sT\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + tT\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + uT\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= (a-c+b)\cdot 3 + b\cdot(-1) + (c-b)\cdot 0 + d\cdot 0 \\ &= 3a + 2b - 3c \end{aligned}$$

5. b. Since T is linear, the given conditions read

$$\begin{aligned} T(\mathbf{v}) + 2T(\mathbf{w}) &= 3\mathbf{v} - \mathbf{w} \\ T(\mathbf{v}) - T(\mathbf{w}) &= 2\mathbf{v} - 4\mathbf{w} \end{aligned}$$

Add twice the second equation to the first to get $3T(\mathbf{v}) = 7\mathbf{v} - 9\mathbf{w}$, $T(\mathbf{v}) = \frac{7}{3}\mathbf{v} - 3\mathbf{w}$. Similarly, subtracting the second from the first gives $3T(\mathbf{w}) = \mathbf{v} + 3\mathbf{w}$, $T(\mathbf{w}) = \frac{1}{3}\mathbf{v} + \mathbf{w}$. [Alternatively, we can use gaussian elimination with constants $3\mathbf{v} - \mathbf{w}$ and $2\mathbf{v} - 4\mathbf{w}$.]

8. b. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of V, every vector $\mathbf{v}$ in V is a unique linear combination $\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$, r_i in $\mathbb{R}$. Hence, as T is linear,

$$T(\mathbf{v}) = r_1T(\mathbf{v}_1) + \cdots + r_nT(\mathbf{v}_n) = r_1(-\mathbf{v}_1) + \cdots + r_n(-\mathbf{v}_n) = -\mathbf{v} = (-1)\mathbf{v}$$

Since this holds for every $\mathbf{v}$ in V, it shows that $T = -1$, the scalar operator.

12. $\{1\}$ is a basis of the vector space $\mathbb{R}$. If $T : \mathbb{R} \to V$ is a linear transformation, write $T(1) = \mathbf{v}$. Then, for all r in $\mathbb{R}$:

$$T(r) = T(r\cdot 1) = rT(1) = r\mathbf{v}$$

Since $T(r) = r\mathbf{v}$ is linear for each $\mathbf{v}$ in V, this shows that every linear transformation $T : \mathbb{R} \to V$ arises in this way.

15. b. Write $U = \{\mathbf{v} \in V \mid T(\mathbf{v}) \in P\}$. If $\mathbf{v}$ and $\mathbf{v}_1$ are in U, then $T(\mathbf{v})$ and $T(\mathbf{v}_1)$ are in P. As P is a subspace, it follows that $T(\mathbf{v} + \mathbf{v}_1) = T(\mathbf{v}) + T(\mathbf{v}_1)$ and $T(r\mathbf{v}) = rT(\mathbf{v})$ are both in P; that is $\mathbf{v} + \mathbf{v}_1$ and $r\mathbf{v}$ are in U. Since $\mathbf{0}$ is in U—because $T(\mathbf{0}) = \mathbf{0}$ is in P—it follows that U is a subspace.

18. Assume that $\{\mathbf{v}, T(\mathbf{v})\}$ is independent. Then $T(\mathbf{v}) \neq \mathbf{v}$ (or else $1\mathbf{v} + (-1)T(\mathbf{v}) = \mathbf{0}$) and similarly $T(\mathbf{v}) \neq -\mathbf{v}$.

Conversely, assume that $T(\mathbf{v}) \neq \mathbf{v}$ and $T(\mathbf{v}) \neq -\mathbf{v}$. To verify that $\{\mathbf{v}, T(\mathbf{v})\}$ is independent, let $r\mathbf{v} + sT(\mathbf{v}) = \mathbf{0}$; we must show that $r = s = 0$. If $s \neq 0$, then $T(\mathbf{v}) = a\mathbf{v}$ where $a = -\frac{r}{s}$. Hence $\mathbf{v} = T[T(\mathbf{v})] = T(a\mathbf{v}) = aT(\mathbf{v}) = a^2\mathbf{v}$. Since $\mathbf{v} \neq \mathbf{0}$, this gives $a = \pm 1$, contrary to hypothesis. So $s = 0$, whence $r\mathbf{v} = \mathbf{0}$ and $r = 0$.

21. b. Suppose that $T : \mathbf{P}_n \to \mathbb{R}$ is a linear transformation such that $T(x^k) = T(x)^k$ holds for all $k \geq 0$ (where $x^0 = 1$). Write $T(x) = a$. We have $T(x^k) = T(x)^k = a^k = E_a(x^k)$ for each k by assumption. This gives $T = E_a$ by Theorem 7.1.2 because $\left\{1, x, x^2, \dots, x^i, \dots, x^n\right\}$ is a basis of $\mathbf{P}_n$.

7.2 Kernel and Image of a Linear Transformation

1. b. We have $\ker T_A = \{\mathbf{x} \mid A\mathbf{x} = 0\}$; to determine this space, we use gaussian elimination:

$$\left[\begin{array}{rrrr|r} 2 & 1 & -1 & 3 & 0 \\ 1 & 0 & 3 & 1 & 0 \\ 1 & 1 & -4 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 3 & 1 & 0 \\ 0 & 1 & -7 & 1 & 0 \\ 0 & 1 & -7 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 3 & 1 & 0 \\ 0 & 1 & -7 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence $\ker T_A = \left\{ \left[\begin{array}{c} -3s-t \\ 7s-t \\ s \\ t \end{array}\right] \middle| s,\ t \text{ in } \mathbb{R} \right\} = \operatorname{span}\left\{ \left[\begin{array}{r} -3 \\ 7 \\ 1 \\ 0 \end{array}\right], \left[\begin{array}{r} 1 \\ 1 \\ 0 \\ -1 \end{array}\right] \right\}$. These vectors are independent so nullity of $T_A = \dim(\ker T_A) = 2$. Next

$$\begin{aligned} \operatorname{im} T_A &= \{A\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^4\} \\ &= \left\{ \left[\begin{array}{rrrr} 2 & 1 & -1 & 3 \\ 1 & 0 & 3 & 1 \\ 1 & 1 & -4 & 2 \end{array}\right] \left[\begin{array}{c} r \\ s \\ t \\ y \end{array}\right] \middle| r,\ s,\ t,\ u \text{ in } \mathbb{R} \right\} \\ &= \left\{ r\left[\begin{array}{r} 2 \\ 1 \\ 1 \end{array}\right] + s\left[\begin{array}{r} 1 \\ 0 \\ 1 \end{array}\right] + t\left[\begin{array}{r} -1 \\ 3 \\ -4 \end{array}\right] + u\left[\begin{array}{r} 3 \\ 1 \\ 2 \end{array}\right] \middle| r,\ s,\ t,\ u \text{ in } \mathbb{R} \right\} \end{aligned}$$

Thus $\operatorname{im} T_A = \operatorname{col} A$ as is true in general. Hence $\dim(\operatorname{im} T_A) = \dim(\operatorname{col} A) = \operatorname{rank} A$, and we can compute this by carrying A to row-echelon form:

$$\left[\begin{array}{rrrr} 2 & 1 & -1 & 3 \\ 1 & 0 & 3 & 1 \\ 1 & 1 & -4 & 2 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & 0 & 3 & 1 \\ 0 & 1 & -7 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $\dim(\operatorname{im} T_A) = \operatorname{rank} A = 2$. However, we want a basis of $\operatorname{col} A$, and we obtain this by writing the columns of A as rows and carrying the resulting matrix (it is A^T) to row-echelon form:

$$\left[\begin{array}{rrr} 2 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 3 & -4 \\ 3 & 1 & 2 \end{array}\right] \rightarrow \left[\begin{array}{rrr} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 3 & -3 \\ 0 & 1 & -1 \end{array}\right] \rightarrow \left[\begin{array}{rrr} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Hence, by Lemma 5.4.2, $\left\{ \left[\begin{array}{r} 1 \\ 0 \\ 1 \end{array}\right], \left[\begin{array}{r} 0 \\ 1 \\ -1 \end{array}\right] \right\}$ is a basis of $\operatorname{im} T_A = \operatorname{col} A$. Of course this once again shows that $\operatorname{rank} T_A = \dim(\operatorname{col} A) = 2$.

d. $\ker T_A = \{\mathbf{x} \mid A\mathbf{x} = 0\}$ so, as in (b), we use gaussian elimination:

$$\left[\begin{array}{rrr} 2 & 1 & 0 \\ 1 & -1 & 3 \\ 1 & 2 & -3 \\ 0 & 3 & -6 \end{array}\right] \rightarrow \left[\begin{array}{rrr} 1 & -1 & 3 \\ 0 & 3 & -6 \\ 0 & 3 & -6 \\ 0 & 3 & -6 \end{array}\right] \rightarrow \left[\begin{array}{rrr} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Hence, $\ker T_A = \left\{ \left[\begin{array}{c} -t \\ 2t \\ t \end{array}\right] \middle| t \text{ in } \mathbb{R} \right\} = \operatorname{span}\left\{ \left[\begin{array}{r} -1 \\ 2 \\ 1 \end{array}\right] \right\}$. Thus the nullity of T_A is $\dim(\ker T_A) = 1$. As in (b), $\operatorname{im} T_A = \operatorname{col} A$ and we find a basis by doing gaussian elimination on A^T:

$$\left[\begin{array}{rrrr} 2 & 1 & 1 & 0 \\ 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 3 & -3 & -6 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Hence, $\operatorname{im} T_A = \operatorname{col} A = \operatorname{span}\left\{ \left[\begin{array}{r} 1 \\ 0 \\ 1 \\ 1 \end{array}\right], \left[\begin{array}{r} 0 \\ 1 \\ -1 \\ -2 \end{array}\right] \right\}$, so $\operatorname{rank} T_A = \dim(\operatorname{im} T_A) = 2$.

2. b. Here $T = \mathbf{P}_2 \to \mathbb{R}^2$ given by $T(p(x)) = \left[\begin{array}{cc} p(0) & p(1) \end{array}\right]$. Hence

$$\ker T = \{p(x) \mid p(0) = p(1) = 0\}$$

If $p(x) = a + bx + cx^2$ is in $\ker T$, then $0 = p(0) = a$ and $0 = p(1) = a + b + c$. This means that $p(x) = bx - bx^2$, and so $\ker T = \text{span}\left\{x - x^2\right\}$. Thus $\left\{x - x^2\right\}$ is a basis of $\ker T$. Next, $\text{im } T$ is a subspace of $\mathbb{R}^2$. We have $(1,\ 0) = T(1 - x)$ and $(0,\ 1) = T(x)$ are both in $\text{im } T$, so $\text{im } T = \mathbb{R}^2$. Thus $\{(1,\ 0),\ (0,\ 1)\}$ is a basis of $\text{im } T$.

d. Here $T : \mathbb{R}^3 \to \mathbb{R}^4$ given by $T(x,\ y,\ z) = (x,\ x,\ y,\ y)$. Thus,

$$\ker T = \{(x,\ y,\ z) \mid (x,\ x,\ y,\ y) = (0,\ 0,\ 0,\ 0)\} = \{(0,\ 0,\ z) \mid z \text{ in } \mathbb{R}\} = \text{span}\ \{(0,\ 0,\ 1)\}$$

Hence, $\{(0,\ 0,\ 1)\}$ is a basis of $\ker T$. On the other hand,

$$\text{im } T = \{(x,\ x,\ y,\ y) \mid x,\ y \text{ in } \mathbb{R}\} = \text{span}\ \{(1,\ 1,\ 0,\ 0),\ (0,\ 0,\ 1,\ 1)\}$$

Then $\{(1,\ 1,\ 0,\ 0),\ (0,\ 0,\ 1,\ 1)\}$ is a basis of $\text{im } T$.

f. Here $T : \mathbf{M}_{22} \to \mathbb{R}$ is given by $T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = a + d$. Hence

$$\begin{aligned} \ker T &= \left\{\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] \middle|\, a + d = 0\right\} = \left\{\left[\begin{array}{cc} a & b \\ c & -a \end{array}\right] \middle|\, a,\ b,\ c \text{ in } \mathbb{R}\right\} \\ &= \text{span}\left\{\left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]\right\} \end{aligned}$$

Hence, $\left\{\left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]\right\}$ is a basis of $\ker T$ (being independent). On the other hand,

$$\text{im } T = \left\{a + d \,\middle|\, \left[\begin{array}{cc} a & b \\ c & d \end{array}\right] \text{ in } \mathbf{M}_{22}\right\} = \mathbb{R}$$

So $\{1\}$ is a basis of $\text{im } T$.

h. $T : \mathbb{R}^n \to \mathbb{R}$, $T(r_1,\ r_2,\ \dots,\ r_n) = r_1 + r_2 + \cdots + r_n$. Hence,

$$\begin{aligned} \ker T &= \{(r_1,\ r_2,\ \dots,\ r_n) \mid r_1 + r_2 + \cdots + r_n = 0\} \\ &= \{(r_1,\ r_2,\ \dots,\ r_{n-1},\ -r_1 - \cdots - r_{n-1}) \mid r_i \text{ in } \mathbb{R}\} \\ &= \text{span}\ \{(1,\ 0,\ 0,\ \dots,\ -1),\ (0,\ 1,\ 0,\ \dots,\ -1),\ \dots,\ (0,\ 0,\ 1,\ \dots,\ -1)\} \end{aligned}$$

This is a basis of $\ker T$. On the other hand,

$$\text{im } T = \{r_1 + \cdots + r_n \mid (r_1,\ r_2,\ \dots,\ r_n) \text{ is in } \mathbb{R}^n\} = \mathbb{R}$$

Thus $\{1\}$ is a basis of $\text{im } T$.

j. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ is given by $T(X) = XA$ where $A = \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right]$. Writing $X = \left[\begin{array}{cc} x & y \\ z & w \end{array}\right]$:

$$\begin{aligned} \ker T = \{X \mid XA = 0\} &= \left\{\left[\begin{array}{cc} x & y \\ z & w \end{array}\right] \middle|\, \left[\begin{array}{cc} x & x \\ z & x \end{array}\right] = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right]\right\} = \left\{\left[\begin{array}{cc} 0 & y \\ 0 & w \end{array}\right] \middle|\, y,\ w \text{ in } \mathbb{R}\right\} \\ &= \text{span}\left\{\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\right\} \end{aligned}$$

Thus, $\left\{\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\right\}$ is a basis of $\ker T$ (being independent). On the other hand,

$$\text{im } T = \{XA \mid X \text{ in } \mathbf{M}_{22}\} = \left\{\left[\begin{array}{cc} x & x \\ z & z \end{array}\right] \middle|\, x,\ z \text{ in } \mathbb{R}\right\} = \text{span}\left\{\left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 1 \end{array}\right]\right\}$$

Thus, $\left\{\left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 1 \end{array}\right]\right\}$ is a basis of $\text{im } T$.

3. b. We have $T : V \to \mathbb{R}^2$ given by $T(\mathbf{v}) = (P(\mathbf{v}),\ Q(\mathbf{v}))$ where $P : V \to \mathbb{R}$ and $Q : V \to \mathbb{R}$ are linear transformations. T is linear by (a). Now

$$\begin{aligned}\ker T &= \{\mathbf{v} \mid T(\mathbf{v}) = (\mathbf{0},\ \mathbf{0})\}\\ &= \{\mathbf{v} \mid P(\mathbf{v}) = \mathbf{0} \text{ and } Q(\mathbf{v}) = \mathbf{0}\}\\ &= \{\mathbf{v} \mid P(\mathbf{v}) = \mathbf{0}\} \cap \{\mathbf{v} \mid Q(\mathbf{v}) = \mathbf{0}\}\\ &= \ker P \cap \ker Q\end{aligned}$$

4. b. $\ker T = \{(x,\ y,\ z) \mid x + y + z = 0,\ 2x - y + 3z = 0,\ z - 3y = 0,\ 3x + 4z = 0\}$. Solving:

$$\left[\begin{array}{rrr|r} 1 & 1 & 1 & 0\\ 2 & -1 & 3 & 0\\ 0 & -3 & 1 & 0\\ 3 & 0 & 4 & 0\end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & 1 & 0\\ 0 & -3 & 1 & 0\\ 0 & -3 & 1 & 0\\ 0 & -3 & 1 & 0\end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & \frac{4}{3} & 0\\ 0 & 1 & -\frac{1}{3} & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

Hence, $\ker T = \{(-4t,\ t,\ 3t) \mid t \text{ in } \mathbb{R}\} = \operatorname{span}\{(-4,\ 1,\ 3)\}$. Hence, $\{(1,\ 0,\ 0),\ (0,\ 1,\ 0),\ (-4,\ 1,\ 3)\}$ is one basis of $\mathbb{R}^3$ containing a basis of $\ker T$. Thus

$$\{T(1,\ 0,\ 0),\ T(0,\ 1,\ 0)\} = \{(1,\ 2,\ 0,\ 3),\ (1,\ -1,\ -3,\ 0)\}$$

is a basis of $\operatorname{im} T$ by Theorem 7.2.5.

6. b. Yes. $\dim(\operatorname{im} T) = \dim V - \dim(\ker T) = 5 - 2 = 3$. As $\dim W = 3$ and $\operatorname{im} T$ is a 3-dimensional subspace, $\operatorname{im} T = W$. Thus, T is onto.

   d. No. If $\ker T = V$ then $T(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in V, so $T = 0$ is the zero transformation. But W need not be the zero space. For example, $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x,\ y) = (0,\ 0)$ for all $(x,\ y)$ in $\mathbb{R}^2$.

   f. No. Let $T : R^2 \to R^2$ be defined by $T(x,\ y) = (y,\ 0)$ for all $(x,\ y) \in \mathbb{R}^2$. Then $\ker T = \{(x,\ 0) \mid x \in \mathbb{R}\} = \operatorname{im} T$.

   h. Yes. We always have $\dim(\operatorname{im} T) \leq \dim W$ (because $\operatorname{im} T$ is a subspace of W). Since $\dim(\ker T) \leq \dim W$ also holds in this case:

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T) \leq \dim W + \dim W = 2\dim W$$

   Hence $\dim W \geq \frac{1}{2}\dim V$.

   j. No. $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(x,\ y) = (x,\ 0)$ is not one-to-one (because $\ker T = \{(0,\ y) \mid y \in \mathbb{R}\}$ is not 0).

   l. No. $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(x,\ y) = (x,\ 0)$ is not onto.

   n. No. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x,\ y) = (x,\ 0)$, and let $\mathbf{v}_1 = (1,\ 0)$ and $\mathbf{v}_2 = (0,\ 1)$. Then $\{\mathbf{v}_1,\ \mathbf{v}_2\}$ spans $\mathbb{R}^2$, but $\{T(\mathbf{v}_1),\ T(\mathbf{v}_2)\} = \{\mathbf{v}_1,\ \mathbf{0}\}$ does not span $\mathbb{R}^2$.

7. b. Given $\mathbf{w}$ in W, we must show that it is a linear combination of $T(\mathbf{v}_1),\ \dots,\ T(\mathbf{v}_n)$. As T is onto, $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V. Since $V = \operatorname{span}\{\mathbf{v}_1,\ \dots,\ \mathbf{v}_n\}$ we can write $\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$ where each r_i is in $\mathbb{R}$. Hence

$$\mathbf{w} = T(\mathbf{v}) = T(r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n) = r_1T(\mathbf{v}_1) + \cdots + r_nT(\mathbf{v}_n)$$

8. b. If T is onto, let $\mathbf{v}$ be any vector in V. Then $\mathbf{v} = T(r_1, \dots, r_n)$ for some $(r_1, \dots, r_n)$ in $\mathbb{R}^n$; that is $\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$ is in span $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Thus $V =$ span $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Conversely, if $V =$ span $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, let $\mathbf{v}$ be any vector in V. Then $\mathbf{v}$ is in span $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ so $r_1, \dots, r_n$ exist in $\mathbb{R}$ such that

$$\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n = T(r_1, \dots, r_n)$$

Thus T is onto.

10. The trace map $T : \mathbf{M}_{22} \to \mathbb{R}$ is linear (Example 7.1.2) and it is onto (for example, $r = \text{tr}\,[\text{diag}\,(r, 0, \dots, 0)] = T\,[\text{diag}\,(r, 0, \dots, 0)]$ for any r in $\mathbb{R}$). Hence the dimension theorem gives $\dim(\ker T) = \dim \mathbf{M}_{nn} - \dim(\text{im } T) = n^2 - \dim(\mathbb{R}) = n^2 - 1$.

12. Define $T_A : \mathbb{R}^n \to \mathbb{R}^m$ and $T_B : \mathbb{R}^n \to \mathbb{R}^k$ by $T_A(\mathbf{x}) = A\mathbf{x}$ and $T_B(\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Then the given condition means $\ker T_A \subseteq \ker T_B$, so $\dim(\ker T_A) \leq \dim(\ker T_B)$. Hence

$$\text{rank } A = \dim(\text{im } T_A) = n - \dim(\ker T_A) \geq n - \dim(\ker T_B) = \dim(\text{im } T_B) = \text{rank } B$$

15. b. Write $B = \left\{x - 1, x^2 - 1, \dots, x^n - 1\right\}$. Then $B \subseteq \ker T$ because $T(x^k - 1) = 1 - 1 = 0$ for all k. Hence span $B \subseteq \ker T$. Moreover, the polynomials in B are independent (they have distinct degrees), so $\dim(\text{span } B) = n$. Hence, by Theorem 6.4.2, it suffices to show that $\dim(\ker T) = n$. But $T : \mathbf{P}_n \to \mathbb{R}$ is onto, so the dimension theorem gives $\dim(\ker T) = \dim(\mathbf{P}_n) - \dim(\mathbb{R}) = (n+1) - 1 = n$, as required.

20. If we can find an onto linear transformation $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ with $\ker T = U$ and $\text{im } T = V$, then we are done by the dimension theorem. The condition $\ker T = U$ suggests that we define T by $T(A) = A - A^T$ for all A in $\mathbf{M}_{nn}$. By Example 7.2.3, T is linear, $\ker T = U$, and $\text{im } T = V$. This is what we wanted.

22. Fix a column $\mathbf{y} \neq 0$ in $\mathbb{R}^n$, and define $T : \mathbf{M}_{mn} \to \mathbb{R}^m$ by $T(A) = A\mathbf{y}$ for all A in $\mathbf{M}_{mn}$. This is linear and $\ker T = U$, so the dimension theorem gives

$$mn = \dim(\mathbf{M}_{mn}) = \dim(\ker T) + \dim(\text{im } T) = \dim U + \dim(\text{im } T)$$

Hence, it suffices to show that $\dim(\text{im } T) = m$, equivalently (since $\text{im } T \subseteq \mathbb{R}^m$) that T is onto. So let $\mathbf{x}$ be a column in $\mathbb{R}^m$, we must find a matrix A in $\mathbf{M}_{mn}$ such that $A\mathbf{y} = \mathbf{x}$. Write A in terms of its columns as $A = \left[\begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array}\right]$ and write $\mathbf{y} = \left[\begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array}\right]^T$. Then the requirement that $A\mathbf{y} = \mathbf{x}$ becomes

$$\mathbf{x} = \left[\begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array}\right] \left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array}\right] = y_1C_1 + y_2C_2 + \cdots + y_nC_n \qquad (*)$$

Since $\mathbf{y} \neq \mathbf{0}$, let $y_k \neq 0$. Then $A\mathbf{y} = \mathbf{x}$ if we choose $C_k = y_k^{-1}\mathbf{x}$ and $C_j = \mathbf{0}$ if $j \neq k$. Hence T is onto as required.

29. b. Choose a basis $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ of U and (by Theorem 6.4.1) let $\{\mathbf{u}_1, \dots, \mathbf{u}_m, \dots, \mathbf{u}_n\}$ be a basis of V. By Theorem 7.1.3, there is a linear transformation $S : V \to V$ such that

$$\begin{array}{ll} S(\mathbf{u}_i) = \mathbf{u}_i & \text{if } 1 \leq i \leq m \\ S(\mathbf{u}_i) = \mathbf{0} & \text{if } i > m \end{array}$$

Hence, $\mathbf{u}_i$ is in $\text{im } S$ for $1 \leq i \leq m$, whence $U \subseteq \text{im } S$. On the other hand, if $\mathbf{w}$ is in $\text{im } S$, write $\mathbf{w} = S(\mathbf{v})$, $\mathbf{v}$ in V. Then r_i exist in $\mathbb{R}$ such that

$$\mathbf{v} = r_1\mathbf{u}_1 + \cdots + r_m\mathbf{u}_m + \cdots + r_n\mathbf{u}_n$$

so

$$\begin{aligned} \mathbf{w} &= r_1 S(\mathbf{u}_1) + \cdots + r_m S(\mathbf{u}_m) + \cdots + r_n S(\mathbf{u}_n) \\ &= r_1\mathbf{u}_1 + \cdots + r_m\mathbf{u}_m + \mathbf{0} \end{aligned}$$

It follows that $\mathbf{w}$ is in U, so $\text{im } S \subseteq U$. Then $U = \text{im } S$ as required.

7.3 Isomorphisms and Composition

1. b. T is one-to-one because $T(x, y, z) = (0, 0, 0)$ means $x = 0$, $x + y = 0$ and $x + y + z = 0$, whence $x = y = z = 0$. Now T is onto by Theorem 7.3.3.

   Alternatively: $\{T(1, 0, 0), T(0, 1, 0), T(0, 0, 1)\} = \{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}$ is independent, so T is an isomorphism by Theorem 7.3.1.

   d. T is one-to-one because $T(X) = 0$ implies $UXV = 0$, whence $X = 0$ (as U and V are invertible). Now Theorem 7.3.3 implies that T is onto and so is an isomorphism.

   f. T is one-to-one because $T(\mathbf{v}) = \mathbf{0}$ implies $k\mathbf{v} = \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$ because $k \neq 0$. Hence, T is onto if $\dim V$ is finite (by Theorem 7.3.3) and so is an isomorphism. Alternatively, T is onto because $T(k^{-1}\mathbf{v}) = k(k^{-1}\mathbf{v}) = \mathbf{v}$ holds for all $\mathbf{v}$ in V.

   h. T is onto because $T(A^T) = (A^T)^T = A$ for every $n \times m$ matrix A (note that A^T is in $\mathbf{M}_{mn}$ so $T(A^T)$ makes sense). Since $\dim \mathbf{M}_{mn} = mn = \dim \mathbf{M}_{nm}$, it follows that T is one-to-one by Theorem 7.3.3, and so is an isomorphism. (A direct proof that T is one-to-one: $T(A) = 0$ implies $A^T = 0$, whence $A = 0$.)

4. b. $ST(x, y, z) = S(x + y, 0, y + z) = (x + y, 0, y + z)$; $TS(x, y, z) = T(x, 0, z) = (x, 0, z)$. These are not equal (if $y \neq 0$) so $ST \neq TS$.

   d. $ST\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = S\left[\begin{array}{cc} c & a \\ d & b \end{array}\right] = \left[\begin{array}{cc} c & 0 \\ 0 & b \end{array}\right]$; $TS\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = T\left[\begin{array}{cc} a & 0 \\ 0 & d \end{array}\right] = \left[\begin{array}{cc} 0 & a \\ d & 0 \end{array}\right]$. These are not equal for some values of a, b, c and d (nearly all) so $ST \neq TS$.

5. b. $T^2(x, y) = T[T(x, y)] = T(x + y, 0) = (x + y + 0, 0) = (x + y, 0) = T(x, y)$. This holds for all (x, y), whence $T^2 = T$.

   d. $$\begin{aligned} T^2\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] &= T\left(T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]\right) = T\left(\frac{1}{2}\left[\begin{array}{cc} a+c & b+d \\ a+c & b+d \end{array}\right]\right) = \frac{1}{2}T\left[\begin{array}{cc} a+c & b+d \\ a+c & b+d \end{array}\right] \\ &= \frac{1}{4}\left[\begin{array}{cc} (a+c)+(a+c) & (b+d)+(b+d) \\ (a+c)+(a+c) & (b+d)+(b+d) \end{array}\right] = \frac{1}{2}\left[\begin{array}{cc} a+c & b+d \\ a+c & b+d \end{array}\right] = T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] \end{aligned}$$

   This holds for all $\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$, so $T^2 = T$.

6. b. No inverse. For example $T(1, -1, 1, -1) = (0, 0, 0, 0)$ so $(1, -1, 1, -1)$ is a nonzero vector in $\ker T$. Hence T is not one-to-one, and so has no inverse.

d. T is one-to-one because $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ implies $a+2c = 0 = 3c-a$ and $b+2d = 0 = 3d-b$, whence $a = b = c = d = 0$. Thus T is an isomorphism by Theorem 7.3.3. If $T^{-1}V\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$, then $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = T\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} x+2z & y+2w \\ 3z-x & 3w-y \end{bmatrix}$. Thus

$$\begin{aligned} x+2z &= a \\ -x+3z &= c \\ y+2w &= b \\ -y+3w &= d \end{aligned}$$

The solution is $x = \frac{1}{5}(3a-2c)$, $z = \frac{1}{5}(a+c)$, $y = \frac{1}{5}(3b-2d)$, $w = \frac{1}{5}(b+d)$. Hence

$$T^{-1}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3a-2c & 3b-2d \\ a+c & b+d \end{bmatrix} \tag{$*$}$$

A better way to find T^{-1} is to observe that $T(X) = AX$ where $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$. This matrix is invertible which easily implies that T is one-to-one (and onto), and if $S : \mathbf{M}_{22} \to \mathbf{M}_{22}$ is defined by $S(X) = A^{-1}X$ then $ST = 1_{\mathbf{M}_{22}}$ and $TS = 1_{\mathbf{M}_{22}}$. Hence $S = T^{-1}$ by Theorem 7.3.5. Note that $A^{-1} = \frac{1}{5}\begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix}$ which gives ($*$).

f. T is one-to-one because, if p in $\mathbf{P}_2$ satisfies $T(p) = 0$, then $p(0) = p(1) = p(-1) = 0$. If $p = a+bx+cx^2$, this means $a = 0$, $a+b+c = 0$ and $a-b+c = 0$, whence $a = b = c = 0$, and $p = 0$. Hence, T^{-1} exists by Theorem 7.3.3. If $T^{-1}(a,\ b,\ c) = r+sx+tx^2$, then

$$(a,\ b,\ c) = T(r+sx+tx^2) = (r,\ r+s+t,\ r-s+t)$$

Then $r = a$, $r+s+t = b$, $r-s+t = c$, whence $r = a$, $s = \frac{1}{2}(b-c)$, $t = \frac{1}{2}(b+c-2a)$. Finally

$$T^{-1}(a,\ b,\ c) = a + \tfrac{1}{2}(b-c)x + \tfrac{1}{2}(b+c-2a)x^2$$

7. b. $T^2(x,\ y) = T[T(x,\ y)] = T(ky-x,\ y) = (ky-(ky-x),\ y) = (x,\ y) = 1_{\mathbb{R}^2}(x,\ y)$. Since this holds for all $(x,\ y)$ in $\mathbb{R}^2$, it shows that $T^2 = 1_{\mathbb{R}^2}$. This means that $T^{-1} = T$.

d. It is a routine verification that $A^2 = I$. Hence

$$T^2(\mathbf{x}) = T[T(\mathbf{x})] = A[A\mathbf{x}] = A^2\mathbf{x} = I\mathbf{x} = \mathbf{x} = 1_{\mathbf{M}_{22}}(\mathbf{x})$$

holds for all $\mathbf{x}$ in $\mathbf{M}_{22}$. This means that $T^2 = 1_{\mathbf{M}_{22}}$, and hence that $T^{-1} = T$.

8. b.
$$\begin{aligned} T^2(x,\ y,\ z,\ w) &= T[T(x,\ y,\ z,\ w)] = T(-y,\ x-y,\ z,\ -w) \\ &= (-(x-y),\ -y-(x-y),\ z,\ -(-w)) = (y-x,\ -x,\ z,\ w) \\ T^3(x,\ y,\ z,\ w) &= T\left[T^2(x,\ y,\ z,\ w)\right] = T(y-x,\ -x,\ z,\ w) = (x,\ y,\ z,\ -w) \\ T^6(x,\ y,\ z,\ w) &= T^3\left[T^3(x,\ y,\ z,\ w)\right] = T^3[x,\ y,\ z,\ -w] = (x,\ y,\ z,\ w) = 1_{\mathbb{R}^4}(x,\ y,\ z,\ w) \end{aligned}$$

Hence, $T^6 = 1_{\mathbb{R}^4}$ so $T^{-1} = T^5$. Explicitly:

$$T^{-1}(x,\ y,\ z,\ w) = T^2\left[T^3(x,\ y,\ z,\ w)\right] = T^2(x,\ y,\ z,\ -w) = (y-x,\ -x,\ z,\ -w)$$

9. b. Define $S : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ by $S(A) = U^{-1}A$. Then

$$\begin{aligned} ST(A) &= S(T(A)) = U^{-1}(UA) = A = 1_{\mathbf{M}_{nn}}(A) \text{ so } ST = 1_{\mathbf{M}_{nn}} \\ TS(A) &= T(S(A)) = U(U^{-1}A) = A = 1_{\mathbf{M}_{nn}}(A) \text{ so } TS = 1_{\mathbf{M}_{nn}} \end{aligned}$$

Hence, T is invertible and $T^{-1} = S$.

10. b. Given $V \xrightarrow{T} W \xrightarrow{S} U$ with T and S both onto, we are to show that $ST : V \to U$ is onto. Given $\mathbf{u}$ in U, we have $\mathbf{u} = S(\mathbf{w})$ for some $\mathbf{w}$ in W because S is onto; then $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V because T is onto. Hence,

$$ST(\mathbf{v}) = S[T(\mathbf{v})] = S[\mathbf{w}] = \mathbf{u}$$

This shows that ST is onto.

