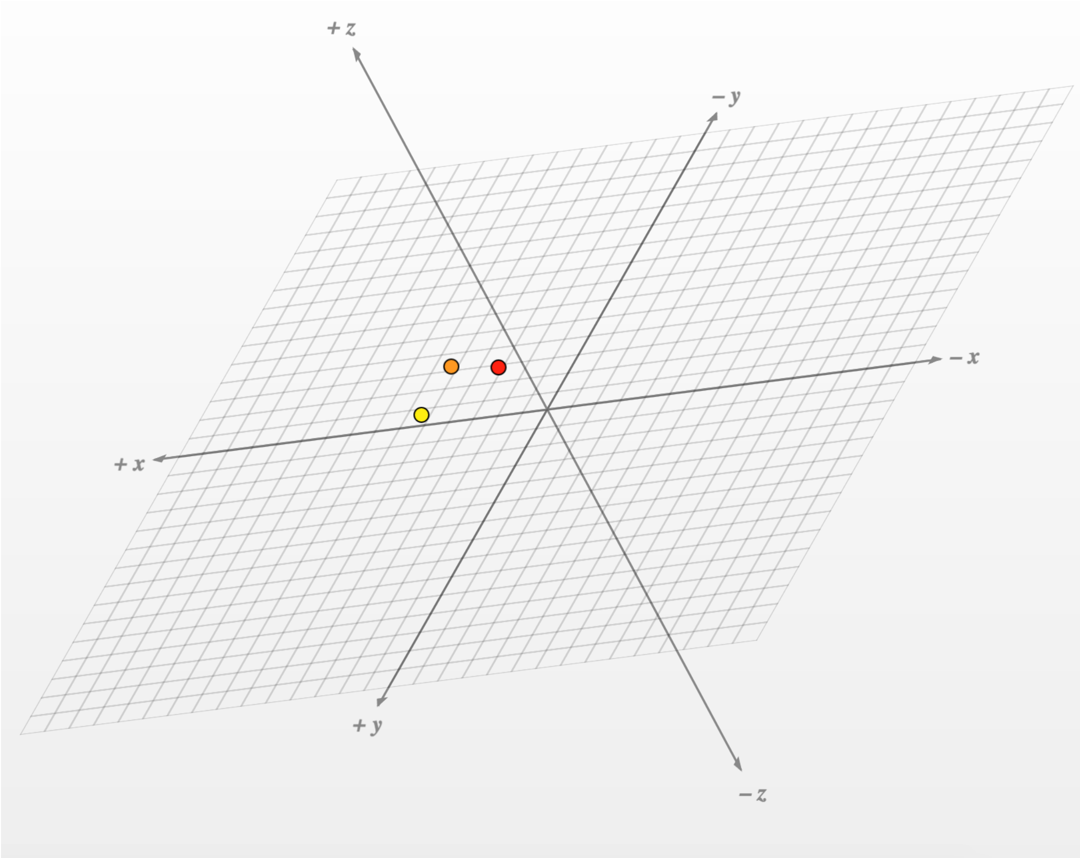
CHAPTER 1

Section 1.5

1. (a)



   (b)

$$|\overrightarrow{P_1P_2}| = \sqrt{(2-1)^2 + (1-0)^2 + (3-2)^2} = \sqrt{3}$$
$$|\overrightarrow{P_1P_3}| = \sqrt{(1-1)^2 + (5-0)^2 + (4-2)^2} = \sqrt{29}$$
$$|\overrightarrow{P_2P_3}| = \sqrt{(1-2)^2 + (5-1)^2 + (4-3)^2} = 3\sqrt{2}$$

   (c)

$$\sphericalangle(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}) = \cos^{-1}\left(\frac{\overrightarrow{P_1P_2} \cdot \overrightarrow{P_1P_3}}{|\overrightarrow{P_1P_2}||\overrightarrow{P_1P_3}|}\right) = \cos^{-1}\left(\frac{7}{\sqrt{87}}\right)$$
$$\sphericalangle(\overrightarrow{P_2P_1}, \overrightarrow{P_2P_3}) = \cos^{-1}\left(\frac{\overrightarrow{P_2P_1} \cdot \overrightarrow{P_2P_3}}{|\overrightarrow{P_2P_1}||\overrightarrow{P_2P_3}|}\right) = \cos^{-1}\left(\frac{-2\sqrt{6}}{9}\right)$$
$$\sphericalangle(\overrightarrow{P_3P_1}, \overrightarrow{P_3P_2}) = \cos^{-1}\left(\frac{\overrightarrow{P_3P_1} \cdot \overrightarrow{P_3P_2}}{|\overrightarrow{P_3P_1}||\overrightarrow{P_3P_2}|}\right) = \cos^{-1}\left(\frac{11\sqrt{58}}{87}\right)$$

(d) The area of the triangle may be found by calculating the absolute value of the cross product between any pair of vectors, say $\overrightarrow{P_1P_2}$, $\overrightarrow{P_1P_3}$, to get the area of the parallelogram with the given vectos as sides. Dividing by 2 then gives the area of the triangle:

$$\frac{1}{2}|\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}| = \frac{1}{2}|(\mathbf{i}+\mathbf{j}+\mathbf{k}) \times (5\mathbf{j}+2\mathbf{k})| = \frac{1}{2}|-3\mathbf{i}-2\mathbf{j}+5\mathbf{k}| = \frac{\sqrt{38}}{2}$$

(e) Let a denote the length of the altitude on side P_1P_2 and $\theta = \sphericalangle(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3})$ be the angle between sides P_1P_2 and P_1P_3. It then holds that

$$a = |\overrightarrow{P_1P_3}| \sin\theta$$

Next, using formula (1.18) for the area of a parallelogram of sides $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_1P_3}$

$$|\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}| = |\overrightarrow{P_1P_2}||\overrightarrow{P_1P_3}| \sin\theta$$

and solving for a using the first equation gives

$$a = \frac{|\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}|}{|\overrightarrow{P_1P_2}|} = \sqrt{\frac{38}{3}}$$

(f) The midpoint of side P_1P_2 is given by $(3/2, 1/2, 5/2)$

(g) The point O where the medians of the triangle meet (i.e. the centroid of the triangle) is given by the average of the three vertices:

$$O = \frac{P_1+P_2+P_3}{3} = \left(\frac{4}{3}, 2, 3\right)$$

2. (a)

$$\mathbf{u}\cdot\mathbf{v} = (\mathbf{i}-\mathbf{j}+2\mathbf{k})\cdot(3\mathbf{i}+\mathbf{j}-\mathbf{k}) = 0$$

(b)

$$\mathbf{u}\times\mathbf{v} = (u_2v_3 - u_3v_2)\,\mathbf{i} + (u_3v_1 - u_1v_3)\,\mathbf{j} + (u_1v_2 - u_2v_1)\,\mathbf{k} = -\mathbf{i}+7\mathbf{j}+4\mathbf{k}$$

(c)

$$|\mathbf{u}| = \sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{1+1+4} = \sqrt{6}$$

(d)

$$\sphericalangle\,(\mathbf{u},\mathbf{v}) = \cos^{-1}\left(\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|}\right) = \frac{\pi}{2}$$

(e)

$$\mathbf{u}\times(2\mathbf{v}+3\mathbf{w}) = -38\mathbf{i}+14\mathbf{j}+26\mathbf{k}$$

(f)

$$
\begin{aligned}
\mathbf{u}\cdot\mathbf{v}\times\mathbf{w} &= \mathbf{u}\cdot(7\mathbf{i}-7\mathbf{j}+14\mathbf{k}) = (\mathbf{i}-\mathbf{j}+2\mathbf{k})\cdot(7\mathbf{i}-7\mathbf{j}+14\mathbf{k}) = 42\\
\mathbf{v}\times\mathbf{u}\cdot\mathbf{w} &= (\mathbf{i}-7\mathbf{j}-4\mathbf{k})\cdot\mathbf{w} \quad = (\mathbf{i}-7\mathbf{j}-4\mathbf{k})\cdot(\mathbf{i}+5\mathbf{j}+2\mathbf{k}) \quad = -42
\end{aligned}
$$

(g) Let $\mathbf{z}=\mathbf{v}\times\mathbf{w}$, then

$$
\mathbf{z} = (v_2w_3 - v_3w_2)\,\mathbf{i} + (v_3w_1 - v_1w_3)\,\mathbf{j} + (v_1w_2 - v_2w_1)\,\mathbf{k}
$$

from which then follows

$$
\begin{aligned}
\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = \mathbf{u}\times\mathbf{z} &= (u_2z_3 - u_3z_2)\,\mathbf{i} + (u_3z_1 - u_1z_3)\,\mathbf{j} + (u_1z_2 - u_2z_1)\,\mathbf{k}\\
&= [u_2\,(v_1w_2 - v_2w_1) - u_3\,(v_3w_1 - v_1w_3)]\,\mathbf{i}\\
&\quad + [u_3\,(v_2w_3 - v_3w_2) - u_1\,(v_1w_2 - v_2w_1)]\,\mathbf{j}\\
&\quad + [u_1\,(v_3w_1 - v_1w_3) - u_2\,(v_2w_3 - v_3w_2)]\,\mathbf{k}\\
&= 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}\\
&= \mathbf{0}
\end{aligned}
$$

(h)

$$
\cos\alpha = \frac{u_x}{|\mathbf{u}|} = \frac{1}{\sqrt{6}}, \qquad \cos\beta = \frac{u_y}{|\mathbf{u}|} = -\frac{1}{\sqrt{6}}, \qquad \cos\gamma = \frac{u_z}{|\mathbf{u}|} = \frac{2}{\sqrt{6}}
$$

(i) Using the fact that $\mathbf{v}\cdot\mathbf{w} = |\mathbf{v}||\mathbf{w}|\cos\theta$ we may express $\text{comp}_w\mathbf{v}$ as

$$
\text{comp}_w\mathbf{v} = |\mathbf{v}|\cos\theta = \frac{\mathbf{v}\cdot\mathbf{w}}{|\mathbf{w}|} = \frac{\sqrt{30}}{5}
$$

3. (a) Let $\overrightarrow{P_1P_2} = (x_2-x_1)\mathbf{i}+(y_2-y_1)\mathbf{j}+(z_2-z_1)\mathbf{k}$ and $\overrightarrow{P_1P_3} = (x_3-x_1)\mathbf{i}+(y_3-y_1)\mathbf{j}+(z_3-z_1)\mathbf{k}$ be two vectors lying in the plane, originating from the same point (x_1, y_1, z_1) and let $\mathbf{n} = \overrightarrow{P_1P_2}\times\overrightarrow{P_1P_3}$ be a nonzero vector perpendicular to the plane. Here we have made use of the fact that the cross product of two vectors produces a vector which is perpendicular to both vectors appearing in the cross product. Then, following a similar reasoning, the vector $\overrightarrow{P_1P} = (x-x_1)\mathbf{i} + (y-y_1)\mathbf{j} + (z-z_1)\mathbf{k}$ and hence the point $P : (x, y, z)$ is in the plane precisely when

$$
\mathbf{n}\cdot\overrightarrow{P_1P} = \overrightarrow{P_1P_2}\times\overrightarrow{P_1P_3}\cdot\overrightarrow{P_1P} = 0
$$

which according to (1.33) and (1.34) may also be written as

$$
\overrightarrow{P_1P_2}\times\overrightarrow{P_1P_3}\cdot\overrightarrow{P_1P} = \begin{vmatrix} x-x_1 & y-y_1 & z-z_1\\ x_2-x_1 & y_2-y_1 & z_2-z_1\\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix} = D = 0
$$

(b) Evaluating the determinant from part (a) results in the linear equation

$$A\,(x - x_1) + B\,(y - y_1) + C\,(z - z_1) = 0$$

where

$$\begin{aligned} A &= (y_2 - y_1)\,(z_3 - z_1) - (z_2 - z_1)\,(y_3 - y_1) \\ B &= (z_2 - z_1)\,(x_3 - x_1) - (x_2 - x_1)\,(z_3 - z_1) \\ C &= (x_2 - x_1)\,(y_3 - y_1) - (y_2 - y_1)\,(x_3 - x_1) \end{aligned}$$

4. (a) If the points would lie in a plane, then at least one of the three vectors $\overrightarrow{P_1P_2}$, $\overrightarrow{P_1P_3}$, $\overrightarrow{P_1P_4}$ could be written as a linear combination of the other two vectors. Hence, assuming this is the case it would hold that

$$c_1\overrightarrow{P_1P_2} + c_2\overrightarrow{P_1P_3} + c_3\overrightarrow{P_1P_4} = \mathbf{0}$$

for some scalars c_1, c_2 and c_3 not all equal to 0, which results in the system of linear equations

$$\begin{aligned} c_1 + 4c_3 &= 0 \\ -2c_1 - 3c_3 &= 0 \\ 5c_1 + \quad 8c_2 + 3c_3 &= 0 \end{aligned}$$

Adding the first two equations gives $c_1 = c_3$. By substituting back into both equations we can conclude that the only way for these two equations to be satisfied is when $c_1 = c_3 = 0$, from which automatically follows that $c_2 = 0$ as well. Hence, we have proven that the three vectors are in fact linear independent and so cannot all lie in the same plane.

(b) The formula for the volume of the tetrahedron is given by

$$V = \frac{|\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_4} \cdot \overrightarrow{P_1P_3}|}{6} = \frac{20}{3}$$

which follows from the fact that 6 tetrahedrons fit into the volume of a parallelpiped with edges $\overrightarrow{P_1P_2}$, $\overrightarrow{P_1P_3}$ and $\overrightarrow{P_1P_4}$

5. (a) If $\mathbf{u}$ and $\mathbf{v}$ are linear independent, then it must hold that

$$c\mathbf{u} + d\mathbf{v} = 0$$

only for $c = d = 0$. Expanding the vector equation above gives

$$\begin{aligned} 15c + 20d &= 0 \\ -21c - 28d &= 0 \end{aligned}$$

Adding the equations then gives $c = (-4/3)d$. From substituting back into the two equations we may at once conclude that the equations are satisfied for any scalar $c = (-4/3)d$ and hence, the two vectors are linear dependent.

(b) If $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are linear independent, then it must hold that

$$c\mathbf{u} + d\mathbf{v} + e\mathbf{w} = 0$$

only for $c = d = e = 0$. Expanding the vector equation above gives

$$\begin{aligned} c + 2d + 7e &= 0 \\ c + d + 5e &= 0 \\ -c + d - e &= 0 \end{aligned}$$

Subtracting the second equation from the first gives $d = -2e$. Subtracting the third equation from the second gives $c = -3e$. Multiplying the third equation by 5 and adding it to the second gives $c = (3/2)d$. From substituting back into the three equations we may at once conclude that the equations are satisfied for any scalar $c = (3/2)d = -3e$ and hence, the three vectors are linearly dependent.

(c) If $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{p}$ are linearly independent, then it must hold that

$$c\mathbf{u} + d\mathbf{v} + e\mathbf{w} + f\mathbf{p} = 0$$

only for $c = d = e = f = 0$. Expanding the vector equations above gives

$$\begin{aligned} 2c + 2e + f &= 0 \\ c + 2d + f &= 0 \\ d + e + f &= 0 \end{aligned}$$

Subtracting the second equation from the first and the third from the second gives

$$\begin{aligned} c + 2e - 2d &= 0 \\ c + d - e &= 0 \end{aligned}$$

Subtracting the second equation from the first gives $e = d$, from which follows $c = 0$. Multiplying the third equation by 2 and subtracting it from the first and the second equation respectively gives

$$\begin{aligned} 2c - 2d - f &= 0 \\ c - 2e - f &= 0 \end{aligned}$$

Multiplying the second of these equations by 2 subtracting it from the first and substituting for e in terms of d gives $f = -2d$. From substituting back into the three equations we may at once conclude that the equations are satisfied for any scalar $f = -2d = -2e$ and $c = 0$. Hence, the four vectors are linearly dependent.

6. Let the vector $\mathbf{v} = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k} = \mathbf{i} + \mathbf{j} + 5\mathbf{k}$ be a vector along the line, then the point $P : (x, y, z)$ is on the same line when $\overrightarrow{P_1P} \times t\mathbf{v} = 0$. Hence, $\overrightarrow{P_1P} = t\mathbf{v}$, where $-\infty < t < \infty$.

(a)
$$x = 2 + t, \qquad y = 1 + t, \qquad z = 5t$$

(b) From the three given parametric equations for the line passing through the point $(2, 1, 0)$ follows that
$$\mathbf{v} = -5\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$$
From this it follows that the parametric equations for a line parallel to the given line passing through the point $(1, 1, 2)$ is given by
$$x = 1 - 5t, \qquad y = 1 + 2t, \qquad z = 2 + 3t$$

(c) Using the fact that $\mathbf{n} \cdot \overrightarrow{P_1P} = 0$ for a vector $\mathbf{n}$ normal to a plane in which the vector $\overrightarrow{P_1P}$ lies, we may re-write the equation for the plane as
$$5x - y + z = 5x_1 - y_1 + z_1 = 2$$
from which we may infer $\mathbf{n} = 5\mathbf{i} - \mathbf{j} + \mathbf{k}$. The parametric equation for a line perpendicular to the plane passing through the point $P : (0, 0, 0)$ is then given by
$$x = 5t, \qquad y = -t, \qquad z = t$$

(d) Let the vectors $\overrightarrow{P_1P_2}$ and $\overrightarrow{P_1P_3}$ be vectors along the two lines. We may then obtain a vector perpendicular to both of these two vectors, and hence the lines, using the vector cross product:
$$\mathbf{n} = \overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3} = -6\mathbf{i} - 3\mathbf{j} + 3\mathbf{k}$$
The parametric equations for a line perpendicular to the given lines passing through the point $P : (1, 2, 2)$ is then given by
$$x = 1 - 6t, \qquad y = 2 - 3t, \qquad z = 2 + 3t$$

7. (a) Let $\mathbf{u} = \mathbf{i} + 2\mathbf{j} + 5\mathbf{k}$ be a vector pointing from the origin to the point $P_1 : (1, 2, 5)$, assumed to be lying in the plane. If the plane is supposed to pass through the z-axis, then the unit vector $\mathbf{k}$ will also lie in the plane. Hence, the normal to the plane may be calculated from the vector product
$$\mathbf{n} = \mathbf{u} \times \mathbf{k} = 2\mathbf{i} - \mathbf{j}$$
The equation for the plane may then be determined from $\mathbf{n} \cdot \overrightarrow{P_1P} = 0$ or
$$\begin{aligned} (2\mathbf{i} - \mathbf{j}) \cdot [(x-1)\,\mathbf{i} + (y-2)\,\mathbf{j} + (z-5)\,\mathbf{k}] &= 0 \\ 2x - y &= 0 \end{aligned}$$

(b) From the parametric equations of the line perpendicular to the plane we may form the normal

$$\mathbf{n} = \mathbf{i} - 5\mathbf{j} + 4\mathbf{k}$$

The equation $\mathbf{n} \cdot \overrightarrow{P_1P} = 0$ then gives

$$\begin{aligned}(\mathbf{i} - 5\mathbf{j} + 4\mathbf{k}) \cdot [(x-1)\,\mathbf{i} + (y-2)\,\mathbf{j} + (z-2)\,\mathbf{k}] = 0 \\ x - 5y + 4z + 1 = 0\end{aligned}$$

(c) Let us denote the normal of the given plane as $\mathbf{u} = 2\mathbf{i} + \mathbf{j} - \mathbf{k}$, which follows immediately from (1.23). Since the two planes are assumed to be perpendicular this normal will be a vector lying wholly in the sought plane. Let $\mathbf{v} = \mathbf{i} + 3\mathbf{j} + 5\mathbf{k}$ be another vector lying in the sought plane, where we have used he parametric equations for the line lying in the plane together with (1.27). Having found two vectors lying in the plane, we may calculate its normal using the vector product

$$\mathbf{n} = \mathbf{u} \times \mathbf{v} = 8\mathbf{i} - 11\mathbf{j} + 5\mathbf{k}$$

The equation for the plane may then be found from

$$\mathbf{n} \cdot \overrightarrow{P_1P} = (8\mathbf{i} - 11\mathbf{j} + 5\mathbf{k}) \cdot [(x-1)\,\mathbf{i} + (y-2)\,\mathbf{j} + (z-1)\,\mathbf{k}] = 8x - 11y + 5z + 9 = 0$$

8. (a) If $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{0}$ then the two lines are parallel, since the angle $\theta = \sphericalangle(\mathbf{v}_1, \mathbf{v}_2) = 0$ or π. However, they are not necessarily coincident. When $\mathbf{u} \times \mathbf{v}_1 = \mathbf{0}$ as well, then the angle $\phi \sphericalangle(\mathbf{u}, \mathbf{v}_1) = 0$ or π and hence, the two lines L_1 and L_2 coincide precisely.

(b) By the same argument as given in (a), L_1 and L_2 are parallel when $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{0}$ or $\theta = \sphericalangle(\mathbf{v}_1, \mathbf{v}_2) = 0$ or π. However, they are noncoincident precisely when $\mathbf{u} \times \mathbf{v}_1 \neq \mathbf{0}$, for then $\phi \sphericalangle(\mathbf{u}, \mathbf{v}_1) \neq 0$ or π, which will cause $\mathbf{v}_1$ and $\mathbf{v}_2$ to be offset from each other in space.

(c) If $\mathbf{v}_1 \times \mathbf{v}_2 \neq \mathbf{0}$, then the two lines L_1 and L_2 are not parallel, and hence will never be able to coincide in more than one point. However, this condition alone is not enough to assure that the two lines do in fact intersect, as they could be skewed. In order to guarantee that the lines do intersect we need to apply the additional condition $\mathbf{u} \cdot \mathbf{v}_1 \times \mathbf{v}_2 = 0$. To prove why this is the case, let us first note that $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{n}$ is some new vector which is perpendicular to both $\mathbf{v}_1$ and $\mathbf{v}_2$. Now assuming for the moment that L_1 and L_2 are in fact coinciding at a common point P_0, then $\mathbf{n} \cdot \overrightarrow{P_0P_1} = \mathbf{n} \cdot t\mathbf{v}_1 = 0$ and $\mathbf{n} \cdot \overrightarrow{P_0P_2} = \mathbf{n} \cdot t\mathbf{v}_2 = 0$, which is the equation for a plane. Since the vector $\mathbf{u} = \overrightarrow{P_1P_2}$ will lie in this plane also, it follows immediately that $\mathbf{u} \cdot \mathbf{n} = 0$, which thus concludes the prove.