12. b. If $\mathbf{u}$ lies in $\operatorname{im} RT$ write $\mathbf{u} = RT(\mathbf{v})$, $\mathbf{v}$ in V. Thus $\mathbf{u} = R[T(\mathbf{v})]$, where $T(\mathbf{v})$ in W, so $\mathbf{u}$ is in $\operatorname{im} R$.

13. b. Given $V \xrightarrow{T} U \xrightarrow{S} W$ with ST onto, let $\mathbf{w}$ be a vector in W. Then $\mathbf{w} = ST(\mathbf{v})$ for some $\mathbf{v}$ in V because ST is onto, whence $\mathbf{w} = S[T(\mathbf{v})]$ where $T(\mathbf{v})$ is in U. This shows that S is onto. Now the dimension theorem applied to S gives

$$\dim U = \dim(\ker S) + \dim(\operatorname{im} S) = \dim(\ker S) + \dim W$$

because $\operatorname{im} S = W$ (S is onto). As $\dim(\ker S) \geq 0$, this gives $\dim U \geq \dim W$.

14. If $T^2 = 1_V$ then $TT = 1_V$ so T is invertible and $T^{-1} = T$ by the definition of the inverse of a transformation. Conversely, if $T^{-1} = T$ then $T^2 = TT^{-1} = 1_V$.

16. Theorem 7.2.5 shows that $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_r)\}$ is a basis of $\operatorname{im} T$. Write

$U = \operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$. Then $B = \{\mathbf{e}_1, \dots, \mathbf{e}_r\}$ is a basis of U, and $T : U \to \operatorname{im} T$ carries B to the basis $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_r)\}$. Thus $T : U \to \operatorname{im} T$ is itself an isomorphism. Note that $T : V \to W$ may not be an isomorphism, but restricting T to the subspace U of V does result in an isomorphism in this case.

19. b. We have $V = \{(x, y) \mid x, y \text{ in } \mathbb{R}\}$ with a new addition and scalar multiplication:

$$\begin{aligned} (x, y) \oplus (x_1, y_1) &= (x + x_1, y + y_1 + 1) \\ a \odot (x, y) &= (ax, ay + a - 1) \end{aligned}$$

We use the notation $\oplus$ and $\odot$ for clarity. Define

$$T : V \to \mathbb{R}^2 \text{ by } T(x, y) = (x, y + 1)$$

Then T is a linear transformation because:

$$\begin{aligned} T[(x, y) \oplus (x_1, y_1)] &= T(x + x_1, y + y_1 + 1) \\ &= (x + x_1, (y + y_1 + 1) + 1) \\ &= (x, y + 1) + (x_1, y_1 + 1) \\ &= T(x, y) + T(x_1, y_1) \end{aligned}$$

$$\begin{aligned} T(a \odot (x,\ y)] &= T(ax,\ ay + a - 1) \\ &= (ax,\ ay + a) \\ &= a(x,\ y + 1) \\ &= aT(x,\ y) \end{aligned}$$

Moreover T is one-to-one because $T(x,\ y) = (0,\ 0)$ means $x = 0 = y + 1$, so $(x,\ y) = (0,\ 1)$, the zero vector of V. (Alternatively, $T(x,\ y) = T(x_1,\ y_1)$ implies $(x,\ y + 1) = (x_1,\ y_1 + 1)$, whence $x = x_1,\ y = y_1$.) As T is clearly onto $\mathbb{R}^2$, it is an isomorphism.

24. b. $TS[x_0,\ x_1,\ \dots) = T[0,\ x_0,\ x_1,\ \dots) = [x_0,\ x_1,\ \dots)$ so $TS = 1_V$. Hence TS is both onto and one-to-one, so T is onto and S is one-to-one by Exercise 10. But $[1,\ 0,\ 0,\ \dots)$ is in $\ker T$ while $[1,\ 0,\ 0,\ \dots)$ is not in $\operatorname{im} S$.

26. b. If $p(x)$ is in $\ker T$, then $p(x) = -xp'(x)$. If we write $p(x) = a_0 + a_1x + \cdots + a_nx^n$, this becomes

$$a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n = -a_1x - 2a_2x^2 - \cdots - na_nx^n$$

Equating coefficients gives $a_0 = 0,\ a_1 = -a_1,\ a_2 = -2a_2,\ \dots,\ a_n = -na_n$. Hence we have, $a_0 = a_1 = \cdots = a_n = 0$, so $p(x) = 0$. Thus $\ker T = \{0\}$, so T is one-to-one. As $T : \mathbf{P}_n \to \mathbf{P}_n$ and $\dim \mathbf{P}_n$ is finite, this implies that T is also onto, and so is an isomorphism.

27. b. If $TS = 1_W$ then, given $\mathbf{w}$ in W, $T[S(\mathbf{w})] = \mathbf{w}$, so T is onto. Conversely, if T is onto, choose a basis $\{\mathbf{e}_1,\ \dots,\ \mathbf{e}_r,\ \mathbf{e}_{r+1},\ \dots,\ \mathbf{e}_n\}$ of V such that $\{\mathbf{e}_{r+1},\ \dots,\ \mathbf{e}_n\}$ is a basis of $\ker T$. By Theorem 7.2.5, $\{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_n)\}$ is a basis of $\operatorname{im} T = W$ (as T is onto). Hence, a linear transformation $S : W \to V$ exists such that $S[T(\mathbf{e}_i)] = \mathbf{e}_i$ for $i = 1,\ 2,\ \dots,\ r$. We claim that $TS = 1_W$, and we show this by verifying that these transformations agree on the basis $\{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_r)\}$ of W. Indeed

$$TS[T(\mathbf{e}_i)] = T\{S[T(\mathbf{e}_i)]\} = T(\mathbf{e}_i) = 1_W[T(\mathbf{e}_i)]$$

for $i = 1,\ 2,\ \dots,\ n$.

28. b. If $T = SR$, then every vector $T(\mathbf{v})$ in $\operatorname{im} T$ has the form $T(\mathbf{v}) = S[R(\mathbf{v})]$, whence $\operatorname{im} T \subseteq \operatorname{im} S$. Since R is invertible, $S = TR^{-1}$ implies $\operatorname{im} S \subseteq \operatorname{im} T$, so $\operatorname{im} S = \operatorname{im} T$.

Conversely, assume that $\operatorname{im} S = \operatorname{im} T$. The dimension theorem gives

$$\dim(\ker S) = n - \dim(\operatorname{im} S) = n - \dim(\operatorname{im} T) = \dim(\ker T)$$

Hence, let $\{\mathbf{e}_1,\ \dots,\ \mathbf{e}_r,\ \dots,\ \mathbf{e}_n\}$ and $\{\mathbf{f}_1,\ \dots,\ \mathbf{f}_r,\ \dots,\ \mathbf{f}_n\}$ be bases of V such that $\{\mathbf{e}_{r+1},\ \dots,\ \mathbf{e}_n\}$ and $\{\mathbf{f}_{r+1},\ \dots,\ \mathbf{f}_n\}$ are bases of $\ker S$ and $\ker T$, respectively. By Theorem 7.2.5, $\{S(\mathbf{e}_1),\ \dots,\ S(\mathbf{e}_r)\}$ and $\{T(\mathbf{f}_1),\ \dots,\ T(\mathbf{f}_r)\}$ are both bases of $\operatorname{im} S = \operatorname{im} T$. So let $\mathbf{g}_1,\ \dots,\ \mathbf{g}_r$ in V be such that

$$S(\mathbf{e}_i) = T(\mathbf{g}_i)$$

for each $i = 1,\ 2,\ \dots,\ r$.

Claim: $B = \{\mathbf{g}_1,\ \dots,\ \mathbf{g}_r,\ \mathbf{f}_{r+1},\ \dots,\ \mathbf{f}_n\}$ is a basis of V.

Proof. It suffices (by Theorem 6.4.4) to show that B is independent. If

$$a_1\mathbf{g}_1 + \cdots + a_r\mathbf{g}_r + b_{r+1}\mathbf{f}_{r+1} + \cdots + b_n\mathbf{f}_n = \mathbf{0}$$

apply T to get

$$\begin{aligned}\mathbf{0} &= a_1T(\mathbf{g}_1)+\cdots+a_rT(\mathbf{g}_r)+b_{r+1}T(\mathbf{f}_{r+1})+\cdots+b_nT(\mathbf{f}_n)\\ &= a_1T(\mathbf{g}_1)+\cdots+a_rT(\mathbf{g}_r)+\mathbf{0}\end{aligned}$$

because $T(\mathbf{f}_j)=0$ if $j>r$. Hence $a_1=\cdots=a_r=0$; whence $\mathbf{0}=b_{r+1}\mathbf{f}_{r+1}+\cdots+b_n\mathbf{f}_n$. This gives $b_{r+1}=\cdots=b_n=0$ and so proves the claim.

By the claim, we can define $R: V\to V$ by

$$\begin{aligned}R(\mathbf{g}_i)&=\mathbf{e}_i \quad \text{for } i=1,\ 2,\ \dots,\ r\\ R(\mathbf{f}_j)&=\mathbf{e}_j \quad \text{for } j=r+1,\ \dots,\ n\end{aligned}$$

Then R is an isomorphism by Theorem 7.3.1, and we claim that $SR=T$. We show this by verifying that SR and T have the same effect on the basis B in the claim. The definition of R gives

$$\begin{aligned}SR(\mathbf{g}_i)&=S[R(\mathbf{g}_i)]=S(\mathbf{e}_i)=T(\mathbf{g}_i) && \text{for } i=1,\ 2,\ \dots,\ r\\ SR(\mathbf{f}_j)&=S[\mathbf{e}_j]=\mathbf{0}=T(\mathbf{f}_j) && \text{for } j=r+1,\ \dots,\ n\end{aligned}$$

Hence $SR=T$.

29. As in the hint, let $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_r,\ \dots,\ \mathbf{e}_n\}$ be a basis of V where $\{\mathbf{e}_{r+1},\ \dots,\ \mathbf{e}_n\}$ is a basis of $\ker T$. Then $\{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_r)\}$ is linearly independent by Theorem 7.2.5, so extend it to a basis $\{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_r),\ \mathbf{w}_{r+1},\ \dots,\ \mathbf{w}_n\}$ of V. Then define $S: V\to V$ by

$$\begin{aligned}S[T(\mathbf{e}_i)]&=\mathbf{e}_i \quad \text{for } 1\le i\le r\\ S(\mathbf{w}_j)&=\mathbf{e}_j \quad \text{for } r<j\le n\end{aligned}$$

Then, S is an isomorphism (by Theorem 7.3.1) and we claim that $TST=T$. We verify this by showing that TST and T agree on the basis $\{\mathbf{e}_1,\ \dots,\ \mathbf{e}_r,\ \dots,\ \mathbf{e}_n\}$ of V (and invoking Theorem 7.1.2).

$$\begin{gathered}\text{If } 1\le i\le r\text{: } TST(\mathbf{e}_i)=T\left(S[T(\mathbf{e}_i)]\right)=T(\mathbf{e}_i)\\ \text{If } r+1\le j\le n\text{: } TST(\mathbf{e}_j)=TS[T(\mathbf{e}_j)]=TS[\mathbf{0}]=\mathbf{0}=T(\mathbf{e}_j)\end{gathered}$$

where, at the end, we use the fact that $\mathbf{e}_j$ is in $\ker T$ for $r+1\le j\le n$.

7.4 A Theorem about Differential Equations

This section contains no exercises.

7.5 More on Linear Recurrences

1. b. The associated polynomial is

$$p(x) = x^3 - 7x + 6 = (x-1)(x-2)(x+3)$$

Hence, $\{[1), [2^n), [(-3)^n)\}$ is a basis of the space of all solutions to the recurrence. The general solution is thus,

$$[x_n) = a[1) + b[2^n) + c[(-3)^n)$$

where a, b and c are constants. The requirement that $x_0 = 1$, $x_1 = 2$, $x_2 = 1$ determines a, b, and c. We have

$$x_n = a + b2^n + c(-3)^n$$

for all $n \geq 0$. So taking $n = 0, 1, 2$ gives

$$\begin{aligned} a + b + c &= x_0 = 1 \\ a + 2b - 3c &= x_1 = 2 \\ a + 4b + 9c &= x_2 = 1 \end{aligned}$$

The solution is $a = \frac{15}{20}$, $b = \frac{8}{20}$, $c = -\frac{3}{20}$, so

$$x_n = \tfrac{1}{20}(15 + 2^{n+3} + (-3)^{n+1}) \quad n \geq 0$$

2. b. The associated polynomial is

$$p(x) = x^3 - 3x + 2 = (x-1)^2(x+2)$$

As 1 is a double root of $p(x)$, $[1^n) = [1)$ and $[n1^n) = [n)$ are solutions to the recurrence by Theorem 7.5.3. Similarly, $[(-2)^n)$ is a solution, so $\{[1), [n), [(-2)^n)\}$ is a basis for the space of solutions by Theorem 7.5.4. The required sequence has the form

$$[x_n) = a[1) + b[n) + c[(-2)^n)$$

for constants a, b, c. Thus, $x_n = a + bn + c(-2)^n$ for $n \geq 0$, so taking $n = 0, 1, 2$, we get

$$\begin{array}{ccccccccc} a & + & & & c & = & x_0 & = & 1 \\ a & + & b & - & 2c & = & x_1 & = & -1 \\ a & + & 2b & + & 4c & = & x_2 & = & 1 \end{array}$$

The solution is $a = \frac{5}{9}$, $b = -\frac{6}{9}$, $c = \frac{4}{9}$, so

$$x_n = \tfrac{1}{9}\left[5 - 6n + (-2)^{n+2}\right] \quad n \geq 0$$

d. The associated polynomial is

$$p(x) = x^3 - 3x^2 + 3x - 1 = (x-1)^3$$

Hence, $[1^n) = [1)$, $[n1^n) = [n)$ and $[n^2 1^n) = [n^2)$ are solutions and so $\{[1),\ [n),\ [n^2)\}$ is a basis for the space of solutions. Thus

$$x_n = a \cdot 1 + bn + cn^2$$

a, b, c constants. As $x_0 = 1$, $x_1 = -1$, $x_2 = 1$, we obtain

$$\begin{array}{rcrcrcrcr} a & & & & & = & x_0 & = & 1 \\ a & + & b & + & c & = & x_1 & = & -1 \\ a & + & 2b & + & 4c & = & x_2 & = & 1 \end{array}$$

The solution is $a = 1$, $b = -4$, $c = 2$, so

$$x_n = 1 - 4n + 2n^2 \quad n \geq 0$$

This can be written

$$x_n = 2(n-1)^2 - 1$$

3. b. The associated polynomial is

$$p(x) = x^2 - (a+b)x + ab = (x-a)(x-b)$$

Hence, as $a \neq b$, $\{[a^n),\ [b^n)\}$ is a basis for the space of solutions.

4. b. The recurrence $x_{n+4} = -x_{n+2} + 2x_{n+3}$ has $r_0 = 0$ as there is no term x_n. If we write $y_n = x_{n+2}$, the recurrence becomes

$$y_{n+2} = -y_n + 2y_{n+1}$$

Now the associated polynomial is $x^2 - 2x + 1 = (x-1)^2$ so basis sequences for the solution space for y_n are $[1^n) = [1,\ 1,\ 1,\ 1,\ \dots)$ and $[n1^n) = [0,\ 1,\ 2,\ 3,\ \dots)$. As $y_n = x_{n+2}$, corresponding basis sequences for x_n are $[0,\ 0,\ 1,\ 1,\ 1,\ 1,\ \dots)$ and $[0,\ 0,\ 0,\ 1,\ 2,\ 3,\ \dots)$. Also, $[1,\ 0,\ 0,\ 0,\ 0,\ 0,\ \dots)$ and $[0,\ 1,\ 0,\ 0,\ 0,\ 0,\ \dots)$ are solutions for x_n, so these four sequences form a basis for the solution space for x_n.

7. The sequence has length 2 and associated polynomial $x^2 + 1$. The roots are nonreal: $\lambda_1 = i$ and $\lambda_2 = -i$. Hence, by Remark 2,

$$[i^n + (-i)^n) = [2,\ 0,\ -2,\ 0,\ 2,\ 0,\ -2,\ 0,\ \dots) \text{ and } [i(i^n - (-i)^n)) = [0,\ -2,\ 0,\ 2,\ 0,\ -2,\ 0,\ 2,\ \dots)$$

are solutions. They are independent as is easily verified, so they are a basis for the space of solutions.

8. Orthogonality

8.1 Orthogonal Complements and Projections

1. b. Write $\mathbf{x}_1 = (2, 1)$ and $\mathbf{x}_2 = (1, 2)$. The Gram-Schmidt algorithm gives

$$\begin{aligned}
\mathbf{e}_1 &= \mathbf{x}_1 = (2, 1) \\
\mathbf{e}_2 &= \mathbf{x}_2 - \tfrac{\mathbf{x}_2 \cdot \mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 \\
&= (1, 2) - \tfrac{4}{5}(2, 1) \\
&= \tfrac{1}{5}\{(5, 10) - (8, 4)\} \\
&= \tfrac{3}{5}(-1, 2)
\end{aligned}$$

In hand calculations, $\{(2, 1), (-1, 2)\}$ may be a more convenient orthogonal basis.

d. If $\mathbf{x}_1 = (0, 1, 1)$, $\mathbf{x}_2 = (1, 1, 1)$, $\mathbf{x}_3 = (1, -2, 2)$ then
$\mathbf{e}_1 = \mathbf{x}_1 = (0, 1, 1)$
$\mathbf{e}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 = (1, 1, 1) - \frac{2}{2}(0, 1, 1) = (1, 0, 0)$
$\mathbf{e}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 = (1, -2, 2) - \frac{0}{2}(0, 1, 1) - \frac{1}{1}(1, 0, 0) = (0, -2, 2)$

2. b. Write $\mathbf{e}_1 = (3, -1, 2)$ and $\mathbf{e}_2 = (2, 0, -3)$. Then $\{\mathbf{e}_1, \mathbf{e}_2\}$ is orthogonal and so is an orthogonal basis of $U = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}$. Now $\mathbf{x} = (2, 1, 6)$ so take

$$\begin{aligned}
\mathbf{x}_1 = \text{proj}_U \mathbf{x} &= \tfrac{\mathbf{x} \cdot \mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \tfrac{\mathbf{x} \cdot \mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 \\
&= \tfrac{17}{14}(3, -1, 2) - \tfrac{14}{13}(2, 0, -3) \\
&= \tfrac{1}{182}(271, -221, 1030)
\end{aligned}$$

Then $\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 = \frac{1}{182}(93, 402, 62)$. As a check: $\mathbf{x}_2$ is orthogonal to both $\mathbf{e}_1$ and $\mathbf{e}_2$ (and so is in $U^\perp$).

d. If $\mathbf{e}_1 = (1, 1, 1, 1)$, $\mathbf{e}_2 = (1, 1, -1, -1)$, $\mathbf{e}_3 = (1, -1, 1, -1)$ and $\mathbf{x} = (2, 0, 1, 6)$, then $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is orthogonal so take

$$\begin{aligned}
\mathbf{x}_1 = \text{proj}_U \mathbf{x} &= \tfrac{\mathbf{x} \cdot \mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \tfrac{\mathbf{x} \cdot \mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 + \tfrac{\mathbf{x} \cdot \mathbf{e}_3}{\|\mathbf{e}_3\|^2}\mathbf{e}_3 \\
&= \tfrac{9}{4}(1, 1, 1, 1) - \tfrac{5}{4}(1, 1, -1, -1) - \tfrac{3}{4}(1, -1, 1, -1) \\
&= \tfrac{1}{4}(1, 7, 11, 17)
\end{aligned}$$

Then, $\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 = \frac{1}{4}(7, -7, -7, 7) = \frac{7}{4}(1, -1, -1, 1)$. Check: $\mathbf{x}_2$ is orthogonal to each $\mathbf{e}_i$, hence $\mathbf{x}_2$ is in $U^\perp$.

f. If $\mathbf{e}_1 = (1, -1, 2, 0)$ and $\mathbf{e}_2 = (-1, 1, 1, 1)$ then (as $\mathbf{x} = (a, b, c, d)$)

$$\begin{aligned}
\mathbf{x}_1 = \text{proj}_U\, \mathbf{x} &= \tfrac{a-b+2c}{6}(1, -1, 2, 0) + \tfrac{-a+b+c+d}{4}(-1, 1, 1, 1) \\
&= \left(\tfrac{5a-5b+c-3d}{12}, \tfrac{-5a+5b-c+3d}{12}, \tfrac{a-b+11c+3d}{12}, \tfrac{-3a+3b+3c+3d}{12}\right) \\
\mathbf{x}_2 = \mathbf{x} - \mathbf{x}_1 &= \left(\tfrac{7a+5b-c+3d}{12}, \tfrac{5a+7b+c-3d}{12}, \tfrac{-a+b+c-3d}{12}, \tfrac{3a-3b-3c+9d}{12}\right)
\end{aligned}$$

3. a. Write $\mathbf{e}_1 = (2, 1, 3, -4)$ and $\mathbf{e}_2 = (1, 2, 0, 1)$, so $\{\mathbf{e}_1, \mathbf{e}_2\}$ is orthogonal. As $\mathbf{x} = (1, -2, 1, 6)$

$$\begin{aligned}
\text{proj}_U\, \mathbf{x} &= \tfrac{\mathbf{x}\cdot\mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \tfrac{\mathbf{x}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 \\
&= -\tfrac{21}{30}(2, 1, 3, -4) + \tfrac{3}{6}(1, 2, 0, 1) = \tfrac{3}{10}(-3, 1, -7, 11)
\end{aligned}$$

c. $\text{proj}_U\, \mathbf{x} = -\frac{15}{14}(1, 0, 2, -3) + \frac{3}{70}(4, 7, 1, 2) = \frac{3}{10}(-3, 1, -7, 11)$.

4. b. $U = \text{span}\,\{(1, -1, 0), (-1, 0, 1)\}$ but this basis is not orthogonal. By Gram-Schmidt:

$$\begin{aligned}
\mathbf{e}_1 &= (1, -1, 0) \\
\mathbf{e}_2 &= (-1, 0, 1) - \tfrac{(-1, 0, 1)\cdot(1, -1, 0)}{\|(1, -1, 0)\|^2}(1, -1, 0) = -\tfrac{1}{2}(1, 1, -2)
\end{aligned}$$

So we use $U = \text{span}\,\{(1, -1, 0), (1, 1, -2)\}$. Then the vector $\mathbf{x}_1$ in U closest to $\mathbf{x} = (2, 1, 0)$ is

$$\mathbf{x}_1 = \text{proj}_U\, \mathbf{x} = \tfrac{2-1+0}{2}(1, -1, 0) + \tfrac{2+1+0}{6}(1, 1, -2) = (1, 0, -1)$$

d. The given basis of U is not orthogonal. The Gram-Schmidt algorithm gives

$$\begin{aligned}
\mathbf{e}_1 &= (1, -1, 0, 1) \\
\mathbf{e}_2 &= (1, 1, 0, 0) = (1, 1, 0, 0) - \tfrac{0}{3}\mathbf{e}_1 = (1, 1, 0, 0) \\
\mathbf{e}_3 &= (1, 1, 0, 1) - \tfrac{1}{3}(1, -1, 0, 1) - \tfrac{2}{2}(1, 1, 0, 0) = \tfrac{1}{3}(-1, 1, 0, 2)
\end{aligned}$$

Given $\mathbf{x} = (2, 0, 3, 1)$, we get (using $\mathbf{e}_3' = (-1, 1, 0, 2)$ for convenience)
$\text{proj}_U\, \mathbf{x} = \frac{3}{3}(1, -1, 0, 1) + \frac{2}{2}(1, 1, 0, 0) + \frac{0}{6}(-1, 1, 0, 2) = (2, 0, 0, 1)$.

5. b. Here $A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ 1 & 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -3 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -3 & 0 \end{bmatrix}$. Hence, $A\mathbf{x}^T = 0$ has $\mathbf{x} = (s-t, 3s, s, t) = s(1, 3, 1, 0) + t(-1, 0, 0, 1)$.
Thus $U^\perp = \text{span}\,\{(1, 3, 1, 0), (-1, 0, 0, 1)\}$.

8. If $\mathbf{x} = \text{proj}_U\, \mathbf{x}$ then $\mathbf{x}$ is in U by Theorem 8.1.3. Conversely, if $\mathbf{x}$ is in U, let $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ be an orthogonal basis of U. Then the expansion theorem (applied to the space U) gives $\mathbf{x} = \sum_i \frac{\mathbf{x}\cdot\mathbf{f}_i}{\|\mathbf{f}_i\|^2}\mathbf{f}_i = \text{proj}_U\, \mathbf{x}$ by the definition of the projection.

10. Let $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ be an orthonormal basis of U. If $\mathbf{x}$ is in U then, since $\|\mathbf{f}_i\| = 1$ for each i, so $\mathbf{x} = (\mathbf{x}\cdot\mathbf{f}_1)\mathbf{f}_1 + \cdots + (\mathbf{x}\cdot\mathbf{f}_m)\mathbf{f}_m = \text{proj}_U\, \mathbf{x}$ by the expansion theorem (applied to the space U).

14. If $\{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ is a basis of $U^\perp$, take $A = \begin{bmatrix} \mathbf{y}_1^T \\ \vdots \\ \mathbf{y}_m^T \\ \mathbf{0} \end{bmatrix}$. Then $A\mathbf{x} = \mathbf{0}$ if and only if $\mathbf{y}_i^T\mathbf{x} = \mathbf{0}$ for each i; if and only if $\mathbf{y}_i \cdot \mathbf{x} = \mathbf{0}$ for each i; if and only if $\mathbf{x}$ is in $(U^\perp)^\perp = U^{\perp\perp} = U$. This shows that $U = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$.

17. d. If AA^T is invertible and $E = A^T(AA^T)^{-1}A$, then

$$E^2 = A^T(AA^T)^{-1}A \cdot A^T(AA^T)^{-1}A = A^T I (AA^T)^{-1}A = E$$
$$\begin{aligned} E^T &= \left[A^T(AA^T)^{-1}A\right]^T = A^T\left[(AA^T)^{-1}\right]^T(A^T)^T \\ &= A^T\left[(AA^T)^T\right]^{-1}A = A^T\left[(A^T)^T A^T\right]^{-1}A \\ &= A^T\left[AA^T\right]^{-1}A = E \end{aligned}$$

Thus, $E^2 = E = E^T$.

8.2 Orthogonal Diagonalization

1. b. Since $3^2 + 4^2 = 5^2$, each row has length 5. So $\left[\begin{array}{cc} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{array}\right] = \frac{1}{5}\left[\begin{array}{rr} 3 & -4 \\ 4 & 3 \end{array}\right]$ is orthogonal.

   d. Each row has length $\sqrt{a^2+b^2} \neq 0$, so $\frac{1}{\sqrt{a^2+b^2}}\left[\begin{array}{rr} a & b \\ -b & a \end{array}\right]$ is orthogonal.

   f. The rows have length $\sqrt{6}$, $\sqrt{3}$, $\sqrt{2}$ respectively, so

$$\left[\begin{array}{ccc} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right] = \frac{1}{\sqrt{6}}\left[\begin{array}{ccc} 2 & 1 & -1 \\ \sqrt{2} & -\sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} & \sqrt{3} \end{array}\right]$$

   is orthogonal.

   h. Each row has length $\sqrt{4+36+9} = \sqrt{49} = 7$. Hence

$$\left[\begin{array}{ccc} \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \\ \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \end{array}\right] = \frac{1}{7}\left[\begin{array}{rrr} 2 & 6 & -3 \\ 3 & 2 & 6 \\ -6 & 3 & 2 \end{array}\right]$$

   is orthogonal.

2. Let P be orthogonal, so $P^{-1} = P^T$. If P is upper triangular, so also is P^{-1}, so $P^{-1} = P^T$ is both upper triangular (P^{-1}) and lower triangular P^T). Hence, $P^{-1} = P^T$ is diagonal, whence $P = (P^{-1})^{-1}$ is diagonal. In particular, P is symmetric so $P^{-1} = P^T = P$. Thus $P^2 = I$. Since P is diagonal, this implies that all diagonal entries are ± 1.

5. b. $c_A(x) = \left|\begin{array}{cc} x-1 & 1 \\ 1 & x-1 \end{array}\right| = x(x-2)$.

   Hence the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 2$.

   $\lambda_1 = 0$: $\left[\begin{array}{rr} -1 & 1 \\ 1 & -1 \end{array}\right] \to \left[\begin{array}{rr} 1 & -1 \\ 0 & 0 \end{array}\right]$; $E_0(A) = \text{span}\left\{\left[\begin{array}{c} 1 \\ 1 \end{array}\right]\right\}$.

   $\lambda_2 = 2$: $\left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right] \to \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right]$; $E_2(A) = \text{span}\left\{\left[\begin{array}{r} -1 \\ 1 \end{array}\right]\right\}$.

   Note that these eigenvectors are orthogonal (as Theorem 8.2.4 asserts). Normalizing them gives an orthogonal matrix

$$P = \left[\begin{array}{cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right] = \frac{1}{\sqrt{2}}\left[\begin{array}{rr} 1 & -1 \\ 1 & 1 \end{array}\right]$$

   Then $P^{-1} = P^T$ and $P^TAP = \left[\begin{array}{cc} 0 & 0 \\ 0 & 2 \end{array}\right]$.

d. $c_A(x) = \begin{vmatrix} x-3 & 0 & -7 \\ 0 & x-5 & 0 \\ -7 & 0 & x-3 \end{vmatrix} = (x-5)(x^2-6x-40) = (x-5)(x+4)(x-10)$. Hence the eigenvalues are $\lambda_1 = 5$, $\lambda_2 = 10$, $\lambda_3 = -4$.

$\lambda_1 = 5$: $\begin{bmatrix} 2 & 0 & -7 \\ 0 & 0 & 0 \\ -7 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; $E_5(A) = \text{span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$.

$\lambda_2 = 10$: $\begin{bmatrix} 7 & 0 & -7 \\ 0 & 5 & 0 \\ -7 & 0 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $E_{10}(A) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$.

$\lambda_3 = -4$: $\begin{bmatrix} -7 & 0 & -7 \\ 0 & -9 & 0 \\ -7 & 0 & -7 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $E_{-4}(A) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right\}$.

Note that the three eigenvectors are pairwise orthogonal (as Theorem 8.2.4 asserts). Normalizing them gives an orthogonal matrix

$$P = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 \\ \sqrt{2} & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

Then $P^{-1} = P^T$ and $P^TAP = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & -4 \end{bmatrix}$.

f. $$\begin{aligned} c_A(x) &= \begin{vmatrix} x-5 & 2 & 4 \\ 2 & x-8 & 2 \\ 4 & 2 & x-5 \end{vmatrix} = \begin{vmatrix} x-9 & 0 & 9-x \\ 2 & x-8 & 2 \\ 4 & 2 & x-5 \end{vmatrix} = \begin{vmatrix} x-9 & 0 & 0 \\ 2 & x-8 & 4 \\ 4 & 2 & x-1 \end{vmatrix} \\ &= (x-9)\begin{vmatrix} x-8 & 4 \\ 2 & x-1 \end{vmatrix} = (x-9)(x^2-9x) = x(x-9)^2. \end{aligned}$$

The eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 9$.

$\lambda_1 = 0$: $\begin{bmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & -4 & 1 \\ 0 & -18 & 9 \\ 0 & 18 & -9 \end{bmatrix} \to \begin{bmatrix} 1 & -4 & 1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$; $E_0(A) = \text{span}\left\{\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}\right\}$.

$\lambda_2 = 9$: $\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $E_9(A) = \text{span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}\right\}$.

However, these are not orthogonal and the Gram-Schmidt algorithm replaces $\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$ with $Z_2 = \begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix}$. Hence $P = \begin{bmatrix} \frac{2}{3} & \frac{-1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & \frac{-4}{3\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \end{bmatrix} = \frac{1}{3\sqrt{2}} \begin{bmatrix} 2\sqrt{2} & -3 & 1 \\ \sqrt{2} & 0 & -4 \\ 2\sqrt{2} & 3 & 1 \end{bmatrix}$ is orthogonal and satisfies $P^TAP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

We note in passing that $\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$ are another orthogonal basis of $E_9(A)$, so $Q = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$ also satisfies $Q^TAQ = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

h. To evaluate $c_A(x)$, we begin adding rows 2, 3 and 4 to row 1.

$$\begin{aligned} c_A(x) &= \begin{vmatrix} x-3 & -5 & 1 & -1 \\ -5 & x-3 & -1 & 1 \\ 1 & -1 & x-3 & -5 \\ -1 & 1 & -5 & x-3 \end{vmatrix} = \begin{vmatrix} x-8 & x-8 & x-8 & x-8 \\ -5 & x-3 & -1 & 1 \\ 1 & -1 & x-3 & -5 \\ -1 & 1 & -5 & x-3 \end{vmatrix} \\ &= \begin{vmatrix} x-8 & 0 & 0 & 0 \\ -5 & x-2 & 4 & 6 \\ 1 & -2 & x-4 & -6 \\ -1 & 2 & -4 & x-2 \end{vmatrix} = (x-8)\begin{vmatrix} x+2 & 4 & 6 \\ -2 & x-4 & -6 \\ 2 & -4 & x-2 \end{vmatrix} \\ &= (x-8)\begin{vmatrix} x+2 & 4 & 6 \\ x & x & 0 \\ 2 & -4 & x-2 \end{vmatrix} = (x-8)\begin{vmatrix} x-2 & 4 & 6 \\ 0 & x & 0 \\ 6 & -4 & x-2 \end{vmatrix} \\ &= x(x-8)\begin{vmatrix} x-2 & 6 \\ 6 & x-2 \end{vmatrix} = x(x-8)(x^2-4x-32) = x(x+4)(x-8)^2 \end{aligned}$$

$$\lambda_1 = 0: \quad \begin{bmatrix} -3 & -5 & 1 & -1 \\ -5 & -3 & -1 & 1 \\ 1 & -1 & -3 & -5 \\ -1 & 1 & -5 & -3 \end{bmatrix} \to \begin{bmatrix} -3 & -5 & 1 & -1 \\ -8 & -8 & 0 & 0 \\ 1 & -1 & -3 & -5 \\ 0 & 0 & -8 & -8 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & -3 & -5 \\ 0 & -8 & -8 & -16 \\ 0 & -16 & -24 & -40 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & -2 & -3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \ E_0(A) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

$$\lambda_2 = -4: \quad \begin{bmatrix} -7 & -5 & 1 & -1 \\ -5 & -7 & -1 & 1 \\ 1 & -1 & -7 & -5 \\ -1 & 1 & -5 & -7 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & -7 & -5 \\ 0 & -12 & -48 & -36 \\ 0 & -12 & -36 & -24 \\ 0 & 0 & -12 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -3 & -2 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \ E_{-4}(A) = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

$$\lambda_3 = 8: \quad \begin{bmatrix} 5 & -5 & 1 & -1 \\ -5 & 5 & -1 & 1 \\ 1 & -1 & 5 & -5 \\ -1 & 1 & -5 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 5 & -5 \\ 0 & 0 & -24 & 24 \\ 0 & 0 & 24 & -24 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$$E_8(A) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$\text{Hence, } P = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & \sqrt{2} & 0 \\ -1 & 1 & \sqrt{2} & 0 \\ -1 & -1 & 0 & \sqrt{2} \\ 1 & 1 & 0 & \sqrt{2} \end{bmatrix} \text{ gives } P^TAP = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

6. $c_A(x) = \begin{vmatrix} x & -a & 0 \\ -a & x & -c \\ 0 & -c & x \end{vmatrix} = x \begin{vmatrix} x & -c \\ -c & x \end{vmatrix} + a \begin{vmatrix} -a & 0 \\ -c & x \end{vmatrix} = x(x^2 - c^2) - a^2x = x(x^2 - k^2).$

Hence $c_A(x) = x(x-k)(x+k)$, where $k^2 = a^2 + c^2$, so the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = k$, $\lambda_3 = -k$. They are all distinct ($k \neq 0$, and $a \neq 0$ or $c \neq 0$) so the eigenspaces are all one dimensional.

$$\lambda_1 = 0: \begin{bmatrix} 0 & -a & 0 \\ -a & 0 & -c \\ 0 & -c & 0 \end{bmatrix} \begin{bmatrix} c \\ 0 \\ -a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; E_0(A) = \operatorname{span}\left\{ \begin{bmatrix} c \\ 0 \\ -a \end{bmatrix} \right\}$$

$$\lambda_2 = k: \begin{bmatrix} k & -a & 0 \\ -a & k & -c \\ 0 & -c & k \end{bmatrix} \begin{bmatrix} a \\ k \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; E_k(A) = \operatorname{span}\left\{ \begin{bmatrix} a \\ k \\ c \end{bmatrix} \right\}$$

$$\lambda_3 = -k: \begin{bmatrix} -k & -a & 0 \\ -a & -k & -c \\ 0 & -c & -k \end{bmatrix} \begin{bmatrix} a \\ -k \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}; E_{-k}(A) = \operatorname{span}\left\{ \begin{bmatrix} a \\ -k \\ c \end{bmatrix} \right\}$$

These eigenvalues are orthogonal and have length, k, $\sqrt{2}k$, $\sqrt{2}k$ respectively. Hence,

$P = \frac{1}{\sqrt{2}k} \begin{bmatrix} c\sqrt{2} & a & a \\ 0 & k & -k \\ -a\sqrt{2} & c & c \end{bmatrix}$ is orthogonal and $P^TAP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & -k \end{bmatrix}$.

10. Similar to Example 8.2.6, q has matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ with eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 2$ and corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ respectively. Hence $P = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$ is orthogonal and $P^TAP = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$. Let

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{y} = P^T\mathbf{x} = \frac{1}{\sqrt{5}} \begin{bmatrix} -x_1 + 2x_2 \\ 2x_1 + x_2 \end{bmatrix}; \text{ so } y_1 = \frac{1}{\sqrt{5}}(-x_1 + 2x_2) \text{ and } y_2 = \frac{1}{\sqrt{5}}(2x_1 + x_2).$$

Then $q = -3y_1^2 + 2y_2^2$ is diagonalized by these variables.

11. (c) $\Rightarrow$ (a). By Theorem 8.2.1 let $P^{-1}AP = D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ where the λ_i are the eigenvalues of A. By (c) we have $\lambda_i = \pm 1$ for each i. It follows that

$$D^2 = \operatorname{diag}(\lambda_1^2, \dots, \lambda_n^2) = \operatorname{diag}(1, \dots, 1) = I$$

Since $A = PDP^{-1}$, we obtain $A^2 = (PDP^{-1})^2 = PD^2P^{-1} = PIP^{-1} = I$. Since A is symmetric, this proves (a).

13. b. Let A and B be orthogonally similar, say $B = P^TAP$ where $P^T = P^{-1}$. Then $B^2 = P^TAPP^TAP = P^TAIAP = P^TA^2P$. Hence A^2 and B^2 are orthogonally similar.

15. Assume that $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot A\mathbf{y}$ for all columns $\mathbf{x}$ and $\mathbf{y}$; we must show that $A^T = A$. We have $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x}^TA^T\mathbf{y}$ and $\mathbf{x} \cdot A\mathbf{y} = \mathbf{x}^TA\mathbf{y}$, so the given condition asserts that

$$\mathbf{x}^TA^T\mathbf{y} = \mathbf{x}^TA\mathbf{y} \text{ for all columns } \mathbf{x} \text{ and } \mathbf{y}. \tag{$*$}$$

But if E_j denotes column j of the identity matrix, then writing $A = \left[a_{ij}\right]$ we have

$$\mathbf{e}_i^TA\mathbf{e}_j = a_{ij} \text{ for all } i \text{ and } j.$$

Since $(*)$ shows that A^T and A have the same $(i,\ j)$-entry for each i and j. In other words, $A^T = A$.

Note that the same argument shows that if A and B are matrices with the property that $\mathbf{x}^TB\mathbf{y} = \mathbf{x}^TA\mathbf{y}$ for all columns $\mathbf{x}$ and $\mathbf{y}$, then $B = A$.

18. b. If $P = \left[\begin{array}{cc} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{array}\right]$ and $Q = \left[\begin{array}{cc} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{array}\right]$ then P and Q are orthogonal matrices, $\det P = 1$ and $\det Q = -1$. (We note that every 2×2 orthogonal matrix has the form of P or Q for some θ.)

d. Since P is orthogonal, $P^T = P^{-1}$. Hence

$$P^T(I-P) = P^T - P^TP = P^T - I = -(I-P^T) = -(I-P)^T$$

Since P is $n \times n$, taking determinants gives

$$\det P^T \det(I-P) = (-1)^n \det[(I-P)^T] = (-1)^n \det(I-P)$$

Hence, if $I-P$ is invertible, then $\det(I-P) \neq 0$ so this gives $\det P^T = (-1)^n$; that is $\det P = (-1)^n$, contrary to assumption.

21. By the definition of matrix multiplication, the $(i,\ j)$-entry of AA^T is $\mathbf{r}_i \cdot \mathbf{r}_j$. This is zero if $i \neq j$, and equals $\|\mathbf{r}_i\|^2$ if $i = j$. Hence, $AA^T = D = \text{diag}\left(\|\mathbf{r}_1\|^2,\ \|\mathbf{r}_2\|^2,\ \dots,\ \|\mathbf{r}_n\|^2\right)$. Since D is invertible ($\|\mathbf{r}_i\|^2 \neq 0$ for each i), it follows that A is invertible and, since row i of A^T is $\left[\begin{array}{ccccccc} a_{1i} & a_{2i} & \cdots & a_{ji} & \cdots & a_{ni} \end{array}\right]$

$$A^{-1} = A^TD^{-1} = \left[\begin{array}{ccccc} \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{1i} & \cdots & a_{ji} & \cdots & a_{ni} \\ \vdots & \cdots & \vdots & \cdots & \vdots \end{array}\right] \left[\begin{array}{cccc} \frac{1}{\|\mathbf{r}_1\|^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\|\mathbf{r}_2\|^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\|\mathbf{r}_n\|^2} \end{array}\right]$$

Thus, the $(i,\ j)$-entry of A^{-1} is $\frac{a_{ji}}{\|\mathbf{r}_j\|^2}$.

23. b. Observe first that $I-A$ and $I+A$ commute, whence $I-A$ and $(I+A)^{-1}$ commute. Moreover, $\left[(I+A)^{-1}\right]^T = \left[(I+A)^T\right]^{-1} = (I^T+A^T)^{-1} = (I-A)^{-1}$. Hence,

$$PP^T = (I-A)(I+A)^{-1}[(I-A)(I+A)^{-1}]^T$$

$$
\begin{aligned}
&= (I-A)(I+A)^{-1}[(I+A)^{-1}]^T(I-A)^T \\
&= (I-A)(I+A)^{-1}(I-A)^{-1}(I+A) \\
&= (I+A)^{-1}(I-A)(I-A)^{-1}(I+A) \\
&= (I+A)^{-1}I(I+A) \\
&= I
\end{aligned}
$$

8.3 Positive Definite Matrices

1. b. $\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & -1 \\ 0 & \frac{1}{2} \end{bmatrix}$. Then $A = U^TU$ where $U = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{\sqrt{2}}{2}\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$.

   d. $\begin{bmatrix} 20 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 5 \end{bmatrix} \to \begin{bmatrix} 20 & 4 & 5 \\ 0 & \frac{6}{5} & 2 \\ 0 & 2 & \frac{15}{4} \end{bmatrix} \to \begin{bmatrix} 20 & 4 & 5 \\ 0 & \frac{6}{5} & 2 \\ 0 & 0 & \frac{5}{12} \end{bmatrix}$.

   Hence, $U = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{5}{2\sqrt{5}} \\ 0 & \frac{6}{\sqrt{30}} & \frac{10}{\sqrt{30}} \\ 0 & 0 & \frac{5}{2\sqrt{15}} \end{bmatrix} = \frac{1}{30}\begin{bmatrix} 60\sqrt{5} & 12\sqrt{5} & 15\sqrt{5} \\ 0 & 6\sqrt{30} & 10\sqrt{30} \\ 0 & 0 & 5\sqrt{15} \end{bmatrix}$, and $A = U^TU$.

2. b. If λ^k is positive and k is odd, then λ is positive.

4. Assume $\mathbf{x} \neq \mathbf{0}$ is a column. If A and B are positive definite then $\mathbf{x}^TA\mathbf{x} > 0$ and $\mathbf{x}^TB\mathbf{x} > 0$ so

$$\mathbf{x}^T(A+B)\mathbf{x} = \mathbf{x}^TA\mathbf{x} + \mathbf{x}^TB\mathbf{x} > 0 + 0 = 0$$

   Thus $A+B$ is positive definite. Now suppose $r > 0$. Then $\mathbf{x}^T(rA)\mathbf{x} = r(\mathbf{x}^TA\mathbf{x}) > 0$, proving that rA is positive definite.

6. Given $\mathbf{x}$ in $\mathbb{R}^n$, $\mathbf{x}^T(U^TAU)\mathbf{x} = (U\mathbf{x})^TA(U\mathbf{x}) > 0$ provided $U\mathbf{x} \neq \mathbf{0}$ (because A is positive definite). Write $U = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_m \end{bmatrix}$ where $\mathbf{c}_j$ in $\mathbb{R}^n$ is column j of U. If $0 \neq \mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}^T$, then $U\mathbf{x} = \sum x_j\mathbf{c}_j \neq \mathbf{0}$ because the $\mathbf{c}_j$ are independent [rank of U is m].

10. Since A is symmetric, the principal axis theorem asserts that an orthogonal matrix P exists such that $P^TAP = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ where the λ_i are the eigenvalues of A. Since each $\lambda_i > 0$, $\sqrt{\lambda_i}$ is real and positive, so define $B = \text{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}\right)$. Then $B^2 = D$. As $A = PDP^T$, take $C = PBP^T$. Then

$$C^2 = PBP^TPBP^T = PB^2P^T = PDP^T = A$$

   Finally, C is symmetric because B is symmetric $\left(C^T = P^{TT}B^TP^T = PBP^T = C\right)$ and C has eigenvalues $\sqrt{\lambda_i} > 0$ (C is similar to B). Hence C is positive definite.

12. b. Suppose that A is positive definite so $A = U_0^TU_0$ where U_0 is upper triangular with positive diagonal entries $d_1, d_2, \dots, d_n$. Put $D_0 = \text{diag}(d_1, d_2, \dots, d_n)$. Then $L = U_0^TD_0^{-1}$ is lower triangular with 1's on the diagonal, $U = D_0^{-1}U_0$ is upper triangular with 1's on the diagonal, and $A = LD_0^2U$. Take $D = D_0^2$.

    Conversely, if $A = LDU$ as in (a), then $A^T = U^TDL^T$. Hence, $A^T = A$ implies that $U^TDL^T = LDU$, so $U^T = L$ and $L^T = U$ by (a). Hence, $A = U^TDU$. If $D = \text{diag}(d_1, d_2, \dots, d_n)$, let $D_1 = \text{diag}\left(\sqrt{d_1}, \sqrt{d_2}, \dots, \sqrt{d_n}\right)$. Then $D = D_1^2$ so $A = U^TD_1^2U = (D_1U)^T(D_1U)$. Hence, A is positive definite.

8.4 QR-Factorization

1. b. The columns of A are $\mathbf{c}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. First apply the Gram-Schmidt algorithm

$$\mathbf{f}_1 = \mathbf{c}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$\mathbf{f}_2 = \mathbf{c}_2 - \frac{\mathbf{c}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{3}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} \\ \frac{2}{5} \end{bmatrix}$$

Now normalize to obtain

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$\mathbf{q}_2 = \frac{1}{\|\mathbf{f}_2\|}\mathbf{f}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

Hence $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$ is an orthogonal matrix. We obtain R from equation (8.5) preceding Theorem 8.4.1:

$$L = \begin{bmatrix} \|\mathbf{f}_1\| & \mathbf{c}_2 \cdot \mathbf{q}_1 \\ 0 & \|\mathbf{f}_2\| \end{bmatrix} = \begin{bmatrix} \sqrt{5} & \frac{3}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 5 & 3 \\ 0 & 1 \end{bmatrix}$$

Then $A = QR$.

d. The columns of A are $\mathbf{c}_1 = \begin{bmatrix} 1 & -1 & 0 & 1 \end{bmatrix}^T$, $\mathbf{c}_2 = \begin{bmatrix} 1 & 0 & 1 & -1 \end{bmatrix}^T$ and $\mathbf{c}_3 = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}^T$. Apply the Gram-Schmidt algorithm

$$\mathbf{f}_1 = \mathbf{c}_1 = \begin{bmatrix} 1 & -1 & 0 & 1 \end{bmatrix}^T$$
$$\mathbf{f}_2 = \mathbf{c}_2 - \frac{\mathbf{c}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = \begin{bmatrix} 1 & 0 & 1 & -1 \end{bmatrix}^T - \frac{0}{3}F_1 = \begin{bmatrix} 1 & 0 & 1 & -1 \end{bmatrix}^T$$
$$\begin{aligned} \mathbf{f}_3 &= \mathbf{c}_3 - \frac{\mathbf{c}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{c}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\ &= \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}^T - \frac{-1}{3}\begin{bmatrix} 1 & -1 & 0 & 1 \end{bmatrix}^T - \frac{1}{3}\begin{bmatrix} 1 & 0 & 1 & -1 \end{bmatrix}^T \\ &= \frac{2}{3}\begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}^T \end{aligned}$$

Normalize

$$Q_1 = \frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & -1 & 0 & 1 \end{bmatrix}^T$$
$$Q_2 = \frac{1}{\|\mathbf{f}_2\|}\mathbf{f}_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 0 & 1 & -1 \end{bmatrix}^T$$
$$Q_3 = \frac{1}{\|\mathbf{f}_3\|}\mathbf{f}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}^T$$

Hence $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{bmatrix} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ has orthonormal columns. We obtain R from equation (8.5) preceding Theorem 8.4.1:

$$R = \begin{bmatrix} \|\mathbf{f}_1\| & \mathbf{c}_2 \cdot \mathbf{q}_1 & \mathbf{c}_3 \cdot \mathbf{q}_1 \\ 0 & \|\mathbf{f}_2\| & \mathbf{c}_3 \cdot \mathbf{q}_2 \\ 0 & 0 & \|\mathbf{f}_3\| \end{bmatrix} = \begin{bmatrix} \sqrt{3} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ 0 & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix} = \frac{1}{\sqrt{3}}\begin{bmatrix} 3 & 0 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Then $A = QR$.

2. b. If $A = QR$ is a QR-factorization of A, then R has independent columns (it is invertible) as does Q (its columns are orthonormal). Hence A has independent columns by (a). The converse is by Theorem 8.4.1.

8.5 Computing Eigenvalues

1. b. $A = \begin{bmatrix} 5 & 2 \\ -3 & -2 \end{bmatrix}$. Then $c_A(x) = \begin{vmatrix} x-5 & -2 \\ 3 & x+2 \end{vmatrix} = (x+1)(x-4)$, so $\lambda_1 = -1$, $\lambda_2 = 4$.
If $\lambda_1 = -1$: $\begin{bmatrix} -6 & -2 \\ 3 & 1 \end{bmatrix} \to \begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix}$; eigenvector $= \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.
If $\lambda_2 = 4$: $\begin{bmatrix} -1 & -2 \\ 3 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$; dominant eigenvector $= \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.
Starting with $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, the power method gives $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1, \dots$:

$$\mathbf{x}_1 = \begin{bmatrix} 7 \\ -5 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 25 \\ -11 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 103 \\ -53 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 409 \\ -203 \end{bmatrix}$$

These are approaching (scalar multiples of) the dominant eigenvector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. The Rayleigh quotients are $r_k = \frac{\mathbf{x}_k \cdot \mathbf{x}_{k+1}}{\|\mathbf{x}_k\|^2}$, $k = 0, 1, 2, \dots$, so $r_0 = 1$, $r_1 = 3.29$, $r_2 = 4.23$, $r_3 = 3.94$. These are approaching the dominant eigenvalue 4.

d. $A = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}$; $c_A(x) = \begin{vmatrix} x-3 & -1 \\ -1 & x \end{vmatrix} = x^2 - 3x - 1$, so the eigenvalues are $\lambda_1 = \frac{1}{2}(3+\sqrt{13})$, $\lambda_2 = \frac{1}{2}(3-\sqrt{13})$. Thus the dominant eigenvalue is $\lambda_1 = \frac{1}{2}(3+\sqrt{13})$. Since $\lambda_1\lambda_2 = -1$ and $\lambda_1 + \lambda_2 = 3$, we get

$$\begin{bmatrix} \lambda_1 - 3 & -1 \\ -1 & \lambda_1 \end{bmatrix} \to \begin{bmatrix} 1 & -\lambda_1 \\ 0 & 0 \end{bmatrix}$$

so a dominant eigenvector is $\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$. We start with $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
Then $\mathbf{x}_{k+1} = A\mathbf{x}_k$, $k = 0, 1, \dots$ gives

$$\mathbf{x}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 13 \\ 4 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 43 \\ 13 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 142 \\ 43 \end{bmatrix}$$

These are approaching scalar multiples of the dominant eigenvector $\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.302776 \\ 1 \end{bmatrix}$. The Rayleigh quotients are $r_k = \frac{\mathbf{x}_k \cdot \mathbf{x}_{k+1}}{\|\mathbf{x}_k\|^2}$:

$$r_0 = 2.5,\ r_1 = 3.29,\ r_2 = 3.30270,\ r_3 = 3.30278$$

These are rapidly approaching the dominant eigenvalue $\lambda_1 = 3.302776$.

2. b. $A = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}$; $c_A(x) = \begin{vmatrix} x-3 & -1 \\ -1 & x \end{vmatrix} = x^2 - 3x - 3$; $\lambda_1 = \frac{1}{2}\left[3+\sqrt{13}\right] = 3.302776$ and $\lambda_2 = \frac{1}{2}\left[3-\sqrt{13}\right] = -0.302776$. The QR-algorithm proceeds as follows:
$A_1 = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} = Q_1R_1$ where $Q_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}$, $R_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 10 & 3 \\ 0 & 1 \end{bmatrix}$.

$A_2 = R_1Q_1 = \frac{1}{10}\begin{bmatrix} 33 & 1 \\ 1 & -3 \end{bmatrix} = Q_2R_2$ where $Q_2 = \frac{1}{\sqrt{1090}}\begin{bmatrix} 33 & 1 \\ 1 & -33 \end{bmatrix}$, $R_2 = \frac{1}{\sqrt{1090}}\begin{bmatrix} 109 & 3 \\ 0 & 10 \end{bmatrix}$.
$A_3 = R_2Q_2 = \frac{1}{109}\begin{bmatrix} 360 & 1 \\ 1 & -33 \end{bmatrix} = \begin{bmatrix} 3.302752 & 0.009174 \\ 0.009174 & -0.302752 \end{bmatrix}$.
The diagonal entries already approximate λ_1 and λ_2 to 4 decimal places.

4. We prove that $A_k^T = A_k$ for each k by induction in k. If $k = 1$, then $A_1 = A$ is symmetric by hypothesis, so assume $A_k^T = A_k$ for some $k \geq 1$. We have $A_k = Q_kR_k$ so $R_k = Q_k^{-1}A_k = Q_k^TA_k$ because Q_k is orthogonal. Hence

$$A_{k+1} = R_kQ_k = Q_k^TA_kQ_k$$

so

$$A_{k+1}^T = (Q_k^TA_kQ_k)^T = Q_k^TA_k^TQ_k^{TT} = Q_k^TA_kQ_k = A_{k+1}$$

The eigenvalues of A are all real as A is symmetric, so the QR-algorithm asserts that the A_k converge to an upper triangular matrix T. But T is symmetric (it is the limit of symmetric matrices), so it is diagonal.

8.6 Singular Value Decomposition

4. b. $t\sigma_1, \ldots, t\sigma_r$.

7. If $A = U\Sigma V^T$ then Σ is invertible, so $A^{-1} = V\Sigma^{-1}U^T$ is a SVD.

8. b. First $A^TA = I_n$ so $\Sigma_A = I_n$.

$$\begin{aligned} A &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

9. b.

$$A = F$$
$$= \frac{1}{5}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} 20 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

13. b. If $\mathbf{x} \in \mathbb{R}^n$ then $\mathbf{x}^T(G+H)\mathbf{x} = \mathbf{x}^TG\mathbf{x} + \mathbf{x}^TH\mathbf{x} \geq 0 + 0 = 0$.

17. b. $\begin{bmatrix} \frac{1}{4} & 0 & \frac{1}{4} \\ -\frac{1}{4} & 0 & -\frac{1}{4} \end{bmatrix}$

8.7 Complex Matrices

1. b. $\sqrt{|1-i|^2+|1+i|^2+1^2+(-1)^2}=\sqrt{(1+1)+(1+1)+1+1}=\sqrt{6}$

   d. $\sqrt{4+|-i|^2+|1+i|^2+|1-i|^2+|2i|^2}=\sqrt{4+1+(1+1)+(1+1)+4}=\sqrt{13}$

2. b. Not orthogonal: $\langle(i,\ -i,\ 2+i),\ (i,\ i,\ 2-i)\rangle=i(-i)+(-i)(-i)+(2+i)(2+i)=3+4i$

   d. Orthogonal: $\langle(4+4i,\ 2+i,\ 2i),\ (-1+i,\ 2,\ 3-2i)\rangle=(4+4i)(-1-i)+(2+i)2+(2i)(3+2i)=(-8i)+(4+2i)+(-4+6i)=0$.

3. b. Not a subspace. For example, $i(0,\ 0,\ 1)=(0,\ 0,\ i)$ is not in U.

   d. If $\mathbf{v}=(v+w,\ v-2w,\ v)$ and $\mathbf{w}=(v'+w',\ v'-2w',\ v')$ are in U then

$$\begin{aligned}\mathbf{v}+\mathbf{w}&=((v+v')+(w+w'),\ (v+v')-2(w+w'),\ (v+v'))\text{ is in }U\\ z\mathbf{v}&=(zv+zw,\ zv-2zw,\ zv)\text{ is in }U\\ \mathbf{0}&=(0+0,\ 0-20,\ 0)\text{ is in }U\end{aligned}$$

   Hence U is a subspace.

4. b. Here $U=\{(iv+w,\ 0,\ 2v-w)\mid v,\ w\in\mathbb{C}\}=\{v(i,\ 0,\ 2)+w(1,\ 0,\ -1)\mid v,\ w\in\mathbb{C}\}=\operatorname{span}\{(i,\ 0,\ 2),\ (1,\ 0,\ -1)\}$.
   If $z(i,\ 0,\ 2)+t(1,\ 0,\ -1)=(0,\ 0,\ 0)$ with $z,\ t\in\mathbb{C}$, then $iz+t=0$, $2z-t=0$. Adding gives $(2+i)z=0$, so $z=0$; and so $t=-iz=0$. Thus $\{(i,\ 0,\ 2),\ (1,\ 0,\ -1)\}$ is independent over $\mathbb{C}$, and so is a basis of U. Hence $\dim_{\mathbb{C}}U=2$.

   d. $U=\{(u,\ v,\ w)\mid 2u+(1+i)v-iw=0;u,\ v,\ w\in\mathbb{C})\}$. The condition is $w=-2iu+(1-i)v$, so

$$U=\{(u,\ v,\ -2iu+(1-i)v)\mid u,\ v\in\mathbb{C}\}=\operatorname{span}\{(1,\ 0,\ -2i),\ (0,\ 1,\ 1-i)\}$$

   If $z(1,\ 0,\ -2i)+t(0,\ 1,\ i-1)=(0,\ 0,\ 0)$ then components 1 and 2 give $z=0$ and $t=0$. Thus $\{(1,\ 0,\ -2i),\ (0,\ 1,\ 1-i)\}$ is independent over $\mathbb{C}$, and so is a basis of U. Hence $\dim_{\mathbb{C}}U=2$.

5. b. $A=\begin{bmatrix}2&3\\-3&2\end{bmatrix}$, $A^H=A^T=\begin{bmatrix}2&-3\\3&2\end{bmatrix}$, $A^{-1}=\frac{1}{13}\begin{bmatrix}2&-3\\3&2\end{bmatrix}$. Hence, A is not hermitian ($A\neq A^H$) and not unitary ($A^{-1}\neq A^H$). However, $AA^H=13I=A^HA$, so A is normal.

   d. $A=\begin{bmatrix}1&-i\\i&-1\end{bmatrix}$, $A^H=(\overline{A})^T=\begin{bmatrix}1&i\\-i&-1\end{bmatrix}^T=\begin{bmatrix}1&-i\\i&-1\end{bmatrix}=A$. Thus A is hermitian and so is normal. But, $AA^H=A^2=2I$ so A is not unitary.

   f. $A=\begin{bmatrix}1&1+i\\1+i&i\end{bmatrix}$. Here $A=A^T$ so $A^H=\overline{A}=\begin{bmatrix}1&1-i\\1-i&-i\end{bmatrix}\neq A$ (thus A is not hermitian). Next, $AA^H=\begin{bmatrix}3&2-2i\\2+i&3\end{bmatrix}\neq I$ so A is not unitary. Finally, $A^HA=\begin{bmatrix}3&2+2i\\2-2i&3\end{bmatrix}\neq AA^H$, so A is not normal.

   h. $A=\frac{1}{\sqrt{2}|z|}\begin{bmatrix}z&z\\\overline{z}&-\overline{z}\end{bmatrix}$. Here $\overline{A}=\frac{1}{\sqrt{2}|z|}\begin{bmatrix}\overline{z}&\overline{z}\\z&-z\end{bmatrix}$ so $A^H=\frac{1}{\sqrt{2}|z|}\begin{bmatrix}\overline{z}&z\\\overline{z}&-z\end{bmatrix}$. Thus $A=A^H$ if and only if $z=\overline{z}$; that is A is hermitian if and only if z is real. We have $AA^H=\frac{1}{2|z|^2}\begin{bmatrix}2|z|^2&0\\0&2|z|^2\end{bmatrix}=I$, and similarly, $A^HA=I$. Thus it is unitary (and hence normal).

8. b. $A = \begin{bmatrix} 4 & 3-i \\ 3+i & 1 \end{bmatrix}$, $c_A(x) = \begin{bmatrix} x-4 & -3+i \\ -3-i & x-1 \end{bmatrix} = x^2 - 5x - 6 = (x+1)(x-6)$.

Eigenvectors for $\lambda_1 = -1$: $\begin{bmatrix} -5 & -3+i \\ -3-i & 1 \end{bmatrix} \to \begin{bmatrix} 3+i & 2 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} -2 \\ 3+i \end{bmatrix}$.

Eigenvectors for $\lambda_2 = 6$: $\begin{bmatrix} 2 & -3+i \\ -3-i & 5 \end{bmatrix} \to \begin{bmatrix} 2 & -3+i \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 3-i \\ 2 \end{bmatrix}$.

As $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal and $\|\mathbf{x}_1\| = \|\mathbf{x}_2\| = \sqrt{14}$, $U = \frac{1}{\sqrt{14}} \begin{bmatrix} -2 & 3-i \\ 3+i & 2 \end{bmatrix}$ is unitary and $U^H AU = \begin{bmatrix} -1 & 0 \\ 0 & 6 \end{bmatrix}$.

d. $A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}$; $c_A(x) = \begin{vmatrix} x-2 & -1-i \\ -1+i & x-3 \end{vmatrix} = x^2 - 5x + 4 = (x-1)(x-4)$.

Eigenvectors for $\lambda_1 = 1$: $\begin{bmatrix} -1 & -1-i \\ -1+i & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1+i \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1+i \\ -1 \end{bmatrix}$.

Eigenvectors for $\lambda_2 = 4$: $\begin{bmatrix} 2 & -1-i \\ -1+i & 1 \end{bmatrix} \to \begin{bmatrix} -1+i & 1 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1-i \end{bmatrix}$.

Since $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal and $\|\mathbf{x}_1\| = \|\mathbf{x}_2\| = \sqrt{3}$, $U = \frac{1}{\sqrt{3}} \begin{bmatrix} 1+i & 1 \\ -1 & 1-i \end{bmatrix}$ is unitary and $U^H AU = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$.

f. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1+i \\ 0 & 1-i & 2 \end{bmatrix}$; $c_A(x) = \begin{vmatrix} x-1 & 0 & 0 \\ 0 & x-1 & -1-i \\ 0 & -1+i & x-2 \end{vmatrix} = (x-1)(x^2-3x) = (x-1)x(x-3)$.

Eigenvectors for $\lambda_1 = 1$: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1-i \\ 0 & -1+i & -1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

If $\lambda_2 = 0$: $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & -1-i \\ 0 & -1+i & -i \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1+i \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1+i \\ -1 \end{bmatrix}$.

Eigenvectors for $\lambda_3 = 3$: $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & -1-i \\ 0 & -1+i & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1+i & 1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1-i \end{bmatrix}$.

Since $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is orthogonal and $\|\mathbf{x}_2\| = \|\mathbf{x}_3\| = \sqrt{3}$, $U = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1+i & 1 \\ 0 & -1 & 1-i \end{bmatrix}$ is orthogonal and $U^* AU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

10. b. (1) If $\mathbf{z} = (z_1, z_2, \dots, z_n)$ then $\|\mathbf{z}\|^2 = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2$. Thus $\|\mathbf{z}\| = 0$ if and only if $|z_1| = \cdots = |z_n| = 0$, if and only if $\mathbf{z} = (0, 0, \dots, 0)$.

(2) By Theorem 8.7.1, we have $\langle \lambda\mathbf{z}, \mathbf{w} \rangle = \lambda \langle \mathbf{z}, \mathbf{w} \rangle$ and $\langle \mathbf{z}, \lambda\mathbf{w} \rangle = \overline{\lambda} \langle \mathbf{z}, \mathbf{w} \rangle$. Hence

$$\|\lambda\mathbf{z}\|^2 = \langle \lambda\mathbf{z}, \lambda\mathbf{z} \rangle = \lambda \langle \mathbf{z}, \lambda\mathbf{z} \rangle = \lambda\overline{\lambda} \langle \mathbf{z}, \mathbf{z} \rangle = |\lambda|^2 \|\mathbf{z}\|^2$$

Taking positive square roots gives $\|\lambda\mathbf{z}\| = |\lambda| \|\mathbf{z}\|$.

11. b. If A is hermitian then $\overline{A} = A^T$. If $A = \left[a_{ij}\right]$, the (k, k)-entry of $\overline{A}$ is $\overline{a}_{kk}$, and the (k, k)-entry of A^T is a_{kk}. Thus, $\overline{A} = A^T$ implies that $\overline{a}_{kk} = a_{kk}$ for each k; that is a_{kk} is real.