(d) To show that L_1 and L_2 are skewed precisely when $\mathbf{u} \cdot \mathbf{v}_1 \times \mathbf{v}_2 \neq 0$ is easy, as we can refer back to the result found in (c). There it was found that in order for the two lines to intersect in one point precisely and thus for them not to be skewed, required the condition $\mathbf{u} \cdot \mathbf{v}_1 \times \mathbf{v}_2 = 0$. The shortest distance d between the two skewed lines is given by a straight line (assuming Euclidian space) which is perpendicular to both skewed lines. Let $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2$ be a vector which is perpendicular to both $\mathbf{v}_1$ and $\mathbf{v}_2$ as before. The distance between the two points of intersection of $\mathbf{n}$ with $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively may then be expressed in terms of the absolute value of the component of the vector $\mathbf{u}$ in the direction of $\mathbf{n}$:

$$d = |\mathrm{comp}_n \mathbf{u}| = \frac{|\mathbf{u} \cdot \mathbf{v}_1 \times \mathbf{v}_2|}{|\mathbf{v}_1 \times \mathbf{v}_2|}$$

where in the last step we have used the same notation for expressing $\mathrm{comp}_n \mathbf{u}$ as was used for the solution of (2.i).

9. (a) The trisection points of the segment P_1P_2 are given by

$$P_1 + \frac{1}{3} P_1 P_2 = (4, 5, 7)$$
$$P_1 + \frac{2}{3} P_1 P_2 = (7, 8, 11)$$

(b) The point on L that is not on P_1P_2 and is two units away from P_2 is given by

$$P = P_2 + kP_1P_2 = (10 + 9k, 11 + 9k, 15 + 12k)$$

where $k = 2/|\overrightarrow{P_1P_2}| = \sqrt{34}/51$

10. (a)

$$\begin{vmatrix} 3 & 5 \\ 1 & -4 \end{vmatrix} = (3)(-4) - (1)(5) = -17$$

(b)

$$\begin{vmatrix} 2 & 1 \\ 7 & 0 \end{vmatrix} = (2)(0) - (7)(1) = -7$$

(c)

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

(d)

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

(e)

$$\begin{vmatrix} 5 & 1 & 7 & 2 \\ 3 & 1 & 4 & 1 \\ 2 & 1 & -2 & 3 \\ 0 & 1 & 4 & 1 \end{vmatrix} = 5 \begin{vmatrix} 1 & 4 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 1 \end{vmatrix} - \begin{vmatrix} 3 & 4 & 1 \\ 2 & -2 & 3 \\ 0 & 4 & 1 \end{vmatrix} + 7 \begin{vmatrix} 3 & 1 & 1 \\ 2 & 1 & 3 \\ 0 & 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 3 & 1 & 4 \\ 2 & 1 & -2 \\ 0 & 1 & 4 \end{vmatrix} = -36$$

(f)

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix} = a\left(b^2c^3 - b^3c^2\right) - b\left(a^2c^3 - a^3c^2\right) + c\left(a^2b^3 - a^3b^2\right)$$
$$= abc\,(b-a)\,(c-a)\,(c-b)$$

11. Using the formula $D = \mathbf{u} \times \mathbf{v} \cdot \mathbf{w}$ in conjunction with the fact that $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ form a positive or negative triple when the determinant is positive or negative respectively:

(a)

$$D = -5 \qquad \Longrightarrow \qquad \text{negative}$$

(b)

$$D = 0 \qquad \Longrightarrow \qquad \text{neither, linear dependence}$$

12. (a)

$$\begin{aligned}
|\mathbf{u} \times \mathbf{v}|^2 = \left(\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix}\right)^2 &= \left[(u_yv_z - u_zv_y)\,\mathbf{i} + (u_zv_x - u_xv_z)\,\mathbf{j} + (u_xv_y - u_yv_x)\,\mathbf{k}\right]^2 \\
&= (u_yv_z - u_zv_y)^2 + (u_zv_x - u_xv_z)^2 + (u_xv_y - u_yv_x)^2 \\
&= u_y^2v_z^2 + u_z^2v_y^2 - 2u_yv_zu_zv_y + u_z^2v_x^2 + u_x^2v_z^2 - 2u_zv_xu_xv_z \\
&\quad + u_x^2v_y^2 + u_y^2v_x^2 - 2u_xv_yu_yv_x \\
&= u_y^2v_z^2 + u_z^2v_y^2 - 2u_yv_zu_zv_y + u_z^2v_x^2 + u_x^2v_z^2 - 2u_zv_xu_xv_z \\
&\quad + u_x^2v_y^2 + u_y^2v_x^2 - 2u_xv_yu_yv_x + u_x^2v_x^2 - u_x^2v_x^2 \\
&\quad + u_y^2v_y^2 - u_y^2v_y^2 + u_z^2v_z^2 - u_z^2v_z^2 \\
&= \left(u_x^2 + u_y^2 + u_z^2\right)\left(v_x^2 + v_y^2 + v_z^2\right) - (u_xv_x + u_yv_y + u_zv_z)^2 \\
&= (\mathbf{u} \cdot \mathbf{u})\,(\mathbf{v} \cdot \mathbf{v}) - (\mathbf{u} \cdot \mathbf{v})^2 \\
&= (\mathbf{u} \cdot \mathbf{u})\,(\mathbf{v} \cdot \mathbf{v}) - (\mathbf{v} \cdot \mathbf{u})\,(\mathbf{u} \cdot \mathbf{v}) \\
&= \begin{vmatrix} \mathbf{u} \cdot \mathbf{u} & \mathbf{u} \cdot \mathbf{v} \\ \mathbf{v} \cdot \mathbf{u} & \mathbf{v} \cdot \mathbf{v} \end{vmatrix}
\end{aligned}$$

(b)

$$
\begin{aligned}
(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{w} \times \mathbf{z}) &= \left[(u_y v_z - u_z v_y)\,\mathbf{i} + (u_z v_x - u_x v_z)\,\mathbf{j} + (u_x v_y - u_y v_x)\,\mathbf{k}\right] \\
&\quad \cdot \left[(w_y z_z - w_z z_y)\,\mathbf{i} + (w_z z_x - w_x z_z)\,\mathbf{j} + (w_x z_y - w_y z_x)\,\mathbf{k}\right] \\
&= (u_y v_z - u_z v_y)(w_y z_z - w_z z_y) + (u_z v_x - u_x v_z)(w_z z_x - w_x z_z) \\
&\quad + (u_x v_y - u_y v_x)(w_x z_y - w_y z_x) \\
&= (u_y v_z - u_z v_y)(w_y z_z - w_z z_y) + (u_z v_x - u_x v_z)(w_z z_x - w_x z_z) \\
&\quad + (u_x v_y - u_y v_x)(w_x z_y - w_y z_x) + (u_x w_x - u_x w_x)\, v_x z_x \\
&\quad + (u_y w_y - u_y w_y)\, v_y z_y + (u_z w_z - u_z w_z)\, v_z z_z \\
&= (u_x w_x + u_y w_y + u_z w_z)(v_x z_x + v_y z_y + v_z z_z) \\
&\quad - (u_x z_x + u_y z_y + u_z z_z)(v_x w_x + v_y w_y + v_z w_z) \\
&= (\mathbf{u} \cdot \mathbf{w})(\mathbf{v} \cdot \mathbf{z}) - (\mathbf{u} \cdot \mathbf{z})(\mathbf{v} \cdot \mathbf{w})
\end{aligned}
$$

(c)

$$
\begin{aligned}
(\mathbf{u} \times \mathbf{v}) \times (\mathbf{u} \times \mathbf{w}) &= \left[(u_y v_z - u_z v_y)\,\mathbf{i} + (u_z v_x - u_x v_z)\,\mathbf{j} + (u_x v_y - u_y v_x)\,\mathbf{k}\right] \\
&\quad \times \left[(u_y w_z - u_z w_y)\,\mathbf{i} + (u_z w_x - u_x w_z)\,\mathbf{j} + (u_x w_y - u_y w_x)\,\mathbf{k}\right] \\
&= \left[(u_z v_x - u_x v_z)(u_x w_y - u_y w_x) - (u_x v_y - u_y v_x)(u_z w_x - u_x w_z)\right]\mathbf{i} \\
&\quad + \left[(u_x v_y - u_y v_x)(u_y w_z - u_z w_y) - (u_y v_z - u_z v_y)(u_x w_y - u_y w_x)\right]\mathbf{j} \\
&\quad + \left[(u_y v_z - u_z v_y)(u_z w_x - u_x w_z) - (u_z v_x - u_x v_z)(u_y w_z - u_z w_y)\right]\mathbf{k} \\
&= u_x\left[u_x (v_y w_z - v_z w_y) + u_y (v_z w_x - v_x w_z) + u_z (v_x w_y - v_y w_x)\right]\mathbf{i} \\
&\quad + u_y\left[u_x (v_y w_z - v_z w_y) + u_y (v_z w_x - v_x w_z) + u_z (v_x w_y - v_y w_x)\right]\mathbf{j} \\
&\quad + u_z\left[u_x (v_y w_z - v_z w_y) + u_y (v_z w_x - v_x w_z) + u_z (v_x w_y - v_y w_x)\right]\mathbf{k} \\
&= \left[u_x (v_y w_z - v_z w_y) + u_y (v_z w_x - v_x w_z) + u_z (v_x w_y - v_y w_x)\right]\mathbf{u} \\
&= \left\{\mathbf{u} \cdot \left[(v_y w_z - v_z w_y)\,\mathbf{i} + (v_z w_x - v_x w_z)\,\mathbf{j} + (v_x w_y - v_y w_x)\,\mathbf{k}\right]\right\}\mathbf{u} \\
&= (\mathbf{u} \cdot \mathbf{v} \times \mathbf{w})\,\mathbf{u}
\end{aligned}
$$

(d)

$$
\begin{aligned}
\left[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})\right]_x &= u_y (v_x w_y - v_y w_x) - u_z (v_z w_x - v_x w_z) \\
\left[\mathbf{v} \times (\mathbf{w} \times \mathbf{u})\right]_x &= v_y (w_x u_y - w_y u_x) - v_z (w_z u_x - w_x u_z) \\
\left[\mathbf{w} \times (\mathbf{u} \times \mathbf{v})\right]_x &= w_y (u_x v_y - u_y v_x) - w_z (u_z v_x - u_x v_z)
\end{aligned}
$$

$$
\begin{aligned}
\left[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})\right]_y &= u_z (v_y w_z - v_z w_y) - u_x (v_x w_y - v_y w_x) \\
\left[\mathbf{v} \times (\mathbf{w} \times \mathbf{u})\right]_y &= v_z (w_y u_z - w_z u_y) - v_x (w_x u_y - w_y u_x) \\
\left[\mathbf{w} \times (\mathbf{u} \times \mathbf{v})\right]_y &= w_z (u_y v_z - u_z v_y) - w_x (u_x v_y - u_y v_x)
\end{aligned}
$$

$$
\begin{aligned}
\left[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})\right]_z &= u_x (v_z w_x - v_x w_z) - u_y (v_y w_z - v_z w_y) \\
\left[\mathbf{v} \times (\mathbf{w} \times \mathbf{u})\right]_z &= v_x (w_z u_x - w_x u_z) - v_y (w_y u_z - w_z u_y) \\
\left[\mathbf{w} \times (\mathbf{u} \times \mathbf{v})\right]_z &= w_x (u_z v_x - u_x v_z) - w_y (u_y v_z - u_z v_y)
\end{aligned}
$$

For each component it then follows that

$$[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})]_x + [\mathbf{v} \times (\mathbf{w} \times \mathbf{u})]_x + [\mathbf{w} \times (\mathbf{u} \times \mathbf{v})]_x = 0$$
$$[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})]_y + [\mathbf{v} \times (\mathbf{w} \times \mathbf{u})]_y + [\mathbf{w} \times (\mathbf{u} \times \mathbf{v})]_y = 0$$
$$[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})]_z + [\mathbf{v} \times (\mathbf{w} \times \mathbf{u})]_z + [\mathbf{w} \times (\mathbf{u} \times \mathbf{v})]_z = 0$$

Which may be written more succinctly as

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0}$$

13. (a) Adding the two equations and solving for x gives $x = 2$. Substituting for x in either one of the two equations then gives $y = 1$.

(b) Multiplying the second equation by 5 and subtracting it from the first gives $y = 1$. Substituting for y in either of the two equations then gives $x = 1$.

(c) Using Cramer's rule the solution of the system of three linear equations may be written as

$$x = \frac{D_1}{D} = \frac{3}{2}, \qquad y = \frac{D_2}{D} = -2, \qquad z = \frac{D_3}{D} = -\frac{5}{2}$$

(d) Since $D = 0$, this homogeneous system of three linear equations has infinitely many solutions. If we introduce the three vectors

$$\mathbf{v}_1 = \mathbf{i} + 3\mathbf{j} + 6\mathbf{k}$$
$$\mathbf{v}_2 = -\mathbf{i} - \mathbf{j} - 4\mathbf{k}$$
$$\mathbf{v}_3 = \mathbf{i} + 2\mathbf{j} + 5\mathbf{k}$$

the system of equations may be written as the vector equation

$$x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3 = \mathbf{0}$$

It follows that the three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly dependent, as the determinant $\mathbf{v}_1 \cdot \mathbf{v}_2 \times \mathbf{v}_3$ equals D with rows and columns interchanged (see section 1.4). Hence, constants c_1, c_2 and c_3, not all equal to 0, can be found such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$. We thus have

$$x = c_1 t, \qquad y = c_2 t, \qquad z = c_3 t$$

where $-\infty < t < \infty$ and $c_1 = 1$, $c_2 = -1$ and $c_3 = -2$.

14. (a) If $D \neq 0$ then according to (1.46) the system of linear equations has the unique solution

$$x = \frac{D_1}{D} = \frac{\begin{vmatrix} k_1 & a_{12} \\ k_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{k_1 a_{22} - k_2 a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, \quad y = \frac{D_2}{D} = \frac{\begin{vmatrix} a_{11} & k_1 \\ a_{21} & k_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{k_2 a_{11} - k_1 a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$$

Geometrically, this may be interpreted as the ratio of the area of two parallelograms:

$$x = \frac{|\mathbf{w} \times \mathbf{v}_2|}{|\mathbf{v}_1 \times \mathbf{v}_2|}, \qquad y = \frac{|\mathbf{v}_1 \times \mathbf{w}|}{|\mathbf{v}_1 \times \mathbf{v}_2|}$$

where $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{w}$ are vectors having components

$$\mathbf{v}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}, \qquad \mathbf{w} = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$$

(b) If $D = 0$, but $D_1 \neq 0$ then the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent and thus $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{0}$. This would imply that $\mathbf{w} \times \mathbf{v}_1$ has to be the zero vector as well, which clearly contradicts the assumption $D_1 \neq 0$. Hence, no solution exists in this case.

(c) If $D = 0$, $D_1 = 0$ and $D_2 = 0$, a number of different cases can arise geometrically. Firstly, in the case $\mathbf{w} = \mathbf{0}$ there exist infinitely many solutions, as in this case the system of linear equations is homogeneous and can be written in the vector form

$$x\mathbf{v}_1 + y\mathbf{v}_2 = \mathbf{0}$$

For x and y not both equal to zero this implies that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent (so that $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{0}$) and hence constants c_1 and c_2 may be found such that $x = c_1 t$, $y = c_2 t$ for $-\infty < t < \infty$. If $\mathbf{w} \neq \mathbf{0}$ then two potential cases need to be examined. First, lets consider the case where both $\mathbf{v}_1 = \mathbf{0}$ and $\mathbf{v}_2 = \mathbf{0}$. This would satisfy the conditions $x|\mathbf{v}_1 \times \mathbf{v}_2| = |\mathbf{w} \times \mathbf{v}_2| = \mathbf{0}$ and $y|\mathbf{v}_1 \times \mathbf{v}_2| = |\mathbf{v}_1 \times \mathbf{w}| = \mathbf{0}$. However, this would contradict the fact that

$$a_{11}x + a_{12}y = k_1, \qquad a_{21}x + a_{22}y = k_2$$

since it is assumed that $a_{11} = a_{12} = a_{21} = a_{22} = 0$, while $k_1 \neq 0$ and $k_2 \neq 0$. Hence, no solution exists. In the second case, lets assume that $\mathbf{v}_1 \neq \mathbf{0}$ and $\mathbf{v}_2 \neq \mathbf{0}$. Then for all determinants to equal zero, it is a necessary condition that the three vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{w}$ are parallel to each other so that their cross products vanish. This will yield infinitely many solutions for $-\infty < x < \infty$, $-\infty < y < \infty$.

15. After substituting for x, y and z in terms of D, D_1, D_2 and D_3, (1.36) can be re-written as

$$\begin{aligned} a_{11}D_1 + a_{12}D_2 + a_{13}D_3 &= k_1 D \\ a_{21}D_1 + a_{22}D_2 + a_{23}D_3 &= k_2 D \\ a_{31}D_1 + a_{32}D_2 + a_{33}D_3 &= k_3 D \end{aligned}$$

Focusing on the first of the three equations and substituting for D_1, D_2 and D_3 using (1.37), the left-hand side may be written as

$$a_{11} \begin{vmatrix} k_1 & a_{12} & a_{13} \\ k_2 & a_{22} & a_{23} \\ k_3 & a_{32} & a_{33} \end{vmatrix} + a_{12} \begin{vmatrix} a_{11} & k_1 & a_{13} \\ a_{21} & k_2 & a_{23} \\ a_{31} & k_3 & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{11} & a_{12} & k_1 \\ a_{21} & a_{22} & k_2 \\ a_{31} & a_{32} & k_3 \end{vmatrix}$$

Next, let Δ_{ij} denote the minor determinant corresponding to the i^{th} row and j^{th} column of the determinants in the equation above, which gives

$$a_{11}\left(k_1\Delta_{11} - k_2\Delta_{21} + k_3\Delta_{31}\right) + a_{12}\left(-k_1\Delta_{12} + k_2\Delta_{22} - k_3\Delta_{32}\right) \\ + a_{13}\left(k_1\Delta_{13} - k_2\Delta_{23} + k_3\Delta_{33}\right)$$

Re-arranging terms gives

$$k_1\left(a_{11}\Delta_{11} - a_{12}\Delta_{12} + a_{13}\Delta_{13}\right) + k_2\left(-a_{11}\Delta_{21} + a_{12}\Delta_{22} - a_{13}\Delta_{23}\right) \\ + k_3\left(a_{11}\Delta_{31} - a_{12}\Delta_{32} + a_{13}\Delta_{33}\right)$$

The last two terms cancel out and hence we are left with the equation

$$a_{11}\Delta_{11} - a_{12}\Delta_{12} + a_{13}\Delta_{13} = D$$

which immediately can be verified to be valid. This concludes the prove that Cramer's rule provides a solution to (1.36) when $D \neq 0$, as the former procedure will be identical for the remaining two equations.

16. (a) If $a_{ij} = \delta_{ij}$, (1.45) takes the form

$$D = \begin{vmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{vmatrix} = \begin{vmatrix} \delta_{11} & & 0 \\ & \ddots & \\ 0 & & \delta_{nn} \end{vmatrix} = \begin{vmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{vmatrix}$$

To show that $\det(\mathbf{I}) = 1$, first let us re-write the expression for D as the sum

$$D = \sum_{j=1}^{n} a_{1j}\left(-1\right)^{1+j} \bar{I}_{1j} = \sum_{j=1}^{n} \delta_{1j}\left(-1\right)^{1+j} \bar{I}_{1j} = \bar{I}$$

where the minor $\bar{I}$ is the $(n-1) \times (n-1)$ identity matrix obtained from $\mathbf{I}$ by removing its first row and column. Repeating the same process recursively to each subsequent minor then ultimately gives $D = \det(\mathbf{I}) = 1$.

(b) It should be noted that when $a_{ij} = 1$ if $i + j = n + 1$ and $a_{ij} = 0$ otherwise gives $D = -1$ only for $n = 2, 3, \ldots, 6, 7, \ldots, 10, 11, \ldots$ In this case the determinant takes the form

$$D = \begin{vmatrix} 0 & & a_{1n} \\ & \ddots & \\ a_{n1} & & 0 \end{vmatrix} = \begin{vmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{vmatrix}$$

The easiest way to establish the sign of D is to determine for each n given how many row (or column) interchanges are necessary to transform the determinant above into the determinant of the identity matrix, as according to rule I of section (1.4) interchanging two rows (or columns) multiplies the determinant by -1. If the number of row or column interchanges is even $D = 1$, if it is odd $D = -1$.

(c) If $a_{ij} = 0$ for $i > j$ the determinant is of the form

$$D = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix}$$

Then the determinant expressed as a sum takes the form

$$D = \sum_{j=1}^{n} a_{1j} A_{1j} = a_{11} A_{11}$$

Here A_{11} is the $(1,1)$ cofactor associated with A, which alternatively may be expressed as

$$A_{11} = \det(M_{11}) = \begin{vmatrix} a_{22} & a_{23} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{vmatrix}$$

where M_{11} is the $(n-1) \times (n-1)$ matrix obtained from A by removing its first row and column. The last step follows from the fact that D is upper triangular and hence, apart from the $(1,1)^{th}$ cofactor, the first column of each cofactor will contain zero entrees only (and thus will result in a zero valued determinant). By induction, it then follows that $D = a_{11} a_{22} \dots a_{nn}$.

(d) For $n = 2$ we have

$$D = \sum \epsilon_{j_1 j_2} a_{1 j_1} a_{2 j_2} = a_{11} a_{22} - a_{12} a_{21}$$

Which we immediately can verify to be correct. For $n = 3$ we have

$$\begin{aligned} D = \sum \epsilon_{j_1 j_2 j_3} a_{1 j_1} a_{2 j_2} a_{3 j_3} &= \epsilon_{123} a_{11} a_{22} a_{33} + \epsilon_{132} a_{11} a_{23} a_{32} + \epsilon_{213} a_{12} a_{21} a_{33} \\ &\quad + \epsilon_{231} a_{12} a_{23} a_{31} + \epsilon_{312} a_{13} a_{21} a_{32} + \epsilon_{321} a_{13} a_{22} a_{31} \\ &= a_{11} \left(a_{22} a_{33} - a_{23} a_{32} \right) - a_{12} \left(a_{21} a_{32} - a_{23} a_{31} \right) \\ &\quad + a_{13} \left(a_{21} a_{32} - a_{22} a_{31} \right) \end{aligned}$$

which from the last step we again can verify to be correct.

Section 1.7

1. (a) Number of rows of **A**: 2, number of columns of **A**: 1.
   Number of rows of **F**: 2, number of columns of **F**: 3.
   Number of rows of **H**: 1, number of columns of **H**: 3.
   Number of rows of **L**: 3, number of columns of **L**: 3.
   Number of rows of **P**: 3, number of columns of **P**: 2.