14. b. Let B be skew-hermitian, that is $B^H = -B$. Then Theorem 8.7.3 gives

$$(B^2)^H = (B^H)^2 = (-B)^2 = B^2, \text{ so } B^2 \text{ is hermitian}$$
$$(iB)^H = (-i)B^H = (-i)(-B) = iB, \text{ so } iB \text{ is hermitian.}$$

d. If $Z = A + B$ where $A^H = A$ and $B^H = -B$, then $Z^H = A^H + B^H = V - B$. Solving gives $Z + Z^H = 2V$ and $Z - Z^H = 2B$, so $V = \frac{1}{2}(Z + Z^H)$ and $S = \frac{1}{2}(Z + Z^H)$. Hence the matrices A

and B are uniquely determined by the conditions $Z = A + B$, $A^H = A$, $B^H = -B$, provided such A and B exist. But always,

$$Z = \frac{1}{2}(Z + Z^H) + \frac{1}{2}(Z - Z^H)$$

and the matrices $A = \frac{1}{2}(Z + Z^H)$ and $B = \frac{1}{2}(Z - Z^H)$ are hermitian and skew-hermitian respectively:

$$\begin{aligned} A^H &= \tfrac{1}{2}(Z^H + Z^{HH}) = \tfrac{1}{2}(Z^H + Z) = A \\ B^H &= \tfrac{1}{2}(Z^H - Z^{HH}) = \tfrac{1}{2}(Z^H - Z) = -B \end{aligned}$$

16. b. If U is unitary, then $U^{-1} = U^H$. We must show that U^{-1} is unitary, that is $(U^{-1})^{-1} = (U^{-1})^H$. But

$$(U^{-1})^{-1} = U = (U^H)^H = (U^{-1})^H$$

18. b. If $V = \left[\begin{array}{cc} 1 & i \\ -i & 0 \end{array}\right]$ then V is hermitian because $\overline{V} = \left[\begin{array}{cc} 1 & -i \\ i & 0 \end{array}\right] = V^T$, but $iV = \left[\begin{array}{cc} i & -1 \\ 1 & 0 \end{array}\right]$ is not hermitian (it has a nonreal entry on the main diagonal).

21. b. Given $A = \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right]$, let $U = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ be invertible and real, and assume that $U^{-1}AU = \left[\begin{array}{cc} \lambda & \mu \\ 0 & \nu \end{array}\right]$. Thus, $AU = U\left[\begin{array}{cc} \lambda & \mu \\ 0 & \nu \end{array}\right]$ so

$$\left[\begin{array}{cc} c & d \\ -a & -b \end{array}\right] = \left[\begin{array}{cc} a\lambda & a\mu + b\nu \\ c\lambda & c\mu + d\nu \end{array}\right]$$

Equating first column entries gives $c = a\lambda$ and $-a = c\lambda$. Thus, $-a = (a\lambda)\lambda = a\lambda^2$ so $(1 + \lambda^2)a = 0$. Now λ is real (a and c are not both zero so either $\lambda = \frac{c}{a}$ or $\lambda = -\frac{a}{c}$), so $1 + \lambda^2 \neq 0$. Thus $a = 0$ (because $(1 + \lambda^2)a = 0$) whence $c = a\lambda = 0$. This contradicts the assumption that A is invertible.

8.8 An Application to Linear Codes over Finite Fields

1. b. The elements with inverses are 1, 3, 7, 9: 1 and 9 are self-inverse; 3 and 7 are inverses of each other. As for the rest, $2 \cdot 5 = 4 \cdot 5 = 6 \cdot 5 = 8 \cdot 5 = 0$ in $\mathbb{Z}_{10}$ so 2, 5, 4, 6 and 8 do not have inverses in $\mathbb{Z}_{10}$.

   d. The powers of 2 computed in $\mathbb{Z}_{10}$ are: 2, 4, 8, $16 = 6$, $32 = 2$, $\dots$, so the sequence repeats: 2, 4, 8, 16, 2, 4, 8, 16, $\dots$.

2. b. If $2a = 0$ in $\mathbb{Z}_{10}$ then $2a = 10k$ for some integer k. Thus $a = 5k$ so $a = 0$ or $a = 5$ in $\mathbb{Z}_{10}$. Conversely, it is clear that $2a = 0$ in $\mathbb{Z}_{10}$ if $a = 0$ or $a = 5$.

3. b. We want a number a in $\mathbb{Z}_{19}$ such that $11a = 1$. We could try all 19 elements in $\mathbb{Z}_{19}$, the one that works is $a = 7$. However the euclidean algorithm is a systematic method for finding a. As in Example 8.8.2, first divide 19 by 11 to get

$$19 = 1 \cdot 11 + 8$$

Then divide 11 by 8 to get

$$11 = 1 \cdot 8 + 3$$

Now divide 8 by 3 to get

$$8 = 2 \cdot 3 + 2$$

Finally divide 3 by 2 to get

$$3 = 1 \cdot 2 + 1$$

The process stops here since a remainder of 1 has been reached. Now eliminate remainders from the bottom up:

$$\begin{aligned} 1 &= 3 - 1 \cdot 2 = 3 - (8 - 2 \cdot 3) = 3 \cdot 3 - 8 \\ &= 3(11 - 1 \cdot 8) - 8 = 3 \cdot 11 - 4 \cdot 8 \\ &= 3 \cdot 11 - 4(19 - 1 \cdot 11) = 7 \cdot 11 - 4 \cdot 19 \end{aligned}$$

Hence $1 = 7 \cdot 11 - 4 \cdot 19 = 7 \cdot 11$ in $\mathbb{Z}_{19}$ because $19 = 0$ in $\mathbb{Z}_{19}$.

6. b. Working in $\mathbb{Z}_7$, we have $\det A = 15 - 24 = 1 + 4 = 5 \neq 0$, so A^{-1} exists. Since $5^{-1} = 3$ in $\mathbb{Z}_7$, $A^{-1} = 3\left[\begin{array}{rr} 3 & -6 \\ -4 & 5 \end{array}\right] = 3\left[\begin{array}{rr} 3 & 1 \\ 3 & 5 \end{array}\right] = \left[\begin{array}{rr} 2 & 3 \\ 2 & 1 \end{array}\right]$.

7. b. Gaussian elimination works over any field F in the same way that we have been using it over $\mathbb{R}$. In this case we have $F = \mathbb{Z}_7$, and we reduce the augmented matrix of the system as follows. We have $5 \cdot 3 = 1$ in $\mathbb{Z}_7$ so the first step in the reduction is to multiply row 1 by 5 in $\mathbb{Z}_7$:

$$\left[\begin{array}{rrrr} 3 & 1 & 4 & 3 \\ 4 & 3 & 1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & 5 & 6 & 1 \\ 4 & 3 & 1 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & 5 & 6 & 1 \\ 0 & 4 & 5 & 4 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & 5 & 6 & 1 \\ 0 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrrr} 1 & 0 & 5 & 3 \\ 0 & 1 & 3 & 1 \end{array}\right]$$

Hence x and y are the leading variables, and the non-leading variable z is assigned as a parameter, say $z = t$. Then, exactly as in the real case, we obtain $x = 3 + 2t$, $y = 1 + 4t$, $z = t$ where t is arbitrary in $\mathbb{Z}_7$.

9. b. If the inverse is $a + bt$ then $1 = (1 + t)(a + bt) = (a - b) + (a + b)t$. This certainly holds if $a - b = 1$ and $a + b = 0$. Adding gives $2a = 1$, that is $-a = 1$ in Z_3, that is $a = -1 = 2$. Hence $a + b = 0$ gives $b = -a = 1$, so $a + bt = 2 + t$; that is $(1 + t)^{-1} = 2 + t$. Of course it is easily checked directly that $(1 + t)(2 + t) = 1$.

10. b. The minimum weight of C is 5, so it detects 4 errors and corrects 2 errors by Theorem 8.8.5.

11. b. The linear (5, 2)-code $\{00000,\ 01110,\ 10011,\ 11101\}$ has minimum weight 3 so it corrects 1 error by Theorem 8.8.5.

12. b. The code is $\{0000000000$, 1001111000, 0101100110, 0011010111, 1100011110, 1010101111, 0110110001, $1111001001\}$.

    This has minimum distance 5 and so corrects 2 errors.

13. b. $C = \{00000,\ 10110,\ 01101,\ 11011\}$ is a linear $(5,\ 2)$-code of minimal weight 3, so it corrects single errors.

14. b. $G = \left[\begin{array}{cc} 1 & \mathbf{u} \end{array}\right]$ where $\mathbf{u}$ is any nonzero vector in the code. $H = \left[\begin{array}{c} \mathbf{u} \\ I_{n-1} \end{array}\right]$.

8.9 An Application to Quadratic Forms

1. b. $A = \begin{bmatrix} 1 & \frac{1}{2}(1-1) \\ \frac{1}{2}(-1+1) & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

   d. $A = \begin{bmatrix} 1 & \frac{1}{2}(2+4) & \frac{1}{2}(-1+5) \\ \frac{1}{2}(4+2) & 1 & \frac{1}{2}(0-2) \\ \frac{1}{2}(5-1) & \frac{1}{2}(-2+0) & 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

2. b. $q = \mathbf{x}^T A\mathbf{x}$ where $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. $c_A(x) = \begin{vmatrix} x-1 & -2 \\ -2 & x-1 \end{vmatrix} = x^2 - 2x - 3 = (x+1)(x-3)$

   $\lambda_1 = 3$: $\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$; so an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

   $\lambda_2 = -1$: $\begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$; so an eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

   Hence, $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is orthogonal and $P^TAP = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$.

   As in Theorem 8.9.1, take $\mathbf{y} = P^T\mathbf{x} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x_1+x_2 \\ x_1-x_2 \end{bmatrix}$. Then

$$y_1 = \frac{1}{\sqrt{2}}(x_1+x_2) \text{ and } y_2 = \frac{1}{\sqrt{2}}(x_1-x_2)$$

   Finally, $q = 3y_1^2 - y_2^2$, the index of q is 1 (the number of positive eigenvalues) and the rank of q is 2 (the number of nonzero eigenvalues).

   d. $q = \mathbf{x}^T A\mathbf{x}$ where $A = \begin{bmatrix} 7 & 4 & 4 \\ 4 & 1 & -8 \\ 4 & -8 & 1 \end{bmatrix}$. To find $c_A(x)$, subtract row 2 from row 3:

$$c_A(x) = \begin{vmatrix} x-7 & -4 & -4 \\ -4 & x-1 & 8 \\ -4 & 8 & x-1 \end{vmatrix} = \begin{vmatrix} x-7 & -4 & -4 \\ -4 & x-1 & 8 \\ 0 & -x+9 & x-9 \end{vmatrix}$$
$$= \begin{vmatrix} x-7 & -8 & -4 \\ -4 & x+7 & 8 \\ 0 & 0 & x-9 \end{vmatrix} = (x-9)^2(x+9)$$

   $\lambda_1 = 9$: $\begin{bmatrix} 2 & -4 & -4 \\ -4 & 8 & 8 \\ -4 & 8 & 8 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; orthogonal eigenvectors $\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$.

   $\lambda_2 = -9$: $\begin{bmatrix} -16 & -4 & -4 \\ -4 & -10 & 8 \\ -4 & 8 & -10 \end{bmatrix} \to \begin{bmatrix} 4 & 1 & 1 \\ 0 & -9 & 9 \\ 0 & 9 & -9 \end{bmatrix} \to \begin{bmatrix} 4 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$; eigenvector $\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$. These eigenvectors are orthogonal and each has length 3. Hence, $P = \frac{1}{3}\begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$ is orthogonal and $P^TAP = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -9 \end{bmatrix}$. Thus

$$\mathbf{y} = P^T\mathbf{x} = \frac{1}{3}\begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2x_1+2x_2-x_3 \\ 2x_1-x_2+2x_3 \\ -x_1+2x_2+2x_3 \end{bmatrix}$$

   so

$$y_1 = \frac{1}{3}[2x_1+2x_2-x_3]$$
$$y_2 = \frac{1}{3}[2x_1-x_2+2x_3]$$
$$y_3 = \frac{1}{3}[-x_1+2x_2+2x_3]$$

   will give $q = 9y_1^2 + 9y_2^2 - 9y_3^2$. The index of q is 2 and the rank of q is 3.

f. $q = \mathbf{x}^T A\mathbf{x}$ where $A = \begin{bmatrix} 5 & -2 & -4 \\ -2 & 8 & -2 \\ -4 & -2 & 5 \end{bmatrix}$. To find $c_A(x)$, subtract row 3 from row 1:

$$\begin{aligned} c_A(x) &= \begin{vmatrix} x-5 & 2 & 4 \\ 2 & x-8 & 2 \\ 4 & 2 & x-5 \end{vmatrix} = \begin{vmatrix} x-9 & 0 & -x+9 \\ 2 & x-8 & 2 \\ 4 & 2 & x-5 \end{vmatrix} \\ &= \begin{vmatrix} x-9 & 0 & 0 \\ 2 & x-8 & 4 \\ 4 & 2 & x-1 \end{vmatrix} = x(x-9)^2 \end{aligned}$$

$\lambda_1 = 9$: $\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; orthogonal eigenvectors are $\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$.

$\lambda_2 = 0$: $\begin{bmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & -4 & 1 \\ 0 & -18 & 9 \\ 0 & 18 & -9 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$.

These eigenvectors are orthogonal and each has length 3. Hence $P = \frac{1}{3}\begin{bmatrix} -2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & -2 & 2 \end{bmatrix}$ is orthogonal and $P^T AP = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. If

$$\mathbf{y} = P^T\mathbf{x} = \frac{1}{3}\begin{bmatrix} -2 & 2 & 1 \\ 1 & 2 & -2 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

then

$$\begin{aligned} y_1 &= \tfrac{1}{3}(-2x_1 + 2x_2 + x_3) \\ y_2 &= \tfrac{1}{3}(x_1 + 2x_2 - 2x_3) \\ y_3 &= \tfrac{1}{3}(2x_1 + x_2 + 2x_3) \end{aligned}$$

gives $q = 9y_1^2 + 9y_2^2$. The rank and index of q are both 2.

h. $q = \mathbf{x}^T A\mathbf{x}$ where $A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. To find $c_A(x)$, add row 3 to row 1:

$$\begin{aligned} c_A(x) &= \begin{vmatrix} x-1 & 1 & 0 \\ 1 & x & -1 \\ 0 & -1 & x-1 \end{vmatrix} = \begin{vmatrix} x-1 & 0 & x-1 \\ 1 & x & -1 \\ 0 & -1 & x-1 \end{vmatrix} \\ &= \begin{vmatrix} x-1 & 0 & 0 \\ 1 & x & -2 \\ 0 & -1 & x-1 \end{vmatrix} = (x-1)(x-2)(x+1) \end{aligned}$$

$\lambda_1 = 2$: $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$.

$\lambda_2 = 1$: $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

$\lambda_3 = -1$: $\begin{bmatrix} -2 & 1 & 0 \\ 1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & -1 \\ 0 & -1 & -2 \\ 0 & -1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.

Hence,

$$P = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix} -\sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & 0 & 2 \\ \sqrt{2} & \sqrt{3} & -1 \end{bmatrix}$$

is orthogonal and $P^TAP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. If

$$\mathbf{y} = P^T\mathbf{x} = \frac{1}{\sqrt{6}} \begin{bmatrix} -\sqrt{2} & \sqrt{2} & \sqrt{2} \\ \sqrt{3} & 0 & \sqrt{3} \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

then

$$\begin{aligned} y_1 &= \tfrac{1}{\sqrt{3}}(-x_1 + x_2 + x_3) \\ y_2 &= \tfrac{1}{\sqrt{2}}(x_1 + x_3) \\ y_3 &= \tfrac{1}{\sqrt{6}}(x_1 + 2x_2 - x_3) \end{aligned}$$

gives $q = 2y_1^2 + y_2^2 - y_3^2$. Here q has index 2 and rank 3.

3. b. $q = 3x^2 - 4xy = \mathbf{x}^TA\mathbf{x}$ where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $A = \begin{bmatrix} 3 & -2 \\ -2 & 0 \end{bmatrix}$. $c_A(t) = \begin{vmatrix} t-3 & 2 \\ 2 & t \end{vmatrix} = (t-4)(t+1)$
$\lambda_1 = 4$: $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$.
$\lambda_2 = -1$: $\begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \to \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.
Hence, $P = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$ gives $P^TAP = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$. If $\mathbf{y} = P^T\mathbf{x} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, then $x_1 = \frac{1}{\sqrt{5}}(2x - y)$ and $y_1 = \frac{1}{\sqrt{5}}(x + 2y)$. The equation $q = 2$ becomes $4x_1^2 - y_1^2 = 2$, a hyperbola.

d. $q = 2x^2 + 4xy + 5y^2 = \mathbf{x}^TA\mathbf{x}$ where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$.
In this case $c_A(t) = \begin{vmatrix} t-2 & -2 \\ -2 & t-5 \end{vmatrix} = (t-1)(t-6)$.
$\lambda_1 = 6$: $\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.
$\lambda_2 = 1$: $\begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$.
Hence, $P = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$ gives $P^TAP = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$. If $\mathbf{y} = P^T\mathbf{x} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, then $x_1 = \frac{1}{\sqrt{5}}(x + 2y)$, $y_1 = \frac{1}{\sqrt{5}}(2x - y)$ and $q = 1$ becomes $6x_1^2 + y_1^2 = 1$. This is an ellipse.

4. After the rotation, the new variables $\mathbf{x}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ are related to $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ by $\mathbf{x} = A\mathbf{x}_1$ where $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ (this is equation (8.8) preceding Theorem 8.9.2, or see Theorem 2.6.4). Thus $x = x_1\cos\theta - y_1\sin\theta$ and $y = x_1\sin\theta + y_1\cos\theta$. If these are substituted in the equation $ax^2 + bxy + cy^2 = d$, the coefficient of x_1y_1 is

$$-2a\sin\theta\cos\theta + b(\cos^2\theta - \sin^2\theta) + 2c\sin\theta\cos\theta = b\cos 2\theta - (a-c)\sin 2\theta.$$

This is zero if θ is chosen so that

$$\cos 2\theta = \frac{a-c}{\sqrt{b^2+(a-c)^2}} \quad \text{and} \quad \sin 2\theta = \frac{b}{\sqrt{b^2+(a-c)^2}}$$

Such an angle 2θ exists because $\left[\frac{a-c}{\sqrt{b^2+(a-c)^2}}\right]^2 + \left[\frac{b}{\sqrt{b^2(a-c)^2}}\right]^2 = 1$.

7. b. The equation is $\mathbf{x}^T A\mathbf{x} + B\mathbf{x} = 7$ where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $A = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \\ -2 & 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 0 & -6 \end{bmatrix}$.

$$\begin{aligned} c_A(x) &= \begin{vmatrix} t-1 & -2 & 2 \\ -2 & t-3 & 0 \\ 2 & 0 & t-3 \end{vmatrix} = \begin{vmatrix} t-1 & -2 & 2 \\ -2 & t-3 & 0 \\ 0 & t-3 & t-3 \end{vmatrix} = \begin{vmatrix} t-1 & -4 & 2 \\ -2 & t-3 & 0 \\ 0 & 0 & t-3 \end{vmatrix} \\ &= (t-3)(t^2-4t-5) = (t-3)(t-5)(t+1) \end{aligned}$$

$\lambda_1 = 3$: $\begin{bmatrix} 2 & -2 & 2 \\ -2 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

$\lambda_2 = 5$: $\begin{bmatrix} 4 & -2 & 2 \\ -2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} -2 & 2 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

$\lambda_3 = -1$: $\begin{bmatrix} -2 & -2 & 2 \\ -2 & -4 & 0 \\ 2 & 0 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 \\ 0 & -2 & -2 \\ 0 & -2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; an eigenvector is $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$.

Hence, $P = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & +\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & \sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & -1 \\ \sqrt{3} & -\sqrt{2} & 1 \end{bmatrix}$ satisfies $P^T AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. If

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = P^T \mathbf{x} = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3}(x_2 + x_3) \\ \sqrt{2}(x_1 + x_2 - x_3) \\ 2x_1 - x_2 + x_3 \end{bmatrix}$$

then

$$\begin{aligned} y_1 &= \tfrac{1}{\sqrt{2}}(x_2 + x_3) \\ y_2 &= \tfrac{1}{\sqrt{3}}(x_1 + x_2 - x_3) \\ y_3 &= \tfrac{1}{\sqrt{6}}(2x_1 - x_2 + x_3) \end{aligned}$$

As $P^{-1} = P^T$, we have $\mathbf{x} = P\mathbf{y}$ so substitution in $\mathbf{x}^T A\mathbf{x} + B\mathbf{x} = 7$ gives

$$\mathbf{y}^T (P^T AP)\mathbf{y} + (BP)\mathbf{y} = 7$$

As $BP = \frac{1}{\sqrt{6}} \begin{bmatrix} -6\sqrt{3} & 11\sqrt{2} & 4 \end{bmatrix} = \begin{bmatrix} -3\sqrt{2} & \frac{11\sqrt{3}}{3} & \frac{2\sqrt{6}}{3} \end{bmatrix}$, this is

$$3y_1^2 + 5y_2^2 - y_3^2 - (3\sqrt{2})y_1 + (\tfrac{11}{3}\sqrt{3})y_2 + (\tfrac{2}{3}\sqrt{6})y_3 = 7$$

9. b. We have $A = U^T U$ where U is upper triangular with positive diagonal entries. Hence

$$q(\mathbf{x}) = \mathbf{x}^T U^T U\mathbf{x} = (U\mathbf{x})^T (U\mathbf{x}) = \|U\mathbf{x}\|^2$$

So take $\mathbf{y} = U\mathbf{x}$ as the new column of variables.

8.10 An Application to Constrained Optimization

This section contains no exercises.

8.11 An Application to Statistical Principal Component Analysis

This section contains no exercises.

9. Change of Basis

9.1 The Matrix of a Linear Transformation

1. b. $C_B(\mathbf{v}) = \left[\begin{array}{c} a \\ 2b-c \\ c-b \end{array}\right]$ because $\mathbf{v} = ax^2 + bx + c = ax^2 + (2b-c)(x+1) + (c-b)(x+2)$.

   d. $C_B(\mathbf{v}) = \frac{1}{2}\left[\begin{array}{c} a-b \\ a+b \\ -a+3b+2c \end{array}\right]$ because

$$\mathbf{v} = (a,\ b,\ c) = \tfrac{1}{2}[(a-b)(1,\ -1,\ 2) + (a+b)(1,\ 1,\ -1) + (-a+3b+2c)(0,\ 0,\ 1)]$$

2. b. $M_{DB}(T) = \left[\begin{array}{ccc} C_D[T(1)] & C_D[T(x)] & C_D[T(x^2)] \end{array}\right] = \left[\begin{array}{ccc} 2 & 1 & 3 \\ -1 & 0 & -2 \end{array}\right]$. Comparing columns gives

$$C_D[T(1)] = \left[\begin{array}{c} 2 \\ -1 \end{array}\right] \qquad C_D[T(x)] = \left[\begin{array}{c} 1 \\ 0 \end{array}\right] \qquad C_D[T(x^2)] = \left[\begin{array}{c} 3 \\ -2 \end{array}\right]$$

   Hence

$$\begin{aligned} T(1) &= 2(1,\ 1) - (0,\ 1) = (2,\ 1) \\ T(x) &= 1(1,\ 1) + 0(0,\ 1) = (1,\ 1) \\ T(x^2) &= 3(1,\ 1) - 2(0,\ 1) = (3,\ 1) \end{aligned}$$

   Thus

$$\begin{aligned} T(a+bx+cx^2) &= aT(1) + bT(x) + cT(x^2) \\ &= a(2,\ 1) + b(1,\ 1) + c(3,\ 1) \\ &= (2a+b+3c,\ a+b+c) \end{aligned}$$

3. b.
$$\begin{aligned} M_{DB}(T) &= \left[\begin{array}{cccc} C_D\left\{T\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right]\right\} & C_D\left\{T\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]\right\} & C_D\left\{T\left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]\right\} & C_D\left\{T\left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\right\} \end{array}\right] \\ &= \left[\begin{array}{cccc} C_D\left\{\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right]\right\} & C_D\left\{\left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]\right\} & C_D\left\{\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]\right\} & C_D\left\{\left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\right\} \end{array}\right] \\ &= \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \end{aligned}$$

   d.
$$\begin{aligned} M_{DB}(T) &= \left[\begin{array}{ccc} C_D[T(1)] & C_D[T(x)] & C_D[T(x^2)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_D(1) & C_D(x+1) & C_D(x^2+2x+1) \end{array}\right] \\ &= \left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{array}\right] \end{aligned}$$

4. b. $M_{DB}(T) = \begin{bmatrix} C_D[T(1,\ 1)] & C_D[T(1,\ 0)] \end{bmatrix} = \begin{bmatrix} C_D(1,\ 5,\ 4,\ 1) & C_D(2,\ 3,\ 0,\ 1) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 3 \\ 4 & 0 \\ 1 & 1 \end{bmatrix}$.

We have $\mathbf{v} = (a,\ b) = b(1,\ 1) + (a-b)(1,\ 0)$ so $C_B(\mathbf{v}) = \begin{bmatrix} b \\ a-b \end{bmatrix}$. Hence,

$$C_D[T(\mathbf{v})] = M_{DB}(T)C_B(\mathbf{v}) = \begin{bmatrix} 1 & 2 \\ 5 & 3 \\ 4 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} b \\ a-b \end{bmatrix} = \begin{bmatrix} 2a-b \\ 3a+2b \\ 4b \\ a \end{bmatrix}$$

Finally, we recover the action of T:

$$\begin{aligned} T(\mathbf{v}) &= (2a-b)(1,\ 0,\ 0,\ 0) + (3a+2b)(0,\ 1,\ 0,\ 0) + 4b(0,\ 0,\ 1,\ 0) + a(0,\ 0,\ 0,\ 1) \\ &= (2a-b,\ 3a+2b,\ 4b,\ a) \end{aligned}$$

d. $$\begin{aligned} M_{DB}(T) &= \begin{bmatrix} C_D[T(1)] & C_D[T(x)] & C_D[T(x^2)] \end{bmatrix} \\ &= \begin{bmatrix} C_D(1,\ 0) & C_D(1,\ 0) & C_D(0,\ 1) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \end{aligned}$$

We have $\mathbf{v} = a + bx + cx^2$ so $C_B(\mathbf{v}) = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Hence

$$C_D[T(\mathbf{v})] = M_{DB}(T)C_B(\mathbf{v}) = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{2}\begin{bmatrix} a+b-c \\ a+b+c \end{bmatrix}$$

Finally, we recover the action of T:

$$T(\mathbf{v}) = \tfrac{1}{2}(a+b-c)(1,\ -1) + \tfrac{1}{2}(a+b+c)(1,\ 1) = (a+b,\ c).$$

f. $$\begin{aligned} M_{DB}(T) &= \begin{bmatrix} C_D\left\{T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\} & C_D\left\{T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right\} & C_D\left\{T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\} & C_D\left\{T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\} \end{bmatrix} \\ &= \begin{bmatrix} C_D\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & C_D\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & C_D\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & C_D\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

We have $\mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$,

so $C_B(\mathbf{v}) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$. Hence $C_D[T(\mathbf{v})] = M_{DB}(T)C_B(\mathbf{v}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a \\ b+c \\ b+c \\ d \end{bmatrix}$.

Finally, we recover the action of T:

$$T(\mathbf{v}) = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (b+c)\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + (b+c)\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b+c \\ b+c & d \end{bmatrix}$$

5. b. Have $\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^4 \xrightarrow{S} \mathbb{R}^2$. Let B, D, E be the standard bases. Then

$$\begin{aligned} M_{ED}(S) &= \begin{bmatrix} C_E[S(1,\ 0,\ 0,\ 0)] & C_E[S(0,\ 1,\ 0,\ 0)] & C_E[S(0,\ 0,\ 1,\ 0)] & C_E[S(0,\ 0,\ 0,\ 1)] \end{bmatrix} \\ &= \begin{bmatrix} C_E(1,\ 0) & C_E(1,\ 0) & C_E(0,\ 1) & C_E(0,\ -1) \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \\
M_{DB}(T) &= \begin{bmatrix} C_D[T(1,\ 0,\ 0)] & C_D[T(0,\ 1,\ 0)] & C_D[T(0,\ 0,\ 1)] \end{bmatrix} \\
&= \begin{bmatrix} C_D(1,\ 0,\ 1,\ -1) & C_D(1,\ 1,\ 0,\ 1) & C_D(0,\ 1,\ 1,\ 0) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}
\end{aligned}
$$

We have $ST(a,\ b,\ c) = S(a+b,\ c+b,\ a+c,\ b-a) = (a+2b+c,\ 2a-b+c)$. Hence

$$
\begin{aligned}
M_{EB}(ST) &= \begin{bmatrix} C_E\,[ST(1,\ 0,\ 0)] & C_E\,[ST(0,\ 1,\ 0)] & C_E\,[ST(0,\ 0,\ 1)] \end{bmatrix} \\
&= \begin{bmatrix} C_E(1,\ 2) & C_E(2,\ -1) & C_E(1,\ 1) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 1 \end{bmatrix}
\end{aligned}
$$

With this we confirm Theorem 9.1.3 as follows:

$$
M_{ED}(S)M_{DB}(T) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 1 \end{bmatrix} = M_{EB}(ST)
$$

d. Have $\mathbb{R}^3 \xrightarrow{T} \mathbf{P}_2 \xrightarrow{S} \mathbb{R}^2$ with bases $B = \{(1,\ 0,\ 0),\ (0,\ 1,\ 0),\ (0,\ 0,\ 1)\}$, $D = \{1,\ x,\ x^2\}$, $E = \{(1,\ 0),\ (0,\ 1)\}$.

$$
\begin{aligned}
M_{ED}(S) &= \begin{bmatrix} C_E[S(1)] & C_E[S(x)] & C_E[S(x^2)] \end{bmatrix} \\
&= \begin{bmatrix} C_E(1,\ 0) & C_E(-1,\ 0) & C_E(0,\ 1) \end{bmatrix} \\
&= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
M_{DB}(T) &= \begin{bmatrix} C_D[T(1,\ 0,\ 0)] & C_D[T(0,\ 1,\ 0)] & C_D[T(0,\ 0,\ 1)] \end{bmatrix} \\
&= \begin{bmatrix} C_D(1-x) & C_D(-1+x^2) & C_D(x) \end{bmatrix} \\
&= \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}
\end{aligned}
$$

The action of ST is $ST(a,\ b,\ c) = S\left[(a-b)+(c-a)x+bx^2\right] = (2a-b-c,\ b)$. Hence,

$$
\begin{aligned}
M_{EB}(ST) &= \begin{bmatrix} C_E[ST(1,\ 0,\ 0)] & C_E[ST(0,\ 1,\ 0)] & C_E[ST(0,\ 0,\ 1)] \end{bmatrix} \\
&= \begin{bmatrix} C_E(2,\ 0) & C_E(-1,\ 1) & C_E(-1,\ 0) \end{bmatrix} \\
&= \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix}
\end{aligned}
$$

Hence, we verify Theorem 9.1.3 as follows:

$$
M_{ED}(S)M_{DB}(T) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix} = M_{EB}(ST)
$$

7. b. $$\begin{aligned} M_{DB}(T) &= \left[\begin{array}{ccc} C_D[T(1,\,0,\,0)] & C_D[T(0,\,1,\,0)] & C_D[T(0,\,0,\,1)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_D(0,\,1,\,1) & C_D(1,\,0,\,1) & C_D(1,\,1,\,0) \end{array}\right] \\ &= \left[\begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{array}\right] \end{aligned}$$

If $T^{-1}(a,\,b,\,c) = (x,\,y,\,z)$ then $(a,\,b,\,c) = T(x,\,y,\,z) = (y+z,\,x+z,\,x+y)$. Hence, $y+z=a$, $x+z=b$, $x+y=c$. The solution is

$$T^{-1}(a,\,b,\,c) = (x,\,y,\,z) = \tfrac{1}{2}(b+c-a,\,a+c-b,\,a+b-c)$$

Hence,

$$\begin{aligned} M_{BD}(T^{-1}) &= \left[\begin{array}{ccc} C_B[T^{-1}(1,\,0,\,0)] & C_B[T^{-1}(0,\,1,\,0)] & C_B[T^{-1}(0,\,0,\,1)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_B\left(-\frac{1}{2},\,\frac{1}{2},\,\frac{1}{2}\right) & C_B\left(\frac{1}{2},\,-\frac{1}{2},\,\frac{1}{2}\right) & C_B\left(\frac{1}{2},\,\frac{1}{2},\,-\frac{1}{2}\right) \end{array}\right] \\ &= \frac{1}{2}\left[\begin{array}{rrr} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{array}\right] \end{aligned}$$

This matrix is $M_{DB}(T)^{-1}$ as Theorem 9.1.4 asserts.

d. $$\begin{aligned} M_{DB}(T) &= \left[\begin{array}{ccc} C_D[T(1)] & C_D[T(x)] & C[T(x^2)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_D(1,\,0,\,0) & C_D(1,\,1,\,0) & C_D(1,\,1,\,1) \end{array}\right] \\ &= \left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right] \end{aligned}$$

If $T^{-1}(a,\,b,\,c) = r+sx+tx^2$, then $(a,\,b,\,c) = T(r+sx+tx^2) = (r+s+t,\,s+t,\,t)$. Hence, $r+s+t=a$, $s+t=b$, $t=c$; the solution is $t=c$, $s=b-c$, $r=a-b$. Thus,

$$T^{-1}(a,\,b,\,c) = r+sx+tx^2 = (a-b)+(b-c)x+cx^2$$

Hence,

$$\begin{aligned} M_{BD}(T^{-1}) &= \left[\begin{array}{ccc} C_B[T^{-1}(1,\,0,\,0)] & C_B[T^{-1}(0,\,1,\,0)] & C_B[T^{-1}(0,\,0,\,1)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_B(1) & C_B(-1+x) & C_B(-x+x^2) \end{array}\right] \\ &= \left[\begin{array}{rrr} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right] \end{aligned}$$

This matrix is $M_{DB}(T)^{-1}$ as Theorem 9.1.4 asserts.

8. b. $$\begin{aligned} M_{DB}(T) &= \left[\begin{array}{cccc} C_D\left\{T\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right]\right\} & C_D\left\{T\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]\right\} & C_D\left\{T\left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]\right\} & C_D\left\{T\left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\right\} \end{array}\right] \\ &= \left[\begin{array}{cccc} C_D(1,\,0,\,0,\,0) & C_D(1,\,1,\,0,\,0) & C_D(1,\,1,\,1,\,0) & C_D(0,\,0,\,0,\,1) \end{array}\right] \\ &= \left[\begin{array}{cccc} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \end{aligned}$$

This is invertible and the matrix inversion algorithm (and Theorem 9.1.4) gives

$$M_{DB}(T^{-1}) = [M_{DB}(T)]^{-1} = \left[\begin{array}{rrrr} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

If $\mathbf{v} = (a,\,b,\,c,\,d)$ then

$$C_B[T^{-1}(\mathbf{v})] = M_{DB}(T^{-1})C_D(\mathbf{v}) = \left[\begin{array}{rrrr} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]\left[\begin{array}{c} a \\ b \\ c \\ d \end{array}\right] = \left[\begin{array}{c} a-b \\ b-c \\ c \\ d \end{array}\right]$$

Hence, we get a formula for the action of T^{-1}:

$$T^{-1}(a,\ b,\ c,\ d) = T^{-1}(\mathbf{v}) = (a-b)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (b-c)\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} a-b & b-c \\ c & d \end{bmatrix}$$

12. Since $D = \{T(\mathbf{e}_1),\ \dots,\ T(\mathbf{e}_n)\}$, we have $C_D\left[T(\mathbf{e}_j)\right] = C_j =$ column j of I_n. Hence,

$$\begin{aligned} M_{DB}(T) &= \begin{bmatrix} C_D[T(\mathbf{e}_1)] & C_D[T(\mathbf{e}_2)] & \cdots & C_D[T(\mathbf{e}_n)] \end{bmatrix} \\ &= \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix} = I_n \end{aligned}$$

16. b. Define $T : \mathbf{P}_n \to \mathbb{R}^{n+1}$ by $T[p(x)] = (p(a_0),\ p(a_1),\ \dots,\ p(a_n))$, where $a_0,\ \dots,\ a_n$ are fixed distinct real numbers. If $B = \{1,\ x,\ \dots,\ x^n\}$ and $D \subseteq \mathbb{R}^{n+1}$ is the standard basis,

$$\begin{aligned} M_{DB}(T) &= \begin{bmatrix} C_D[T(1)] & C_D[T(x)] & C_D\left[T(x^2)\right] & \cdots & C_D[T(x^n)] \end{bmatrix} \\ &= \begin{bmatrix} C_D(1,\ 1,\ \dots,\ 1) & C_D(a_0,\ a_1,\ \dots,\ a_n) & C_D(a_0^2,\ a_1^2,\ \dots,\ a_n^2) & \cdots & C_D(a_0^n,\ a_1^n,\ \dots,\ a_n^n) \end{bmatrix} \\ &= \begin{bmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^n \\ 1 & a_1 & a_1^2 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^n \end{bmatrix} \end{aligned}$$

Since the a_i are distinct, this matrix has nonzero determinant by Theorem 3.2.7. Hence, T is an isomorphism by Theorem 9.1.4.

20. d. Assume that $V \xrightarrow{R} W \xrightarrow{S,\, T} U$. Recall that the sum $S+T : W \to U$ of two operators is defined by $(S+T)(\mathbf{w}) = S(\mathbf{w}) + T(\mathbf{w})$ for all $\mathbf{w}$ in W. Hence, for $\mathbf{v}$ in V:

$$\begin{aligned} [(S+T)R](\mathbf{v}) &= (S+T)[R(\mathbf{v})] \\ &= S[R(\mathbf{v})] + T[R(\mathbf{v})] \\ &= (SR)(\mathbf{v}) + (TR)(\mathbf{v}) \\ &= (SR+TR)(\mathbf{v}) \end{aligned}$$

Since this holds for all $\mathbf{v}$ in V, it shows that $(S+T)R = SR+TR$.

21. b. If P and Q are subspaces of a vector space W, recall that $P+Q = \{p+q \mid p \text{ in } P,\ q \text{ in } Q\}$ is a subspace of W (Exercise 25 Section 6.4). Now let $\mathbf{w}$ be any vector in im $(S+T)$. Then $\mathbf{w} = (S+T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$ for some $\mathbf{v}$ in V, whence $\mathbf{w}$ is in im $S+$ im T. Thus, im $(S+T) \subseteq$ im $S+$ im T.

22. b. If T is in X_1^0, then $T(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in X_1. As $X \subseteq X_1$, this implies that $T(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in X; that is T is in X^0. Hence, $X_1^0 \subseteq X^0$.

24. b. We have $R : V \to \mathbf{L}(\mathbb{R},\ V)$ defined by $R(\mathbf{v}) = S_\mathbf{v}$. Here $S_\mathbf{v} : \mathbb{R} \to V$ is defined by $S_\mathbf{v}(r) = r\mathbf{v}$.

<u>R is a linear transformation</u>: The requirements that $R(\mathbf{v}+\mathbf{w}) = R(\mathbf{v}) + R(\mathbf{w})$ and $R(a\mathbf{v}) = aR(\mathbf{v})$ translate to $S_{\mathbf{v}+\mathbf{w}} = S_\mathbf{v} + S_\mathbf{w}$ and $S_{a\mathbf{v}} = aS_\mathbf{v}$. If r is arbitrary in $\mathbb{R}$:

$$S_{\mathbf{v}+\mathbf{w}}(r) = r(\mathbf{v}+\mathbf{w}) = r\mathbf{v} + r\mathbf{w} = S_\mathbf{v}(r) + S_\mathbf{w}(r) = (S_\mathbf{v} + S_\mathbf{w})(r)$$

$$S_{a\mathbf{v}}(r) = r(a\mathbf{v}) = a(r\mathbf{v}) = a[S_{\mathbf{v}}(r)] = (aS_{\mathbf{v}})(r)$$

Hence, $S_{\mathbf{v}+\mathbf{w}} = S_{\mathbf{v}} + S_{\mathbf{w}}$ and $S_{a\mathbf{v}} = aS_{\mathbf{v}}$ so R is linear.

<u>R is one-to-one</u>: If $R(\mathbf{v}) = \mathbf{0}$ then $S_{\mathbf{v}} = 0$ is the zero transformation $\mathbb{R} \to V$. Hence we have $0 = S_{\mathbf{v}}(r) = r\mathbf{v}$ for all r; taking $r = 1$ gives $\mathbf{v} = \mathbf{0}$. Thus $\ker R = 0$.

<u>R is onto</u>: Given T in $\mathbf{L}(\mathbb{R}, V)$, we must find $\mathbf{v}$ in V such that $T = R(\mathbf{v})$; that is $T = S_{\mathbf{v}}$. Now $T : \mathbb{R} \to V$ is a linear transformation and we take $\mathbf{v} = T(1)$. Then, for r in $\mathbb{R}$:

$$S_{\mathbf{v}}(r) = r\mathbf{v} = rT(1) = T(r \cdot 1) = T(r)$$

Hence, $S_{\mathbf{v}} = T$ as required.

25. b. Given the linear transformation $T : \mathbb{R} \to V$ and an ordered basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ of V, write $T(1) = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n$ where the a_i are in $\mathbb{R}$. We must show that $T = a_1S_1 + a_2S_2 + \cdots + a_nS_n$ where $S_i(r) = r\mathbf{b}_i$ for all r in $\mathbb{R}$. We have

$$\begin{aligned}(a_1S_1 + a_2S_2 + \cdots + a_nS_n)(r) &= a_1S_1(r) + a_2S_2(r) + \cdots + a_nS_n(r)\\ &= a_1(r\mathbf{b}_1) + a_2(r\mathbf{b}_2) + \cdots + a_n(r\mathbf{b}_n)\\ &= rT(1)\\ &= T(r)\end{aligned}$$

for all r in $\mathbb{R}$. Hence $a_1S_1 + a_2S_2 + \cdots + a_nS_n = T$.

27. b. Given $\mathbf{v}$ in V, write $\mathbf{v} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \cdots + r_n\mathbf{b}_n$, r_i in $\mathbb{R}$. We must show that $r_j = E_j(\mathbf{v})$ for each j. To see this, apply the linear transformation E_j:

$$\begin{aligned}E_j(\mathbf{v}) &= E_j(r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \cdots + r_j\mathbf{b}_j + \cdots + r_n\mathbf{b}_n)\\ &= r_1E_j(\mathbf{b}_1) + r_2E_j(\mathbf{b}_2) + \cdots + r_jE_j(\mathbf{b}_j) + \cdots + r_nE_j(\mathbf{b}_n)\\ &= r_1 \cdot 0 + r_2 \cdot 0 + \cdots + r_j \cdot 1 + \cdots + r_n \cdot 0\\ &= r_j\end{aligned}$$

using the definition of E_j.

9.2 Operators and Similarity

1. b. $P_{D \leftarrow B} = \left[\begin{array}{ccc} C_D(x) & C_D(1+x) & C_D(x^2) \end{array}\right] = \left[\begin{array}{ccc} -\frac{3}{2} & -1 & \frac{1}{2} \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] = \frac{1}{2}\left[\begin{array}{ccc} -3 & -2 & 1 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{array}\right]$ because

$$\begin{aligned} x &= -\tfrac{3}{2} \cdot 2 + 1(x+3) + 0(x^2 - 1)\\ 1 + x &= (-1) \cdot 2 + 1(x+3) + 0(x^2 - 1)\\ x^2 &= \tfrac{1}{2} \cdot 2 + 0(x+3) + 1(x^2 - 1)\end{aligned}$$

Given $\mathbf{v} = 1 + x + x^2$, we have

$$C_B(\mathbf{v}) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad C_D(\mathbf{v}) = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 1 \end{bmatrix}$$

because $\mathbf{v} = 0 \cdot x + 1(1+x) + 1 \cdot x^2$ and $\mathbf{v} = -\frac{1}{2} \cdot 2 + 1 \cdot (x+3) + 1(x^2 - 1)$. Hence

$$P_{D \leftarrow B} C_B(\mathbf{v}) = \frac{1}{2} \begin{bmatrix} -3 & -2 & 1 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = C_D(\mathbf{v})$$

as expected.

4. b. $P_{B \leftarrow D} = \begin{bmatrix} C_B(1+x+x^2) & C_B(1-x) & C_B(-1+x^2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

$P_{D \leftarrow B} = \begin{bmatrix} C_D(1) & C_D(x) & C_D(x^2) \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & 2 \end{bmatrix}$ because

$$\begin{aligned} 1 &= \tfrac{1}{3}\left[(1+x+x^2) + (1-x) - (-1+x^2)\right] \\ x &= \tfrac{1}{3}\left[(1+x+x^2) - 2(1-x) - (-1+x^2)\right] \\ x^2 &= \tfrac{1}{3}\left[(1+x+x^2) + (1-x) + 2(-1+x^2)\right] \end{aligned}$$

The fact that $P_{D \leftarrow B} = (P_{B \leftarrow D})^{-1}$ is verified by multiplying these matrices. Next:

$$P_{E \leftarrow D} = \begin{bmatrix} C_E(1+x+x^2) & C_E(1-x) & C_E(-1+x^2) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

$$P_{E \leftarrow B} = \begin{bmatrix} C_E(1) & C_E(x) & C_E(x^2) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

where we note the order of the vectors in $E = \{x^2,\ x,\ 1\}$. Finally, matrix multiplication verifies that $P_{E \leftarrow D} P_{D \leftarrow B} = P_{E \leftarrow B}$.

5. b. Let $B = \{(1,\ 2,\ -1),\ (2,\ 3,\ 0),\ (1,\ 0,\ 2)\}$ be the basis formed by the transposes of the columns of A. Since D is the standard basis:

$$P_{D \leftarrow B} = \begin{bmatrix} C_D(1,\ 2,\ -1) & C_D(2,\ 3,\ 0) & C_D(1,\ 0,\ 2) \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ -1 & 0 & 2 \end{bmatrix} = A$$

Hence Theorem 9.2.2 gives

$$A^{-1} = (P_{D \leftarrow B})^{-1} = P_{B \leftarrow D} = \begin{bmatrix} C_B(1,\ 0,\ 0) & C_B(0,\ 1,\ 0) & C_B(0,\ 0,\ 1) \end{bmatrix} = \begin{bmatrix} 6 & -4 & -3 \\ -4 & 3 & 2 \\ 3 & -2 & -1 \end{bmatrix}$$

because

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

7. b. Since $B_0 = \{1,\ x,\ x^2\}$, we have

$$P = P_{B_0 \leftarrow B} = \left[\begin{array}{ccc} C_{B_0}(1-x^2) & C_{B_0}(1+x) & C_{B_0}(2x+x^2) \end{array}\right] = \left[\begin{array}{rrr} 1 & 1 & 0 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{array}\right]$$

$$\begin{aligned} M_{B_0}(T) &= \left[\begin{array}{ccc} C_{B_0}[T(1)] & C_{B_0}[T(x)] & C_{B_0}[T(x^2)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_{B_0}(1+x^2) & C_{B_0}(1+x) & C_{B_0}(x+x^2) \end{array}\right] \\ &= \left[\begin{array}{rrr} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{array}\right] \end{aligned}$$

Finally

$$\begin{aligned} M_B(T) &= \left[\begin{array}{ccc} C_B[T(1-x^2)] & C_B[T(1+x)] & C_B[T(2x+x^2)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_B(1-x) & C_B(2+x+x^2) & C_B(2+3x+x^2) \end{array}\right] \\ &= \left[\begin{array}{rrr} -2 & -3 & -1 \\ 3 & 5 & 3 \\ -2 & -2 & 0 \end{array}\right] \end{aligned}$$

because

$$\begin{aligned} 1-x &= -2(1-x^2)+3(1+x)-2(2x+x^2) \\ 2+x+x^2 &= -3(1-x^2)+5(1+x)-2(2x+x^2) \\ 2+3x+x^2 &= -1(1-x^2)+3(1+x)+0(2x+x^2) \end{aligned}$$

The verification that $P^{-1}M_{B_0}(T)P = M_B(T)$ is equivalent to checking that $M_{B_0}(T)P = PM_B(T)$, and so can be seen by matrix multiplication.

8. b. $P^{-1}AP = \left[\begin{array}{rr} 5 & -2 \\ -7 & 3 \end{array}\right]\left[\begin{array}{rr} 29 & -12 \\ 70 & -29 \end{array}\right]\left[\begin{array}{rr} 3 & 2 \\ 7 & 5 \end{array}\right] = \left[\begin{array}{rr} 5 & -2 \\ 7 & -3 \end{array}\right]\left[\begin{array}{rr} 3 & 2 \\ 7 & 5 \end{array}\right] = \left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right]$.

Let $B = \left\{\left[\begin{array}{c} 3 \\ 7 \end{array}\right], \left[\begin{array}{c} 2 \\ 5 \end{array}\right]\right\}$ consist of the columns of P. These are eigenvectors of A corresponding to the eigenvalues 1, -1 respectively. Hence,

$$M_B(T_A) = \left[\begin{array}{cc} C_B\left(T_A\left[\begin{array}{c} 3 \\ 7 \end{array}\right]\right) & C_B\left(T_A\left[\begin{array}{c} 2 \\ 5 \end{array}\right]\right) \end{array}\right] = \left[\begin{array}{cc} C_B\left[\begin{array}{c} 3 \\ 7 \end{array}\right] & C_B\left[\begin{array}{r} -2 \\ -5 \end{array}\right] \end{array}\right] = \left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right]$$

9. b. Choose a basis of $\mathbb{R}^2$, say $B = \{(1,\ 0),\ (0,\ 1)\}$, and compute

$$M_B(T) = \left[\begin{array}{cc} C_B[T(1,\ 0)] & C_B[T(0,\ 1)] \end{array}\right] = \left[\begin{array}{cc} C_B(3,\ 2) & C_B(5,\ 3) \end{array}\right] = \left[\begin{array}{cc} 3 & 5 \\ 2 & 3 \end{array}\right]$$

Hence, $c_T(x) = c_{M_B(T)}(x) = \left|\begin{array}{cc} x-3 & -5 \\ -2 & x-3 \end{array}\right| = x^2 - 6x - 1$. Note that the calculation is easy because B is the standard basis, but any basis could be used.

d. Use the basis $B = \{1,\ x,\ x^2\}$ of $\mathbf{P}_2$ and compute

$$\begin{aligned} M_B(T) &= \left[\begin{array}{ccc} C_B[T(1)] & C_B[T(x)] & C_B[T(x^2)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_B(1+x-2x^2) & C_B(1-2x+x^2) & C_B(-2+x) \end{array}\right] \end{aligned}$$

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

Hence,

$$c_T(x) = c_{M_B(T)}(x) = \begin{vmatrix} x-1 & -1 & 2 \\ -1 & x+2 & -1 \\ 2 & -1 & x \end{vmatrix} = \begin{vmatrix} x-1 & -1 & 2 \\ -1 & x+2 & -1 \\ -x+3 & 0 & x-2 \end{vmatrix}$$
$$= x^3 + x^2 - 8x - 3$$

f. Use $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ and compute

$$M_B(T) = \left[\begin{array}{cccc} C_B\left\{T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\} & C_B\left\{T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right\} & C_B\left\{T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\} & C_B\left\{T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\} \end{array} \right]$$
$$= \left[\begin{array}{cccc} C_B\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} & C_B\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} & C_B\begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} & C_B\begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix} \end{array} \right]$$
$$= \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Hence,

$$c_T(x) = c_{M_B(T)}(x) = \begin{vmatrix} x-1 & 0 & 1 & 0 \\ 0 & x-1 & 0 & 1 \\ -1 & 0 & x+1 & 0 \\ 0 & -1 & 0 & x+1 \end{vmatrix}$$
$$= (x-1)\begin{vmatrix} x-1 & 0 & 1 \\ 0 & x+1 & 0 \\ -1 & 0 & x+1 \end{vmatrix} + \begin{vmatrix} 0 & x-1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & x+1 \end{vmatrix} = x^4$$

12. Assume that A and B are both $n \times n$ and that null $A =$ null B. Define $T_A : \mathbb{R}^n \to \mathbb{R}^n$ by $T_A(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$; similarly for T_B. Then let $T = T_A$ and $S = T_B$. Then ker $S =$ null B and ker $T =$ null A so, by Exercise 28 Section 7.3 there is an isomorphism $R : \mathbb{R}^n \to \mathbb{R}^n$ such that $T = RS$. If B_0 is the standard basis of $\mathbb{R}^n$, we have

$$A = M_{B_0}(T) = M_{B_0}(RS) = M_{B_0}(R)M_{B_0}(S) = UB$$

where $U = M_{B_0}(R)$. This is what we wanted because U is invertible by Theorem 9.1.4.

Conversely, assume that $A = UB$ with U invertible. If $\mathbf{x}$ is in null A then $A\mathbf{x} = \mathbf{0}$, so $UB\mathbf{x} = \mathbf{0}$, whence $B\mathbf{x} = \mathbf{0}$ (because U is invertible), that is $\mathbf{x}$ is in null B. In other words null $A \subseteq$ null B. But $B = U^{-1}A$ so null $B \subseteq$ null A by the same argument. Hence null $A =$ null B.

16. b. We verify first that S is linear. Showing $S(w+v) = S(w)+S(v)$ means showing that $M_B(T_{w+v}) = M_B(T_w) + M_B(T_v)$. If $B = \{b_1, b_2\}$ then column j of $M_B(T_{w+v})$ is

$$C_B[T_{w+v}(b_j)] = C_B[(w+v)b_j] = C_B(wb_j + vb_j) = C_B(wb_j) + C_B(vb_j)$$

because C_B is linear. This is column j of $M_B(T_w) + M_B(T_v)$, which shows that $S(w+v) = S(w) + S(v)$. A similar argument shows that $M_B(T_{aw}) = aM_B(T_w)$, so $S(aw) = aS(w)$, and hence that S is linear.

To see that S is one-to-one, let $S(w) = 0$; by Theorem 7.2.2 we must show that $w = 0$. We have $M_B(T_w) = S(w) = 0$ so, comparing j^{th} columns, we see that $C_B[T_w(b_j)] = C_B[wb_j] = 0$ for

$j = 1, 2$. As C_B is an isomorphism, this means that $wb_j = 0$ for each j. But B is a basis of $\mathbb{C}$ and 1 is in $\mathbb{C}$, so there exist r and s in $\mathbb{R}$ such that $1 = rb_1 + sb_2$. Hence $w = w1 = rwb_1 + swb_2 = 0$, as required.

Finally, to show that $S(wv) = S(w)S(v)$ we first show that $T_wT_v = T_{wv}$. Indeed, given z in $\mathbb{C}$, we have

$$(T_wT_v)(z) = T_w(T_v(z)) = w(vz) = (wv)z = T_{wv}(z)$$

Since this holds for all z in C, it shows that $T_wT_v = T_{wv}$. But then Theorem 9.2.1 shows that

$$S(wv) = M_B(T_wT_v) = M_B(T_w)M_B(T_v) = S(w)S(v)$$

This is what we wanted.

9.3 Invariant Subspaces and Direct Sums

2. b. Let $\mathbf{v} \in T(U)$, say $\mathbf{v} = T(\mathbf{u})$ where $\mathbf{u} \in U$. Then $T(\mathbf{v}) = T[T(\mathbf{u})] \in T(U)$ because $T(\mathbf{u}) \in U$. This shows that $T(U)$ is T-invariant.

3. b. Given $\mathbf{v}$ in $S(U)$, we must show that $T(\mathbf{v})$ is also in $S(U)$. We have $\mathbf{v} = S(\mathbf{u})$ for some $\mathbf{u}$ in U. As $ST = TS$, we compute:

$$T(\mathbf{v}) = T[S(\mathbf{u})] = (TS)(\mathbf{u}) = (ST)(\mathbf{u}) = S[T(\mathbf{u})]$$

   As $T(\mathbf{u})$ is in U (because U is T-invariant), this shows that $T(\mathbf{v}) = S[T(\mathbf{u})]$ is in $S(U)$.

6. Suppose that a subspace U of V is T-invariant for every linear operator $T : V \to V$; we must show that either $U = 0$ or $U = V$. Assume that $U \neq 0$; we must show that $U = V$. Choose $\mathbf{u} \neq \mathbf{0}$ in U, and (by Theorem 6.4.1) extend $\{\mathbf{u}\}$ to a basis $\{\mathbf{u}, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of V. Now let $\mathbf{v}$ be any vector in V. Then (by Theorem 7.1.3) there is a linear transformation $T : V \to V$ such that $T(\mathbf{u}) = \mathbf{v}$ and $T(\mathbf{e}_i) = \mathbf{0}$ for each i. Then $\mathbf{v} = T(\mathbf{u})$ lies in U because U is T-invariant. As $\mathbf{v}$ was an arbitrary vector in V, it follows that $V = U$.

   [Remark: The only place we used the hypothesis that V is finite dimensional is in extending $\{\mathbf{u}\}$ to a basis of V. In fact, this is true for any vector space, even of infinite dimension.]

8. b. We have $U = \text{span}\left\{1 - 2x^2,\ x + x^2\right\}$. To show that U is T-invariant, it suffices (by Example 9.3.3) to show that $T(1 - 2x^2)$ and $T(x + x^2)$ both lie in U. We have

$$\left.\begin{array}{l} T(1-2x^2) = 3 + 3x - 3x^2 = 3(1-2x^2) + 3(x+x^2) \\ T(x+x^2) = -1 + 2x^2 = -(1-2x^2) \end{array}\right\} \qquad (*)$$

   So both $T(1 - 2x^2)$ and $T(x + x^2)$, so U is T-invariant. To get a block triangular matrix for T extend the basis $\left\{1 - 2x^2,\ x + x^2\right\}$ of U to a basis B of V in any way at all, say

$$B = \left\{1 - 2x^2,\ x + x^2,\ x^2\right\}$$

Then, using $(*)$, we have

$$M_B(T) = \left[\begin{array}{ccc} C_B\left[T(1-2x^2)\right] & C_B\left[T(x+x^2)\right] & C_B\left[T(x^2)\right] \end{array}\right] = \left[\begin{array}{rrr} 3 & -1 & 1 \\ 3 & 0 & 1 \\ 0 & 0 & 3 \end{array}\right]$$

where the last column is because $T(x^2) = 1 + x + 2x^2 = (1-2x^2) + (x+x^2) + 3(x^2)$. Finally,

$$c_T(x) = \left|\begin{array}{ccc} x-3 & 1 & -1 \\ -3 & x & -1 \\ 0 & 0 & x-3 \end{array}\right| = (x-3)\left|\begin{array}{cc} x-3 & 1 \\ -3 & x \end{array}\right| = (x-3)(x^2-3x+3)$$

9. b. *Algebraic Solution*. If U is T_A-invariant and $U \neq \{\mathbf{0}\}$, $U \neq \mathbb{R}^2$, then $\dim U = 1$. Thus $U = \mathbb{R}\mathbf{u}$ where $\mathbf{u} \neq \mathbf{0}$. Thus $T_A(\mathbf{u})$ is in $\mathbb{R}\mathbf{u}$ (because U is T-invariant), say $T_A(\mathbf{u}) = r\mathbf{u}$, that is $A\mathbf{u} = r\mathbf{u}$, whence $(rI - A)\mathbf{u} = \mathbf{0}$. But

$$\det(rI - A) = \left|\begin{array}{cc} r-\cos\theta & -\sin\theta \\ \sin\theta & r-\cos\theta \end{array}\right| = (r-\cos\theta)^2 + \sin^2\theta \neq 0 \text{ as } \sin\theta \neq 0 (0 < \theta < \pi)$$

Hence, $(rI - A)\mathbf{u} = 0$ implies $\mathbf{u} = \mathbf{0}$, a contradiction. So $U = 0$ or $U = \mathbb{R}^2$.

Geometric Solution. If we view $\mathbb{R}^2$ as the euclidean plane, and $U \neq 0$, $\mathbb{R}^2$, is a T_A-invariant subspace, then U must have dimension 1 and so be a line through the origin (Example 5.2.13). But T_A is rotation through θ counterclockwise about the origin (Theorem 2.6.4), so it will move the line U unless $\theta = 0$ or $\theta = \pi$, contrary to our assumption that $0 < \theta < \pi$. So no such line U can exist.

10. b. If $\mathbf{v}$ is in $U \cap W$, then $\mathbf{v} = (a, a, b, b) = (c, d, c, -d)$ for some a, b, c, d. Hence $a = c$, $a = d$, $b = c$ and $b = -d$. It follows that $d = -d$ so $a = b = c = d = 0$; that is $U \cap W = \{\mathbf{0}\}$. To see that $\mathbb{R}^4 = U + W$, we have (after solving systems of equations)

$$\begin{aligned}
(1, 0, 0, 0) &= \tfrac{1}{2}(1, 1, -1, -1) + \tfrac{1}{2}(1, -1, 1, 1) \text{ is in } U+W \\
(0, 1, 0, 0) &= \tfrac{1}{2}(1, 1, 1, 1) + \tfrac{1}{2}(-1, 1, -1, -1) \text{ is in } U+W \\
(0, 0, 1, 0) &= \tfrac{1}{2}(-1, -1, 1, 1) + \tfrac{1}{2}(1, 1, 1, -1) \text{ is in } U+W \\
(0, 0, 0, 1) &= \tfrac{1}{2}(1, 1, 1, 1) + \tfrac{1}{2}(-1, -1, -1, 1) \text{ is in } U+W
\end{aligned}$$

Hence, $\mathbb{R}^4 = U + W$. A simpler argument is as follows. As $\dim U = 2 = \dim W$, the subspace $U \oplus W$ has dimension $2 + 2 = 4$ by Theorem 9.3.6. Hence $U \oplus W = \mathbb{R}^4$ because $\dim \mathbb{R}^4 = 4$.

d. If A is in $U \cap W$, then $A = \left[\begin{array}{cc} a & a \\ b & b \end{array}\right] = \left[\begin{array}{cc} c & d \\ -c & d \end{array}\right]$ for some a, b, c, d, whence $a = b = c = d = 0$. Thus, $U \cap W = \{0\}$. Thus, by Theorem 9.3.7

$$\dim(U \oplus W) = \dim U + \dim W = 2 + 2 = 4$$

Since $\dim \mathbf{M}_{22} = 4$, we have $U \oplus W = \mathbf{M}_{22}$. Again, as in (b), we could show directly that each of $\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right]$, $\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]$, $\left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]$, $\left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$ is in $U + W$.

14. First U is a subspace because $0E = 0$, and $AE = A$ and $A_1E = A_1$ implies that

$$(A + A_1)E = AE + A_1E = A + A_1 \text{ and } (rA)E = r(AE) = rA \text{ for all } r \in \mathbb{R}$$

Similarly, W is a subspace because $0E = 0$, and $BE = 0 = B_1E$ implies that we have $(B + B_1)E = BE + B_1E = 0 + 0 = 0$ and $(rB)E = r(BE) = r0 = 0$ for all $r \in \mathbb{R}$.

These calculations hold for *any* matrix E; but if $E^2 = E$ we get $\mathbf{M}_{nn} = U \oplus W$. First $U \cap W = \{0\}$ because X in $U \cap W$ implies $X = XE$ because X is in U and $XE = 0$ because X is in W, so $X = XE = 0$. To prove that $U + W = \mathbf{M}_{nn}$ let X be any matrix in $\mathbf{M}_{nn}$. Then:

$$\begin{array}{rll} XE \text{ is in } U & \text{because} & (XE)E = XE^2 = XE \\ X - XE \text{ is in } W & \text{because} & (X - XE)E = XE - XE^2 = XE - XE = 0. \end{array}$$

Hence $X = XE + (X - XE)$ where XE is in U and $(X - XE)$ is in W; that is X is in $U + W$. Thus $\mathbf{M}_{nn} = U + W$.

17. By Theorem 6.4.5, we have $\dim(U \cap W) + \dim(U + W) = \dim U + \dim W = n$ by hypothesis. So if $U + W = V$ then $\dim(U + W) = n$, whence $\dim(U \cap W) = 0$. This means that $U \cap W = \{0\}$ so, since $U + W = V$, we have proved that $V = U \oplus W$.

18. b. First, $\ker T_A$ is T_A-invariant by Exercise 2. Now suppose that U is *any* T_A-invariant subspace, $U \neq 0$, $U \neq \mathbb{R}^2$. Then $\dim U = 1$, say $U = \mathbb{R}\mathbf{p}$, $\mathbf{p} \neq \mathbf{0}$. Thus $\mathbf{p}$ is in U so $A\mathbf{p} = T_A(\mathbf{p})$ is in U, say $A\mathbf{p} = \lambda\mathbf{p}$ where λ is a real number. Applying A again, we get $A^2\mathbf{p} = \lambda A\mathbf{p} = \lambda^2\mathbf{p}$. But $A^2 = 0$, so this gives $\mathbf{0} = \lambda^2\mathbf{p}$. Thus $\lambda^2 = 0$, whence $\lambda = 0$ and $A\mathbf{p} = \lambda\mathbf{p} = \mathbf{0}$. Hence $\mathbf{p}$ is in $\ker T_A$, whence $U \subseteq \ker T_A$. But $\dim U = 1 = \dim(\ker T_A)$, so $U = \ker T_A$.

20. Let B_1 be a basis of U and extend it (using Theorem 6.4.1) to a basis B of V. Then $M_B(T) = \left[\begin{array}{cc} M_{B_1}(T) & Y \\ 0 & Z \end{array}\right]$ by Theorem 9.3.1. Since we are writing T_1 for the restriction of T to U, $M_{B_1}(T) = M_{B_1}(T_1)$. Hence,

$$\begin{aligned} c_T(x) &= \det[xI - M_B(T)] = \det\left[\begin{array}{cc} xI - M_{B_1}(T) & -x \\ 0 & xI - Z \end{array}\right] \\ &= \det[xI - M_{B_1}(T_1)]\det[xI - Z] = c_{T_1}(x) \cdot q(x) \end{aligned}$$

where $q(x) = \det[xI - Z]$.