(b)

$$a_{11} = 1,\ a_{21} = 3,\ c_{21} = 4,\ c_{22} = 1,\ d_{12} = -1,\ e_{21} = 2,\ f_{11} = 1$$
$$g_{23} = -1,\ g_{21} = -1,\ h_{12} = 0,\ m_{23} = 1$$

(c)

$$\mathbf{c}_1 = (2,3)\,,\ \mathbf{c}_2 = (4,1)\,,\ \mathbf{g}_1 = (3,1,4)\,,\ \mathbf{g}_2 = (-1,0,-1)$$
$$\mathbf{l}_1 = (3,1,0)\,,\ \mathbf{l}_2 = (2,5,6)\,,\ \mathbf{l}_3 = (1,4,3)$$
$$\mathbf{p}_1 = (2,2)\,,\ \mathbf{p}_2 = (-1,-1)\,,\ \mathbf{p}_3 = (3,3)$$

(d)

$$\mathbf{d}_1 = \begin{pmatrix}1\\2\end{pmatrix},\ \mathbf{d}_2 = \begin{pmatrix}-1\\0\end{pmatrix},\ \mathbf{f}_1 = \begin{pmatrix}1\\2\end{pmatrix},\ \mathbf{f}_2 = \begin{pmatrix}4\\0\end{pmatrix},\ \mathbf{f}_3 = \begin{pmatrix}5\\7\end{pmatrix}$$
$$\mathbf{l}_1 = \begin{pmatrix}3\\2\\1\end{pmatrix},\ \mathbf{l}_2 = \begin{pmatrix}1\\5\\4\end{pmatrix},\ \mathbf{l}_3 = \begin{pmatrix}0\\6\\3\end{pmatrix},\ \mathbf{n}_1 = \begin{pmatrix}1\\0\\7\end{pmatrix},\ \mathbf{n}_2 = \begin{pmatrix}4\\3\\1\end{pmatrix}$$

2. (a)

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix}1\\3\end{bmatrix} + \begin{bmatrix}2\\0\end{bmatrix} = \begin{bmatrix}3\\3\end{bmatrix}$$

(b)

$$\mathbf{C} + \mathbf{D} = \begin{bmatrix}2 & 3\\4 & 1\end{bmatrix} + \begin{bmatrix}1 & -1\\2 & 0\end{bmatrix} = \begin{bmatrix}3 & 2\\6 & 1\end{bmatrix}$$

(c) meaningless

(d)

$$\mathbf{L} + \mathbf{M} = \begin{bmatrix}3 & 1 & 0\\2 & 5 & 6\\1 & 4 & 3\end{bmatrix} + \begin{bmatrix}2 & -1 & 0\\1 & 2 & 1\\3 & 2 & -1\end{bmatrix} = \begin{bmatrix}5 & 0 & 0\\3 & 7 & 7\\4 & 6 & 2\end{bmatrix}$$

(e)

$$\mathbf{N} - \mathbf{P} = \begin{bmatrix}1 & 4\\0 & 3\\7 & 1\end{bmatrix} + \begin{bmatrix}2 & 2\\-1 & -1\\3 & 3\end{bmatrix} = \begin{bmatrix}-1 & 2\\1 & 4\\4 & -2\end{bmatrix}$$

(f)

$$\mathbf{G} - \mathbf{F} = \begin{bmatrix}3 & 1 & 4\\-1 & 0 & -1\end{bmatrix} + \begin{bmatrix}1 & 4 & 5\\2 & 0 & 7\end{bmatrix} = \begin{bmatrix}2 & -3 & -1\\-3 & 0 & -8\end{bmatrix}$$

(g)

$$5\mathbf{C} = \begin{bmatrix}10 & 15\\20 & 5\end{bmatrix}$$

(h)

$$2\mathbf{E} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$$

(i)

$$3\mathbf{E} + 4\mathbf{D} = \begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix} + \begin{bmatrix} 4 & -4 \\ 8 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 14 & 12 \end{bmatrix}$$

(j)

$$2\mathbf{C} + \mathbf{D} - \mathbf{E} = \begin{bmatrix} 4 & 6 \\ 8 & 2 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 8 & -2 \end{bmatrix}$$

(k) meaningless

3. (a)

$$\mathbf{X} = \mathbf{D} - \mathbf{C} = \begin{bmatrix} -1 & -4 \\ -2 & -1 \end{bmatrix}$$

(b)

$$\mathbf{X} = \frac{1}{5}\left(\mathbf{F} - \mathbf{G}\right) = \begin{bmatrix} -2/5 & 3/5 & 1/5 \\ 3/5 & 0 & 8/5 \end{bmatrix}$$

4. (a)

$$\mathbf{X} = \frac{1}{2}\left(\mathbf{N} + \mathbf{P}\right) = \begin{bmatrix} 3/2 & 3 \\ -1/2 & 1 \\ 5 & 2 \end{bmatrix}, \qquad \mathbf{Y} = \frac{1}{2}\left(\mathbf{N} - \mathbf{P}\right) = \begin{bmatrix} -1/2 & 1 \\ 1/2 & 2 \\ 2 & -1 \end{bmatrix}$$

(b)

$$\mathbf{X} = 2\mathbf{L} - 3\mathbf{M} = \begin{bmatrix} 0 & 5 & 0 \\ 1 & 4 & 9 \\ -7 & 2 & 9 \end{bmatrix}, \qquad \mathbf{Y} = \mathbf{L} - 2\mathbf{M} = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 1 & 4 \\ -5 & 0 & 5 \end{bmatrix}$$

5. (a)

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = a_{ij} + (b_{ij} + c_{ij}) = a_{ij} + b_{ij} + c_{ij} = (a_{ij} + b_{ij}) + c_{ij} = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$$

(b)

$$c\left(\mathbf{A} + \mathbf{B}\right) = c\left(a_{ij} + b_{ij}\right) = ca_{ij} + cb_{ij} = c\mathbf{A} + c\mathbf{B}$$

(c)

$$\left(a + b\right)\mathbf{C} = \left(a + b\right)c_{ij} = ac_{ij} + bc_{ij} = a\mathbf{C} + b\mathbf{C}$$

(d)

$$a\left(b\mathbf{C}\right) = a\left(bc_{ij}\right) = abc_{ij} = \left(ab\right)c_{ij} = \left(ab\right)\mathbf{C}$$

(e)

$$1\mathbf{A} = 1a_{ij} = a_{ij} = \mathbf{A}$$

(f)
$$0\mathbf{A} = 0a_{ij} = \mathbf{0}$$

(g)
$$\mathbf{A} + \mathbf{0} = a_{ij} + 0 = a_{ij} = \mathbf{A}$$

(h)
$$\begin{aligned} \mathbf{A} + \mathbf{C} = \mathbf{B} \quad &\Longrightarrow \quad a_{ij} + c_{ij} = b_{ij} \\ c_{ij} = b_{ij} - a_{ij} \quad &\Longrightarrow \quad \mathbf{C} = \mathbf{B} - \mathbf{A} \end{aligned}$$

Section 1.8

1. (a) meaningless

(b)
$$\mathbf{CA} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 7 \end{bmatrix}$$

(c) meaningless

(d)
$$\mathbf{CD} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 8 & -2 \\ 6 & -4 \end{bmatrix}, \quad \mathbf{DC} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 4 & 6 \end{bmatrix}$$

(e)
$$\mathbf{CE} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 8 & 16 \\ 6 & 12 \end{bmatrix}, \quad \mathbf{EC} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 5 \\ 20 & 10 \end{bmatrix}$$

(f)
$$\mathbf{AI} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} = \mathbf{A}$$

(g)
$$\mathbf{IL} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ 2 & 5 & 6 \\ 1 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 2 & 5 & 6 \\ 1 & 4 & 3 \end{bmatrix} = \mathbf{L}$$

(h) undefined

(i)

$$\begin{aligned}
\mathbf{C}^2 - 3\mathbf{C} - 10\mathbf{C}^0 &= \left(\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}\right)^2 - 3\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - 10\left(\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}\right)^0 \\
&= \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - 3\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - 10\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 16 & 9 \\ 12 & 13 \end{bmatrix} - \begin{bmatrix} 6 & 9 \\ 12 & 3 \end{bmatrix} - \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \mathbf{0}
\end{aligned}$$

(j)

$$\begin{aligned}
\mathbf{E}\left(\mathbf{E} - 5\mathbf{I}\right) &= \mathbf{E}^2 - 5\mathbf{E}\mathbf{I} \\
&= \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - 5\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} - \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \\
&= \mathbf{0}
\end{aligned}$$

(k)

$$\mathbf{LNI} = \begin{bmatrix} 3 & 1 & 0 \\ 2 & 5 & 6 \\ 1 & 4 & 3 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & 3 \\ 7 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 15 \\ 44 & 29 \\ 22 & 19 \end{bmatrix}$$

(l)

$$\mathbf{MPG} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 3 & 2 & -1 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ -1 & -1 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} 3 & 1 & 4 \\ -1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 10 & 5 & 15 \\ 6 & 3 & 9 \\ 2 & 1 & 3 \end{bmatrix}$$

(m)

$$\mathbf{HL} + \mathbf{J} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 & 0 \\ 2 & 5 & 6 \\ 1 & 4 & 3 \end{bmatrix} + \begin{bmatrix} 3 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 10 & 5 \end{bmatrix}$$

(n)

$$\mathbf{KA} = \begin{bmatrix} 3 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = 18$$

(o)

$$\mathbf{0C} + \mathbf{N} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 4 \\ 0 & 3 \\ 7 & 1 \end{bmatrix} = \mathbf{N}$$

(p)

$$\mathbf{E}^2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} = 5 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = 5\mathbf{E}$$

(q)

$$\mathbf{E}^3 = \mathbf{E}^2\mathbf{E} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 25 & 50 \\ 50 & 100 \end{bmatrix} = 25 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = 25\mathbf{E}$$

(r)

$$\mathbf{E}^4 = \mathbf{E}^3\mathbf{E} = \begin{bmatrix} 25 & 50 \\ 50 & 100 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 125 & 250 \\ 250 & 500 \end{bmatrix} = 125 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = 125\mathbf{E}$$

(s) meaningless

2. (a)

$$\mathbf{RS} = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 9 & 1 & 9 \\ 5 & 2 & 16 & 2 & 16 \end{bmatrix}$$

(b)

$$\mathbf{RS} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 2 & 1 \\ 3 & 2 & 2 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 9 & 9 & 14 & 9 \\ 7 & 4 & 4 & 7 & 4 \\ 10 & 5 & 5 & 10 & 5 \end{bmatrix}$$

3. (a)

$$\begin{array}{rcrcl} 3u_1 & + & 2u_2 & = & y_1 \\ 5u_1 & + & 6u_2 & = & y_2 \\ 3\,(5x_1 - x_2) & + & 2\,(x_1 + 2x_2) & = & y_1 \\ 5\,(5x_1 - x_2) & + & 6\,(x_1 + 2x_2) & = & y_2 \\ 17x_1 & + & x_2 & = & y_1 \\ 31x_1 & + & 7x_2 & = & y_2 \end{array}$$

$$\begin{aligned} \begin{bmatrix} 3 & 2 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ \begin{bmatrix} 3 & 2 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ \begin{bmatrix} 17 & 1 \\ 31 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \end{aligned}$$

(b)

$$\begin{array}{rcrcrcl} & & 2u_1 & - & u_2 & = & y_1 \\ & & 5u_1 & + & u_2 & = & y_2 \end{array}$$

$$\begin{array}{rcrcl} 2\left(x_1+2x_2-x_3\right) & - & \left(2x_1+3x_2+x_3\right) & = & y_1 \\ 5\left(x_1+2x_2-x_3\right) & + & \left(2x_1+3x_2+x_3\right) & = & y_2 \end{array}$$

$$\begin{array}{rcrcrcl} & & x_2 & - & 3x_3 & = & y_1 \\ 7x_1 & + & 13x_2 & - & 4x_3 & = & y_2 \end{array}$$

$$\begin{bmatrix} 2 & -1 \\ 5 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -1 \\ 5 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & -3 \\ 7 & 13 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

4. (a)

$$\mathbf{IA} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}\begin{bmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mp} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mp} \end{bmatrix} = \mathbf{A}$$

(b) Let $\mathbf{D} = c\left(\mathbf{AB}\right)$ then

$$d_{ij} = c\left(a_{i1}b_{1j}\right) + \ldots + c\left(a_{ip}b_{pj}\right) = a_{i1}\left(cb_{1j}\right) + \ldots + a_{ip}\left(cb_{pj}\right)$$

and hence $\mathbf{D} = \mathbf{A}\left(c\mathbf{B}\right)$.

(c) Assuming that $\mathbf{A}$ is a square matrix of order n and starting from the definitions $\mathbf{A}^0 = \mathbf{I}$ and $\mathbf{A}^1 = \mathbf{A}$, the product $\mathbf{AA}$ may be written as $\mathbf{A}^2$. Similarly, $\mathbf{A}^3 = \mathbf{A}^2\mathbf{A}$, $\mathbf{A}^4 = \mathbf{A}^3\mathbf{A}, \ldots, \mathbf{A}^{k+1} = \mathbf{A}^k\mathbf{A}$. But this is just the special case where $l = 1$. By induction, it then follows immediately that more generally $\mathbf{A}^k\left(\mathbf{A}^{l-1}\mathbf{A}\right) = \mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l}$.

(d) By definition $\left(\mathbf{A}^k\right)^1 = \mathbf{A}^k$. From this it follows that $\left(\mathbf{A}^k\right)^2 = \mathbf{A}^k\mathbf{A}^k = \mathbf{A}^{2k}$. Similarly $\left(\mathbf{A}^k\right)^3 = \mathbf{A}^k\mathbf{A}^k\mathbf{A}^k = \mathbf{A}^{2k}\mathbf{A}^k = \mathbf{A}^{3k}$. By induction, it then follows that $\left(\mathbf{A}^k\right)^4 = \mathbf{A}^{4k}, \ldots, \left(\mathbf{A}^k\right)^l = \mathbf{A}^{kl}$.

5. (a) Using rule 13, the matrix product $\left(\mathbf{A}+\mathbf{I}\right)\left(\mathbf{A}-\mathbf{I}\right)$ may be written as

$$\left(\mathbf{A}+\mathbf{I}\right)\left(\mathbf{A}-\mathbf{I}\right) = \mathbf{A}^2 - \mathbf{I}^2 + \mathbf{A} - \mathbf{A} = \mathbf{A}^2 - \mathbf{I}$$

where the last step is follows from the fact that $\mathbf{I}^n = \mathbf{I}$.

(b)

$$(\mathbf{A}-\mathbf{I})\left(\mathbf{A}^2+\mathbf{A}+\mathbf{I}\right)=\mathbf{A}^3-\mathbf{A}^2+\mathbf{A}^2-\mathbf{A}+\mathbf{A}-\mathbf{I}^2=\mathbf{A}^3-\mathbf{I}$$

(c)

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

(d)

$$(2\mathbf{A}+\mathbf{I})(3\mathbf{A}-2\mathbf{I})=6\mathbf{A}^2-4\mathbf{A}\mathbf{I}+3\mathbf{I}\mathbf{A}-2\mathbf{I}^2=6\mathbf{A}^2-\mathbf{A}-2\mathbf{I}$$

6. If $\mathbf{A}$ and $\mathbf{B}$ are $n\times n$ matrices, it is not necessarily true that $\mathbf{A}^2-\mathbf{B}^2=(\mathbf{A}-\mathbf{B})(\mathbf{A}+\mathbf{B})$, since in general order matters when multiplying matrices. Expanding out the product gives
$$(\mathbf{A}-\mathbf{B})(\mathbf{A}+\mathbf{B})=\mathbf{A}^2-\mathbf{B}\mathbf{A}+\mathbf{A}\mathbf{B}-\mathbf{B}^2$$
Hence, since order matters it is not necessarily true that $\mathbf{B}\mathbf{A}=\mathbf{A}\mathbf{B}$, and in fact generally it is not. When it is true however, it is said that the matrices $\mathbf{A}$ and $\mathbf{B}$ commute.

7. In order to find two non-zero matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A}^2+\mathbf{B}^2=\mathbf{0}$, we require that the square of each matrix equals $\mathbf{0}$. Let us focus on the matrix $\mathbf{A}$ for now. We then require that
$$\mathbf{A}^2=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}^2=\begin{bmatrix}a_{11}^2+a_{12}a_{21} & a_{12}(a_{11}+a_{22})\\ a_{21}(a_{11}+a_{22}) & a_{12}a_{21}+a_{22}^2\end{bmatrix}=\mathbf{0}$$
From the diagonal elements it follows that $a_{22}^2=a_{11}^2$, hence $a_{22}=\pm a_{11}$. However, if we want a_{12} and a_{21} to be non-zero scalars we require that $a_{22}=-a_{11}$, as can be deduced from looking at (one of) the two off-diagonal elements. Lets assume the coefficient $a_{22}=cd$ (i.e. the product of two arbitrary scalars not equal to zero). Substituting for a_{11}^2 or a_{22}^2 in one of the diagonal elements then gives $c^2d^2+a_{12}a_{21}=0$ from which follows that $a_{12}a_{21}=-c^2d^2$. Hence, we may conclude that $\mathbf{A}$ and $\mathbf{B}$ are matrices which are of the form
$$\mathbf{A}=\mathbf{B}=\lambda\begin{bmatrix}-cd & -c^2\\ d^2 & cd\end{bmatrix}$$
where λ, c and d are arbitrary scalars.

8. Writing out $\mathbf{A}\mathbf{B}=\mathbf{B}\mathbf{A}$ explicitly, and adding the second and third equations results in the system of equations
$$\begin{aligned} a_{12}b_{21}-a_{21}b_{12} &= 0\\ (a_{21}-a_{12})(b_{11}-b_{22})+(b_{21}-b_{12})(a_{22}-a_{11}) &= 0\end{aligned}$$
Taking into consideration that $\mathbf{B}$ could be any matrix (i.e. $b_{21}-b_{12}$ is not necessarily equal to zero), it then follows that the second term of the second equation will vanish if we choose $a_{22}=a_{11}=c$. We apply the same logic to the first term of the second equation and set $a_{21}=a_{12}=d$. Substituting for a_{12} and a_{21} in the first equation then gives $d(b_{21}-b_{12})=0$. Again, since we require that $\mathbf{B}$ can be any matrix, this equation is only satisfied for $d=0$. Thus, we may conclude that for any 2×2 matrix $\mathbf{B}$, the equation $\mathbf{A}\mathbf{B}=\mathbf{B}\mathbf{A}$ is satisfied whenever $\mathbf{A}$ is the 2×2 matrix $\mathbf{A}=c\mathbf{I}$.

Section 1.9

1. (a) For the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix}$$

the simultaneous equations are

$$3x_1 + 5x_2 = y_1, \qquad 2x_1 + 4x_2 = y_2$$

Solving for x_1 and x_2 gives

$$x_1 = \frac{1}{2}\left(4y_1 - 5y_2\right), \qquad x_2 = \frac{1}{2}\left(-2y_1 + 3y_2\right)$$

and hence,

$$\mathbf{A}^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}$$

(b) For the matrix

$$\mathbf{B} = \begin{bmatrix} 4 & 7 \\ 1 & 6 \end{bmatrix}$$

the simultaneous equations are

$$4x_1 + 7x_2 = y_1, \qquad x_1 + 6x_2 = y_2$$

Solving for x_1 and x_2 gives

$$x_1 = \frac{1}{17}\left(6y_1 - 7y_2\right), \qquad x_2 = \frac{1}{17}\left(-y_1 + 4y_2\right)$$

and hence,

$$\mathbf{B}^{-1} = \frac{1}{17}\begin{bmatrix} 6 & -7 \\ -1 & 4 \end{bmatrix}$$

(c) For the matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

the simultaneous equations are

$$x_1 + x_3 = y_1, \qquad 2x_1 + 2x_2 + x_3 = y_2, \qquad x_2 - x_3 = y_3$$

Solving for x_1, x_2 and x_3 gives

$$x_1 = 3y_1 - y_2 + 2y_3, \qquad x_2 = -2y_1 + y_2 - y_3, \qquad x_3 = -2y_1 + y_2 - 2y_3$$

and hence,

$$\mathbf{C}^{-1} = \begin{bmatrix} 3 & -1 & 2 \\ -2 & 1 & -1 \\ -2 & 1 & -2 \end{bmatrix}$$

(d) For the matrix

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

the simultaneous equations are

$$2x_1 + x_3 = y_1, \qquad 3x_1 + x_2 + 2x_3 = y_2, \qquad 4x_1 + 3x_3 = y_3$$

Solving for x_1, x_2 and x_3 gives

$$x_1 = \frac{1}{2}\left(3y_1 - y_3\right), \qquad x_2 = \frac{1}{2}\left(-y_1 + 2y_2 - y_3\right), \qquad x_3 = \frac{1}{2}\left(-4y_1 + 2y_3\right)$$

and hence,

$$\mathbf{D}^{-1} = \frac{1}{2}\begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & -1 \\ -4 & 0 & 2 \end{bmatrix}$$

2. (a)

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{A}\mathbf{X} = \frac{1}{2}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -5 & 4 \\ 3 & -2 \end{bmatrix}$$

(b)

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{A}\mathbf{X} = \frac{1}{2}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 5 \\ 6 & 2 \end{bmatrix} = \begin{bmatrix} -13 & 5 \\ 8 & -2 \end{bmatrix}$$

(c)

$$\mathbf{X} = \mathbf{B}^{-1}\mathbf{B}\mathbf{X}\mathbf{A}\mathbf{A}^{-1} = \frac{1}{34}\begin{bmatrix} 6 & -7 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 7 & 2 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix} = \frac{1}{34}\begin{bmatrix} 214 & -279 \\ -64 & 89 \end{bmatrix}$$

(d)

$$\mathbf{X} = \left(\mathbf{B}^{-1}\mathbf{B}\right)^2 \mathbf{X}\left(\mathbf{A}\mathbf{A}^{-1}\right)^2 = \frac{1}{1156}\begin{bmatrix} -1714 & 2323 \\ 654 & -883 \end{bmatrix}$$

(e)

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \frac{1}{2}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(f)

$$\mathbf{x} = \mathbf{B}^{-1}\mathbf{B}\mathbf{x} = \frac{1}{17}\begin{bmatrix} 6 & -7 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 7 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} -43 \\ 27 \end{bmatrix}$$