22. b. We have $T : \mathbf{P}_3 \to \mathbf{P}_3$ given by $T[p(x)] = p(-x)$ for all $p(x)$ in $\mathbf{P}_3$. We leave it to the reader to verify that T is a linear operator. We have

$$T^2[p(x)] = T\{T[p(x)]\} = T[p(-x)] = p(-(-x)) = p(x) = 1_{\mathbf{P}_3}(p(x))$$

Hence, $T^2 = 1_{\mathbf{P}_3}$. As in Example 9.3.10, let

$$\begin{aligned} U_1 &= \{p(x) \mid T[p(x)] = p(x)\} = \{p(x) \mid p(-x) = p(x)\} \\ U_2 &= \{p(x) \mid T[p(x)] = -p(x)\} = \{p(x) \mid p(-x) = -p(x)\} \end{aligned}$$

These are the subspaces of even and odd polynomials in $\mathbf{P}_3$, respectively, and have bases $B_1 = \{1, x^2\}$ and $B_2 = \{x, x^3\}$. Hence, use the ordered basis $B = \{1, x^2, x, x^3\}$ of $\mathbf{P}_3$. Then

$$M_B(T) = \left[\begin{array}{cc} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{array}\right] = \left[\begin{array}{cc} I_2 & 0 \\ 0 & -I_2 \end{array}\right]$$

as in Example 9.3.10. More explicitly,

$$M_B(T) = \left[\begin{array}{cccc} C_B[T(1)] & C_B[T(x^2)] & C_B[T(x)] & C_B[T(x^3)] \end{array}\right]$$

$$= \begin{bmatrix} C_B(1) & C_B(x^2) & C_B(-x) & C_B(-x^3) \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}$$

d. Here $T^2(a,\ b,\ c) = [-(-a+2b+c)+2((b+c)+(-c),\ (b+c)-c),\ -(-c)] = (a,\ b,\ c)$, so $T^2 = 2_{\mathbb{R}^3}$.
Note that $T(1,\ 1,\ 0) = (1,\ 1,\ 0)$, while $T(1,\ 0,\ 0) = -(1,\ 0,\ 0)$ and $T(0,\ 1,\ -2) = -(0,\ 1,\ -2)$. Let $B_1 = \{(1,\ 1,\ 0)\}$ and $B_2 = \{(1,\ 0,\ 0),\ (0,\ -1,\ 2)\}$. These are bases of $U_1 = \mathbb{R}(1,\ 1,\ 0)$ and $U_2 = \mathbb{R}(1,\ 0,\ 0) + \mathbb{R}(0,\ 1,\ -2)$, respectively. So if we take $B = \{(1,\ 1,\ 0),\ (1,\ 0,\ 0),\ (0,\ -1,\ 2)\}$ then $M_{B_1}(T) = [1]$ and $M_{B_2}(T) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Hence $M_B(T) = \begin{bmatrix} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

23. b. Given $\mathbf{v}$, $T[\mathbf{v} - T(\mathbf{v})] = T(\mathbf{v}) - T^2(\mathbf{v}) = T(\mathbf{v}) - T(\mathbf{v}) = \mathbf{0}$, so $\mathbf{v} - T(\mathbf{v})$ lies in $\ker T$. Hence $\mathbf{v} = (\mathbf{v} - T(\mathbf{v})) + T(\mathbf{v})$ is in $\ker T + \operatorname{im} T$ for all $\mathbf{v}$, that is $V = \ker T + \operatorname{im} T$. If $\mathbf{v}$ lies in $\ker T \cap \operatorname{im} T$, write $\mathbf{v} = T(\mathbf{w})$, $\mathbf{w}$ in V. Then $\mathbf{0} = T(\mathbf{v}) = T^2(\mathbf{w}) = T(\mathbf{w}) = \mathbf{v}$, so $\ker T \cap \operatorname{im} T = \mathbf{0}$.

25. b. We first verify that $T^2 = T$. Given $(a,\ b,\ c)$ in $\mathbb{R}^3$, we have

$$T^2(a,\ b,\ c) = T(a+2b,\ 0,\ 4b+c) = (a+2b,\ 0,\ 4b+c) = T(a,\ b,\ c)$$

Hence $T^2 = T$. As in the preceding exercise, write

$$U_1 = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{v}\} \qquad \text{and} \qquad U_2 = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{0}\} = \ker(T).$$

Then we claim that $\mathbb{R}^3 = U_1 \oplus U_2$. To show $\mathbb{R}^3 = U_1 + U_2$, observe that $\mathbf{v} = T(\mathbf{v}) + [\mathbf{v} - T(\mathbf{v})]$ for each $\mathbf{v}$ in $\mathbb{R}^3$, and $T(\mathbf{v})$ is in U_1 [because $T[T(\mathbf{v})] = T^2(\mathbf{v}) = T(\mathbf{v})$] while $\mathbf{v} - T(\mathbf{v})$ is in U_2 [because $T[\mathbf{v} - T(\mathbf{v})] = T(\mathbf{v}) - T^2(\mathbf{v}) = \mathbf{0}$]. Finally we show that $U_1 \cap U_2 = \{\mathbf{0}\}$. For if $\mathbf{v}$ is in $U_1 \cap U_2$ then $T(\mathbf{v}) = \mathbf{v}$ and $T(\mathbf{v}) = \mathbf{0}$ so certainly $\mathbf{v} = 0$.

Next, we show that U_1 and U_2 are T-invariant. If $\mathbf{v}$ is in U_1 then $T(\mathbf{v})$ is also in U_1 because $T[T(\mathbf{v})] = T^2(\mathbf{v}) = T(\mathbf{v})$. Similarly U_2 is T-invariant because, if $\mathbf{v}$ is in U_2, that is $T(\mathbf{v}) = \mathbf{0}$, then $T[T(\mathbf{v})] = T^2(\mathbf{v}) = R(\mathbf{v}) = \mathbf{0}$; that is $T(\mathbf{v})$ is also in U_2.

It is clear that $T(a,\ b,\ c) = (a,\ b,\ c)$ if and only if $b = 0$; that is $U_1 = \{(a,\ 0,\ c) \mid b,\ c \text{ in } \mathbb{R}\}$, so $B_1 = \{(1,\ 0,\ 0),\ (0,\ 0,\ 1)\}$ is a basis of U_1. Since $T(v) = v$ for all v in U_1 the restriction of T to U_1 is the identity transformation on U_1, and so has matrix I_2.

Similarly, $T(a,\ b,\ c) = (0,\ 0,\ 0)$ holds if and only if $a = -2b$ and $c = -4b$ for some b, so $U_1 = \mathbb{R}(2,\ -1,\ 4)$ and $B_2 = \{(2,\ -1,\ 4)\}$ is a basis of U_2. Clearly the restriction of T to U_2 is the zero transformation, and so has matrix 0_2 — a 1×1 matrix.

Finally then, $B = B_1 \cup B_2 = \{(1,\ 0,\ 0),\ (0,\ 0,\ 1),\ (2,\ -1,\ 4)\}$ is a basis of $\mathbb{R}^3$ (since we have shown that $\mathbb{R}^3 = U_1 \oplus U_2$), so T has matrix $\begin{bmatrix} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ 0 & 0_1 \end{bmatrix}$.

29. b. We have $T^2_{f,\ \mathbf{z}}[\mathbf{v}] = T_{f,\ \mathbf{z}}[T_{f,\ \mathbf{z}}(\mathbf{v})] = T_{f,\ \mathbf{z}}[f(\mathbf{v})\mathbf{z}] = f[f(\mathbf{v})\mathbf{z}]\mathbf{z} = f(\mathbf{v})f(\mathbf{z})\mathbf{z}$. This expression equals $T_{f,\ \mathbf{z}}(\mathbf{v}) = f(\mathbf{v})\mathbf{z}$ for all $\mathbf{v}$ if and only if

$$f(\mathbf{v})(\mathbf{z} - f(\mathbf{z})\mathbf{z}) = \mathbf{0}$$

for all $\mathbf{v}$. Since $f \neq 0$, $f(\mathbf{v}) \neq 0$ for some $\mathbf{v}$, so this holds if and only if

$$\mathbf{z} = f(\mathbf{z})\mathbf{z}$$

As $\mathbf{z} \neq \mathbf{0}$, this holds if and only if $f(\mathbf{z}) = 1$.

30. b. Let λ be an eigenvalue of T. If A is in $E_\lambda(T)$ then $T(A) = \lambda A$; that is $UA = \lambda A$. If we write $A = \left[\begin{array}{cccc} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{array} \right]$ in terms of its columns $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$, then $UA = \lambda A$ becomes

$$U \left[\begin{array}{cccc} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{array} \right] = \lambda \left[\begin{array}{cccc} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{array} \right]$$
$$\left[\begin{array}{cccc} U\mathbf{p}_1 & U\mathbf{p}_2 & \cdots & U\mathbf{p}_n \end{array} \right] = \left[\begin{array}{cccc} \lambda\mathbf{p}_1 & \lambda\mathbf{p}_2 & \cdots & \lambda\mathbf{p}_n \end{array} \right]$$

Comparing columns gives $U\mathbf{p}_i = \lambda \mathbf{p}_i$ for each i; that is $\mathbf{p}_i$ is in $E_\lambda(U)$ for each i. Conversely, if $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ are all in $E_\lambda(U)$ then $U\mathbf{p}_i = \lambda \mathbf{p}_i$ for each i, so $T(A) = UA = \lambda A$ as above. Thus A is in $E_\lambda(T)$.

10. Inner Product Spaces

10.1 Inner Products and Norms

1. b. $P5$ fails: $\langle (0,\, 1,\, 0),\, (0,\, 1,\, 0)\rangle = -1$
   The other axioms hold. Write $\mathbf{x} = (x_1,\, x_2,\, x_3)$, $\mathbf{y} = (y_1,\, y_2,\, y_3)$ and $\mathbf{z} = (z_1,\, z_2,\, z_3)$.
   $P1$ holds: $\langle \mathbf{x},\, \mathbf{y}\rangle = x_1y_1 - x_2y_2 + x_3y_3$ is real for all $x,\, y$ in $\mathbb{R}^n$.
   $P2$ holds: $\langle \mathbf{x},\, \mathbf{y}\rangle = x_1y_1 - x_2y_2 + x_3y_3 = y_1x_1 - y_2x_2 + y_3x_3 = \langle \mathbf{y},\, \mathbf{x}\rangle$
   $P3$ holds: $$\begin{aligned}\langle \mathbf{x}+\mathbf{y},\, \mathbf{z}\rangle &= (x_1+y_1)z_1 - (x_2+y_2)z_2 + (x_3+y_3)z_3 \\ &= (x_1z_1 - x_2z_2 + x_3z_3) + (y_1z_1 - y_2z_2 + y_3z_3) = \langle \mathbf{x},\, \mathbf{z}\rangle + \langle \mathbf{y},\, \mathbf{z}\rangle\end{aligned}$$
   $P4$ holds: $\langle r\mathbf{x},\, \mathbf{y}\rangle = (rx_1)y_1 - (rx_2)y_2 + (rx_3)y_3 = r(x_1y_1 - x_2y_2 + x_3y_3) = r\langle \mathbf{x},\, \mathbf{y}\rangle$

   d. $P5$ fails: $\langle x-1,\, x-1\rangle = 0\cdot 0 = 0$
   $P1$ holds: $\langle p(x),\, q(x)\rangle = p(1)q(1)$ is real.
   $P2$ holds: $\langle p(x),\, q(x)\rangle = p(1)q(1) = q(1)p(1) = \langle q(x),\, p(x)\rangle$
   $P3$ holds: $$\begin{aligned}\langle p(x)+r(x),\, q(x)\rangle &= [p(1)+r(1)]q(1) = p(1)q(1) + r(1)q(1) \\ &= \langle p(x),\, q(x)\rangle + \langle r(x),\, q(x)\rangle\end{aligned}$$
   $P4$ holds: $\langle rp(x),\, q(x)\rangle = [rp(1)]q(1) = r[p(1)q(1)] = r\langle p(x),\, q(x)\rangle$

   f. $P5$ fails: Here $\langle f,\, f\rangle = 2f(0)f(1)$ for any f, so if $f(x) = x - \frac{1}{2}$ then $\langle f,\, f\rangle = -\frac{1}{2}$.
   $P1$ holds: $\langle f,\, g\rangle = f(1)g(0) + f(0)g(1)$ is real.
   $P2$ holds: $\langle f,\, g\rangle = f(1)g(0) + f(0)g(1) = g(1)f(0) + g(0)f(1) = \langle g,\, f\rangle$
   $P3$ holds: $$\begin{aligned}\langle f+h,\, g\rangle &= (f+h)(1)g(0) + (f+h)(0)g(1) \\ &= [f(1)+h(1)]g(0) + [f(0)+h(0)]g(1) \\ &= [f(1)g(0)+f(0)g(1)] + [h(1)g(0)+h(0)g(1)] = \langle f,\, g\rangle + \langle h,\, g\rangle\end{aligned}$$
   $P4$ holds: $$\begin{aligned}\langle rf,\, h\rangle &= (rf)(1)g(0) + (rf)(0)g(1) = [r\cdot f(1)]g(0) + [rf(0)]g(1) \\ &= r[f(1)g(0) + f(0)g(1)] = r\langle f,\, g\rangle\end{aligned}$$

2. If $\langle\ ,\ \rangle$ denotes the inner product on V, then $\langle \mathbf{u}_1,\, \mathbf{u}_2\rangle$ is a real number for all $\mathbf{u}_1$ and $\mathbf{u}_2$ in U. Moreover, the axioms $P1 - P5$ hold for the space U because they hold for V and U is a subset of V. So $\langle\ ,\ \rangle$ is an inner product for the vector space U.

3. b. $\|f\|^2 = \int_{-\pi}^{\pi} \cos^2 x\,dx = \int_{-\pi}^{\pi} \frac{1}{2}[1+\cos(2x)]\,dx = \frac{1}{2}\left[x + \frac{1}{2}\sin(2x)\right]_{-\pi}^{\pi} = \pi$. Hence $\hat{f} = \frac{1}{\sqrt{\pi}}f$ is a unit vector.

   d. $\|\mathbf{v}\|^2 = \langle \mathbf{v},\, \mathbf{v}\rangle = \mathbf{v}^T \left[\begin{array}{rr} 1 & -1 \\ -1 & 2 \end{array}\right] \mathbf{v} = \left[\begin{array}{rr} 3 & -1 \end{array}\right] \left[\begin{array}{rr} 1 & -1 \\ -1 & 2 \end{array}\right] \left[\begin{array}{r} 3 \\ -1 \end{array}\right] = 17$.

   Hence $\frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{\sqrt{17}} \left[\begin{array}{r} 3 \\ -1 \end{array}\right]$ is a unit vector in this space.

4. b. $d(\mathbf{u},\, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|(1,\, 2,\, -1,\, 2) - (2,\, 1,\, -1,\, 3)\| = \|(-1,\, 1,\, 0,\, -1)\| = \sqrt{3}$.

d. $\|f-g\|^2 = \int_{-\pi}^{\pi}(1-\cos x)^2\,dx = \int_{-\pi}^{\pi}\left[\frac{3}{2}-2\cos x+\frac{1}{2}\cos(2x)\right]dx$ because we have $\cos^2(x) = \frac{1}{2}[1+\cos(2x)]$. Hence $\|f-g\|^2 = \left[\frac{3}{2}x-2\sin(x)+\frac{1}{4}\sin(2x)\right]_{-\pi}^{\pi} = \frac{3}{2}[\pi-(-\pi)] = 3\pi$. Hence $d(f,\, g) = \sqrt{3\pi}$.

8. The space $\mathbf{D}_n$ uses pointwise addition and scalar multiplication:

$$(f+g)(k) = f(k)+g(k) \quad \text{and} \quad (rf)(k) = rf(k)$$

for all $k = 1,\, 2,\, \dots,\, n$.
$P1$. $\langle f,\, g\rangle = f(1)g(1)+f(2)g(2)+\cdots+f(n)g(n)$ is real.
$P2$. $\langle f,\, g\rangle = f(1)g(1)+f(2)g(2)+\cdots+f(n)g(n)) = g(1)f(1)+g(2)f(2)+\cdots+g(n)f(n) = \langle g,\, f\rangle$

$P3$.
$$\begin{aligned}\langle f+h,\, g\rangle &= (f+h)(1)g(1)+(f+h)(2)g(2)+\cdots+(f+h)(n)g(n)\\ &= [f(1)+h(1)]g(1)+[f(2)+h(2)]g(2)+\cdots+[f(n)+h(n)]g(n)\\ &= [f(1)g(1)+f(2)g(2)+\cdots+f(n)g(n)]+[h(1)g(1)+h(2)g(2)+\cdots+h(n)g(n)]\\ &= \langle f,\, g\rangle+\langle h,\, g\rangle\end{aligned}$$

$P4$.
$$\begin{aligned}\langle rf,\, g\rangle &= (rf)(1)g(1)+(rf)(2)g(2)+\cdots+(rf)(n)g(n)\\ &= [rf(1)]g(1)+[rf(2)]g(2)+\cdots+[rf(n)]g(n)\\ &= r[f(1)g(1)+f(2)g(2)+\cdots+f(n)g(n)] = r\langle f,\, g\rangle\end{aligned}$$

$P5$. $\langle f,\, f\rangle = f(1)^2+f(2)^2+\cdots+f(n)^2 \geq 0$ for all f. If $\langle f,\, f\rangle = 0$ then

$$f(1) = f(2) = \cdots = f(n) = 0 \text{ (as the } f(k) \text{ are real numbers) so } f = 0$$

12. b. We need only verify $P5$. [$P1-P4$ hold for any symmetric matrix A by (the discussion preceding) Theorem 10.1.2.] If $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$:

$$\begin{aligned}\langle \mathbf{v},\, \mathbf{v}\rangle = \mathbf{v}^T A\mathbf{v} &= \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\\ &= 5v_1^2-6v_1v_2+2v_2^2\\ &= 5\left[v_1^2-\tfrac{6}{5}v_1v_2+\tfrac{9}{25}v_2^2\right]-\tfrac{9}{5}v_2^2+2v_2^2\\ &= 5\left(v_1-\tfrac{3}{5}v_2\right)^2+\tfrac{1}{5}v_2^2\\ &= \tfrac{1}{5}\left[(5v_1-3v_2)^2+v_2^2\right]\end{aligned}$$

Thus, $\langle \mathbf{v},\, \mathbf{v}\rangle \geq 0$ for all $\mathbf{v}$; and $\langle \mathbf{v},\, \mathbf{v}\rangle = 0$ if and only if $5v_1-3v_2 = 0 = v_2$; that is if and only if $v_1 = v_2 = 0$ (i.e. $\mathbf{v} = \mathbf{0}$). So $P5$ holds.

d. As in (b), consider $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

$$\begin{aligned}\langle \mathbf{v},\, \mathbf{v}\rangle &= \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\\ &= 3v_1^2+8v_1v_2+6v_2^2\\ &= 3(v_1^2+\tfrac{8}{3}v_1v_2+\tfrac{16}{9}v_2^2)-\tfrac{16}{3}v_2^2+6v_2^2\\ &= 3\left(v_1+\tfrac{4}{3}v_2\right)^2+\tfrac{2}{3}v_2^2\\ &= \tfrac{1}{3}\left[(3v_1+4v_2)^2+2v_2^2\right]\end{aligned}$$

Thus, $\langle \mathbf{v},\, \mathbf{v}\rangle \geq 0$ for all $\mathbf{v}$; and $(\mathbf{v},\, \mathbf{v}) = 0$ if and only if $3v_1+4v_2 = 0 = v_2$; that is if and only if $\mathbf{v} = \mathbf{0}$. Hence $P5$ holds. The other axioms hold because A is symmetric.

13. b. If $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, then a_{ij} is the coefficient of $v_i w_j$ in $\langle \mathbf{v},\ \mathbf{w} \rangle$. Here $a_{11} = 1$, $a_{12} = -1 = a_{21}$, and $a_{22} = 2$. Thus, $A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$. Note that $a_{12} = a_{21}$, so A is symmetric.

d. As in (b): $A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 5 \end{bmatrix}$.

14. As in the hint, write $\langle \mathbf{x},\ \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$. Since A is symmetric, this satisfies axioms $P1$, $P2$, $P3$ and $P4$ for an inner product on $\mathbb{R}^n$—(and only $P2$ requires that A be symmetric). Then it follows that

$$0 = \langle \mathbf{x}+\mathbf{y},\ \mathbf{x}+\mathbf{y} \rangle = \langle \mathbf{x},\ \mathbf{x} \rangle + \langle \mathbf{x},\ \mathbf{y} \rangle + \langle \mathbf{y},\ \mathbf{x} \rangle + \langle \mathbf{y},\ \mathbf{y} \rangle = 2\langle \mathbf{x},\ \mathbf{y} \rangle \text{ for all } \mathbf{x},\ \mathbf{y} \text{ in } \mathbb{R}^n.$$

Hence $\langle \mathbf{x},\ \mathbf{y} \rangle = 0$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. But if $\mathbf{e}_j$ denotes column j of I_n, then $\langle \mathbf{e}_i,\ \mathbf{e}_j \rangle = \mathbf{e}_i^T A \mathbf{e}_j$ is the $(i,\ j)$-entry of A. It follows that $A = 0$.

16. b.
$$\begin{aligned} \langle \mathbf{u}-2\mathbf{v}-\mathbf{w},\ 3\mathbf{w}-\mathbf{v} \rangle &= 3\langle \mathbf{u},\ \mathbf{w} \rangle - 6\langle \mathbf{v},\ \mathbf{w} \rangle - 3\langle \mathbf{w},\ \mathbf{w} \rangle - \langle \mathbf{u},\ \mathbf{v} \rangle + 2\langle \mathbf{v},\ \mathbf{v} \rangle + \langle \mathbf{w},\ \mathbf{v} \rangle \\ &= 3\langle \mathbf{u},\ \mathbf{w} \rangle - 5\langle \mathbf{v},\ \mathbf{w} \rangle - 3\,\|\mathbf{w}\|^2 - \langle \mathbf{u},\ \mathbf{v} \rangle + 2\,\|\mathbf{v}\|^2 \\ &= 3\cdot 0 - 5\cdot 3 - 3\cdot 3 - (-1) + 2\cdot 4 \\ &= -15 \end{aligned}$$

20. (1) $\langle \mathbf{u},\ \mathbf{v}+\mathbf{w} \rangle \overset{P2}{=} \langle \mathbf{v}+\mathbf{w},\ \mathbf{u} \rangle \overset{P3}{=} \langle \mathbf{v},\ \mathbf{u} \rangle + \langle \mathbf{w},\ \mathbf{u} \rangle \overset{P2}{=} \langle \mathbf{u},\ \mathbf{v} \rangle + \langle \mathbf{u},\ \mathbf{w} \rangle$

(2) $\langle \mathbf{v},\ r\mathbf{w} \rangle \overset{P2}{=} \langle r\mathbf{w},\ \mathbf{v} \rangle \overset{P4}{=} r\langle \mathbf{w},\ \mathbf{v} \rangle \overset{P2}{=} r\langle \mathbf{v},\ \mathbf{w} \rangle$

(3) By (1): $\langle \mathbf{v},\ \mathbf{0} \rangle = \langle \mathbf{v},\ \mathbf{0}+\mathbf{0} \rangle \overset{(1)}{=} \langle \mathbf{v},\ \mathbf{0} \rangle + \langle \mathbf{v},\ \mathbf{0} \rangle$. Hence $\langle \mathbf{v},\ \mathbf{0} \rangle = 0$. Now $\langle \mathbf{0},\ \mathbf{v} \rangle = 0$ by P2.

(4) If $\mathbf{v} = \mathbf{0}$ then $\langle \mathbf{v},\ \mathbf{v} \rangle = \langle \mathbf{0},\ \mathbf{0} \rangle = 0$ by (3). If $\langle \mathbf{v},\ \mathbf{v} \rangle = 0$ then it is impossible that $\mathbf{v} \neq \mathbf{0}$ by $P5$, so $\mathbf{v} = \mathbf{0}$.

22. b.
$$\begin{aligned} \langle 3\mathbf{u}-4\mathbf{v},\ 5\mathbf{u}+\mathbf{v} \rangle &= 15\langle \mathbf{u},\ \mathbf{u} \rangle + 3\langle \mathbf{u},\ \mathbf{v} \rangle - 20\langle \mathbf{v},\ \mathbf{u} \rangle - 4\langle \mathbf{v},\ \mathbf{v} \rangle \\ &= 15\,\|\mathbf{u}\|^2 - 17\langle \mathbf{u},\ \mathbf{v} \rangle - 4\,\|\mathbf{v}\|^2 \end{aligned}$$

d.
$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 = \langle \mathbf{u}+\mathbf{v},\ \mathbf{u}+\mathbf{v} \rangle &= \langle \mathbf{u},\ \mathbf{u} \rangle + \mathbf{u},\ \mathbf{v} \rangle + \langle \mathbf{v},\ \mathbf{u} \rangle + 4\langle \mathbf{v},\ \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u},\ \mathbf{v} \rangle + \|\mathbf{v}\|^2 \end{aligned}$$

26. b. Here

$$\begin{aligned} W &= \left\{ \mathbf{w} \mid \mathbf{w} \text{ in } \mathbb{R}^3 \text{ and } \mathbf{v}\cdot\mathbf{w} = 0 \right\} \\ &= \{(x,\ y,\ z) \mid x - y + 2z = 0\} \\ &= \{(s,\ s+2t,\ t) \mid s,\ t \text{ in } \mathbb{R}\} \\ &= \operatorname{span} B \end{aligned}$$

where $B = \{(1,\ 1,\ 0),\ (0,\ 2,\ 1)\}$. Then B is the desired basis because B is independent [In fact, if $s(1,\ 1,\ 0) + t(0,\ 2,\ 1) = (s,\ s+2t,\ t) = (0,\ 0,\ 0)$ then $s = t = 0$].

28. Write $\mathbf{u} = \mathbf{v} - \mathbf{w}$; we show that $\mathbf{u} = \mathbf{0}$. We are given that

$$\langle \mathbf{u},\ \mathbf{v}_i \rangle = \langle \mathbf{v}-\mathbf{w},\ \mathbf{v}_i \rangle = \langle \mathbf{v},\ \mathbf{v}_i \rangle - \langle \mathbf{w},\ \mathbf{v}_i \rangle = 0$$

for each i. As $V = \operatorname{span}\{\mathbf{v}_1,\ \dots,\ \mathbf{v}_n\}$, write $\mathbf{u} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$, r_i in $\mathbb{R}$. Then

$$\|\mathbf{u}\|^2 = \langle \mathbf{u},\ \mathbf{u} \rangle = \langle \mathbf{u},\ r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n \rangle$$

$$\begin{aligned} &= r_1\langle \mathbf{u},\ \mathbf{v}_1\rangle + \cdots + r_n\langle \mathbf{u},\ \mathbf{v}_n\rangle \\ &= r_1 \cdot 0 + \cdots + r_n \cdot 0 \\ &= 0 \end{aligned}$$

Thus, $\|\mathbf{u}\| = 0$, so $\mathbf{u} = \mathbf{0}$.

29. b. If $\mathbf{u} = (\cos\theta,\ \sin\theta)$ in $\mathbb{R}^2$ (with the dot product), then $\|\mathbf{u}\| = 1$. If $\mathbf{v} = (x,\ y)$ the Schwarz inequality (Theorem 10.1.4) gives

$$\langle \mathbf{u},\ \mathbf{v}\rangle^2 \leq \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \leq 1 \cdot \|\mathbf{v}\|^2 = \|\mathbf{v}\|^2$$

This is what we wanted.

10.2 Orthogonal Sets of Vectors

1. b. B is an orthogonal set because (writing $\mathbf{f}_1 = \left[\begin{array}{c} 1 \\ 1 \\ 1 \end{array}\right]$, $\mathbf{f}_2 = \left[\begin{array}{c} -1 \\ 0 \\ 1 \end{array}\right]$ and $\mathbf{f}_3 = \left[\begin{array}{c} 1 \\ -6 \\ 1 \end{array}\right]$)

$$\langle \mathbf{e}_1,\ \mathbf{e}_2\rangle = \left[\begin{array}{ccc} 1 & 1 & 1 \end{array}\right]\left[\begin{array}{ccc} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{array}\right]\left[\begin{array}{c} -1 \\ 0 \\ 1 \end{array}\right] = \left[\begin{array}{ccc} 3 & 1 & 3 \end{array}\right]\left[\begin{array}{c} -1 \\ 0 \\ 1 \end{array}\right] = 0$$

$$\langle \mathbf{f}_1,\ \mathbf{f}_3\rangle = \left[\begin{array}{ccc} 1 & 1 & 1 \end{array}\right]\left[\begin{array}{ccc} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{array}\right]\left[\begin{array}{c} 1 \\ -6 \\ 1 \end{array}\right] = \left[\begin{array}{ccc} 3 & 1 & 3 \end{array}\right]\left[\begin{array}{c} 1 \\ -6 \\ 1 \end{array}\right] = 0$$

$$\langle \mathbf{f}_2,\ \mathbf{f}_3\rangle = \left[\begin{array}{ccc} -1 & 0 & 1 \end{array}\right]\left[\begin{array}{ccc} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{array}\right]\left[\begin{array}{c} 1 \\ -6 \\ 1 \end{array}\right] = \left[\begin{array}{ccc} -1 & 0 & 1 \end{array}\right]\left[\begin{array}{c} 1 \\ -6 \\ 1 \end{array}\right] = 0$$

Thus, B is an orthogonal basis of V and the expansion theorem gives

$$\begin{aligned} \mathbf{v} &= \tfrac{\langle \mathbf{v},\ \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \tfrac{\langle \mathbf{v},\ \mathbf{f}_2\rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \tfrac{\langle \mathbf{v},\ \mathbf{f}_3\rangle}{\|\mathbf{f}_3\|^2}\mathbf{f}_3 \\ &= \tfrac{3a+b+3c}{7}\mathbf{e}_1 + \tfrac{c-a}{2}\mathbf{e}_2 + \tfrac{3a-6b+3c}{42}\mathbf{e}_3 \\ &= \tfrac{1}{14}[(6a+2b+6c)\mathbf{e}_1 + (7c-7a)\mathbf{e}_2 + (a-2b+c)\mathbf{e}_3] \end{aligned}$$

d. Observe first that $\left\langle \left[\begin{array}{cc} a & b \\ c & d \end{array}\right],\ \left[\begin{array}{cc} a' & b' \\ c' & d' \end{array}\right] \right\rangle = aa' + bb' + cc' + dd'$. Now write $B = \{\mathbf{f}_1,\ \mathbf{f}_2,\ \mathbf{f}_3,\ \mathbf{f}_4\}$ where $\mathbf{f}_1 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$, $\mathbf{f}_2 = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]$, $\mathbf{f}_3 = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$, $\mathbf{f}_4 = \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right]$. Then B is orthogonal because

$$\begin{array}{ll} \langle \mathbf{f}_1,\ \mathbf{f}_2\rangle = 1+0+0-1 = 0 & \langle \mathbf{f}_2,\ \mathbf{f}_3\rangle = 0+0+0+0 = 0 \\ \langle \mathbf{f}_1,\ \mathbf{f}_3\rangle = 0+0+0+0 = 0 & \langle \mathbf{f}_2,\ \mathbf{f}_4\rangle = 0+0+0+0 = 0 \\ \langle \mathbf{f}_1,\ \mathbf{f}_4\rangle = 0+0+0+0 = 0 & \langle \mathbf{f}_3,\ \mathbf{f}_4\rangle = 0+1+0+-1 = 0 \end{array}$$

The expansion theorem gives

$$\begin{aligned} \mathbf{v} &= \tfrac{\langle \mathbf{v},\ \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \tfrac{\langle \mathbf{v},\ \mathbf{f}_2\rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \tfrac{\langle \mathbf{v},\ \mathbf{f}_3\rangle}{\|\mathbf{f}_3\|^2}\mathbf{f}_3 + \tfrac{\langle \mathbf{v},\ \mathbf{f}_4\rangle}{\|\mathbf{f}_4\|^2}\mathbf{f}_4 \\ &= \left(\tfrac{a+d}{2}\right)\mathbf{f}_1 + \left(\tfrac{a-d}{2}\right)\mathbf{f}_2 + \left(\tfrac{b+c}{2}\right)\mathbf{f}_3 + \left(\tfrac{b-c}{2}\right)\mathbf{f}_4 \end{aligned}$$

2. b. Write $\mathbf{b}_1 = (1, 1, 1)$, $\mathbf{b}_2 = (1, -1, 1)$, $\mathbf{b}_3 = (1, 1, 0)$. Note that in the Gram-Schmidt algorithm we may multiply each $\mathbf{e}_i$ by a nonzero constant and not change the subsequent $\mathbf{e}_i$. This avoids fractions.

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{b}_1 = (1, 1, 1) \\
\mathbf{f}_2 &= \mathbf{b}_2 - \tfrac{\langle \mathbf{b}_2, \mathbf{e}_1 \rangle}{\|\mathbf{e}_1\|^2}\mathbf{f}_1 \\
&= (1, -1, 1) - \tfrac{4}{6}(1, 1, 1) \\
&= \tfrac{1}{3}(1, -5, 1); \text{ use } \mathbf{e}_3 = (1, -5, 1) \text{ with no loss of generality} \\
\mathbf{f}_3 &= \mathbf{b}_3 - \tfrac{\langle \mathbf{b}_3, \mathbf{f}_1 \rangle}{\|\mathbf{e}_1\|^2}\mathbf{f}_1 - \tfrac{\langle \mathbf{b}_3, \mathbf{f}_2 \rangle}{\|\mathbf{e}_2\|^2}\mathbf{f}_2 \\
&= (1, 1, 0) - \tfrac{3}{6}(1, 1, 1) - \tfrac{(-3)}{(30)} \cdot (1, -5, 1) \\
&= \tfrac{1}{10}[(10, 10, 9) - (5, 5, 5) + (1, -5, 1)] \\
&= \tfrac{1}{5}(3, 0, -2); \text{ use } \mathbf{f}_3 = (3, 0, -2) \text{ with no loss of generality}
\end{aligned}$$

So the orthogonal basis is $\{(1, 1, 1), (1, -5, 1), (3, 0, -2)\}$.

3. b. Note that $\left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \right\rangle = aa' + bb' + cc' + dd'$. For convenience write $\mathbf{b}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{b}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then:

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{b}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\
\mathbf{f}_2 &= \mathbf{b}_2 - \tfrac{\langle \mathbf{b}_2, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
&= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} - \tfrac{2}{3}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \tfrac{1}{3}\begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}
\end{aligned}$$

For the rest of the algorithm, use $\mathbf{f}_2 = \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}$, the result is the same.

$$\begin{aligned}
\mathbf{f}_3 &= \mathbf{b}_3 - \tfrac{\langle \mathbf{b}_3, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \tfrac{\langle \mathbf{b}_3, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \tfrac{2}{3}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \tfrac{2}{15}\begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} \\
&= \tfrac{1}{5}\begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}
\end{aligned}$$

Now use $\mathbf{f}_4 = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$, the results are unchanged.

$$\begin{aligned}
\mathbf{f}_4 &= \mathbf{b}_4 - \tfrac{\langle \mathbf{b}_4, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \tfrac{\langle \mathbf{b}_4, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \tfrac{\langle \mathbf{b}_4, \mathbf{f}_3 \rangle}{\|\mathbf{f}_3\|^2}\mathbf{f}_3 \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \tfrac{1}{3}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \tfrac{1}{15}\begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} - \tfrac{1}{10}\begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \\
&= \tfrac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned}$$

Use $\mathbf{f}_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ for convenience. Hence, finally, the Gram-Schmidt algorithm gives the orthogonal basis $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\}$.