(g)

$$\mathbf{X} = \mathbf{B}^{-1}\mathbf{B}\mathbf{X} = \frac{1}{17}\begin{bmatrix} 6 & -7 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 1 & 6 & 2 \\ 5 & 0 & -3 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} -29 & 36 & 33 \\ 19 & -6 & -14 \end{bmatrix}$$

(h)

$$\mathbf{X} = \mathbf{X}\mathbf{C}\mathbf{C}^{-1} = \begin{bmatrix} 0 & 5 & 4 \\ 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 2 \\ -2 & 1 & -1 \\ -2 & 1 & -2 \end{bmatrix} = \begin{bmatrix} -18 & 9 & -13 \\ 13 & -4 & 8 \end{bmatrix}$$

(i)

$$\mathbf{x} = \mathbf{x}\mathbf{C}\mathbf{C}^{-1} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & -1 & 2 \\ -2 & 1 & -1 \\ -2 & 1 & -2 \end{bmatrix} = \begin{bmatrix} -7 & 4 & -6 \end{bmatrix}$$

(j)

$$\mathbf{x} = \mathbf{x}\mathbf{D}\mathbf{D}^{-1} = \frac{1}{2} \begin{bmatrix} 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & -1 \\ -1 & 2 & -1 \\ -4 & 0 & 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5 & 2 & -3 \end{bmatrix}$$

3. (a)

$$\left[(\mathbf{AB})^{-1} \mathbf{A}^{-1} \right]^{-1} = \left(\mathbf{B}^{-1}\mathbf{A}^{-2} \right)^{-1} = \mathbf{A}^2\mathbf{B}$$

(b)

$$(\mathbf{ABC})^{-1} \left(\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1} \right)^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{ABC} = \mathbf{I}$$

(c)

$$\left\{ \left[\left(\mathbf{A}^{-1}\right)^2 \mathbf{B} \right]^{-1} \mathbf{A}^{-2}\mathbf{B}^{-1} \right\}^{-2} = \left[\mathbf{B}\mathbf{A}^2 \left(\mathbf{A}^{-1}\right)^2 \mathbf{B} \right]^2 = \left[\mathbf{B}\mathbf{A}^2\mathbf{A}^{-2}\mathbf{B} \right]^2 = \mathbf{B}^4$$

4. (a) If $\mathbf{A}$ and $\mathbf{B}$ are nonsingular $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ exist. Furthermore, if $\mathbf{AB} = \mathbf{BA}$ then taking the inverse of both sides gives

$$(\mathbf{AB})^{-1} = (\mathbf{BA})^{-1} \qquad \Longrightarrow \qquad \mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}$$

(b) if $\mathbf{ABC} = \mathbf{I}$ then

$$\mathbf{A}^{-1}\mathbf{ABCA} = \mathbf{A}^{-1}\mathbf{IA} \qquad \Longrightarrow \qquad \mathbf{BCA} = \mathbf{I}$$

Similarly

$$\mathbf{CABCC}^{-1} = \mathbf{CIC}^{-1} \qquad \Longrightarrow \qquad \mathbf{CAB} = \mathbf{I}$$

5. To prove that $(c\mathbf{A})^{-1} = c^{-1}\mathbf{A}^{-1}$ $(c \neq 0)$ we multiply both sides with $c\mathbf{A}$ to get

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

It follows directly that the right-hand side equals $\mathbf{I}$, since

$$c^{-1}\mathbf{A}^{-1}c\mathbf{A} = c^{-1}c\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

It remains to be proven that the left-hand side equals $\mathbf{I}$ as well. This is easily done by defining $c\mathbf{A} = \mathbf{B}$, such that $b_{ij} = ca_{ij}$. Then since $\mathbf{A}$ is invertible, $\mathbf{B}$ is invertible and it follows immediately that $\mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$ and hence $(c\mathbf{A})^{-1} c\mathbf{A} = \mathbf{I}$. This concludes the prove.

6. For $s \geq 0,\ t \geq 0$ it holds that

$$(\mathbf{A}^s)^{-t} = (\mathbf{A}\mathbf{A}\ldots\mathbf{A})^{-t} = \left(\mathbf{A}^{-1}\mathbf{A}^{-1}\ldots\mathbf{A}^{-1}\right)^t = \left(\mathbf{A}^{-s}\right)^t = \mathbf{A}^{-s}\mathbf{A}^{-s}\ldots\mathbf{A}^{-s} = \mathbf{A}^{-st}$$

Similarly, we have

$$\left(\mathbf{A}^{-s}\right)^{-t} = \left(\mathbf{A}^{-1}\mathbf{A}^{-1}\ldots\mathbf{A}^{-1}\right)^{-t} = (\mathbf{A}\mathbf{A}\ldots\mathbf{A})^t = (\mathbf{A}^s)^t = \mathbf{A}^s\mathbf{A}^s\ldots\mathbf{A}^s = \mathbf{A}^{st}$$

The prove for non-negative powers is given in section 1.8 (also see the solution to problem 4d of secion 1.8).

7. (a) In order to prove the equation for $n = 2$, let us introduce the matrix

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

We now re-write the equation as

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{\det \mathbf{A}}\mathbf{B} \\ \mathbf{A}\mathbf{A}^{-1} &= \frac{1}{\det \mathbf{A}}\mathbf{A}\mathbf{B} \\ (\det \mathbf{A})\,\mathbf{I} &= \mathbf{A}\mathbf{B} \end{aligned}$$

Hence, what remains to be verified is if the product $\mathbf{AB}$ is indeed equal to the diagonal matrix diag $(a_{11}a_{22} - a_{12}a_{21})$:

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (\det \mathbf{A})\,\mathbf{I}$$

This concludes the prove.

(b) For $n = 3$ we have

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{11} \\ a_{32} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} \\ &= \begin{bmatrix} \det \mathbf{A} & 0 & 0 \\ 0 & \det \mathbf{A} & 0 \\ 0 & 0 & \det \mathbf{A} \end{bmatrix} = (\det \mathbf{A})\,\mathbf{I} \end{aligned}$$

where

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

(c) In the most general case we want to verify that $(\det \mathbf{A})\,\mathbf{I} = \mathbf{A}\mathrm{adj}\,\mathbf{A}$, where $\mathrm{adj}\,\mathbf{A} = \mathbf{B}$ with $b_{ij} = (-1)^{i+j} A_{ji}$. Writing out $\mathbf{A}\mathrm{adj}\,\mathbf{A}$ explicitly gives

$$\begin{aligned}
\mathbf{A}\mathrm{adj}\,\mathbf{A} &= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^{n} a_{1k} b_{k1} & \cdots & \sum_{k=1}^{n} a_{1k} b_{kn} \\ \vdots & & \vdots \\ \sum_{k=1}^{n} a_{nk} b_{k1} & \cdots & \sum_{k=1}^{n} a_{nk} b_{kn} \end{bmatrix}
\end{aligned}$$

Each sum appearing along the diagonal of the final matrix (i.e. all b_{ij}'s where $i = j$) is nothing other than the expansion of $\det \mathbf{A}$ by the cofactors of $\mathbf{A}$ (i.e. the b_{ij}'s). Hence, all entrees along the diagonal evaluate to $\det \mathbf{A}$ as we require. It remains to be proven that the off-diagonal elements (i.e. all b_{ij}'s where $i \neq j$) sum to zero. This follows directly from rule *IV* of section 1.4, which states that if two rows or columns of a determinant are proportional, the determinant equals zero. For instance, for b_{21} we have

$$\begin{aligned}
b_{21} &= \sum_{k=1}^{n} a_{2k} b_{k1} = \sum_{k=1}^{n} (-1)^{k+1} a_{2k} A_{1k} \\
&= a_{21} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} - a_{22} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} + \dots \\
&\quad + (-1)^{k+1} a_{2n-1} \begin{vmatrix} a_{21} & \dots & a_{2n-2} & a_{2n} \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{nn-2} & a_{nn} \end{vmatrix} + (-1)^{k+1} a_{2n} \begin{vmatrix} a_{21} & \dots & a_{2n-1} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn-1} \end{vmatrix} \\
&= \begin{vmatrix} a_{21} & \dots & a_{2n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = 0
\end{aligned}$$

The same holds for any of the other off-diagonal matrix elements and hence, we have finally shown that

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \mathrm{adj}\,\mathbf{A}$$

as claimed.

8. All matrices whose inverses appear in the solutions are assumed to be nonsingular.

(a)

$$\mathbf{X} = \frac{1}{2}(\mathbf{A}+\mathbf{B}) \qquad\qquad \mathbf{Y} = \frac{1}{2}(\mathbf{A}-\mathbf{B})$$

(b) Multiplying the first equation by $\mathbf{B}$ and subtracting off the second gives

$$\mathbf{X} = (\mathbf{B}-\mathbf{I})^{-1}(\mathbf{BA}-\mathbf{C})$$

Subtracting the second equation from the first gives

$$\mathbf{Y} = (\mathbf{I}-\mathbf{B})^{-1}(\mathbf{A}-\mathbf{C})$$

(c) Solving the first equation for $\mathbf{Y}$ gives

$$\mathbf{Y} = \mathbf{C}^{-1}(\mathbf{D}-\mathbf{X})$$

Substituting for $\mathbf{Y}$ in the first equation gives

$$\mathbf{X} = \left(\mathbf{A}^{-1}-\mathbf{C}^{-1}\right)^{-1}\left(\mathbf{A}^{-1}\mathbf{B}-\mathbf{C}^{-1}\mathbf{D}\right)$$

Substituting for $\mathbf{X}$ in the first equation gives

$$\mathbf{Y} = \mathbf{A}^{-1}\mathbf{B}-\mathbf{A}^{-1}\left(\mathbf{A}^{-1}-\mathbf{C}^{-1}\right)^{-1}\left(\mathbf{A}^{-1}\mathbf{B}-\mathbf{C}^{-1}\mathbf{D}\right)$$

(d) Solving the second equation for $\mathbf{Y}$ gives

$$\mathbf{Y} = \mathbf{E}^{-1}(\mathbf{F}-\mathbf{DX})$$

Substituting in the first equation then gives

$$\mathbf{X} = \left(\mathbf{B}^{-1}\mathbf{A}-\mathbf{E}^{-1}\mathbf{D}\right)^{-1}\left(\mathbf{B}^{-1}\mathbf{C}-\mathbf{E}^{-1}\mathbf{F}\right)$$

Solving the first equation for $\mathbf{X}$ gives

$$\mathbf{X} = \mathbf{A}^{-1}(\mathbf{C}-\mathbf{BY})$$

Substituting in the second equation then gives

$$\mathbf{Y} = \left(\mathbf{A}^{-1}\mathbf{B}-\mathbf{D}^{-1}\mathbf{E}\right)^{-1}\left(\mathbf{A}^{-1}\mathbf{C}-\mathbf{D}^{-1}\mathbf{F}\right)$$

(e) Solving the second equation for $\mathbf{Y}$ gives

$$\mathbf{Y} = (\mathbf{F}-\mathbf{XD})\,\mathbf{E}^{-1}$$

Substituting back in the first equation and solving for $\mathbf{X}$ gives

$$\mathbf{X} = \left(\mathbf{CB}^{-1}-\mathbf{FE}^{-1}\right)\left(\mathbf{AB}^{-1}-\mathbf{DE}^{-1}\right)^{-1}$$

Solving the first equation for $\mathbf{X}$ gives

$$\mathbf{X} = (\mathbf{C}-\mathbf{YB})\,\mathbf{A}^{-1}$$

Substituting back in the second equation and solving for $\mathbf{Y}$ gives

$$\mathbf{Y} = \left(\mathbf{CA}^{-1}-\mathbf{FD}^{-1}\right)\left(\mathbf{BA}^{-1}-\mathbf{ED}^{-1}\right)^{-1}$$

9. (a) Let

$$D = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} = \det \mathbf{AB}$$

then

$$\begin{aligned}
D &= \underbrace{\begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix}}_{\text{rule V, section 1.4}} \\
&= \begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{vmatrix} + \begin{vmatrix} a_{11}b_{11} & a_{12}b_{22} \\ a_{21}b_{11} & a_{22}b_{22} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{11}b_{12} \\ a_{22}b_{21} & a_{21}b_{12} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{vmatrix} \\
&= b_{11}b_{12} \underbrace{\begin{vmatrix} a_{11} & a_{11} \\ a_{21} & a_{21} \end{vmatrix}}_{\text{rule IV, section 1.4}} + b_{11}b_{22} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + b_{21}b_{12} \begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} + b_{21}b_{22} \underbrace{\begin{vmatrix} a_{12} & a_{12} \\ a_{22} & a_{22} \end{vmatrix}}_{\text{rule IV, section 1.4}} \\
&= b_{11}b_{22} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} - b_{21}b_{12} \underbrace{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}_{\text{rule II, section 1.4}} \\
&= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} (b_{11}b_{22} - b_{21}b_{12}) \\
&= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}
\end{aligned}$$

(b) For 3×3 matrices the determinant $D = \det \mathbf{AB}$ is given by

$$D = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{vmatrix}$$

Although the approach taken to solve part (a) of the problem can be applied to the 3×3 case as well, it will be cleaner to follow an alternative derivation, as already for the 3×3 case, writing out all the steps explicitly will be very lengthy and hence, becomes unmanageable. Instead let us use the alternative way of writing the determinant introduced in problem 16 of section 1.5. Applying this notation to D gives

$$\begin{aligned}
D &= \sum \epsilon_{j_1 j_2 j_3} \sum_{k=1}^{3} \sum_{l=1}^{3} \sum_{m=1}^{3} a_{1k} b_{kj_1} a_{2l} b_{lj_2} a_{3m} b_{mj_3} \\
&= \sum_{k=1}^{3} \sum_{l=1}^{3} \sum_{m=1}^{3} a_{1k} a_{2l} a_{3m} \sum \epsilon_{j_1 j_2 j_3} b_{kj_1} b_{lj_2} b_{mj_3}
\end{aligned}$$

where the outer, respectively inner sum is over all permutations (j_1, j_2, j_3) of $(1, 2, 3)$ and $\epsilon_{j_1 j_2 j_3}$ is 1 for a permutation which is even (even number of two index interchanges) and is -1 for an odd number of index interchanges. This is equivalent to recursively applying rule V from section 1.4 and writing out the resulting determinants explicitly. For instance, for $k = l = m = 1$ we have

$$\begin{aligned}\sum \epsilon_{j_1 j_2 j_3} a_{11} b_{1j_1} a_{21} b_{1j_2} a_{31} b_{1j_3} &= a_{11}b_{11}a_{21}b_{12}a_{31}b_{13} - a_{11}b_{12}a_{21}b_{11}a_{31}b_{13} \\ &\quad + a_{11}b_{13}a_{21}b_{11}a_{31}b_{12} - a_{11}b_{13}a_{21}b_{12}a_{31}b_{11} \\ &\quad + a_{11}b_{13}a_{21}b_{11}a_{31}b_{12} - a_{11}b_{11}a_{21}b_{13}a_{31}b_{12} \\ &= 0\end{aligned}$$

Note that for fixed $k, l, m \in \{1, 2, 3\}$ the inner sum $\sum \epsilon_{j_1 j_2 j_3} b_{kj_1} b_{lj_2} b_{mj_3}$ is the determinant of a matrix composed of rows k, l, m of $\mathbf{B}$. Hence, it follows from rule IV of section 1.4 that this expression will vanish unless k, l, m are pairwise distinct, as is also demonstrated in the example above for $k = l = m = 1$. Thus, the total sum may be written as

$$D = \sum_{k \neq lm}^{3} \sum_{l \neq km}^{3} \sum_{m \neq kl}^{3} a_{1k} a_{2l} a_{3m} \sum \epsilon_{j_1 j_2 j_3} b_{kj_1} b_{lj_2} b_{mj_3}$$

for k, l, m a positive integer smaller than or equal to 3. This last expression then becomes

$$(\det \mathbf{A}) \det \mathbf{B} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

(c) In the most general $n \times n$ case the determinant $D = \det \mathbf{AB}$ is given by

$$D = \begin{vmatrix} a_{11}b_{11} + \dots + a_{1n}b_{n1} & \dots & a_{11}b_{1n} + \dots + a_{1n}b_{nn} \\ \vdots & & \vdots \\ a_{n1}b_{11} + \dots + a_{nn}b_{n1} & \dots & a_{n1}b_{1n} + \dots + a_{nn}b_{nn} \end{vmatrix}$$

Using the permutation notation, this may also be written as

$$\begin{aligned} D &= \sum \epsilon_{j_1 j_2 \dots j_n} \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} a_{1k_1} b_{k_1 j_1} a_{2k_2} b_{k_2 j_2} \dots a_{nk_n} b_{k_n j_n} \\ &= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} a_{1k_1} a_{2k_2} \dots a_{nk_n} \sum \epsilon_{j_1 j_2 \dots j_n} b_{k_1 j_1} b_{k_2 j_2} \dots b_{k_n j_n} \end{aligned}$$

where the outer, respectively inner sum is over all permutations $(j_1, \dots, j_n)$ of $(1, 2, \dots, n)$ and $\epsilon_{j_1 j_2 \dots j_n}$ is 1 for a permutation which is even (even number of two index interchanges) and is -1 for an odd number of index interchanges. This is equivalent to recursively applying rule V from section 1.4 and writing out the resulting determinants explicitly. Note that for fixed $k_1, \dots, k_n \in \{1, \dots, n\}$ the inner sum $\sum \epsilon_{j_1 j_2 \dots j_n} b_{k_1 j_1} b_{k_2 j_2} \dots b_{k_n j_n}$ is the determinant of a matrix composed of rows $k_1, \dots, k_n$ of $\mathbf{B}$. Hence, it follows from rule IV of section 1.4 that this expression will vanish unless the k_i are pairwise distinct. Thus, the total sum may be written as

$$D = \left(\sum_{k_1 \neq k_2 k_3 \dots k_n}^{n} \sum_{k_2 \neq k_1 k_3 \dots k_n}^{n} \cdots \sum_{k_n \neq k_1 \dots k_{n-2} k_{n-1}}^{n} a_{1k_1} a_{2k_2} \dots a_{nk_n} \right) \times \sum \epsilon_{j_1 j_2 \dots j_n} b_{k_1 j_1} b_{k_2 j_2} \dots b_{k_n j_n}$$

for $k_1, \dots, k_n$ a positive integer smaller than or equal to n. This last expression then becomes

$$(\det \mathbf{A}) \det \mathbf{B} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \begin{vmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{vmatrix}$$

10. (a) The matrix product $\mathbf{AB}$, where $\mathbf{A} = \text{col}(u_1, \dots, u_n)$ and $\mathbf{B} = (v_1, \dots, v_n)$ is given by

$$\mathbf{AB} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \dots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \dots & u_2 v_n \\ \vdots & \vdots & & \vdots \\ u_n v_1 & u_n v_2 & \dots & u_n v_n \end{bmatrix}$$

It then follows directly from rule IV of section 1.4 that $\det \mathbf{AB} = 0$ and hence, the matrix $\mathbf{AB}$ is singular. Alternatively, we can derive the same outcome by completing both $\mathbf{A}$ and $\mathbf{B}$ to $n \times n$ matrices by adding extra columns and rows of zeros. Let $\mathbf{A}_1$ and $\mathbf{B}_1$ be these expanded matrices. Then $\mathbf{AB} = \mathbf{A}_1 \mathbf{B}_1$, since the added zeros contribute nothing to the product. It follows directly that the determinant of both $\mathbf{A}_1$ and $\mathbf{B}_1$ is zero (expanding the determinant along any row or column gives 0). Hence, $\mathbf{A}_1 \mathbf{B}_1 = \mathbf{AB}$ is singular, since $\det \mathbf{A}_1 \mathbf{B}_1 = (\det \mathbf{A}_1) \det \mathbf{B}_1 = 0$.

(b) The matrix product $\mathbf{BA}$ takes the form

$$\mathbf{BA} = \underbrace{\begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}}_{1 \times n} \underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}}_{n \times 1} = \underbrace{\begin{bmatrix} v_1 u_1 + v_2 u_2 + \dots + v_n u_n \end{bmatrix}}_{1 \times 1} = v_1 u_1 + v_2 u_2 + \dots + v_n u_n$$

11. (a) Interchanging any two rows of the identity matrix results in an $n \times n$ symmetric matrix $\mathbf{P}_{ij}$. It is because of this symmetry and the fact that there is only one nonzero element appearing in each row (and column) that $\mathbf{P}_{ij}^2 = \mathbf{I}$. The matrix $\mathbf{P}_{ij}$ is an example of a symmetric orthogonal matrix; a matrix which has the property that $\mathbf{P}_{ij} = \mathbf{P}_{ij}^{-1}$.

(b) The $n \times n$ identity matrix $\mathbf{I}$ may be written as

$$\mathbf{I} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{bmatrix}$$

Here the $\mathbf{e}_k^{th}$ row vector has value 1 for the k^{th} entree and zeros everywhere else. Since $\mathbf{I}$ is symmetric it follows directly that we could have written $\mathbf{I}$ as

$$\mathbf{I} = \begin{bmatrix} \mathrm{col}(\mathbf{e}_1) & \mathrm{col}(\mathbf{e}_2) & \ldots & \mathrm{col}(\mathbf{e}_n) \end{bmatrix}$$

where $\mathrm{col}(\mathbf{e}_i)$ is identical to $\mathbf{e}_i$, the only difference being that it is arranged as a column vector. We can now form the matrix product $\mathbf{P}_{ij}\bar{\mathbf{P}}_{ij}$, where $\mathbf{P}_{ij}$ is the permutation matrix obtained from $\mathbf{I}$ by swapping the i^{th} and j^{th} row and $\bar{\mathbf{P}}_{ij}$ is the permutation matrix obtained from $\mathbf{I}$ by swapping the i^{th} and j^{th} column. The kl^{th} $(1 \leq k, l \leq n)$ entree of this matrix product is the result from taking the inner product $\mathbf{e}_k \cdot \mathrm{col}(\mathbf{e}_l)$, which is 1 when $k = l$ and 0 otherwise. This proves that $\mathbf{P}_{ij}\bar{\mathbf{P}}_{ij} = \mathbf{I}$. Since $\mathbf{P}_{ij}^2 = \mathbf{P}_{ij}\mathbf{P}_{ij} = \mathbf{I}$, it follows directly that $\bar{\mathbf{P}}_{ij} = \mathbf{P}_{ij} = \mathbf{P}_{ij}^{-1}$. In conclusion, swapping the i^{th} and j^{th} row of $\mathbf{I}$ is equivalent to swapping the i^{th} and j^{th} column of $\mathbf{I}$ and hence both operations produce the same permutation matrix $\mathbf{P}_{ij}$.

(c) Let $\mathbf{P}_{ij}$ be the permutation matrix corresponding to swapping the i^{th} and j^{th} row of $\mathbf{I}$. Then the matrix product $\mathbf{P}_{ij}\mathbf{A} = \bar{\mathbf{A}}$ may be interpreted as taking the dot product between each row of $\mathbf{P}_{ij}$ and each column of $\mathbf{A}$:

$$\bar{a}_{kl} = \mathbf{e}_k \cdot \mathrm{col}(\mathbf{a}_l)$$

The row vector $\mathbf{e}_k$ will contain a 1 at its k^{th} entree and zeros otherwise for the rows of $\mathbf{P}_{ij}$ that were not swapped and a 1 at its i^{th}, respectively j^{th} entree and zeros otherwise for the two rows of $\mathbf{P}_{ij}$ that were swapped. Hence, the dot product will simply pick out the a_{kl}th element of $\mathbf{A}$, where the row index k maps onto a specific permutation (i.e. re-ordering of two row vectors $\mathbf{e}_i$ and $\mathbf{e}_j$ of $\mathbf{I}$). The matrix product $\mathbf{A}\mathbf{P}_{ij} = \bar{\mathbf{A}}$ in turn may be interpreted as taking the dot product between each row of $\mathbf{A}$ and each column of $\mathbf{P}_{ij}$:

$$\bar{a}_{kl} = \mathbf{a}_k \cdot \mathrm{col}(\mathbf{e}_l)$$

Again, the dot product will pick out the a_{kl}th element of $\mathbf{A}$, but now the column index l maps onto a specific permutation (i.e. re-ordering of two column vectors $\mathrm{col}(\mathbf{e}_i)$ and $\mathrm{col}(\mathbf{e}_j)$ of $\mathbf{I}$).

12. To show that $\mathbf{P}_{ij}$ is nonsingular is equivalent to proving that $\mathbf{P}_{ij}$ has an inverse. Since $\mathbf{P}_{ij}^2 = \mathbf{P}_{ij}\mathbf{P}_{ij} = \mathbf{I}$, it follows directly that $\mathbf{P}_{ij} = \mathbf{P}_{ij}^{-1}$ and hence, that $\mathbf{P}_{ij}$ is nonsingular.

Section 1.10

1. (a) Gaussian elimination gives

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

and so $x = -6$, $y = 12$, $z = 5$.

(b)

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

so $x = 3$, $y = 1$, $z = -1$.

(c)

$$\left[\begin{array}{rrr|r} 1 & 1 & -1 & 3 \\ 2 & 0 & 1 & 4 \\ 0 & 1 & 2 & 2 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & -1 & 3 \\ 0 & 1 & -3/2 & 1 \\ 0 & 1 & 2 & 2 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & -1 & 3 \\ 0 & 1 & -3/2 & 1 \\ 0 & 0 & 1 & 2/7 \end{array}\right]$$

so $x = 13/7$, $y = 10/7$, $z = 2/7$.

(d)

$$\left[\begin{array}{rrr|r} -2 & 3 & -4 & -3 \\ 2 & 1 & 3 & 6 \\ 2 & 1 & 5 & 8 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & -3/2 & 2 & 3/2 \\ 0 & 1 & -1/4 & 3/4 \\ 0 & 1 & 1/4 & 5/4 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & -3/2 & 2 & 3/2 \\ 0 & 1 & -1/4 & 3/4 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

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

2. (a)

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 5 \\ 3 & 0 & 2 \end{vmatrix} \to -3 \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & -6 & -1 \end{vmatrix} \to -3 \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -7 \end{vmatrix} = 21$$

(b)

$$\begin{vmatrix} 1 & 2 & 1 & 2 \\ 3 & 1 & 3 & 1 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \end{vmatrix} \to - \begin{vmatrix} 1 & 2 & 1 & 2 \\ 0 & -5 & 0 & -5 \\ 0 & -3 & -1 & -2 \\ 0 & 0 & 1 & -1 \end{vmatrix} \to 5 \begin{vmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{vmatrix} \to -5 \begin{vmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$

3. (a)

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

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

so $x = 2$, $y - k = -1$, $z = k$, where k can be any real number.

(b)

$$\left[\begin{array}{rrrr|r} 2 & -1 & 1 & -1 & 7 \\ 1 & 1 & -1 & 1 & 5 \\ 1 & 3 & -1 & -1 & 5 \\ 1 & -1 & -1 & 2 & 4 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 1 & -1 & 1 & 5 \\ 0 & -2 & 0 & 1 & -1 \\ 0 & -3 & 3 & -3 & -3 \\ 0 & 2 & 0 & -2 & 0 \end{array}\right]$$

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

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

so $x = 4$, $y = 1$, $z = 1$, $w = 1$.

4. (a)

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 4 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 7/4 & 1/2 & -1/4 & 0 \\ 0 & -6 & -10 & -4 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 7/4 & 1/2 & -1/4 & 0 \\ 0 & 0 & 1 & -2 & -3 & 2 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 1 & 2 & 0 & 7 & 9 & -6 \\ 0 & 1 & 0 & 4 & 5 & -7/2 \\ 0 & 0 & 1 & -2 & -3 & 2 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & 4 & 5 & -7/2 \\ 0 & 0 & 1 & -2 & -3 & 2 \end{array}\right]$$

(b)

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

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

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

(c)

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

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

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

$$\to \left[\begin{array}{rrrr|rrrr} 1 & 0 & 0 & -1/3 & 0 & -1/3 & 2/3 & 0 \\ 0 & 1 & 0 & 5/3 & 1 & -1/3 & -1/3 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3/7 & -2/7 & -2/7 & 3/7 \end{array}\right]$$

$$\to \left[\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 1/7 & -3/7 & 4/7 & 1/7 \\ 0 & 1 & 0 & 0 & 2/7 & 1/7 & 1/7 & -5/7 \\ 0 & 0 & 1 & 0 & -4/7 & 5/7 & -2/7 & 3/7 \\ 0 & 0 & 0 & 1 & 3/7 & -2/7 & -2/7 & 3/7 \end{array}\right]$$

5. (a)

$$\left[\begin{array}{rrr|r} 2 & -1 & 1 & -3 \\ 1 & 2 & -1 & 1 \\ 5 & -5 & 4 & 8 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 0 & -5 & 3 & 1 \\ 0 & -15 & 9 & 3 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 0 & 1 & -3/5 & -1/5 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Hence,

$$x = -\frac{2}{5}(3t-1) + t + 1 = \frac{1}{5}(7-t), \quad y = \frac{1}{5}(3t-1), \quad z = t, \quad -\infty < t < \infty$$

(b)

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

Hence, there is no solution.

(c)

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

$$\to \left[\begin{array}{rrrr|r} 1 & 0 & 1/3 & 1/3 & 0 \\ 0 & 1 & -2/3 & -2/3 & 0 \\ 0 & 0 & 1 & -1/8 & 0 \\ 0 & 0 & 8/3 & -1/3 & 0 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 3/8 & 0 \\ 0 & 1 & 0 & -3/4 & 0 \\ 0 & 0 & 1 & -1/8 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence,

$$x = -3t, \quad y = 6t, \quad z = t, \quad w = 8t, \quad -\infty < t < \infty$$

(d)

$$\left[\begin{array}{rrrr|r} 1 & 1 & -2 & 3 & 0 \\ 2 & -1 & 1 & -1 & 0 \\ 3 & 0 & -1 & 2 & 0 \\ 5 & 2 & -5 & 8 & 0 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 1 & -2 & 3 & 0 \\ 0 & 1 & -5/3 & 7/3 & 0 \\ 0 & -3 & 5 & -7 & 0 \\ 0 & -3 & 5 & -7 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{rrrr|r} 1 & 0 & -1/3 & 2/3 & 0 \\ 0 & 1 & -5/3 & 7/3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Hence,

$$x = t - 2u, \quad y = 5t - 7u, \quad z = 3t, \quad w = 3u, \quad -\infty < t, u < \infty$$

6. (a) Assuming that Gaussian elimination can be carried out, the system of two equations in three unknowns will have the matrix form

$$\left[\begin{array}{ccc|c} a_1 & b_1 & c_1 & k_1 \\ a_2 & b_2 & c_2 & k_2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Then it follows that all the solutions of the system are given by the parametric equations $x = x_1 + at, \quad y = y_1 + bt, \quad z = z_1 + ct$ where

$$x_1 = \frac{b_2k_1 - b_1k_2}{a_1b_2 - a_2b_1} \qquad y_1 = \frac{a_2k_1 - a_1k_2}{a_2b_1 - a_1b_2} \qquad z_1 = 0$$

$$a = -\frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1} \qquad b = -\frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2} \qquad c = 1$$

(b) If the two equations have no solution the planes are parallel and non-intersecting. In the case elimination leads to a second equation $0 = 0$, the two planes are parallel and intersecting.

7. If two straight lines in space have points in common there are only two possible scenarios. The first one is that they cross each other at an arbitrary angle and hence, have only one point in common. Geographically, this would correspond to the four planes intersecting at a single point. In this case the set of solutions is an ordered triple (x, y, z) and the fourth equation is redundant, as it is just a linear combination of two of the other three equations. In the second scenario the two lines would coincide. This would correspond to four different planes intersecting along the same line (a special case being where planes coincide in pairs and the pairs intersect). In this case two of the four equations may be made redundant as they simply are a linear combination of the other two equations.

8. (a) The matrix product $\mathbf{CA}$ may be interpreted as taking an inner product of each row of $\mathbf{C}$ with each column of $\mathbf{A}$. Hence, for the h^{th} row of $\mathbf{C}$ and the j^{th} column of $\mathbf{A}$ we will get $\mathbf{c}_h \cdot \mathrm{col}(\mathbf{a}_j) = c_{hh}a_{hj}$, since all other terms in this inner product will be zero due to the fact that c_{hh} is the only non-zero entree in $\mathbf{c}_h$. The h^{th} row of the matrix product $\mathbf{CA}$ is thus given by $(\mathbf{CA})_h = \left[c_{hh}a_{h1} \ldots c_{hh}a_{hn}\right] = \left[\lambda a_{h1} \ldots \lambda a_{hn}\right]$. Replacing c_{hh} by 1 and applying the same reasoning, it is then easy to see that the inner products between the other rows of $\mathbf{C}$, i.e. $\mathbf{c}_i$ with the columns $\mathrm{col}(\mathbf{a}_j)$ of $\mathbf{A}$ simply reproduce the a_{ij}th matrix element of $\mathbf{A}$.

   (b) Just as for part (a), the matrix product $\mathbf{BA}$ may be interpreted as taking an inner product of each row of $\mathbf{B}$ with each column of $\mathbf{A}$. Hence, for the h^{th} row of $\mathbf{B}$ (i.e. the row where $\mathbf{B}$ differs from $\mathbf{I}$ only in that $b_{hk} = 1$ in addition to $b_{hh} = 1$) and the j^{th} column of $\mathbf{A}$ we will get $\mathbf{b}_h \cdot \mathrm{col}(\mathbf{a}_j) = a_{hj} + a_{kj}$ where $k \neq h$. The h^{th} row of the matrix product $\mathbf{BA}$ is thus given by $(\mathbf{BA})_h = \left[a_{h1} + a_{k1} \ldots a_{hn} + a_{kn}\right]$. Since the other rows of $\mathbf{B}$ are assumed to be equal to the identity matrix $\mathbf{I}$, the remaining rows of $\mathbf{BA}$ will simply consist of the corresponding rows of $\mathbf{A}$.

   (c) For $\mathbf{B}$ and $\mathbf{C}$ to be singular means their inverse should exist. In order to find the two inverse matrices we can apply the same elimination techniques that are represented by the matrices themselves when they are multiplied with another matrix of appropriate dimensions. For $\mathbf{C}$ this implies multiplying the h^{th} row of $\mathbf{C}$ by $1/\lambda$ so that $\mathbf{CC}^{-1} = \mathbf{I}$. This shows that $\mathbf{C}$ is indeed non-singular. For $\mathbf{B}^{-1}$ consider the same matrix $\mathbf{B}$ with the exception that the h^{th} row contains a -1 instead of 1 as its j^{th} entree. Then $\mathbf{BB}^{-1} = \mathbf{I}$, which thus proves the inverse to $\mathbf{B}$ exists and hence, $\mathbf{B}$ is also non-singular.

9. When solving the n systems we are trying to find unique columns $\mathbf{k}_j$ such that $\mathbf{Ak}_j = \mathbf{q}_j$. To find a particular $\mathbf{k}_j$, we append the $\mathbf{q}_j$th column to the matrix $\mathbf{A}$ and perform Gauss-Jordan elimination on the augmented matrix $\left[\begin{array}{c|c} \mathbf{A} & \mathbf{q}_j \end{array}\right]$ in order to get

the matrix $\left[\,\mathbf{I}\,\middle|\,\mathbf{k}_j\,\right]$. Now when applying the Gauss-Jordan procedure to the augmented matrix $\left[\,\mathbf{A}\,\middle|\,\mathbf{I}\,\right]$ we are doing n simultaneous reductions of exactly the same type. The same steps are performed in each reduction $\left[\,\mathbf{A}\,\middle|\,\mathbf{q}_j\,\right] \to \left[\,\mathbf{I}\,\middle|\,\mathbf{k}_j\,\right]$, since we find the same pivots in each row and the first n columns. Hence, $\left[\,\mathbf{A}\,\middle|\,\mathbf{I}\,\right] = \left[\,\mathbf{A}\,\middle|\,\mathbf{q}_1,\ldots,\mathbf{q}_n\,\right] \to \left[\,\mathbf{I}\,\middle|\,\mathbf{k}_1,\ldots,\mathbf{k}_n\,\right] = \left[\,\mathbf{I}\,\middle|\,\mathbf{H}\,\right]$.

Section 1.11

1. Let all given matrices be denoted by the matrix $\mathbf{A}$. Using equation (1.78) the eigenvalues of the matrix $\mathbf{A}$ are then given by

   (a)

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{v} = \mathbf{0} \quad \Longrightarrow \quad \begin{bmatrix} 3-\lambda & 1 \\ 4 & 3-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad \begin{vmatrix} 3-\lambda & 1 \\ 4 & 3-\lambda \end{vmatrix} = 0$$

   Expanding the determinant then results in the characteristic equation

$$\begin{aligned} (3-\lambda)^2 - 4 &= 0 \\ \lambda^2 - 6\lambda + 5 &= \\ (\lambda-1)(\lambda-5) &= \end{aligned}$$

   resulting in the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 5$. Inserting $\lambda = \lambda_1$ in the matrix equation above and solving for v_1 and v_2 using Gaussian elimination gives

$$\left[\begin{array}{cc|c} 3-\lambda_1 & 1 & 0 \\ 4 & 3-\lambda_1 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 4 & 2 & 0 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1/2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

   from which follows $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}1 & -2\end{bmatrix}\right)$. Repeating the same steps with $\lambda = \lambda_2$ gives $\mathbf{v}_2 = \operatorname{col}\left(k\begin{bmatrix}1 & 2\end{bmatrix}\right)$.

   (b) $\lambda_1 = 0$, $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}3 & -1\end{bmatrix}\right)$ and $\lambda_2 = 7$, $\mathbf{v}_2 = \operatorname{col}\left(k\begin{bmatrix}1 & 2\end{bmatrix}\right)$.

   (c)

$$\begin{aligned} \begin{vmatrix} -\lambda & 1 & -2 \\ 2 & 1-\lambda & 0 \\ 4 & -2 & 5-\lambda \end{vmatrix} &= 0 \\ \lambda^3 - 6\lambda^2 + 11\lambda - 6 &= 0 \\ (\lambda-1)(\lambda-2)(\lambda-3) &= 0 \end{aligned}$$

   Eigenvalues: $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$. Eigenvectors: $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}0 & 2 & 1\end{bmatrix}\right)$, $\mathbf{v}_2 = \operatorname{col}\left(k\begin{bmatrix}1 & 2 & 0\end{bmatrix}\right)$ and $\mathbf{v}_3 = \operatorname{col}\left(k\begin{bmatrix}1 & 1 & -1\end{bmatrix}\right)$.

(d)

$$\begin{vmatrix} 5-\lambda & -2 & 8 \\ -4 & -\lambda & -5 \\ -4 & 2 & -7-\lambda \end{vmatrix} = 0$$
$$\lambda^3 + 2\lambda^2 - \lambda - 2 =$$
$$(\lambda - 1)(\lambda + 1)(\lambda + 2) =$$

Eigenvalues: $\lambda_1 = 1$, $\lambda_2 = -1$ and $\lambda_3 = -2$. Eigenvectors: $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}3 & -2 & -2\end{bmatrix}\right)$, $\mathbf{v}_2 = \operatorname{col}\left(k\begin{bmatrix}-1 & 1 & 1\end{bmatrix}\right)$ and $\mathbf{v}_3 = \operatorname{col}\left(k\begin{bmatrix}2 & -1 & -2\end{bmatrix}\right)$.

2. (a)

$$\mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \qquad \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

(b)

$$\mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} \qquad \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 7 \end{bmatrix}$$

(c)

$$\mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \qquad \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

(d)

$$\mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix} = \begin{bmatrix} 3 & -1 & 2 \\ -2 & 1 & -1 \\ -2 & 1 & -2 \end{bmatrix} \qquad \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

3. (a)

$$\begin{vmatrix} 1-\lambda & -1 \\ 4 & 1-\lambda \end{vmatrix} = 0$$
$$\lambda^2 - 2\lambda + 5 =$$
$$(\lambda - 1 - 2i)(\lambda - 1 + 2i) =$$

Eigenvalues: $\lambda_1 = 1+2i$ and $\lambda_2 = 1-2i$. Eigenvectors: $\mathbf{v}_1 = \operatorname{col}\left(c\begin{bmatrix}1 & -2i\end{bmatrix}\right)$ and $\mathbf{v}_2 = \operatorname{col}\left(c\begin{bmatrix}1 & 2i\end{bmatrix}\right)$, where c is a purely imaginary constant free of choice and

$$\mathbf{B} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1+2i & 0 \\ 0 & 1-2i \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2i & 2i \end{bmatrix}$$

(b) Eigenvalues $\lambda_1 = 4+i$ and $\lambda_2 = 4-i$. Eigenvectors: $\mathbf{v}_1 = \operatorname{col}\left(c\begin{bmatrix}2 & -1-i\end{bmatrix}\right)$ and $\mathbf{v}_2 = \operatorname{col}\left(c\begin{bmatrix}2 & -1+i\end{bmatrix}\right)$, where c is a complex constant free of choice and

$$\mathbf{B} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 4+i & 0 \\ 0 & 4-i \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \begin{bmatrix} 2 & 2 \\ -1-i & -1+i \end{bmatrix}$$

(c) Eigenvalues: $\lambda_1 = 0$, $\lambda = i$ and $\lambda_3 = -i$. Eigenvectors: $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}1 & 0 & 0\end{bmatrix}\right)$, $\mathbf{v}_2 = \operatorname{col}\left(c\begin{bmatrix}0 & 2 & 1-i\end{bmatrix}\right)$ and $\mathbf{v}_3 = \operatorname{col}\left(k\begin{bmatrix}0 & 2 & 1+i\end{bmatrix}\right)$, where k is a real constant and c a complex constant free of choice and

$$\mathbf{B} = \begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix} = \begin{bmatrix}0 & 0 & 0\\ 0 & i & 0\\ 0 & 0 & -i\end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 2\\ 0 & 1-i & 1+i\end{bmatrix}$$

4. (a)

$$\begin{vmatrix}1-\lambda & 0 & 0\\ 0 & 1-\lambda & 0\\ 0 & 0 & 1-\lambda\end{vmatrix} = 0$$
$$(1-\lambda)^3 =$$

Eigenvalues: $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Eigenvectors: all non-zero vectors.

(b)

$$\begin{vmatrix}-\lambda & & \\ & \ddots & \\ & & -\lambda\end{vmatrix} = 0$$
$$\lambda^{44} =$$

Eigenvalues: $\lambda_1 = \lambda_2 = \cdots = \lambda_{44} = 0$. Eigenvectors: all non-zero vectors.

(c)

$$\begin{vmatrix}3-\lambda & -4\\ 4 & -5-\lambda\end{vmatrix} = 0$$
$$1 + 2\lambda + \lambda^2 =$$
$$(1+\lambda)^2 =$$

Eigenvalues: $\lambda_1 = \lambda_2 = -1$. Eigenvectors: $\mathbf{v}_1 = \mathbf{v}_2 = \operatorname{col}\left(k\begin{bmatrix}1 & 1\end{bmatrix}\right)$, where $-\infty < k < \infty$.

(d)

$$\begin{vmatrix}-\lambda & 1 & -2\\ -6 & 5-\lambda & -4\\ 0 & 0 & 3-\lambda\end{vmatrix} = 0$$
$$18 - 21\lambda + 8\lambda^2 - \lambda^3 =$$
$$(2-\lambda)(3-\lambda)^2 =$$

Eigenvalues: $\lambda_1 = 2$ and $\lambda_2 = \lambda_3 = 3$. Eigenvectors: $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}1 & 2 & 0\end{bmatrix}\right)$ and $\mathbf{v}_2 = \mathbf{v}_3 = \operatorname{col}\left(a\begin{bmatrix}1 & 3 & 0\end{bmatrix}\right) + \operatorname{col}\left(b\begin{bmatrix}-2 & 0 & 1\end{bmatrix}\right)$, where $-\infty < a, b, k < \infty$.

5. (a) Let $\lambda \neq \mu$ Solving the eigenvalue equation for λ gives

$$(\mathbf{B} - \lambda \mathbf{I})\,\mathbf{v} = \left(\begin{bmatrix}\lambda & 0\\ 0 & \mu\end{bmatrix} - \lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & \mu-\lambda\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

Next, Gauss-Jordan elimination gives

$$\left[\begin{array}{cc|c}0 & 0 & 0\\ 0 & \mu-\lambda & 0\end{array}\right] \rightarrow \left[\begin{array}{cc|c}0 & \mu-\lambda & 0\\ 0 & 0 & 0\end{array}\right] \rightarrow \left[\begin{array}{cc|c}0 & 1 & 0\\ 0 & 0 & 0\end{array}\right]$$

Hence, $v_2 = 0$. Note that since v_1 doesn't appear anywhere the equation is satisfied for any v_1. Hence the eigenvector associated with the eigenvalue λ is $\mathbf{v}_1 = \operatorname{col}\left(k\begin{bmatrix}1 & 0\end{bmatrix}\right)$. Similarly, for μ we get

$$(\mathbf{B} - \mu \mathbf{I})\,\mathbf{v} = \left(\begin{bmatrix}\lambda & 0\\ 0 & \mu\end{bmatrix} - \mu\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \begin{bmatrix}\lambda-\mu & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

Next, Gauss-Jordan elimination gives

$$\left[\begin{array}{cc|c}\lambda-\mu & 0 & 0\\ 0 & 0 & 0\end{array}\right] \rightarrow \left[\begin{array}{cc|c}1 & 0 & 0\\ 0 & 0 & 0\end{array}\right]$$

Hence, $v_1 = 0$. Note that since v_2 doesn't appear anywhere the equation is satisfied for any v_2. Hence the eigenvector associated with the eigenvalue μ is $\mathbf{v}_2 = \operatorname{col}\left(k\begin{bmatrix}0 & 1\end{bmatrix}\right)$, where $-\infty < k < \infty$. Let $\lambda = \mu$. Then

$$(\mathbf{B} - \lambda \mathbf{I})\,\mathbf{v} = \left(\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix} - \lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

This equation is satisfied for all non-zero vectors $\mathbf{v} = \operatorname{col}\left(\begin{bmatrix}v_1 & v_2\end{bmatrix}\right)$.