4. b. $\mathbf{f}_1 = 1$
$\mathbf{f}_2 = x - \frac{\langle x,\, \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = x - \frac{2}{2}\cdot 1 = x - 1$
$\mathbf{f}_3 = x^2 - \frac{\langle x^2,\, \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle x^2,\, \mathbf{f}_2\rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 = x^2 - \frac{8/3}{2}\cdot 1 - \frac{4/3}{2/3}\cdot(x-1) = x^2 - 2x + \frac{2}{3}.$

6. b. $\left[\begin{array}{cccc} x & y & z & w \end{array}\right]$ is in $U^\perp$ if and only if

$$x + y = \left[\begin{array}{cccc} x & y & z & w \end{array}\right] \cdot \left[\begin{array}{cccc} 1 & 1 & 0 & 0 \end{array}\right] = 0$$

Thus $y = -x$ and

$$\begin{aligned} U^\perp &= \left\{\left[\begin{array}{cccc} x & -x & z & w \end{array}\right] \mid x,\, z,\, w \text{ in } \mathbb{R}\right\} \\ &= \text{span}\left\{\left[\begin{array}{cccc} 1 & -1 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \end{array}\right], \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \end{array}\right]\right\} \end{aligned}$$

Hence $\dim U^\perp = 3$ and $U = 1$.

d. If $p(x) = a + bx + cx^2$, p is in $U^\perp$ if and only if

$$0 = \langle p,\, x\rangle = \int_0^1 (a + bx + cx^2)x\,dx = \frac{a}{2} + \frac{b}{3} + \frac{c}{4}$$

Thus $a = 2s + t$, $b = -3s$, $c = -2t$ where s and t are in $\mathbb{R}$, so $p(x) = (2s + t) - 3sx - 2tx^2$. Hence, $U^\perp = \text{span}\left\{2 - 3x,\ 1 - 2x^2\right\}$ and $\dim U^\perp = 2$, $\dim U = 1$.

f. $\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ is in U if and only if

$$\begin{aligned} 0 &= \left\langle \left[\begin{array}{cc} a & b \\ c & d \end{array}\right], \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right] \right\rangle = a + b \\ 0 &= \left\langle \left[\begin{array}{cc} a & b \\ c & d \end{array}\right], \left[\begin{array}{cc} 1 & 0 \\ 1 & 0 \end{array}\right] \right\rangle = a + c \\ 0 &= \left\langle \left[\begin{array}{cc} a & b \\ c & d \end{array}\right], \left[\begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array}\right] \right\rangle = a + c + d \end{aligned}$$

The solution $d = 0$, $b = c = -a$, so $U^\perp = \left\{\left[\begin{array}{cc} a & -a \\ -a & 0 \end{array}\right] \middle| a \text{ in } \mathbb{R}\right\} = \text{span}\left\{\left[\begin{array}{cc} 1 & -1 \\ -1 & 0 \end{array}\right]\right\}$. Thus $\dim U^\perp = 1$ and $\dim U = 3$.

7. b. Write $\mathbf{b}_1 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$, $\mathbf{b}_2 = \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right]$, and $\mathbf{b}_3 = \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right]$. Then $\{\mathbf{b}_1,\ \mathbf{b}_2,\ \mathbf{b}_3\}$ is independent but not orthogonal. The Gram-Schmidt algorithm gives

$$\begin{aligned} \mathbf{f}_1 &= \mathbf{b}_1 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \\ \mathbf{f}_2 &= \mathbf{b}_2 - \frac{\langle \mathbf{b}_2,\, \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right] - \frac{0}{2}\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] = \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right] \\ \mathbf{f}_3 &= \mathbf{b}_3 - \frac{\langle \mathbf{b}_3,\, \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{b}_3,\, \mathbf{f}_2\rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\ &= \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right] - \frac{1}{2}\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] - \frac{2}{4}\left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right] \\ &= \frac{1}{2}\left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right] \end{aligned}$$

If $E_3' = \left[\begin{array}{rr} 0 & 1 \\ -1 & 0 \end{array}\right]$ then $\{E_1,\ E_2,\ E_3'\}$ is an orthogonal basis of U. If $A = \left[\begin{array}{cc} 2 & 1 \\ 3 & 2 \end{array}\right]$ then

$$\begin{aligned} \operatorname{proj}_U A &= \tfrac{\langle A,\ E_1\rangle}{\|E_1\|^2}E_1 + \tfrac{\langle A,\ E_2\rangle}{\|E_2\|^2}E_2 + \tfrac{\langle A,\ E_3'\rangle}{\|E_3'\|^2}E_3' \\ &= \tfrac{4}{2}\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] + \tfrac{4}{4}\left[\begin{array}{rr} 1 & 1 \\ 1 & -1 \end{array}\right] + \tfrac{-2}{2}\left[\begin{array}{rr} 0 & 1 \\ -1 & 0 \end{array}\right] \\ &= \left[\begin{array}{cc} 3 & 0 \\ 2 & 1 \end{array}\right] \end{aligned}$$

is the vector in U closest to A.

8. b. We are given $U = \operatorname{span}\{1,\ 1+x^2\}$, and applying the Gram-Schmidt algorithm gives an orthogonal basis consisting of 1 and

$$(1+x^2) - \tfrac{\langle 1+x^2,\ 1\rangle}{\|1\|^2}1 = (1+x^2) - \tfrac{(1+0^2)1+(1+1^2)1+(1+2^2)1}{1+1+1} = -\tfrac{5}{3} + x^2$$

We use $U = \operatorname{span}\{1,\ 5-3x^2\}$. Then Theorem 10.2.8 asserts that the closest vector in U to x is

$$\operatorname{proj}_U x = \tfrac{\langle x,\ 1\rangle}{\|1\|^2}1 + \tfrac{\langle x,\ 5-3x^2\rangle}{\|5-3x^2\|^2}(5-3x^2) = \tfrac{3}{3} + \tfrac{-12}{78}(5-3x^2) = \tfrac{3}{13}(1+2x^2)$$

Here, for example $\langle x,\ 5-3x^2\rangle = 0(5)+1(2)+2(-7) = -12$, and the other calculations are similar.

9. b. $\{1,\ 2x-1\}$ is an orthogonal basis of U because $\langle 1,\ 2x-1\rangle = \int_0^1 (2x-1)dx = 0$. Thus

$$\begin{aligned} \operatorname{proj}_U\ (x^2+1) &= \tfrac{\langle x^2+1,\ 1\rangle}{\|1\|^2}1 + \tfrac{\langle x^2+1,\ 2x-1\rangle}{\|2x-1\|^2}(2x-1) \\ &= \tfrac{3/4}{1}1 + \tfrac{1/6}{1/3}(2x-1) \\ &= x + \tfrac{5}{6} \end{aligned}$$

Hence, $x^2+1 = (x+\frac{5}{6}) + (x^2-x+\frac{1}{6})$ is the required decomposition. Check: $x^2-x+\frac{1}{6}$ is in $U^\perp$ because

$$\begin{aligned} \langle x^2-x+\tfrac{1}{6},\ 1\rangle &= \int_0^1 \left(x^2-x+\tfrac{1}{6}\right)dx = 0 \\ \langle x^2-x+\tfrac{1}{6},\ 2x-1\rangle &= \int_0^1 \left(x^2-x+\tfrac{1}{6}\right)(2x-1)dx = 0 \end{aligned}$$

11. b. We have $\langle \mathbf{v}+\mathbf{w},\ \mathbf{v}-\mathbf{w}\rangle = \langle \mathbf{v},\ \mathbf{v}\rangle - \langle \mathbf{v},\ \mathbf{w}\rangle + \langle \mathbf{w},\ \mathbf{v}\rangle - \langle \mathbf{u},\ \mathbf{u}\rangle = \|\mathbf{v}\|^2 - \|\mathbf{w}\|^2$. But this means that $\langle \mathbf{v}+\mathbf{w},\ \mathbf{v}-\mathbf{u}\rangle = 0$ if and only if $\|\mathbf{v}\| = \|\mathbf{w}\|$. This is what we wanted.

14. b. If $\mathbf{v}$ is in $U^\perp$ then $\langle \mathbf{v},\ \mathbf{u}\rangle = 0$ for all $\mathbf{u}$ in U. In particular, $\langle \mathbf{v},\ \mathbf{u}_i\rangle = 0$ for $1 \le i \le n$, so $\mathbf{v}$ is in $\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m\}^\perp$. This shows that $U^\perp \subseteq \{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m\}^\perp$. Conversely, if $\mathbf{v}$ is in $\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m\}^\perp$ then $\langle \mathbf{v},\ \mathbf{u}_i\rangle = 0$ for each i. If $\mathbf{u}$ is in U, write $\mathbf{u} = r_1\mathbf{u}_1 + \cdots + r_m\mathbf{u}_m$, r_i in $\mathbb{R}$. Then

$$\begin{aligned} \langle \mathbf{v},\ \mathbf{u}\rangle &= \langle \mathbf{v},\ r_1\mathbf{u}_1 + \cdots + r_m\mathbf{u}_m\rangle \\ &= r_1\langle \mathbf{v},\ \mathbf{u}_1\rangle + \cdots + r_m\langle \mathbf{v},\ \mathbf{u}_m\rangle \\ &= r_1 \cdot 0 + \cdots + r_m \cdot 0 \\ &= 0 \end{aligned}$$

As $\mathbf{u}$ was arbitrary in U, this shows that $\mathbf{v}$ is in $U^\perp$; that is $\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_m\}^\perp \subseteq U^\perp$.

18. b. Write $\mathbf{e}_1 = (3,\ -2,\ 5)$ and $\mathbf{e}_2 = (-1,\ 1,\ 1)$, write $B = \{\mathbf{e}_1,\ \mathbf{e}_2\}$, and write $U = \text{span } B$. Then B is orthogonal and so is an orthogonal basis of U. Thus if $\mathbf{v} = (-5,\ 4,\ -3)$ then

$$\begin{aligned}\text{proj}_U\ \mathbf{v} &= \tfrac{\mathbf{v}\cdot\mathbf{e}_1}{\|\mathbf{e}_1\|^2}\mathbf{e}_1 + \tfrac{\mathbf{v}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 \\ &= \tfrac{-38}{38}(3,\ -2,\ 5) + \tfrac{6}{3}(-1,\ 1,\ 1) \\ &= (-5,\ 4,\ -3) \\ &= \mathbf{v}\end{aligned}$$

Thus, $\mathbf{v}$ is in U. However, if $\mathbf{v}_1 = (-1,\ 0,\ 2)$ then

$$\begin{aligned}\text{proj}_U\ \mathbf{v}_1 &= \tfrac{\mathbf{v}_1\cdot\mathbf{e}_1}{\|\mathbf{e}_2\|^2}\mathbf{e}_1 + \tfrac{\mathbf{v}\cdot\mathbf{e}_2}{\|\mathbf{e}_2\|^2}\mathbf{e}_2 \\ &= \tfrac{7}{38}(3,\ -2,\ 5) + \tfrac{3}{3}(-1,\ 1,\ 1) \\ &= \tfrac{1}{38}(-17,\ 24,\ 73)\end{aligned}$$

As $\mathbf{v}_1 \neq \text{proj}_U\ \mathbf{v}_1$, $\mathbf{v}_1$ is not in U by (a).

19. b. The plane is $U = \{\mathbf{x} \mid \mathbf{x}\cdot\mathbf{n} = \mathbf{0}\}$, so span $\left\{\mathbf{n}\times\mathbf{w},\ \mathbf{w} - \left(\frac{\mathbf{n}\cdot\mathbf{w}}{\|\mathbf{n}\|^2}\right)\mathbf{n}\right\} \subseteq U$. Since $\dim U = 2$, it suffices to show that $B = \left\{\mathbf{n}\times\mathbf{w},\ \mathbf{w} - \left(\frac{\mathbf{n}\cdot\mathbf{w}}{\|\mathbf{n}\|^2}\right)\mathbf{n}\right\}$ is independent. These two vectors are orthogonal (because $(\mathbf{n}\times\mathbf{w})\cdot\mathbf{n} = \mathbf{0} = (\mathbf{n}\times\mathbf{w})\cdot\mathbf{w}$). Hence B is orthogonal (and so independent) provided each of the vectors is nonzero. But: $\mathbf{n}\times\mathbf{w} \neq \mathbf{0}$ because $\mathbf{n}$ and $\mathbf{w}$ are not parallel, and $\mathbf{w} - \frac{\mathbf{n}\cdot\mathbf{w}}{\|\mathbf{n}\|^2}\mathbf{n}$ is nonzero because $\mathbf{w}$ and $\mathbf{n}$ are not parallel, and $\mathbf{n}\cdot(\mathbf{w} - \frac{\mathbf{n}\cdot\mathbf{w}}{\|\mathbf{n}\|^2}\mathbf{n}) = 0$.

20. b. $C_E(\mathbf{b}_i)$ is column i of P. Since $C_E(\mathbf{b}_i)\cdot C_E(\mathbf{b}_j) = \langle\mathbf{b}_i,\ \mathbf{b}_j\rangle$ by (a), the result follows.

23. b. Let V be an inner product space, and let U be a subspace of V. If $U = \text{span}\{\mathbf{f}_1,\ \dots,\ \mathbf{f}_m\}$, then $\text{proj}_U\ \mathbf{v} = \sum_{i=1}^{m} \frac{\langle\mathbf{v},\ \mathbf{f}_i\rangle}{\|\mathbf{f}_i\|^2}\mathbf{f}_i$ by Theorem 10.2.7 so $\|\text{proj}_U\ \mathbf{v}\|^2 = \sum_{i=1}^{m} \frac{\langle\mathbf{v},\ \mathbf{f}_i\rangle^2}{\|\mathbf{f}_i\|^2}$ by Pythagoras' theorem. So it suffices to show that $\|\text{proj}_U\ \mathbf{v}\|^2 \leq \|\mathbf{v}\|^2$.
Given $\mathbf{v}$ in V, write $\mathbf{v} = \mathbf{u} + \mathbf{w}$ where $\mathbf{u} = \text{proj}_U\ \mathbf{v}$ is in U and $\mathbf{w}$ is in $U^\perp$. Since $\mathbf{u}$ and $\mathbf{w}$ are orthogonal, Pythagoras' theorem (again) gives

$$\|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2 \geq \|\mathbf{u}\|^2 = \|\text{proj}_U\ \mathbf{v}\|^2$$

This is what we wanted.

10.3 Orthogonal Diagonalization

1. b. If $B = \{E_1,\ E_2,\ E_3,\ E_4\}$ where $E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, then B is an orthonormal basis for $\mathbf{M}_{22}$ and

$$T(E_1) = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} = -E_1 + E_3$$

$$T(E_2) = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} = -E_2 + E_4$$
$$T(E_3) = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} = E_1 + 2E_3$$
$$T(E_4) = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = E_2 + 2E_4$$

Hence,

$$\begin{aligned} M_B(T) &= \begin{bmatrix} C_B[T(E_1)] & C_B[T(E_2)] & C_B[T(E_3)] & C_B[T(E_4)] \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} \end{aligned}$$

As $M_B(T)$ is symmetric, T is a symmetric operator.

4. b. If T is symmetric then $\langle \mathbf{v},\ T(\mathbf{w})\rangle = \langle T(\mathbf{v}),\ \mathbf{w}\rangle$ holds for all $\mathbf{v}$ and $\mathbf{w}$ in V. Given r in $\mathbb{R}$:

$$\langle \mathbf{v},\ (rT)(\mathbf{w})\rangle = \langle \mathbf{v},\ rT(\mathbf{w})\rangle = r\langle \mathbf{v},\ T(\mathbf{w})\rangle = r\langle T(\mathbf{v}),\ \mathbf{w}\rangle = \langle rT(\mathbf{v}),\ \mathbf{w}\rangle = \langle (rT)(\mathbf{v}),\ \mathbf{w}\rangle$$

for all $\mathbf{v}$ and $\mathbf{w}$ in V. This shows that rT is symmetric.

d. Given $\mathbf{v}$ and $\mathbf{w}$, write $T^{-1}(\mathbf{v}) = \mathbf{v}_1$ and $T^{-1}(\mathbf{w}) = \mathbf{w}_1$. Then

$$\langle T^{-1}(\mathbf{v}),\ \mathbf{w}\rangle = \langle \mathbf{v}_1,\ T(\mathbf{w}_1)\rangle = \langle T(\mathbf{v}_1),\ \mathbf{w}_1\rangle = \langle \mathbf{v},\ T^{-1}(\mathbf{w})\rangle$$

This shows that T^{-1} is a symmetric operator.

5. b. If $E = \{\mathbf{e}_1 = (1,\ 0,\ 0),\ \mathbf{e}_2 = (0,\ 1,\ 0),\ \mathbf{e}_3 = (0,\ 0,\ 1)\}$ is the standard basis of $\mathbb{R}^3$:

$$\begin{aligned} M_E(T) &= \begin{bmatrix} C_E[T(\mathbf{e}_1)] & C_E[T(\mathbf{e}_2)] & C_E[T(\mathbf{e}_3)] \end{bmatrix} \\ &= \begin{bmatrix} C_E(7,\ -1,\ 0) & C_E(-1,\ 7,\ 0) & C_E(0,\ 0,\ 2) \end{bmatrix} \\ &= \begin{bmatrix} 7 & -1 & 0 \\ -1 & 7 & 0 \\ 0 & 0 & 2 \end{bmatrix} \end{aligned}$$

Thus, $c_T(x) = \begin{vmatrix} x-7 & 1 & 0 \\ 1 & x-7 & 0 \\ 0 & 0 & x-2 \end{vmatrix} = (x-6)(x-8)(x-2)$ so the eigenvalues are $\lambda_1 = 6$, $\lambda_2 = 8$, and $\lambda_3 = 2$, (real as $M_{B_0}(T)$ is symmetric). Corresponding (orthogonal) eigenvectors are $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, so

$$\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},\ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix},\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is an orthonormal basis of eigenvectors of $M_E(T)$. These vectors are equal to $C_E\left[\frac{1}{\sqrt{2}}(1,\ 1,\ 0)\right]$, $C_E\left[\frac{1}{\sqrt{2}}(1,\ -1,\ 0)\right]$, and $C_E[(0,\ 0,\ 1)]$ respectively, so

$$\left\{ \frac{1}{\sqrt{2}}(1,\ 1,\ 0),\ \frac{1}{\sqrt{2}}(1,\ -1,\ 0),\ (0,\ 0,\ 1) \right\}$$

is an orthonormal basis of eigenvectors of T.

d. If $B_0 = \{1,\ x,\ x^2\}$ then

$$\begin{aligned} M_{B_0}(T) &= \left[\ C_{B_0}[T(1)] \quad C_{B_0}[T(x)] \quad C_{B_0}[T(x^2)]\ \right] \\ &= \left[\ C_{B_0}(-1+x^2) \quad C_{B_0}(3x) \quad C_{B_0}(1-x^2)\ \right] \\ &= \begin{bmatrix} -1 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix} \end{aligned}$$

Hence, $c_T(x) = \begin{vmatrix} x+1 & 0 & -1 \\ 0 & x-3 & 0 \\ -1 & 0 & x+1 \end{vmatrix} = x(x-3)(x+2)$ so the (real) eigenvalues are $\lambda_1 = 3$, $\lambda_2 = 0$, $\lambda_3 = -2$. Corresponding (orthogonal) eigenvectors are $\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, so $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$ is an orthonormal basis of eigenvectors of $M_{B_0}(T)$. These have the form $C_{B_0}(x)$, $C_{B_0}\left[\frac{1}{\sqrt{2}}(1+x^2)\right]$, and $C_{B_0}\left[\frac{1}{\sqrt{2}}(1-x^2)\right]$, respectively, so

$$\left\{x,\ \tfrac{1}{\sqrt{2}}(1+x^2),\ \tfrac{1}{\sqrt{2}}(1-x^2)\right\}$$

is an orthonormal basis of eigenvectors of T.

7. b. Write $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and compute:

$$\begin{aligned} M_B(T) &= \left[\ C_B\left(T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) \quad C_B\left(T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) \quad C_B\left(T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) \quad C_B\left(T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)\ \right] \\ &= \left[\ C_B\begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \quad C_B\begin{bmatrix} b & 0 \\ d & 0 \end{bmatrix} \quad C_B\begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \quad C_B\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}\ \right] \\ &= \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix} \\ &= \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \end{aligned}$$

Hence,

$$\begin{aligned} c_T(x) &= \det[xI - M_B(T)] = \det\left\{\begin{bmatrix} xI & 0 \\ 0 & xI \end{bmatrix} - \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}\right\} \\ &= \det\begin{bmatrix} xI-A & 0 \\ 0 & xI-A \end{bmatrix} = \det(xI-A)\cdot\det(xI-A) = [c_A(x)]^2 \end{aligned}$$

12. (2) We prove that (1) $\Rightarrow$ (2). If $B = \{\mathbf{f}_1,\ \dots,\ \mathbf{f}_n\}$ is an orthonormal basis of V, then $M_B(T) = [a_{ij}]$ where $a_{ij} = \langle \mathbf{f}_i,\ T(\mathbf{f}_j)\rangle$ by Theorem 10.3.2. If (1) holds then $a_{ji} = \langle \mathbf{f}_j,\ T(\mathbf{f}_i)\rangle = -\langle T(\mathbf{f}_j),\ \mathbf{f}_i\rangle = -\langle \mathbf{f}_i,\ T(\mathbf{f}_j)\rangle = -a_{ij}$. Hence $[M_V(T)]^T = -M_V(T)$, proving (2).

14. c. We have

$$M_B(T') = \left[\ C_B[T'(\mathbf{f}_1)] \quad C_B[T'(\mathbf{f}_2)] \quad \cdots \quad C_B[T'(\mathbf{f}_n)]\ \right]$$

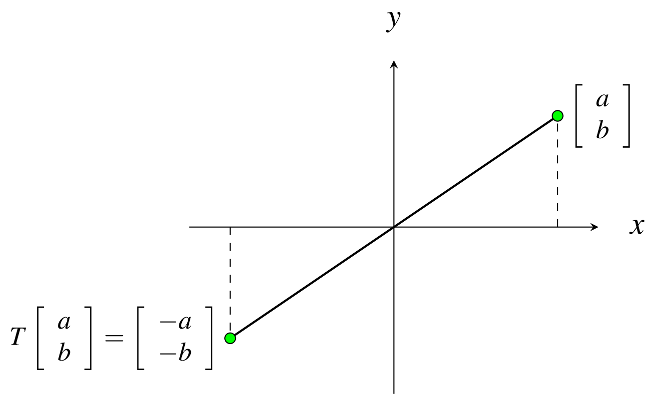
Hence, column j of $M_B(T')$ is

$$C_B(T'(\mathbf{f}_j)) = \begin{bmatrix} \langle \mathbf{f}_j,\ T(\mathbf{f}_1)\rangle \\ \langle \mathbf{f}_j,\ T(\mathbf{f}_2)\rangle \\ \vdots \\ \langle \mathbf{f}_j,\ T(\mathbf{f}_n)\rangle \end{bmatrix}$$

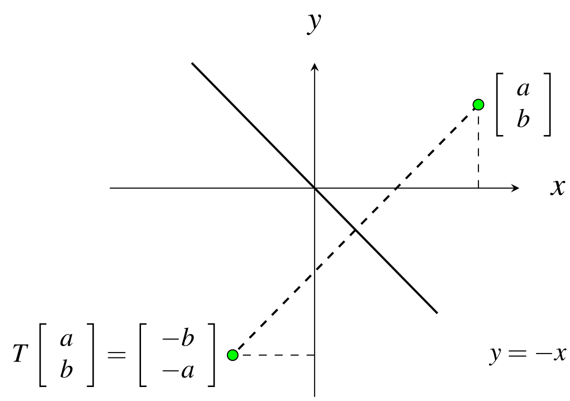
by the definition of T'. Hence the $(i,\ j)$-entry of $M_B(T')$ is $\langle \mathbf{f}_j,\ T(\mathbf{f}_i)\rangle$. But this is the $(j,\ i)$-entry of $M_B(T)$ by Theorem 10.3.2. Thus, $M_B(T')$ is the transpose of $M_B(T)$.

10.4 Isometries

2. b. We have $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -a \\ -b \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$ so T has matrix $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, which is orthogonal. Hence T is an isometry, and $\det T = 1$ so T is a rotation by Theorem 10.4.4. In fact, T is counterclockwise rotation through π. (Rotation through θ has matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ by Theorem 2.6.4; see also the discussion following Theorem 10.4.3). This can also be seen directly from the diagram.



d. We have $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ -a \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$ so T has matrix $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$. This is orthogonal, so T is an isometry. Moreover, $\det T = -1$ so T is a reflection by Theorem 10.4.4. In fact, T is reflection in the line $y = -x$ by Theorem 2.6.5. This can also be seen directly from the diagram.



f. If B_0 is the standard basis of $\mathbb{R}^2$, then

$$\begin{aligned} M_{B_0}(T) &= \begin{bmatrix} C_{B_0}\left(T\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) & C_{B_0}\left(T\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \end{bmatrix} \\ &= \begin{bmatrix} C_{B_0}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) & C_{B_0}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) \end{bmatrix} \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Hence, $\det T = 1$ so T is a rotation. Indeed, (the discussion following) Theorem 10.4.3 shows that T is a rotation through an angle θ where $\cos\theta = \frac{1}{\sqrt{2}}$, $\sin\theta = \frac{1}{\sqrt{2}}$; that is $\theta = \frac{\pi}{4}$.

3. b. $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{3}c - a \\ \sqrt{3}a + c \\ 2b \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 & 0 & \sqrt{3} \\ \sqrt{3} & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$, so T has matrix $\frac{1}{2}\begin{bmatrix} -1 & 0 & \sqrt{3} \\ \sqrt{3} & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$. Thus,

$$c_T(x) = \begin{vmatrix} x+\frac{1}{2} & 0 & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & x & -\frac{1}{2} \\ 0 & -1 & x \end{vmatrix} = \begin{vmatrix} x+\frac{1}{2} & 0 & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & 0 & x^2-\frac{1}{2} \\ 0 & -1 & x \end{vmatrix} = \begin{vmatrix} x+\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & x^2-\frac{1}{2} \end{vmatrix} = (x-1)\left(x^2+\tfrac{3}{2}x+1\right)$$

Hence, we are in (1) of Table 10.1 so T is a rotation about the line $\mathbb{R}\mathbf{e}$ with direction vector $\mathbf{e} = \begin{bmatrix} 1 \\ \sqrt{3} \\ \sqrt{3} \end{bmatrix}$, where $\mathbf{e}$ is an eigenvector corresponding to the eigenvalue 1.

d. $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ -b \\ -c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$, so T has matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. This is orthogonal, so T is an isometry. Since $c_T(x) = (x-1)(x+1)^2$, we are in case (4) of Table 10.1. Then $\mathbf{e} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to 1, so T is a rotation of π about the line $\mathbb{R}\mathbf{e}$ with direction vector $\mathbf{e}$, that is the x-axis.

f. $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} a+c \\ -\sqrt{2}b \\ c-a \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 0 & -\sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$, so T has matrix $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 0 & -\sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}$. Hence,

$$c_T(x) = \begin{vmatrix} x-\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & x+1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & x-\frac{1}{\sqrt{2}} \end{vmatrix} = (x+1)\begin{vmatrix} x-\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & x-\frac{1}{\sqrt{2}} \end{vmatrix} = (x+1)(x^2-\sqrt{2}x+1)$$

Thus we are in case (2) of Table 10.1. Now $\mathbf{e} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue -1, so T is rotation (of $\frac{3\pi}{4}$) about the line $\mathbb{R}\mathbf{e}$ (the y-axis) followed by a reflection in the plane $(\mathbb{R}\mathbf{e})^\perp$ — the xz-plane.

6. Let T be an arbitrary isometry, and let a be a real number. If aT is an isometry then Theorem 10.4.2 gives

$$\|\mathbf{v}\| = \|(aT)(\mathbf{v})\| = \|a(T(\mathbf{v}))\| = |a|\,\|T(\mathbf{v})\| = |a|\|\mathbf{v}\| \text{ holds for all } \mathbf{v}.$$

Thus $|a| = 1$ so, since a is real, $a = \pm 1$. Conversely, if $a = \pm 1$ then $|a| = 1$ so we have $\|(aT)(\mathbf{v})\| = |a|\,\|T(\mathbf{v})\| = 1\,\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v}$. Hence aT is an isometry by Theorem 10.4.2.

12. b. Assume that $S = S_{\mathbf{u}} \circ T$ where $\mathbf{u}$ is in V and T is an isometry of V. Since T is onto (by Theorem 10.4.2), let $\mathbf{u} = T(\mathbf{w})$ where $\mathbf{w} \in V$. Then for any $\mathbf{v} \in V$, we have

$$(T \circ S_{\mathbf{w}})(\mathbf{v}) = T(\mathbf{w}+\mathbf{v}) = T(\mathbf{w}) + T(\mathbf{v}) = S_{T(\mathbf{w})}(T(\mathbf{v})) = (S_{T(\mathbf{w})} \circ T)(\mathbf{v})$$

Since this holds for all $\mathbf{v} \in V$, it follows that $T \circ S_{\mathbf{w}} = S_{T(\mathbf{w})} \circ T$.

10.5 An Application to Fourier Approximation

The integrations involved in the computation of the Fourier coefficients are omitted in 1(b), 1(d), and 2(b).