(b) Let λ_1, λ_2 and λ_3 be distinct eigenvalues. Inserting for λ_1 in the eigenvalue equation gives

$$\begin{aligned}(\mathbf{B} - \lambda_1 \mathbf{I})\,\mathbf{v} &= \left(\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix} - \lambda_1\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}\\ &= \begin{bmatrix}0 & 0 & 0\\ 0 & \lambda_2-\lambda_1 & 0\\ 0 & 0 & \lambda_3-\lambda_1\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}\\ &= \mathbf{0}\end{aligned}$$

Next Gauss-Jordan elimination gives

$$\left[\begin{array}{ccc|c}0 & 0 & 0 & 0\\ 0 & \lambda_2-\lambda_1 & 0 & 0\\ 0 & 0 & \lambda_3-\lambda_1 & 0\end{array}\right] \rightarrow \left[\begin{array}{ccc|c}0 & \lambda_2-\lambda_1 & 0 & 0\\ 0 & 0 & \lambda_3-\lambda_1 & 0\\ 0 & 0 & 0 & 0\end{array}\right] \rightarrow \left[\begin{array}{ccc|c}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

Hence, $v_2 = v_3 = 0$. Note that since v_1 doesn't appear anywhere the equation is satisfied for any v_1. The eigenvector associated with the eigenvalue λ_1 thus is $\operatorname{col}\left(c\begin{bmatrix}1 & 0 & 0\end{bmatrix}\right) = c\mathbf{e}_1$. Applying the same logic to λ_2 and λ_3 gives $\operatorname{col}\left(c\begin{bmatrix}0 & 1 & 0\end{bmatrix}\right) = c\mathbf{e}_2$ and $\operatorname{col}\left(c\begin{bmatrix}0 & 0 & 1\end{bmatrix}\right) = c\mathbf{e}_3$ respectively. Let $\lambda_1 = \lambda_2 \neq \lambda_3$. Inserting for λ_1 in the eigenvalue equation gives

$$\begin{aligned}
(\mathbf{B} - \lambda_1\mathbf{I})\,\mathbf{v} &= \left(\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_3\end{bmatrix} - \lambda_1\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}\\
&= \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \lambda_3 - \lambda_1\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}\\
&= \mathbf{0}
\end{aligned}$$

Next Gauss-Jordan elimination gives

$$\left[\begin{array}{ccc|c}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & \lambda_3 - \lambda_1 & 0\end{array}\right] \rightarrow \left[\begin{array}{ccc|c}0 & 0 & \lambda_3 - \lambda_1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right] \rightarrow \left[\begin{array}{ccc|c}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

Hence, $v_3 = 0$. Note that since v_1 and v_2 don't appear anywhere the equation is satisfied for any linear combination $\mathbf{v}_1 = \mathbf{v}_2 = c_1v_1 + c_2v_2 = \operatorname{col}\left(c_1\begin{bmatrix}1 & 0 & 0\end{bmatrix}\right) + \operatorname{col}\left(c_2\begin{bmatrix}0 & 1 & 0\end{bmatrix}\right) = c_1\mathbf{e}_1 + c_2\mathbf{e}_2$. Since λ_3 is unique we have as before $\mathbf{v}_3 = \operatorname{col}\left(c\begin{bmatrix}0 & 0 & 1\end{bmatrix}\right) = c\mathbf{e}_3$. Lastly, let $\lambda_1 = \lambda_2 = \lambda_3$. The eigenvalue equation now is of the form

$$\begin{aligned}
(\mathbf{B} - \lambda_1\mathbf{I})\,\mathbf{v} &= \left(\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_1\end{bmatrix} - \lambda_1\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}\\
&= \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}\\
&= \mathbf{0}
\end{aligned}$$

Since neither v_1, v_2 or v_3 appears in the equation above, any non-zero vector $\mathbf{v}$ satisfies it.

(c) Let $\mathbf{B} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ be an arbitrary size diagonal matrix and let each λ_k be unique. Then inserting for λ_k in the eigenvalue equation gives

$$(\mathbf{B} - \lambda_k\mathbf{I}) = \left(\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} - \lambda_k\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}\right)\begin{bmatrix}v_1\\ \vdots\\ v_n\end{bmatrix}$$

It should be clear that the matrix subtraction will result in an $n \times n$ matrix that has non-zero entrees along the diagonal everywhere except for the entree in the

k^{th} row and column only. Hence, there are $n-1$ equation of the type $v_i = 0$ for $i \neq k$ and one equation that is identically zero. Thus we may conclude that v_k is free of choice and hence, the eigenvectors associated with λ_k are all vectors $\mathbf{v}_k = \mathrm{col}\left(c\left[0\ldots,0,v_k,0,\ldots,0\right]\right)$, where c is some arbitrary non-zero constant. If the eigenvalue λ_k appears with multiplicity m in $\mathbf{B}$, it follows directly that the eigenvalue equation will give m rows that are identically zero and $l = n - m$ equations of the type $v_i = 0$ for $i \neq k_1, \ldots, k_m$. Now the eigenvectors associated with the eigenvalue λ_k may be written as a linear combination of m vectors

$$\mathbf{v}_k = \sum^{m} \mathrm{col}\left(c_m\left[0,\ldots,0,v_{k_m},0,\ldots,0\right]\right)$$

6. (a) The characteristic equation $p(\lambda)$ of the $n \times n$ matrix $\mathbf{A}$ is given by

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + \lambda^n = 0$$

Since the eigenvalues of $\mathbf{A}$ are the zeros of $p(\lambda)$ this implies that $p(\lambda)$ can be factorised as $p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\ldots(\lambda_n - \lambda) = 0$. Now consider the constant term $c_0 = p(0) = (\lambda_1 - 0)(\lambda_2 - 0)\ldots(\lambda_n - 0) = \lambda_1\lambda_2\ldots\lambda_n = \det(\mathbf{A})$. This proves that the product of the eigenvalues of $\mathbf{A}$ is equal to the determinant of $\mathbf{A}$.

(b) Since it is assumed $\mathbf{A}$ and $\mathbf{B}$ are similar, i.e. $\mathbf{A} = \mathbf{CBC}^{-1}$, we have

$$\begin{aligned}
\det(\mathbf{A} - \lambda\mathbf{I}) &= \det\left(\mathbf{CBC}^{-1} - \lambda\mathbf{I}\right) \\
&= \det\left(\mathbf{CBC}^{-1} - \lambda\mathbf{CIC}^{-1}\right) \\
&= \det\mathbf{C}\left(\mathbf{B} - \lambda\mathbf{I}\right)\mathbf{C}^{-1} \\
&= \det\mathbf{C}\det\left(\mathbf{B} - \lambda\mathbf{I}\right)\det\mathbf{C}^{-1} \\
&= \det\left(\mathbf{B} - \lambda\mathbf{I}\right)
\end{aligned}$$

since $\det\mathbf{C}\det\mathbf{C}^{-1} = 1$. Hence, $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic equation and eigenvalues and so it follows immediately from this and part (a) that $\det(\mathbf{A}) = p(0) = \det(\mathbf{B})$.

(c) Let us consider the coefficient c_{n-1} belonging to the λ^{n-1} term of $p(\lambda)$. As part (a) shows, this may be calculated by expanding $p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda) = (-1)^n\lambda^n + (-1)^n(\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots = 0$. From this we may infer that $c_{n-1} = (-1)^n(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$. Next, the same coefficient may be calculated by expanding $\det(\mathbf{A} - \lambda\mathbf{I})$:

$$\begin{aligned}
\det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} \\
&= (a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda) + c_2'\lambda^{n-2} + c_3'\lambda^{n-1} + \cdots + c_n
\end{aligned}$$

It is trivial to see that the terms other than the leading term are not proportional to λ^n nor λ^{n-1} and hence, are of no interest to us. The leading term is the product of the entrees along the diagonal of $\mathbf{A} - \lambda\mathbf{I}$ and when expanded takes the form.

$$(a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda) = (-1)^n \lambda^n + (-1)^n (a_{11} + a_{22} + \dots + a_{nn}) \lambda^{n-1} + \dots$$

By equating this result to the expansion of $p(\lambda)$ it then follows immediately that

$$\lambda_1 + \lambda_2 + \dots + \lambda_n = a_{11} + a_{22} + \dots + a_{nn}$$

(d) It was shown in part (b) that $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic equation $p(\lambda)$. In other words

$$\det(\mathbf{A} - \lambda\mathbf{I}) = p(\lambda) = \det(\mathbf{B} - \lambda\mathbf{I})$$

This implies that when expanding both $\det(\mathbf{A} - \lambda\mathbf{I})$ and $\det(\mathbf{B} - \lambda\mathbf{I})$, equality should hold on a term by term basis. Hence, from the equality of the λ^{n-1} term it follows directly that

$$a_{11} + a_{22} + \dots + a_{nn} = b_{11} + b_{22} + \dots + b_{nn}$$

7. The determinant of $\mathbf{A}$ is given by

$$\det \mathbf{A} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$$

Part (a) of problem 6 then implies that the product of the eigenvalues of some diagonal matrix similar to $\mathbf{A}$ has to be 1 also. However, this fact alone is not enough to conclude that the only matrix similar to $\mathbf{A}$ is $\mathbf{I}$. Now let $\mathbf{B}$ be a matrix similar to $\mathbf{A}$. Part (d) of problem 6 then introduces the additional constraint $b_{11} + b_{22} = 2$ along with the constraint $b_{11}b_{22} = 1$. This system of two equations in two unknowns is easily solved to give $b_{11} = b_{22} = 1$. Hence, it follows that the matrix $\mathbf{B}$ must be the identity matrix $\mathbf{I}$.

8. (a) Let $\mathbf{A}$ be an $n \times n$ matrix. It then follows immediately that $\mathbf{A}$ is similar to itself, since

$$\mathbf{A} = \mathbf{I}^{-1}\mathbf{A}\mathbf{I}$$

(b) Let $\mathbf{A} = \mathbf{D}^{-1}\mathbf{B}\mathbf{D}$ and $\mathbf{B} = \mathbf{E}^{-1}\mathbf{C}\mathbf{E}$. Substituting for $\mathbf{B}$ in the first equation then gives

$$\mathbf{A} = \mathbf{D}^{-1}\mathbf{E}^{-1}\mathbf{C}\mathbf{E}\mathbf{D} = (\mathbf{E}\mathbf{D})^{-1}\mathbf{C}\mathbf{E}\mathbf{D}$$

Let us denote the result of the matrix product $\mathbf{ED}$ by the matrix $\mathbf{F}$. Then $\mathbf{A} = \mathbf{F}^{-1}\mathbf{C}\mathbf{F}$ and hence, this shows immediately that $\mathbf{A}$ is similar to $\mathbf{C}$.

Section 1.13

1. (a)

$$\begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 3 & 5 \end{bmatrix}$$

   (b)

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

   (c)

$$\operatorname{col}\left(\begin{bmatrix} 1 & 5 & 0 & 4 \end{bmatrix}\right)$$

   (d)

$$\begin{bmatrix} 1 & 0 & 7 \end{bmatrix}$$

2. (a) Solving for a gives

$$2a = 3a - 1 \qquad \Longrightarrow \qquad a = 1$$

   Hence,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

   (b) Solving the system of equation $a = b - a$, $4 + a = b$ gives $a = 4$ and $b = 8$. Hence,

$$\begin{bmatrix} 2 & 4 & 3 \\ 4 & 0 & 8 \\ 3 & 8 & 5 \end{bmatrix}$$

3. (a) Writing the quadratic as $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} = 5x_1^2 + 4x_1x_2 + 3x_2^2$ gives for the coefficient matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}$$

   (b)

$$\mathbf{A} = \begin{bmatrix} 7 & 1 \\ 1 & -1 \end{bmatrix}$$

   (c)

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & -1 \end{bmatrix}$$

   (d)

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

4. (a) Let the matrix $\mathbf{B} = c\mathbf{A} = b_{ij}$, so that

$$b_{ij} = ca_{ij}$$

for all i and j. Then $\mathbf{B}^\top = \mathbf{C} = c_{ij}$, where $c_{ij} = b_{ji}$ for all i and j, or

$$c_{ij} = ca_{ji}$$

Thus $\mathbf{C} = \mathbf{B}^\top = c\mathbf{A}^\top$.

(b) Let the matrix $\mathbf{B} = \mathbf{A}^\top = b_{ij}$, so that

$$b_{ij} = a_{ji}$$

Next, let the matrix $\mathbf{C} = \mathbf{B}^\top = c_{ij}$, so that

$$c_{ij} = b_{ji} = a_{ij}$$

Then $\mathbf{C} = \mathbf{B}^\top = (\mathbf{A}^\top)^\top = \mathbf{A}$.

5. For a matrix $\mathbf{A}$ to be orthogonal implies $\mathbf{A}\mathbf{A}^\top = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$.

(a)

$$\frac{1}{13} \cdot \frac{1}{13} \begin{bmatrix} 5 & 12 \\ -12 & 5 \end{bmatrix} \begin{bmatrix} 5 & -12 \\ 12 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b)

$$\begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix} = \begin{bmatrix} \cos^2\omega + \sin^2\omega & 0 \\ 0 & \sin^2\omega + \cos^2\omega \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

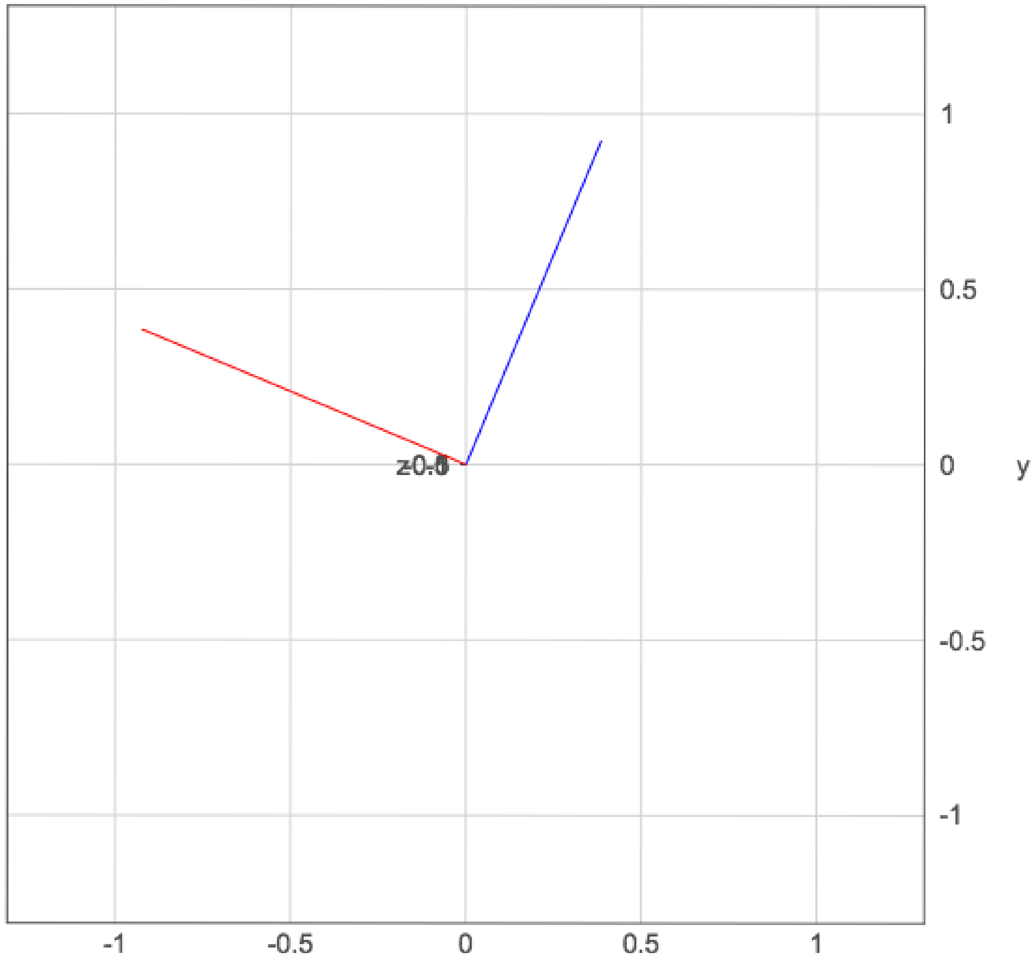
(c)

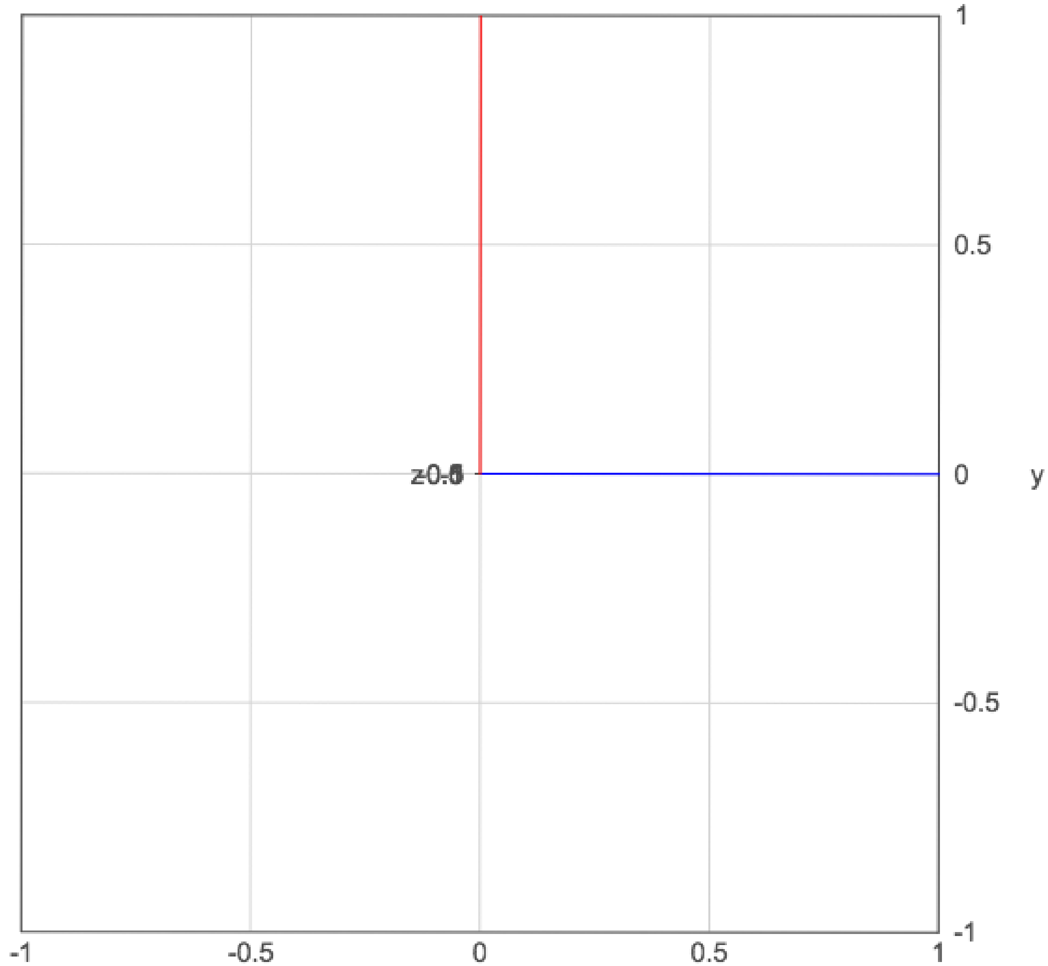
$$\frac{1}{7} \cdot \frac{1}{7} \begin{bmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{bmatrix} \begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(d)

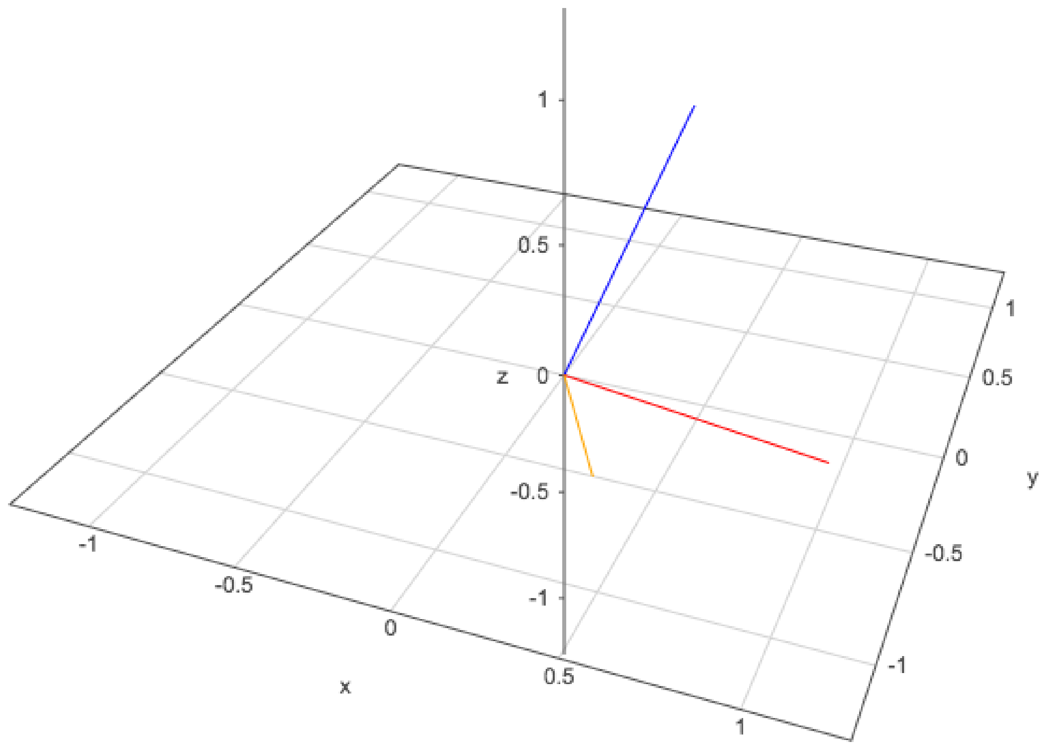
$$\frac{1}{2} \cdot \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\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

6. (a)



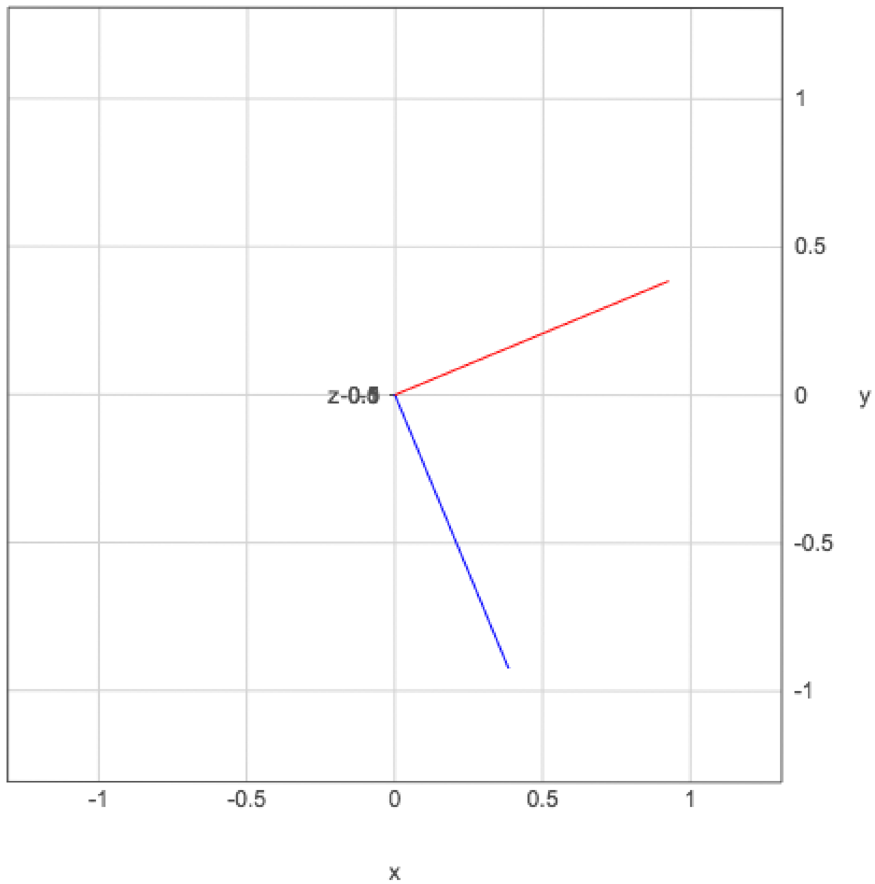


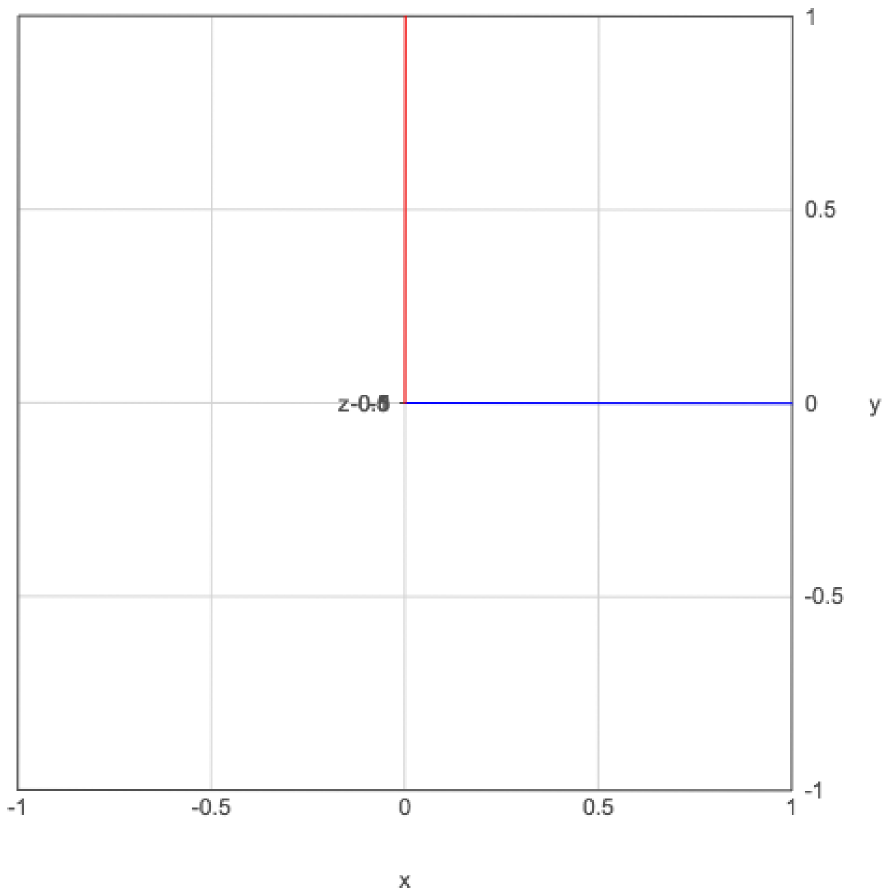
(b) $\omega = 0$



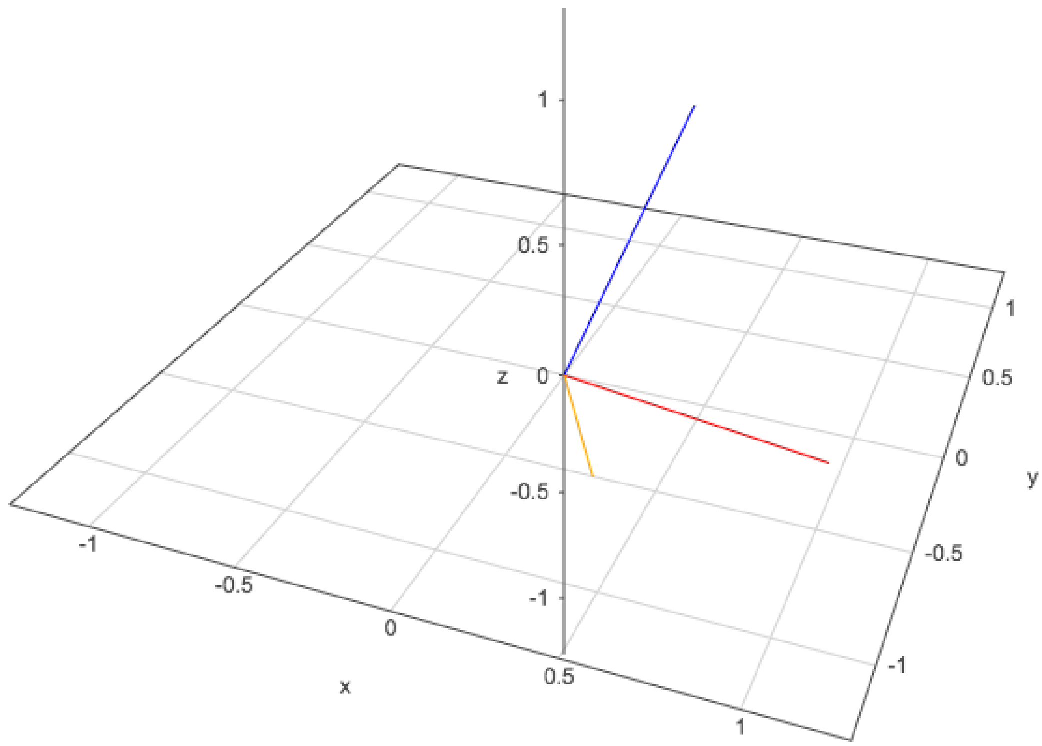
(c)

7. (a)





(b) $\omega = 0$



(c)

8. (a) A matrix $\mathbf{A}$ is orthogonal only if $\mathbf{A}^\top = \mathbf{A}^{-1}$, and hence, $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$. Taking determinants on both sides gives

$$\begin{aligned}\det\left(\mathbf{A}\mathbf{A}^\top\right) &= \det\left(\mathbf{I}\right)\\ \det\left(\mathbf{A}\right)\det\left(\mathbf{A}^\top\right) &= 1\\ \det\left(\mathbf{A}\right)\det\left(\mathbf{A}\right) &= 1\\ \left[\det\left(\mathbf{A}\right)\right]^2 &= 1\\ \det\left(\mathbf{A}\right) &= \pm 1\end{aligned}$$

The fact that $\det(\mathbf{A}^\top) = \det(\mathbf{A})$ follows from rule I of section 1.4.

(b) If $\mathbf{A}$ and $\mathbf{B}$ are orthogonal, then $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$ and $\mathbf{B}\mathbf{B}^\top = \mathbf{I}$. Furthermore, let the matrix $\mathbf{C} = \mathbf{A}\mathbf{B}$. Then it holds that

$$\mathbf{C}\mathbf{C}^\top = \mathbf{A}\mathbf{B}\left(\mathbf{A}\mathbf{B}\right)^\top = \mathbf{A}\mathbf{B}\mathbf{B}^\top\mathbf{A}^\top = \mathbf{A}\mathbf{I}\mathbf{A}^\top = \mathbf{A}\mathbf{A}^\top = \mathbf{I}$$

which thus proves the matrix product $\mathbf{A}\mathbf{B}$ is orthogonal also.

(c) The fact that $\mathbf{A}^\top$ is orthogonal follows from

$$\mathbf{A}^\top\left(\mathbf{A}^\top\right)^\top = \mathbf{A}^\top\mathbf{A} = \left(\mathbf{A}\mathbf{A}^\top\right)^\top = \mathbf{I}^\top = \mathbf{I}$$

As stated for part (a) of this problem, a matrix $\mathbf{A}$ is orthogonal only if $\mathbf{A}^\top = \mathbf{A}^{-1}$. Hence,

$$\mathbf{A}^{-1}\left(\mathbf{A}^{-1}\right)^\top = \mathbf{A}^\top\left(\mathbf{A}^\top\right)^\top = \mathbf{A}^\top\mathbf{A} = \mathbf{I}$$

9. If two $n \times n$ matrices $\mathbf{B}$ and $\mathbf{C}$ are related such that $\mathbf{B} = \mathbf{A}^{-1}\mathbf{C}\mathbf{A}$ for some orthogonal matrix $\mathbf{A}$, then $\mathbf{B}$ is said to be orthogonally congruent to $\mathbf{C}$.

   (a) Setting $\mathbf{A}$ equal to $\mathbf{I}$ gives

   $$\mathbf{B} = \mathbf{I}^{-1}\mathbf{B}\mathbf{I} = \mathbf{B}$$

   which thus proves that every square matrix is orthogonally congruent to itself.

   (b) Multiplying $\mathbf{B}$ by $\mathbf{A}$ from the left and with $\mathbf{A}^{-1}$ from the right gives

   $$\mathbf{C} = \mathbf{A}\mathbf{B}\mathbf{A}^{-1}$$

   which thus proves that $\mathbf{C}$ is orthogonally congruent to $\mathbf{B}$ when $\mathbf{B}$ is orthogonally congruent to $\mathbf{C}$.

   (c) A square matrix $\mathbf{C}$ orthogonally congruent to another square matrix $\mathbf{D}$ is defined as $\mathbf{C} = \mathbf{E}^{-1}\mathbf{D}\mathbf{E}$. Next, substituting for $\mathbf{C}$ in the orthogonal congruence relation between $\mathbf{B}$ and $\mathbf{C}$ gives

   $$\mathbf{B} = \mathbf{A}^{-1}\mathbf{C}\mathbf{A} = \mathbf{A}^{-1}\mathbf{E}^{-1}\mathbf{D}\mathbf{E}\mathbf{A} = (\mathbf{E}\mathbf{A})^{-1}\,\mathbf{D}\mathbf{E}\mathbf{A} = \mathbf{F}^{-1}\mathbf{D}\mathbf{F}$$

   where the matrix $\mathbf{F}$ is the result from the matrix multiplication $\mathbf{E}\mathbf{A}$. Hence, this proves that $\mathbf{B}$ is orthogonally congruent to $\mathbf{D}$.

10. (a) Let $z_1 = a + bi$ and $z_2 = c + di$ be two complex numbers. Then $z_1 + z_2 = a + c + (b + d)i$, and so

   $$\overline{z_1 + z_2} = a + c - (b + d)\,i = a - bi + c - di = \bar{z}_1 + \bar{z}_2$$

   Next, $z_1 z_2 = ac - bd + (ad + bc)i$, and so

   $$\overline{z_1 z_2} = ac - bd - (ad + bc)\,i = (a - bi)\,(c - di) = \bar{z}_1\bar{z}_2$$

   (b) Let $\mathrm{Re}(\mathbf{A}) = \mathrm{Re}(a_{ij})$ and $\mathrm{Im}(\mathbf{A}) = \mathrm{Im}(a_{ij})$ denote the real and imaginary parts of $\mathbf{A}$ and similarly for $\mathbf{B}$. Then $\mathbf{A} + \mathbf{B} = a_{ij} + b_{ij} = \mathrm{Re}(a_{ij}) + \mathrm{Re}(b_{ij}) + [\mathrm{Im}(a_{ij}) + \mathrm{Im}(b_{ij})]\,i$, and so

   $$\begin{aligned}\overline{\mathbf{A} + \mathbf{B}} = \overline{a_{ij} + b_{ij}} &= \mathrm{Re}\,(a_{ij}) + \mathrm{Re}\,(b_{ij}) - [\mathrm{Im}\,(a_{ij}) + \mathrm{Im}\,(b_{ij})] \\ &= \mathrm{Re}\,(a_{ij}) - \mathrm{Im}\,(a_{ij})\,i + \mathrm{Re}\,(b_{ij}) - \mathrm{Im}\,(b_{ij})\,i \\ &= \bar{a}_{ij} + \bar{b}_{ij} = \overline{\mathbf{A}} + \overline{\mathbf{B}}\end{aligned}$$

Next, $\mathbf{AB} = a_{i1}b_{1j} + \cdots + a_{ip}b_{pj} = \alpha_1 + \beta_1 i + \cdots + \alpha_p + \beta_p i$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$ and where

$$\begin{aligned} \alpha_q &= \operatorname{Re}(a_{iq}) \operatorname{Re}(b_{1q}) - \operatorname{Im}(a_{q1}) \operatorname{Im}(b_{qj}) \\ \beta_q &= \operatorname{Re}(a_{iq}) \operatorname{Im}(b_{qj}) + \operatorname{Im}(a_{iq}) \operatorname{Re}(b_{qj}) \end{aligned}$$

for $q = 1, \ldots, p$. Then

$$\begin{aligned} \overline{\mathbf{AB}} &= \overline{a_{i1}b_{1j} + \cdots + a_{ip}b_{pj}} \\ &= \alpha_1 - \beta_1 i + \cdots + \alpha_p - \beta_p i \\ &= \overline{a}_{i1}\overline{b}_{ij} + \cdots + \overline{a}_{ip}\overline{b}_{pj} = \overline{\mathbf{A}}\,\overline{\mathbf{B}} \end{aligned}$$

Lastly, $c\mathbf{A} = ca_{ij} = (\mathrm{Re}(c) + \mathrm{Im}(c)i)(\mathrm{Re}(a_{ij}) + \mathrm{Im}(a_{ij})i) = \mathrm{Re}(c)\mathrm{Re}(a_{ij}) - \mathrm{Im}(c)\mathrm{Im}(a_{ij}) + [\mathrm{Re}(c)\mathrm{Im}(a_{ij}) + \mathrm{Im}(c)\mathrm{Re}(a_{ij})]\, i$, and so

$$\begin{aligned} \overline{c\mathbf{A}} &= \overline{ca_{ij}} \\ &= \operatorname{Re}(c) \operatorname{Re}(a_{ij}) - \operatorname{Im}(c) \operatorname{Im}(a_{ij}) - [\operatorname{Re}(c) \operatorname{Im}(a_{ij}) + \operatorname{Im}(c) \operatorname{Re}(a_{ij})]\, i \\ &= (\operatorname{Re}(c) - \operatorname{Im}(c)\, i) (\operatorname{Re}(a_{ij}) - \operatorname{Im}(a_{ij})\, i) \\ &= \overline{c}\ \overline{a_{ij}} = \overline{c}\overline{\mathbf{A}} \end{aligned}$$

(c) Let the complex matrix $\mathbf{A}$ be given as $\mathbf{A} = a_{ij} = \alpha_{ij} + \beta_{ij} i$. Now let the matrix $\mathbf{A}_1$ represent the real part of $\mathbf{A}$ given by $\mathbf{A}_1 = \mathrm{Re}(\mathbf{A}) = \mathrm{Re}(a_{ij}) = \alpha_{ij}$ and the matrix $\mathbf{A}_2$ represent the imaginart part of $\mathbf{A}$ given by $\mathbf{A}_2 = \mathrm{Im}(\mathbf{A}) = \mathrm{Im}(a_{ij}) = \beta_{ij}$. Then it follows immediately that $\mathbf{A}$ can be written uniquely as a linear combination of $\mathbf{A}_1$ and $\mathbf{A}_2$, namely

$$\begin{aligned} \mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2 &= \frac{1}{2}\left(\mathbf{A} + \overline{\mathbf{A}}\right) + \frac{1}{2}\left(\mathbf{A} - \overline{\mathbf{A}}\right) \\ &= \frac{1}{2}\left(a_{ij} + \overline{a_{ij}}\right) + \frac{1}{2}\left(a_{ij} - \overline{a_{ij}}\right) \\ &= \frac{1}{2}\left(\alpha_{ij} + i\beta_{ij} + \alpha_{ij} - i\beta_{ij}\right) + \frac{1}{2}\left(\alpha_{ij} + i\beta_{ij} - \alpha_{ij} + i\beta_{ij}\right) \\ &= a_{ij} + i\beta_{ij} \end{aligned}$$

(d) From part (c) the complex matrix $\mathbf{A}$ may be written as $\mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2$. Then

$$\overline{\mathbf{A}}^\top = \left(\overline{\mathbf{A}_1 + i\mathbf{A}_2}\right)^\top = (\mathbf{A}_1 - i\mathbf{A}_2)^\top = \mathbf{A}_1^\top - i\mathbf{A}_2^\top = \overline{\left(\mathbf{A}_1^\top + i\mathbf{A}_2^\top\right)} = \overline{\mathbf{A}^\top}$$

Similarly

$$\left(\overline{\mathbf{A}}\right)^{-1} = \left(\overline{\mathbf{A}_1 + i\mathbf{A}_2}\right)^{-1} = (\mathbf{A}_1 - i\mathbf{A}_2)^{-1} = \mathbf{A}_1^{-1} - i\mathbf{A}_2^{-1} = \overline{\left(\mathbf{A}_1^{-1} + i\mathbf{A}_2^{-1}\right)} = \overline{\mathbf{A}^{-1}}$$

(e) If $\mathbf{A}$ is real then according to part (c) $\mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2 = \mathbf{A}_1 + i\mathbf{0} = \mathbf{A}_1 = \alpha_{ij}$. Hence, it follows immediately that

$$\overline{\mathbf{A}} = \overline{\mathbf{A}_1} = \overline{\alpha_{ij}} = \alpha_{ij} = \mathbf{A}_1$$

since the complex conjugate of a real number α_{ij} is just equal to itself. Now if instead $\mathbf{A}$ was a complex matrix, then $\mathbf{A}_2 \neq \mathbf{0}$ and by definition of complex conjugation $\overline{\alpha_{ij} + i\beta_{ij}} = \alpha + i\beta$ only if $\beta_{ij} = 0$.

11. Applying complex conjugation to the equation $\mathbf{Av} = \lambda\mathbf{v}$ gives $\overline{\mathbf{Av}} = \overline{\lambda\mathbf{v}} \implies \mathbf{A}\overline{\mathbf{v}} = \lambda\overline{\mathbf{v}}$, which follows readily from the fact that both $\mathbf{A}$ and λ are assumed to be real. Next, consider the equation $\mathbf{Au} = \lambda\mathbf{u}$, where $\mathbf{u}$ is real and nonzero. This may be obtained easily by adding or subtracting the eigenvalue equation involving $\mathbf{v} = \mathbf{p} + i\mathbf{q}$ and its complex conjugate version involving $\overline{\mathbf{v}} = \mathbf{p} - i\mathbf{q}$. Addition gives

$$\begin{aligned} \mathbf{Av} + \overline{\mathbf{Av}} &= \lambda\mathbf{v} + \overline{\lambda\mathbf{v}} \\ \mathbf{A}\left(\mathbf{v} + \overline{\mathbf{v}}\right) &= \lambda\left(\mathbf{v} + \overline{\mathbf{v}}\right) \\ \mathbf{Ap} &= \lambda\mathbf{p} \end{aligned}$$

whereas subtration gives

$$\begin{aligned} \mathbf{Av} - \overline{\mathbf{Av}} &= \lambda\mathbf{v} - \overline{\lambda\mathbf{v}} \\ \mathbf{A}\left(\mathbf{v} - \overline{\mathbf{v}}\right) &= \lambda\left(\mathbf{v} - \overline{\mathbf{v}}\right) \\ \mathbf{Aq} &= \lambda\mathbf{q} \end{aligned}$$

Hence, we may conclude that a real vector $\mathbf{u}$ may either be obtained by adding to or subtracting off the complex conjugate version of the original eigenvalue equation involving the complex vector $\mathbf{v}$ and real matrix $\mathbf{A}$ and real eigenvalue λ.