1. b. $f_5 = \frac{\pi}{2} - \frac{4}{\pi}\left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2}\right)$

   d. $f_5 = \frac{\pi}{4} + \left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \frac{\sin 5x}{5}\right) - \frac{2}{\pi}\left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2}\right)$

2. b. $\frac{2}{\pi} - \frac{8}{\pi}\left(\frac{\cos 2x}{2^2-1} + \frac{\cos 4x}{4^2-1} + \frac{\cos 6x}{6^2-1}\right)$

4. We use the formula that $\cos(\theta \pm \phi) = \cos\theta\cos\phi \mp \sin\theta\sin\phi$, so that $2\cos\theta\cos\phi = \cos(\theta - \phi)\cos(\theta + \phi)$. Hence:

$$\begin{aligned}
\int_0^\pi \cos(kx)\cos(\ell x)dx &= \frac{1}{2}\int_0^\pi \{\cos[(k-\ell)x] + \cos[(k+\ell)x]\}\,dx \\
&= \frac{1}{2}\left[\frac{\sin[(k+\ell)x]}{k+\ell} + \frac{\sin[(k-\ell)x]}{(k-\ell)}\right]_0^\pi \\
&= 0 \text{ if } k \neq \ell.
\end{aligned}$$

11. Canonical Forms

11.1 Block Triangular Form

1. b. $c_A(x) = \begin{vmatrix} x+5 & -3 & -1 \\ 4 & x-2 & -1 \\ 4 & -3 & x \end{vmatrix} = \begin{vmatrix} x+1 & -x-1 & 0 \\ 4 & x-2 & -1 \\ 4 & -3 & x \end{vmatrix} = \begin{vmatrix} x+1 & 0 & 0 \\ 4 & x+2 & -1 \\ 4 & 1 & x \end{vmatrix} = (x+1)^3$.
Hence, $\lambda_1 = -1$ and we are in case $k = 1$ of the triangulation algorithm.

$$-I - A = \left[\begin{array}{rrr} 4 & -3 & -1 \\ 4 & -3 & -1 \\ 4 & -3 & -1 \end{array}\right] \to \left[\begin{array}{rrr} 4 & -3 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right];\ \mathbf{p}_{11} = \left[\begin{array}{r} 1 \\ 1 \\ 1 \end{array}\right],\ \mathbf{p}_{12} = \left[\begin{array}{r} 0 \\ 1 \\ -3 \end{array}\right]$$

Hence, $\{\mathbf{p}_{11},\ \mathbf{p}_{12}\}$ is a basis of null$\,(-I-A)$. We now expand this to a basis of null $\left[(-I-A)^2\right]$. However, $(-I-A)^2 = 0$ so null $\left[(-I-A)^2\right] = \mathbb{R}^3$. Hence, in this case, we expand $\{\mathbf{p}_{11},\ \mathbf{p}_{12}\}$ to any basis $\{\mathbf{p}_{11},\ \mathbf{p}_{12},\ \mathbf{p}_{13}\}$ of $\mathbb{R}^3$, say by taking $\mathbf{p}_{13} = \left[\begin{array}{r} 0 \\ 0 \\ 1 \end{array}\right]$. Hence

$$P = \left[\begin{array}{ccc} \mathbf{p}_{11} & \mathbf{p}_{12} & \mathbf{p}_{13} \end{array}\right] = \left[\begin{array}{rrr} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -3 & 1 \end{array}\right] \text{ satisfies } P^{-1}AP = \left[\begin{array}{rrr} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{array}\right]$$

as may be verified.

d. $c_A(x) = \begin{vmatrix} x+3 & 1 & 0 \\ -4 & x+1 & -3 \\ -4 & 2 & x-4 \end{vmatrix} = \begin{vmatrix} x+3 & 1 & 0 \\ -4 & x+1 & -3 \\ 0 & -x+1 & x-1 \end{vmatrix} = \begin{vmatrix} x+3 & 1 & 0 \\ -4 & x-2 & -3 \\ 0 & 0 & x-1 \end{vmatrix} = (x-1)^2(x+2)$.
Hence $\lambda_1 = 1$, $\lambda_3 = -2$, and we are in case $k = 2$ of the triangulation algorithm.

$$I - A = \left[\begin{array}{rrr} 4 & 1 & 0 \\ -4 & 2 & -3 \\ -4 & 2 & -3 \end{array}\right] \to \left[\begin{array}{rrr} 4 & 1 & 0 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{array}\right];\ \mathbf{p}_{11} = \left[\begin{array}{r} -1 \\ 4 \\ 4 \end{array}\right]$$

Thus, null$\,(I-A) =$ span $\{\mathbf{p}_{11}\}$. We enlarge $\{\mathbf{p}_{11}\}$ to a basis of null $\left[(I-A)^2\right]$

$$(I-A)^{12} = \left[\begin{array}{rrr} 12 & 6 & -3 \\ -12 & -6 & 3 \\ -12 & -6 & 3 \end{array}\right] \to \left[\begin{array}{rrr} 4 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right];\ \mathbf{p}_{11} = \left[\begin{array}{r} -1 \\ 4 \\ 4 \end{array}\right],\ \mathbf{p}_{12} = \left[\begin{array}{r} 0 \\ 1 \\ 2 \end{array}\right]$$

Thus, null $\left[(I-A)^2\right] =$ span $\{\mathbf{p}_{11},\ \mathbf{p}_{12}\}$. As dim $\left[G_{\lambda_1}(A)\right] = 2$ in this case (by Lemma 11.1.1), we have $G_{\lambda_1}(A) =$ span $\{\mathbf{p}_{11},\ \mathbf{p}_{12}\}$. However, it is instructive to continue the process:

$$(I-A)^2 = 3\left[\begin{array}{rrr} 4 & 2 & -1 \\ -4 & -2 & 1 \\ -4 & -2 & 1 \end{array}\right]$$

whence

$$(I-A)^3 = 9\left[\begin{array}{rrr} 4 & 2 & -1 \\ -4 & -2 & 1 \\ -4 & -2 & 1 \end{array}\right] = 3(I-A)^2$$

This continues to give $(I-A)^4 = 3^2(I-A)^2, \ \dots,$ and in general $(I-A)^k = 3^{k-2}(I-A)^2$ for $k \geq 2$. Thus null $\left[(I-A)^k\right]$ = null $\left[(I-A)^2\right]$ for all $k \geq 2$, so

$$G_{\lambda_1}(A) = \text{null}\left[(I-A)^2\right] = \text{span}\ \{\mathbf{p}_{11},\ \mathbf{p}_{12}\}$$

as we expected. Turning to $\lambda_2 = -2$:

$$-2I - A = \begin{bmatrix} 1 & 1 & 0 \\ -4 & -1 & -3 \\ -4 & 2 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 3 & -3 \\ 0 & 6 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix};\ \mathbf{p}_{21} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Hence, null $[-2I-A]$ = span $\{\mathbf{p}_{21}\}$. We need go no further with this as $\{\mathbf{p}_{11},\ \mathbf{p}_{12},\ \mathbf{p}_{21}\}$ is a basis of $\mathbb{R}^3$. Hence

$$P = \begin{bmatrix} \mathbf{p}_{11} & \mathbf{p}_{12} & \mathbf{p}_{21} \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 \\ 4 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \text{ satisfies } P^{-1}AP = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

as may be verified.

f. To evaluate $c_A(x)$, we begin by adding column 4 to column 1:

$$c_A(x) = \begin{vmatrix} x+3 & -6 & -3 & -2 \\ 2 & x-3 & -2 & -2 \\ 1 & -3 & x & -1 \\ 1 & -1 & -2 & x \end{vmatrix} = \begin{vmatrix} x+1 & -6 & -3 & -2 \\ 0 & x-3 & -2 & -2 \\ 0 & -3 & x & -1 \\ x+1 & -1 & -2 & x \end{vmatrix} = \begin{vmatrix} x+1 & -6 & -3 & -2 \\ 0 & x-3 & -2 & -2 \\ 0 & -3 & x & -1 \\ 0 & 5 & 1 & x+2 \end{vmatrix}$$

$$= (x+1)\begin{vmatrix} x-3 & -2 & -2 \\ -3 & x & -1 \\ 5 & 1 & x+2 \end{vmatrix} = (x+1)\begin{vmatrix} x-3 & -2 & 0 \\ -3 & x & -x-1 \\ 5 & 1 & x+1 \end{vmatrix} = (x+1)\begin{vmatrix} x-3 & -2 & 0 \\ 2 & x+1 & 0 \\ 5 & 1 & x+1 \end{vmatrix}$$

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

Hence, $\lambda_1 = -1,\ \lambda_2 = 1$ and we are in case $k = 2$ of the triangulation algorithm. We omit the details of the row reductions:

$$-I - A = \begin{bmatrix} 2 & -6 & -3 & -2 \\ 2 & -4 & -2 & -2 \\ 1 & -3 & -1 & -1 \\ 1 & -1 & -2 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};\ \mathbf{p}_{11} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$(-I - A)^2 = \begin{bmatrix} -13 & 23 & 13 & 13 \\ -8 & 12 & 8 & 8 \\ -6 & 10 & 6 & 6 \\ -3 & 5 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};\ \mathbf{p}_{11}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},\ \mathbf{p}_{12} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

We have dim $\left[G_{\lambda_1}(A)\right] = 2$ as $\lambda_1 = -1$ has multiplicity 2 in $c_A(x)$, so $G_{\lambda_1}(A)$ = span $\{\mathbf{p}_{11},\ \mathbf{p}_{12}\}$. Turning to $\lambda_2 = 1$:

$$I - A = \begin{bmatrix} 4 & -6 & -3 & -2 \\ 2 & -2 & -2 & -2 \\ 1 & -3 & 1 & -1 \\ 1 & -1 & -2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix};\ \mathbf{p}_{21} = \begin{bmatrix} 5 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

$$(I - A)^2 = \begin{bmatrix} -1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ -2 & -2 & 2 & -6 \\ 1 & 1 & -5 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix};\ \mathbf{p}_{21} = \begin{bmatrix} 5 \\ 2 \\ 2 \\ 1 \end{bmatrix},\ \mathbf{p}_{22} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

Hence, $G_{\lambda_2}(A)$ = span $\{\mathbf{p}_{21},\ \mathbf{p}_{22}\}$ using Lemma 11.1.1. Finally, then

$$P = \begin{bmatrix} \mathbf{p}_{11} & \mathbf{p}_{12} & \mathbf{p}_{21} & \mathbf{p}_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 5 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 1 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \text{ gives } P^{-1}AP = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

as may be verified.

4. Let B be any basis of V and write $A = M_B(T)$. Then $c_T(x) = c_A(x)$ and this is a polynomial: $c_T(x) = a_0 + a_1x + \cdots + a_nx^n$ for some a_i in $\mathbb{R}$. Now recall that $M_B : \mathbf{L}(V, V) \to M_{nn}$ is an isomorphism of vector spaces (Exercise 9.1.26) with the additional property that $M_B(T^k) = M_B(T)^k$ for $k \geq 1$ (Theorem 9.2.1). With this we get

$$\begin{aligned} M_B[c_T(T)] &= M_B[a_0 1_V + a_1 T + \cdots + a_n T^n] \\ &= a_0 M_B(1_V) + a_1 M_B(T) + \cdots + a_n M_B(T)^n \\ &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= c_A(A) \\ &= 0 \end{aligned}$$

by the Cayley-Hamilton theorem. Hence $c_T(T) = 0$ because M_B is one-to-one.

11.2 Jordan Canonical Form

2. $\left[\begin{array}{ccc} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{array}\right]\left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{array}\right] = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{array}\right]\left[\begin{array}{ccc} b & 0 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{array}\right]$, and $\left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{array}\right]$ is invertible.

A. Complex Numbers

1. b. $12+5i=(2+xi)(3-2i)=(6+2x)+(-4+3x)i$. Equating real and imaginary parts gives $6+2x=12$, $-4+3x=5$, so $x=3$.

   d. $5=(2+xi)(2-xi)=(4+x^2)+0i$. Hence $4+x^2=5$, so $x=\pm 1$.

2. b. $(3-2i)(1+i)+|3+4i|=(5+i)+\sqrt{9+16}=10+i$

   d. $$\begin{aligned}\tfrac{3-2i}{1-i}-\tfrac{3-7i}{2-3i} &= \tfrac{(3-2i)(1+i)}{(1-i)(1+i)}-\tfrac{(3-7i)(2+3i)}{(2-3i)(2+3i)}\\ &= \tfrac{5+i}{1+1}-\tfrac{27-5i}{4+9}\\ &= \tfrac{11}{26}+\tfrac{23}{26}i\end{aligned}$$

   f. $(2-i)^3=(2-i)^2(2-i)=(3-4i)(2-i)=2-11i$

   h. $(1-i)^2(2+i)^2=(-2i)(3+4i)=8-6i$

3. b. $iz+1=i+z-6i+3iz=-5i+(1+3i)z$. Hence $1+5i=(1+2i)z$, so

$$z=\tfrac{1+5i}{1+2i}=\tfrac{(1+5i)(1-2i)}{(1+2i)(1-2i)}=\tfrac{11+3i}{1+4}=\tfrac{11}{5}+\tfrac{3}{5}i$$

   d. $z^2=3-4i$. If $z=a+bi$ the condition is $(a^2-b^2)+(2ab)i=3-4i$, whence $a^2-b^2=3$ and $ab=-2$. Thus $b=\frac{-2}{a}$, so $a^2-\frac{4}{a^2}=3$. Hence $a^4-3a^2-4=0$. This factors as $(a^2-4)(a^2+1)=0$, so $a=\pm 2$, whence $b=\mp 1$. Finally, $z=a+bi=\pm(2-i)$.

   f. Write $z=a+bi$. Then the condition reads

$$\begin{aligned}(a+bi)(2-i)&=(a-bi+1)(1+i)\\ (2a+b)+(2b-a)i&=(a+1+b)+(a+1-b)i\end{aligned}$$

   Thus $2a+b=a+1+b$ and $2b-a=a+1-b$; whence $a=1$, $b=1$, so $z=1+i$.

4. b. $x=\frac{1}{2}\left[-(-1)\pm\sqrt{(-1)^2-4}\right]=\frac{1}{2}\left[1\pm i\sqrt{3}\right]$

   d. $x=\frac{1}{4}\left[-(-5)\pm\sqrt{(-5)^2-4\cdot 2\cdot 2}\right]=\frac{1}{4}\left[5\pm\sqrt{9}\right]=2,\ \frac{1}{2}$

5. b. If $x=re^{i\theta}$ then $x^3=-8$ becomes $r^3e^{3i\theta}=8e^{\pi i}$. Thus $r^3=8$ (whence $r=2$) and $3\theta=\pi+2k\pi$. Hence $\theta=\frac{\pi}{3}+k\cdot\frac{2\pi}{3}$, $k=0,\ 1,\ 2$. The roots are

$$\begin{aligned}2e^{i\pi/3}&=1+\sqrt{3}i && (k=0)\\ 2e^{\pi i}&=-2 && (k=1)\\ 2e^{5\pi i/3}&=1-\sqrt{3}i && (k=2)\end{aligned}$$

d. If $x = re^{-i\theta}$ then $x^4 = 64$ becomes $r^4e^{4i\theta} = 64e^{i\cdot 0}$. Hence $r^4 = 64$ (whence $r = 2\sqrt{2}$) and $4\theta = 0 + 2k\pi$; $\theta = k\frac{\pi}{2}$, $k = 0,\ 1,\ 2,\ 3$. The roots are

$$\begin{aligned} 2\sqrt{2}e^{0i} &= 2\sqrt{2} && (k=0)\\ 2\sqrt{2}e^{\pi i/2} &= 2\sqrt{2}i && (k=1)\\ 2\sqrt{2}e^{\pi i} &= -2\sqrt{2} && (k=2)\\ 2\sqrt{2}e^{3\pi i/2} &= -2\sqrt{2}i && (k=3) \end{aligned}$$

6. b. The quadratic is $(x-u)(x-\overline{u}) = x^2 - (u+\overline{u})x + u\overline{u} = x^2 - 4x + 13$. The other root is $\overline{u} = 2+3i$.

   d. The quadratic is $(x-u)(x-\overline{u}) = x^2 - (u+\overline{u})x + u\overline{u} = x^2 - 6x + 25$. The other root is $\overline{u} = 3+4i$.

8. If $u = 2-i$, then u is a root of $(x-u)(x-\overline{u}) = x^2 - (u+\overline{u})x + u\overline{u} = x^2 - 4x + 5$.
   If $v = 3-2i$, then v is a root of $(x-v)(x-\overline{v}) = x^2 - (v+\overline{v})x + v\overline{v} = x^2 - 6x + 13$.
   Hence u and v are roots of

$$(x^2-4x+5)(x^2-6x+13) = x^4 - 10x^3 + 42x^2 - 82x + 65$$

10. b. Taking $x = u = -2$: $x^2 + ix - (4-2i) = 4 - 2i - 4 + 2i = 0$. If v is the other root then $u + v = -i$ (i is the coefficient of x) so $v = -u - i = 2 - i$.

    d. Taking $x = u = -2+i$: $\begin{aligned}(-2+i)^2 &= 3(1-i)(-2+i) - 5i\\ &= (3-ri) + 3(-1+3i) - 5i\\ &= 0.\end{aligned}$

    If v is the other root then $u + v = -3(1-i)$, so $v = -3(1-i) - u = -1 + 2i$.

11. b. $x^2 - x + (1-i) = 0$ gives $x = \frac{1}{2}\left[1 \pm \sqrt{1-4(1-i)}\right] = \frac{1}{2}\left[1 \pm \sqrt{-3+4i}\right]$. Write $w = \sqrt{-3+4i}$ so $w^2 = -3+4i$. If $w = a+bi$ then $w^2 = (a^2-b^2) + (2ab)i$, so $a^2 - b^2 = -3$, $2ab = 4$. Thus $b = \frac{2}{a}$, $a^2 - \frac{4}{a^2} = -3$, $a^4 + 3a^2 - 4 = 0$, $(a^2+4)(a^2-1) = 0$, $a = \pm 1$, $b = \pm 2$, $w = \pm(1+2i)$. Finally the roots are $\frac{1}{2}[1 \pm w] = 1+i,\ -i$.

    d. $x^2 - 3(1-i)x - 5i = 0$ gives $x = \frac{1}{2}\left[3(1-i) \pm \sqrt{9(1-i)^2 + 20i}\right] = \frac{1}{2}\left[3(1-i) \pm \sqrt{2i}\right]$. If $w = \sqrt{2i}$ then $w^2 = 2i$. Write $w = a+bi$ so $(a^2-b^2) + 2abi = 2i$. Hence $a^2 = b^2$ and $ab = 1$; the solution is $a = b = \pm 1$ so $w = \pm(1+i)$. Thus the roots are $x = \frac{1}{2}(3(1-i) \pm w) = 2-i,\ 1-2i$.

12. b. $|z-1| = 2$ means that the distance from z to 1 is 2. Thus the graph is the circle, radius 2, center at 1.

    d. If $z = x+yi$, then $z = -\overline{z}$ becomes $x + yi = -x + yi$. This holds if and only if $x = 0$; that is if and only if $z = yi$. Hence the graph is the imaginary axis.

    f. If $z = x+yi$, then $\operatorname{im} z = m \cdot \operatorname{re} z$ becomes $y = mx$. This is the line through the origin with slope m.

18. b. $-4i = 4e^{3\pi i/2}$

    d. $\left|-4+4\sqrt{3}i\right| = 4\sqrt{1+3} = 8$ and $\cos\varphi = \frac{4}{8} = \frac{1}{2}$. Thus $\varphi = \frac{\pi}{3}$, so $\theta = \frac{2\pi}{3}$ and we have $-4 + 4\sqrt{3}i = 8e^{2\pi i/3}$.

f. $|-6+6i| = 6\sqrt{1+1} = 6\sqrt{2}$ and $\cos\varphi = \frac{6}{6\sqrt{2}} = \frac{1}{\sqrt{2}}$. Thus $\varphi = \frac{\pi}{4}$ so $\theta = \frac{3\pi}{4}$; whence $-6+6i = 6\sqrt{2}e^{-3\pi i/4}$.

19. b. $e^{7\pi i/3} = e^{(\pi/3+2\pi)i} = e^{\pi i/3} = \cos\frac{\pi}{3} + i\sin\frac{\pi}{3} = \frac{1}{2} + \frac{\sqrt{3}}{2}i$

d. $\sqrt{2}e^{-\pi i/4} = \sqrt{2}\left(\cos\left(\frac{-\pi}{4}\right) + i\sin\left(\frac{-\pi}{4}\right)\right) = \sqrt{2}\left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i\right) = 1-i$

f. $2\sqrt{3}e^{-2\pi i/6} = 2\sqrt{3}\left(\cos\left(\frac{-\pi}{3}\right) + i\sin\left(\frac{-\pi}{3}\right)\right) = 2\sqrt{3}\left(\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) = \sqrt{3} - 3i$

20. b.
$$\begin{aligned}(1+\sqrt{3}i)^{-4} &= (2e^{\pi i/3})^{-4} = 2^{-4}e^{-4\pi i/3}\\ &= \tfrac{1}{16}\left[\cos(-4\pi/3) + i\sin(-4\pi/3)\right]\\ &= \tfrac{1}{16}\left(-\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\right)\\ &= -\tfrac{1}{32} + \tfrac{\sqrt{3}}{32}i\end{aligned}$$

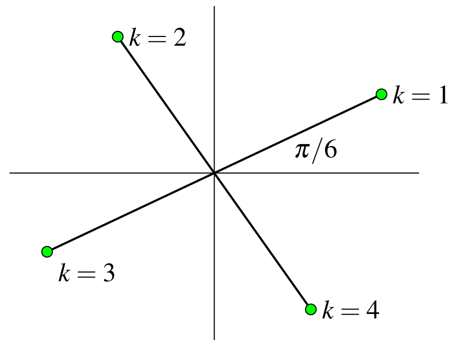
d.
$$\begin{aligned}(1-i)^{10} &= \left[\sqrt{2}e^{-\pi i/4}\right]^{10} = (\sqrt{2})^{10}e^{-5\pi i/2} = (\sqrt{2})^{10}e^{(-\pi/2-2\pi)i}\\ &= (\sqrt{2})^{10}e^{-\pi i/2} = 2^5\left[\cos\left(\tfrac{-\pi}{2}\right) + i\sin\left(\tfrac{-\pi}{2}\right)\right]\\ &= 32(0-i) = -32i\end{aligned}$$

f.
$$\begin{aligned}(\sqrt{3}-i)^9(2-2i)^5 &= \left[2e^{-\pi i/6}]^9\left[2\sqrt{2}e^{-\pi i/4}\right]^5\right]\\ &= 2^9e^{-3\pi i/2}(2\sqrt{2})^5e^{-5\pi i/4}\\ &= 2^9(i)2^5(\sqrt{2})^4\sqrt{2}\left(-\tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}}i\right)\\ &= 2^{16}i(-1+i)\\ &= -2^{16}(1+i)\end{aligned}$$

23. b. Write $z = re^{i\theta}$. Then $z^4 = 2(\sqrt{3}i-1)$ becomes $r^4e^{4i\theta} = 4e^{2\pi i/3}$. Hence $r^4 = 4$, so $r = \sqrt{2}$, and $4\theta = \frac{2\pi}{3} + 2\pi k$; that is
$$\theta = \frac{\pi}{6} + \frac{\pi}{2}k \qquad k = 0,\ 1,\ 2,\ 3$$

The roots are

$$\begin{aligned}\sqrt{2}e^{\pi i/6} &= \sqrt{2}\left(\tfrac{\sqrt{3}}{2} + \tfrac{1}{2}i\right) = \tfrac{\sqrt{2}}{2}\left(\sqrt{3}+i\right)\\ \sqrt{2}e^{4\pi i/6} &= \sqrt{2}\left(-\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\right) = \tfrac{\sqrt{2}}{2}\left(-1+\sqrt{3}i\right)\\ \sqrt{2}e^{7\pi i/6} &= \sqrt{2}\left(-\tfrac{\sqrt{3}}{2} - \tfrac{1}{2}i\right) = -\tfrac{\sqrt{2}}{2}\left(\sqrt{3}+i\right)\\ \sqrt{2}e^{10\pi i/6} &= \sqrt{2}\left(\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i\right) = -\tfrac{\sqrt{2}}{2}\left(-1+\sqrt{3}i\right)\end{aligned}$$

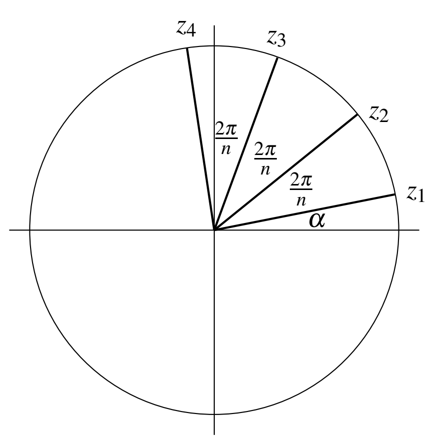


d. Write $z = re^{i\theta}$. Then $z^6 = -64$ becomes $r^6 e^{6i\theta} = 64e^{\pi i}$. Hence $r^6 = 64$, so $r = 2$, and $6\theta = \pi + 2\pi k$; that is $\theta = \frac{\pi}{6} + \frac{\pi}{3}k$ where $k = 0, 1, 2, 3, 4, 5$. The roots are thus $z = 2e^{\pi/6+\pi/3k}$ for these values of k. In cartesian form they are

k	0	1	2	3	4	5
z	$\sqrt{3}+i$	$2i$	$-\sqrt{3}+i$	$-\sqrt{3}-i$	$-2i$	$\sqrt{3}-i$

26. b. Each point on the unit circle has polar form $e^{i\theta}$ for some angle θ. As the n points are equally spaced, the angle between consecutive points is $\frac{2\pi}{n}$. Suppose the first point into the first quadrant is $z_0 = e^{\alpha i}$. Write $w = e^{2\pi i/n}$. If the points are labeled $z_1, z_2, z_3, \dots, z_n$ around the unit circle, they have polar form

$$\begin{aligned}
z_1 &= e^{\alpha i} \\
z_2 &= e^{(\alpha+2\pi/n)i} = e^{\alpha i}e^{2\pi i/n} = z_1 w \\
z_3 &= e^{[\alpha+2(2\pi/n)]i} = e^{\alpha i}e^{4\pi i/n} = z_1 w^2 \\
z_4 &= e^{[\alpha+3(2\pi/n)]i} = e^{\alpha i}e^{6\pi i/n} = z_1 w^3 \\
&\vdots \\
z_n &= e^{[\alpha+(n-1)(2\pi/n)]i} = e^{ai}e^{2(n-1)\pi i/n} = z_1 w^{n-1}
\end{aligned}$$



Hence the sum of the roots is

$$z_1 + z_2 + \cdots + z_n = z_1(1 + w + \cdots + w^{n-1}) \tag{$*$}$$

Now $w^n = \left(e^{2\pi i/n}\right)^n = e^{2\pi i} = 1$ so

$$0 = 1 - w^n = (1 - w)(1 + w + w^2 + \cdots + w^{n-1})$$

As $w \neq 1$, this gives $1 + w + \cdots + w^{n-1} = 0$. Hence $(*)$ gives

$$z_1 + z_2 + \cdots + z_n = z_1 \cdot 0 = 0$$

B. Proofs

1. b. (1). We are to prove that if the statement "*m is even and n is odd*" is true then the statement "*$m+n$ is odd*" is also true.
   If m is even and n is odd, they have the form $m = 2p$ and $n = 2q+1$, where p and q are integers. But then $m+n = 2(p+q)+1$ is odd, as required.

   (2). The converse is false. It states that if $m+n$ is odd then m is even and n is odd; and a counterexample is $m = 1$, $n = 2$.

   d. (1). We are to prove that if the statement "$x^2-5x+6=0$" is true then the statement "$x = 2$ *or* $x = 3$" is also true.
   Observe first that $x^2-5x+6 = (x-2)(x-3)$. So if x is a number satisfying $x^2-5x+6=0$ then $(x-2)(x-3)-0$ so either $x=2$ or $x=3$. [Note that we are using an important fact about real numbers: If the product of two real numbers is zero then one of them is zero.]

   (2). The converse is true. It states that if $x = 2$ or $x = 3$ then x satisfies the equation $x^2-5x+6 = 0$. This is indeed the case as both $x = 2$ or $x = 3$ satisfy this equation.

2. b. The implication here is $p \Rightarrow q$ where p is the statement "*n is any odd integer*", and q is the statement "*$n^2 = 8k+1$ for some integer k*". We are asked to either prove this implication or give a counterexample.

   This implication is true. If p is true then n is odd, say $n = 2t+1$ for some integer t. Then $n^2 = (2t)^2+2(2t)+1 = 4t(t+1)+1$. But $t(t+1)$ is even (because t is either even or odd), say $t(t+1) = 2k$ where k is an integer. Hence $n^2 = 4t(t+1)+1 = 4(2k)+1$, as required.

3. b. The implication here is $p \Rightarrow q$ where p is the statement "*$n+m = 25$, where n and m are integers*", and q is the statement "*one of m and n is greater than* 12" is also true. We are asked to either prove this implication by the method of contradiction, or give a counterexample.

   The implication is true. To prove it by contradiction, we assume that the conclusion q is false, and look for a contradiction. In this case assuming that q is false means both $n \leq 12$ and $m \leq 12$. But then $n+m \leq 24$, contradicting the hypothesis that $n+m = 25$. So the statement is true by the method of proof by contradiction.

   The converse is false. It states that $q \Rightarrow p$, that is if one of m and n is greater than 12 then $n+m = 25$. But $n = 13$ and $m = 13$ is a counterexample.

   d. The implication here is $p \Rightarrow q$ where p is the statement "*mn is even, where n and m are integers*", and q is the statement "*m is even or n is even*". We are asked to either prove this implication by the method of contradiction, or give a counterexample.

   This implication is true. To prove it by contradiction, we assume that the conclusion q is false, and look for a contradiction. In this case assuming that q is false means that m and n are both odd. But then mn is odd (if either were even the product would be even). This contradicts the hypothesis, so the statement is true by the method of proof by contradiction.

   The converse is true. It states that if m or n is even then mn is even, and this is true (if m or n is a multiple of 2, then mn is a multiple of 2).

4. b. The implication here is: "*x is irrational and y is rational*" $\Rightarrow$ "$x+y$ *is irrational*".

   To argue by contradiction, assume that $x+y$ is rational. Then $x=(x+y)-y$ is the difference of two rational numbers, and so is rational, contrary to the hypothesis that x is irrational.

5. b. At first glance the statement does not appear to be an implication. But another way to say it is that if the statement "$n \geq 2$" is true then the statement "$n^3 \geq 2^n$" is also true.

   This is not true. In fact, $n=10$ is a counterexample because $10^3=1000$ while $2^{10}=1024$. It is worth noting that the statement $n^3 \geq 2^n$ *does* hold for $2 \leq n < 9$.

C. Mathematical Induction

6. Write S_n for the statement

$$\tfrac{1}{1\cdot 2} + \tfrac{1}{2\cdot 3} + \cdots + \tfrac{1}{n(n+1)} = \tfrac{n}{n+1} \qquad (S_n)$$

Then S_1 is true: It reads $\frac{1}{1\cdot 2} = \frac{1}{1+1}$, which is true. Now <u>assume</u> S_n is true for some $n \geq 1$. We must use S_n to show that S_{n+1} is also true. The statement S_{n+1} reads as follows:

$$\tfrac{1}{1\cdot 2} + \tfrac{1}{2\cdot 3} + \cdots + \tfrac{1}{(n+1)(n+2)} = \tfrac{n+1}{n+2}$$

The second last term on the left side is $\frac{1}{n(n+1)}$ so we can use S_n:

$$\begin{aligned}
\tfrac{1}{1\cdot 2} + \tfrac{1}{2\cdot 3} + \cdots + \tfrac{1}{(n+1)(n+2)} &= \left[\tfrac{1}{1\cdot 2} + \tfrac{1}{2\cdot 3} + \cdots + \tfrac{1}{n(n+1)}\right] + \tfrac{1}{(n+1)(n+2)} \\
&= \tfrac{n}{n+1} + \tfrac{1}{(n+1)(n+2)} \\
&= \tfrac{n(n+2)+1}{(n+1)(n+2)} \\
&= \tfrac{(n+1)^2}{(n+1)(b+2)} \\
&= \tfrac{n+1}{n+2}
\end{aligned}$$

Thus S_{n+1} is true and the induction is complete.

14. Write S_n for the statement

$$\tfrac{1}{\sqrt{1}} + \tfrac{1}{\sqrt{2}} + \cdots + \tfrac{1}{\sqrt{n}} \leq 2\sqrt{n} - 1 \qquad (S_n)$$

Then S_1 is true as it asserts that $\frac{1}{\sqrt{1}} \leq 2\sqrt{1} - 1$, which is true. Now assume that S_n is true for some $n \geq 1$. We must use S_n to show that S_{n+1} is also true. The statement S_{n+1} reads as follows:

$$\begin{aligned}
\tfrac{1}{\sqrt{1}} + \tfrac{1}{\sqrt{2}} + \cdots + \tfrac{1}{\sqrt{n+1}} &= \left[\tfrac{1}{\sqrt{1}} + \tfrac{1}{\sqrt{2}} + \cdots + \tfrac{1}{\sqrt{n}}\right] + \tfrac{1}{\sqrt{n+1}} \\
&\leq \left[2\sqrt{n} - 1\right] + \tfrac{1}{\sqrt{n+1}} \\
&= \tfrac{2\sqrt{n^2+n}+1}{\sqrt{n+1}} - 1 \\
&< \tfrac{2(n+1)}{\sqrt{n+1}} - 1 \\
&= 2\sqrt{n+1} - 1
\end{aligned}$$

where, at the second last step, we used the fact that $\sqrt{n^2+n} < (n+1)$—this follows by showing that $n^2+n < (n+1)^2$, and taking positive square roots. Thus S_{n+1} is true and the induction is complete.

18. Let S_n stand for the statement

$$n^3 - n \text{ is a multiple of } 3.$$

Clearly S_1 is true. If S_n is true, then $n^3 - n = 3k$ for some integer k. Compute:

$$\begin{aligned}(n+1)^3 - (n+1) &= (n^3 + 3n^2 + 3n + 1) - (n+1) \\ &= 3k + 3n^2 + 3n\end{aligned}$$

which is clearly a multiple of 3. Hence S_{n+1} is true, and so S_n is true for every n by induction.

20. Look at the first few values: $B_1 = 1,\ B_2 = 5,\ B_3 = 23,\ B_4 = 119,\ \dots$. If these are compared to the factorials: $1! = 1,\ 2! = 4,\ 3! = 6,\ 4! = 24,\ 5! = 120,\ \dots$, it is clear that $B_n = (n+1)! - 1$ holds for $n = 1,\ 2,\ 3,\ 4$ and 5. So it seems a reasonable conjecture that

$$B_n = (n+1)! - 1 \text{ for } n \geq 1. \qquad (S_n)$$

This certainly holds for $n = 1$: $B_1 = 1 = 2! - 1$. If this is true for some $n \geq 1$, then

$$\begin{aligned}B_{n+1} &= [1 \cdot 1! + 2 \cdot 2! + \cdots + n \cdot n!] + (n+1)(n+1)! \\ &= [(n+1)! - 1] + (n+1)(n+1)! \\ &= (n+1)![1 + (n+1)] - 1 \\ &= (n+1)![n+2] - 1 \\ &= (n+2)! - 1\end{aligned}$$

Hence S_{n+1} is true and so the induction goes through.

Note that many times mathematical theorems are discovered by "experiment", somewhat as in this example. Several examples are worked out, a pattern is observed and formulated, and the result is proved (often by induction).

22. b. If we know that $S_n \Rightarrow S_{n+8}$ then it is enough to verify that S_1, S_2, S_3, S_4, S_5, S_6, S_7, and S_8 are all true. Then

$$\begin{array}{ccccccccc}
S_1 & \Rightarrow & S_9 & \Rightarrow & S_{17} & \Rightarrow & S_{25} & \Rightarrow & \cdots \\
S_2 & \Rightarrow & S_{10} & \Rightarrow & S_{18} & \Rightarrow & S_{26} & \Rightarrow & \cdots \\
S_3 & \Rightarrow & S_{11} & \Rightarrow & S_{19} & \Rightarrow & S_{27} & \Rightarrow & \cdots \\
S_4 & \Rightarrow & S_{12} & \Rightarrow & S_{20} & \Rightarrow & S_{28} & \Rightarrow & \cdots \\
S_5 & \Rightarrow & S_{13} & \Rightarrow & S_{21} & \Rightarrow & S_{29} & \Rightarrow & \cdots \\
S_6 & \Rightarrow & S_{14} & \Rightarrow & S_{22} & \Rightarrow & S_{30} & \Rightarrow & \cdots \\
S_7 & \Rightarrow & S_{15} & \Rightarrow & S_{23} & \Rightarrow & S_{31} & \Rightarrow & \cdots \\
S_8 & \Rightarrow & S_{16} & \Rightarrow & S_{24} & \Rightarrow & S_{32} & \Rightarrow & \cdots
\end{array}$$

Clearly each S_n will appear in this array, and so will be true.

ADAPTED
OPEN TEXT

FORMATIVE ONLINE
ASSESSMENT

COURSE
SUPPLEMENTS

COURSE LOGISTICS
& SUPPORT

lyryx

advancing learning

LYRYX.COM

CHAMPIONS OF ACCESS TO KNOWLEDGE