12. Let us consider the eigenvalue equation for a real symmetric matrix $\mathbf{A}$

$$\mathbf{Av} = \lambda\mathbf{v}$$

where λ and $\mathbf{v}$ are considered complex for the moment. Next, let $v = \alpha + \beta i$ and let us compute the quantity $Q(\mathbf{v})$ by means of equations (1.84) and (1.85):

$$Q\left(\mathbf{v}\right) = \mathbf{v}^\top \mathbf{A}\overline{\mathbf{v}} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} v_i \overline{v}_j = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\left(\alpha_i + \beta_i i\right)\left(\alpha_j - \beta_j i\right)$$

It remains to be shown that the quantity Q is in fact real. Firstly, let us consider the case $i = j$. Then $v_i\overline{v}_i = \alpha_i^2 + \beta_i^2 = |v_i|^2$, which clearly is a real valued number greater than or equal to zero. Secondly, for $i \neq j$, we can combine the $v_k\overline{v}_l$ and $v_l\overline{v}_k$ terms, where $k \in \{1, \ldots, n\}$ and $l \in \{1, \ldots, j\}$ and $k \neq l$, to get

$$\begin{aligned} Q &= \cdots + a_{kl} v_k \overline{v}_l + a_{lk} v_l \overline{v}_k + \ldots \\ &= \cdots + a_{kl}\left(v_k\overline{v}_l + v_l\overline{v}_k\right) + \ldots \\ &= \cdots + a_{kl}\left[\left(\alpha_k + \beta_k i\right)\left(\alpha_l - \beta_l i\right) + \left(\alpha_l + \beta_l i\right)\left(\alpha_k - \beta_k i\right)\right] + \ldots \\ &= \cdots + 2a_{kl}\left(\alpha_k\alpha_l + \beta_k\beta_l\right) + \ldots \end{aligned}$$

where the second step follows from the fact that $a_{kl} = a_{lk}$, since it is assumed that $\mathbf{A}$ is real and symmetric. Hence, we may conclude that Q is real. Starting from the eigenvalue equation $\mathbf{Av} = \lambda\mathbf{v}$, we can write

$$Q = \mathbf{v}^\top \mathbf{A}\overline{\mathbf{v}} = \lambda \mathbf{v}^\top \overline{\mathbf{v}} = \lambda \left(|v_1|^2 + \cdots + |v_n|^2\right)$$

Since we have established that Q is real, the right hand side of the equation above makes sense only if λ is real (i.e. it should be clear that the term in brackets is a real number greater than or equal to zero).

13. Throughout it is assumed that $\mathbf{x}$ and λ are real. For part (a) of problem 6 it was proven that the product of the eigenvalues of a matrix $\mathbf{A}$ equals the determinant of $\mathbf{A}$, so $\lambda_1\lambda_2 \ldots \lambda_n = \det \mathbf{A}$. Now, let $\mathbf{x}$ be an eigenvector associated with the eigenvalue λ. If $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} = \lambda(x_1^2 + \cdots + x_2^2) > 0$, then clearly $\lambda > 0$. Assuming that $Q(\mathbf{x}) > 0$ for all eigenvectors $\mathbf{x}$ belonging to all eigenvalues λ of the real symmetric matrix $\mathbf{A}$, it follows that all λ's are larger than zero. Hence, $\lambda_i > 0$ for $i \in \{1, \ldots, n\}$ and so $\lambda_1\lambda_2 \ldots \lambda_n = \det \mathbf{A} > 0$.

14. Let $\mathbf{A}$ be a real symmetric matrix and the quantity $Q(\mathbf{x}) > 0$, so $Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} > 0$. Next, let us define the vector $\mathbf{u}$, which relates to $\mathbf{x}$ via the equation $\mathbf{x} = \mathbf{Du}$. Substituting for $\mathbf{x}$ then gives $Q(\mathbf{x}) = (\mathbf{Du})^\top \mathbf{ADu} = \mathbf{u}^\top \mathbf{D}^\top \mathbf{ADu} > 0$. Furthermore, let us assume that $\mathbf{D}$ is non-singular and orthogonal so that $\mathbf{D}^\top = \mathbf{D}^{-1}$ and $Q(\mathbf{x}) = \mathbf{u}^\top \mathbf{D}^{-1}\mathbf{ADu} > 0$. Hence, we may find a diagonal matrix $\mathbf{B} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ that is orthogonally congruent to $\mathbf{A}$: $\mathbf{B} = \mathbf{D}^{-1}\mathbf{AD}$. The items appearing along the diagonal of $\mathbf{B}$ are the (necessarily real) eigenvalues of $\mathbf{A}$ and thus we end up with

$$\begin{aligned} Q\left(\mathbf{x}\right) = \mathbf{u}^\top \mathbf{D}^{-1}\mathbf{ADu} = \mathbf{u}^\top \mathbf{Bu} &= \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij}u_iu_j \\ &= \lambda_1 u_1^2 + \cdots + \lambda_n u_n^2 \end{aligned}$$

Hence, $Q(\mathbf{x})$ can be written in terms of the product of the squares of the coordinates $(u_1, \ldots, u_n)$ and the coefficients $\lambda_1, \ldots, \lambda_n$, which are the eigenvalues of $\mathbf{A}$.

15. Let $\mathbf{A}$ be a real symmetric matrix so that $a_{ij} = a_{ji}$. Then using (1.84) and (1.85)

$$\mathbf{x}^\top \mathbf{Ay} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}x_iy_j = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ji}y_jx_i = \mathbf{y}^\top \mathbf{Ax}$$

Now from $\mathbf{Av}_1 = \lambda_1\mathbf{v}_1$ and $\mathbf{Av}_2 = \lambda_2\mathbf{v}_2$ and the equation above it follows that $\mathbf{v}_2^\top \mathbf{Av}_1 = \mathbf{v}_1^\top \mathbf{Av}_2 = \lambda_1 \mathbf{v}_2^\top \mathbf{v}_1 = \lambda_1 \mathbf{v}_1^\top \mathbf{v}_2$ and $\mathbf{v}_1^\top \mathbf{Av}_2 = \lambda_2 \mathbf{v}_1^\top \mathbf{v}_2$. Subtracting the result and re-arranging then gives

$$\lambda_1 \mathbf{v}_1^\top \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^\top \mathbf{v}_2$$

Since it is assumed that $\lambda_1 \neq \lambda_2$ it must hold that $\mathbf{v}_1^\top \mathbf{v}_2 = 0$ and so $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal.

16. Let again $\mathbf{A}$ be a real symmetric matrix. Then by (1.94) there exists a diagonal matrix $\mathbf{B} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ such that $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$, where the columns of $\mathbf{C}$ are the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\mathbf{A}$. Now if the matrix $\mathbf{C}$ is orthogonal, then $\mathbf{B}$ is said to be orthogonally congruent to $\mathbf{A}$. Hence, it remains to be proven that $\mathbf{C}$ is orthogonal (so that $\mathbf{C}^\top = \mathbf{C}^{-1}$). However, this follows straight from the fact that the columns of $\mathbf{C}$ are the eigenvectors of $\mathbf{A}$ combined with the result of problem 15 that showed that the inner product $\mathbf{v}_i^\top \mathbf{v}_j$ between two eigenvectors of a real symmetric matrix is nonzero only when $i = j$.

Section 1.15

1. (a)
$$\begin{aligned}
\mathbf{u}+\mathbf{v} &= \begin{bmatrix}3 & 2 & 1 & 0\end{bmatrix} + \begin{bmatrix}1 & 0 & 1 & 2\end{bmatrix} = \begin{bmatrix}4 & 2 & 2 & 2\end{bmatrix} \\
\mathbf{u}+\mathbf{w} &= \begin{bmatrix}3 & 2 & 1 & 0\end{bmatrix} + \begin{bmatrix}5 & 4 & 1 & -2\end{bmatrix} = \begin{bmatrix}8 & 6 & 2 & -2\end{bmatrix} \\
2\mathbf{u} &= 2\begin{bmatrix}3 & 2 & 1 & 0\end{bmatrix} = \begin{bmatrix}6 & 4 & 2 & 0\end{bmatrix} \\
-3\mathbf{v} &= -3\begin{bmatrix}1 & 0 & 1 & 2\end{bmatrix} = \begin{bmatrix}-3 & 0 & -3 & -6\end{bmatrix} \\
0\mathbf{w} &= 0\begin{bmatrix}5 & 4 & 1 & -2\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0\end{bmatrix} = \mathbf{0}
\end{aligned}$$

   (b)
$$\begin{aligned}
3\mathbf{u}-2\mathbf{v} &= \begin{bmatrix}7 & 6 & 1 & -4\end{bmatrix} \\
2\mathbf{v}+3\mathbf{w} &= \begin{bmatrix}17 & 12 & 5 & -2\end{bmatrix} \\
\mathbf{u}-\mathbf{w} &= \begin{bmatrix}-2 & -2 & 0 & 2\end{bmatrix} \\
\mathbf{u}+\mathbf{v}-2\mathbf{w} &= \begin{bmatrix}-6 & -6 & 0 & 6\end{bmatrix}
\end{aligned}$$

   (c)
$$\mathbf{u}\cdot\mathbf{v} = \begin{bmatrix}3 & 2 & 1 & 0\end{bmatrix} \cdot \begin{bmatrix}1 & 0 & 1 & 2\end{bmatrix} = (3)(1) + (2)(0) + (1)(1) + (0)(2) = 4$$
$$\begin{aligned}
\mathbf{u}\cdot\mathbf{w} &= 15 + 8 + 1 + 0 = 24 \\
|\mathbf{u}| &= \sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{9+4+1+0} = \sqrt{14} \\
|\mathbf{v}| &= \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{1+0+1+4} = \sqrt{6}
\end{aligned}$$

   (d) The vector $\mathbf{w}$ can be written as $\mathbf{w} = 2\mathbf{u} - \mathbf{v}$

2. (a) Let P denote the point $P = (5, 10, -1, 18)$. Then the vector $\overrightarrow{P_1P}$ may be represented by the tuple $\overrightarrow{P_1P} = (1, 2, 3, 4) - (5, 10, -1, 8) = (4, 8, -4, 4)$. Similarly, the vector $\overrightarrow{P_1P_2}$ may be represented by the tuple $\overrightarrow{P_1P_2} = (7, 14, -3, 10) - (1, 2, 3, 4) = (6, 12, -6, 6)$. Now the point P may be concluded to lie on the line segment P_1P_2 if a scalar t, where $0 \le t \le 1$, can be found so that $\overrightarrow{P_1P} = t\overrightarrow{P_1P_2}$. Hence we should be able to find a t such that
$$(4, 8, -4, 4) = t(6, 12, -6, 6)$$
is satisfied. This is indeed the case for $t = 2/3$.

(b) Going through the same steps as for part (a), we may conclude that the point $P = (2, 8, 1, 6)$ in fact does not lie on the line segment P_1P_2, since there doesn't exist a scalar t, $0 \leq t \leq 1$ such that

$$(1, 1, -1, 1) = t(8, -5, -2, 2)$$

holds.

(c) Let the point P denote the midpoint of the line segment P_1P_2. It then follows immediately that the distance from point P to point P_1 is the same as the distance from point P to P_2, or in other words that

$$\begin{aligned}
|\overrightarrow{P_1P}|^2 - |\overrightarrow{P_2P}|^2 &= 0 \\
(p^{(1)} - 2)^2 + (p^{(2)})^2 + (p^{(3)} - 1)^2 + (p^{(4)} - 3)^2 + (p^{(5)} - 7)^2 &= \\
-\left[(p^{(1)} - 10)^2 + (p^{(2)} - 4)^2 + (p^{(3)} + 3)^2 + (p^{(4)} - 3)^2 + (p^{(5)} - 5)^2\right]&
\end{aligned}$$

Solving term by term (i.e. for each $p^{(n)}$ where $1 \leq n \leq 5$) then gives $P = (p^{(1)}, p^{(2)}, p^{(3)}, p^{(4)}, p^{(5)}) = (6, 2, -1, 3, 6)$.

(d) To trisect the line segment joining P_1P_2 means we are looking for two points P and Q which are equidistant from each other and the end points P_1 and P_2.That is, we are looking for points P and Q such that $(1/3)|\overrightarrow{P_1P}| = (2/3)|\overrightarrow{P_2P}|$ and $(2/3)|\overrightarrow{P_1Q}| = (1/3)|\overrightarrow{P_2Q}|$. For P this means

$$\begin{aligned}
(1/9)|\overrightarrow{P_1P}|^2 - (4/9)|\overrightarrow{P_2P}|^2 &= 0 \\
|\overrightarrow{P_1P}|^2 - 4|\overrightarrow{P_2P}|^2 &= \\
(p^{(1)} - 7)^2 + (p^{(2)} - 1)^2 + (p^{(3)} - 7)^2 + (p^{(4)} - 9)^2 + (p^{(5)} - 6)^2 &= \\
-4\left[(p^{(1)} - 16)^2 + (p^{(2)} + 2)^2 + (p^{(3)} - 10)^2 + (p^{(4)} - 3)^2 + (p^{(5)})^2\right]&
\end{aligned}$$

Solving term by term then gives $P = (13, -1, 9, 5, 2)$. The point Q may be found following the same steps and is given by $Q = (10, 0, 8, 7, 4)$.

(e) If P is on P_1P_2 we have $\overrightarrow{P_1P} = t\overrightarrow{P_1P_2}$, where $0 \leq t \leq 1$ and $\overrightarrow{PP_2} = (1-t)\overrightarrow{P_1P_2}$, so that $\overrightarrow{P_1P} + \overrightarrow{PP_2} = t\overrightarrow{P_1P_2} + (1-t)\overrightarrow{P_1P_2} = \overrightarrow{P_1P_2}$. Taking dot products on each side gives

$$\begin{aligned}
\left(\overrightarrow{P_1P} + \overrightarrow{PP_2}\right) \cdot \left(\overrightarrow{P_1P} + \overrightarrow{PP_2}\right) &= \overrightarrow{P_1P_2} \cdot \overrightarrow{P_1P_2} \\
|\overrightarrow{P_1P}|^2 + |\overrightarrow{P_2P}|^2 + 2\overrightarrow{P_1P} \cdot \overrightarrow{PP_2} &= |\overrightarrow{P_1P_2}|^2 \\
|\overrightarrow{P_1P} + \overrightarrow{PP_2}|^2 &= \\
|\overrightarrow{P_1P} + \overrightarrow{PP_2}| &= |\overrightarrow{P_1P_2}|
\end{aligned}$$

3. We firstly note that $\overrightarrow{P_1P_3} = \overrightarrow{P_1P_2} + \overrightarrow{P_2P_3}$, so that

$$\begin{aligned}|\overrightarrow{P_1P_3}|^2 = \overrightarrow{P_1P_3} \cdot \overrightarrow{P_1P_3} &= \left(\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3}\right) \cdot \left(\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3}\right) \\ &= |\overrightarrow{P_1P_2}|^2 + |\overrightarrow{P_2P_3}|^2 + 2\overrightarrow{P_1P_2} \cdot \overrightarrow{P_2P_3} \\ &= |\overrightarrow{P_1P_2}|^2 + |\overrightarrow{P_2P_3}|^2\end{aligned}$$

Here the last step follows from the assumption that $\overrightarrow{P_1P_2}$ is orthogonal to $\overrightarrow{P_2P_3}$ and so $\overrightarrow{P_1P_2} \cdot \overrightarrow{P_2P_3} = 0$ by the definition of the dot product.

4. Starting from the fact that $\overrightarrow{P_1P_3} = \overrightarrow{P_1P_2} + \overrightarrow{P_2P_3}$ we have

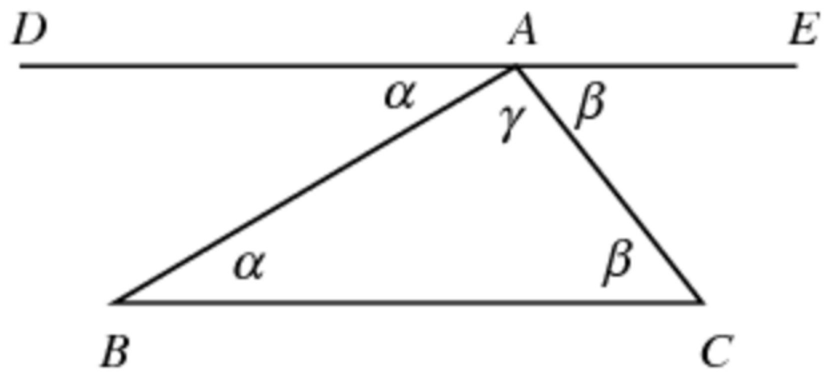
$$\begin{aligned}|\overrightarrow{P_2P_3}|^2 &= \left(\overrightarrow{P_1P_3} - \overrightarrow{P_1P_2}\right) \cdot \left(\overrightarrow{P_1P_3} - \overrightarrow{P_1P_2}\right) \\ &= |\overrightarrow{P_1P_3}|^2 + |\overrightarrow{P_1P_2}|^2 - 2\overrightarrow{P_1P_2} \cdot \overrightarrow{P_1P_3} \\ &= |\overrightarrow{P_1P_3}|^2 + |\overrightarrow{P_1P_2}|^2 - 2|\overrightarrow{P_1P_2}||\overrightarrow{P_1P_3}| \cos\theta\end{aligned}$$

where $\theta = \sphericalangle(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3})$. The last step follows from equation (1.9): $\mathbf{v}\cdot\mathbf{w} = |\mathbf{v}||\mathbf{w}| \cos\theta$, $\theta = \sphericalangle(\mathbf{v}, \mathbf{w})$ with $0 \le \theta \le \pi$.

5. (a) Let $P_1 = (1, 2, 3, 4, 5)$, $P_2 = (5, 4, 3, 2, 1)$ and $P_3 = (2, 2, 2, 2, 2)$, so that $\overrightarrow{P_1P_2} = (4, 2, -1, -1, -4)$, $\overrightarrow{P_1P_3} = (1, 0, -1, -2, -3)$ and $\overrightarrow{P_2P_3} = (-3, -2, 0, -1, 1)$. The sides of the triangle are then given by $|\overrightarrow{P_1P_2}| = \sqrt{38}$, $|\overrightarrow{P_1P_3}| = \sqrt{15}$ and $|\overrightarrow{P_2P_3}| = \sqrt{15}$. The angles between the sides are given by

$$\begin{aligned}\sphericalangle(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}) &= \cos^{-1}\left(\frac{\overrightarrow{P_1P_2} \cdot \overrightarrow{P_1P_3}}{|\overrightarrow{P_1P_2}||\overrightarrow{P_1P_3}|}\right) = \cos^{-1}\left(\frac{19}{\sqrt{570}}\right) \\ \sphericalangle(\overrightarrow{P_1P_2}, \overrightarrow{P_2P_3}) &= \cos^{-1}\left(\frac{\overrightarrow{P_1P_2} \cdot \overrightarrow{P_2P_3}}{|\overrightarrow{P_1P_2}||\overrightarrow{P_2P_3}|}\right) = \cos^{-1}\left(\frac{19}{\sqrt{570}}\right) \\ \sphericalangle(\overrightarrow{P_1P_3}, \overrightarrow{P_2P_3}) &= \cos^{-1}\left(\frac{\overrightarrow{P_1P_3} \cdot \overrightarrow{P_2P_3}}{|\overrightarrow{P_1P_3}||\overrightarrow{P_2P_3}|}\right) = \cos^{-1}\left(-\frac{4}{15}\right)\end{aligned}$$

(b) Let a triangle have sides a, b and c in E^n and let P_1, P_2 and P_3 be the three vertices of the triangle, so that $a = |\overrightarrow{P_1P_2}|$, $b = |\overrightarrow{P_2P_3}|$ and $c = |\overrightarrow{P_1P_3}|$ denotes the distance between the three vertices or alternatively, be the length of the three sides a, b and c. Since $|\overrightarrow{P_1P_2}| = (\overrightarrow{P_1P_2} \cdot \overrightarrow{P_1P_2})^{1/2}$ is a scalar that is independent of dimensionality, we may safely assume that the triangle sides a, b and c are the same in any E^n for $n \ge 2$. Next, by proving that the sum of the angles is π in E^2 (assuming an Euclidian geometry) we can show that it is π in E^n using the law of cosines. As we can see from the picture, the line $DAE \parallel BC$, and so the angles α and β satisfy $\alpha = \angle(DAB) = \angle(ABC)$ and $\beta = \angle(EAC) = \angle(ACB)$. Adding γ then gives



$\alpha + \beta + \gamma = \pi$.
Hence, we have shown for E^2 in Euclidian geometry that indeed the sum of the three triangle angles always equals π. To see, that this carries over to E^n for $n \geq 2$, we can make use of the law of cosines to show that the corresponding angles of the triangle in E^2 and the one in E^n are in fact the same. Let $\alpha = \sphericalangle(\overrightarrow{P_1P_2}, \overrightarrow{P_2P_3})$. Then using the law of cosines, this angle is given by

$$\alpha = \cos^{-1}\left[\frac{1}{2}\left(\frac{|\overrightarrow{P_1P_2}|}{|\overrightarrow{P_2P_3}|} + \frac{|\overrightarrow{P_2P_3}|}{|\overrightarrow{P_1P_2}|} - \frac{|\overrightarrow{P_1P_3}|^2}{|\overrightarrow{P_1P_2}||\overrightarrow{P_2P_3}|}\right)\right]$$

As is apparent immediately, the angle is completely expressible in terms of the length of the triangle sides. Since these sides are the same in any E^n, so will be the angle α. The angles β and γ may be expressed in similar terms, and hence, we may conclude that the corresponding angles of the triangle in E^2 and E^n are equal.

6. (i)

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \ldots, u_n + v_n) = (v_1 + u_1, \ldots, v_n + u_n) = \mathbf{v} + \mathbf{u}$$

(ii)

$$\begin{aligned}(\mathbf{u} + \mathbf{v}) + \mathbf{w} &= (u_1 + v_1, \ldots, u_n + v_n) + (w_1, \ldots, w_n)\\ &= (u_1 + v_1 + w_1, \ldots, u_n + v_n + w_n)\\ &= (u_1, \ldots, u_n) + (v_1 + w_1, \ldots, v_n + w_n)\\ &= \mathbf{u} + (\mathbf{v} + \mathbf{w})\end{aligned}$$

(iii)

$$h\,(\mathbf{u} + \mathbf{v}) = h\,(u_1 + v_1, \ldots, u_n + v_n) = (hu_1 + hv_1, \ldots, hu_n + hv_n) = h\mathbf{u} + h\mathbf{v}$$

(iv)

$$\begin{aligned}(a + b)\,\mathbf{u} = (a + b)\,(u_1, \ldots, u_n) &= ((a + b)u_1, \ldots, (a + b)u_n)\\ &= (au_1 + bu_1, \ldots, au_n + bu_n)\\ &= a\mathbf{u} + b\mathbf{u}\end{aligned}$$

(v)

$$
\begin{aligned}
(ab)\,\mathbf{u} = (ab)\,(u_1,\ldots,u_n) = ((ab)u_1,\ldots,(ab)u_n) &= (abu_1,\ldots,abu_n)\\
&= (a(bu_1),\ldots,a(bu_n))\\
&= a\,((bu_1,\ldots,bu_n))\\
&= a\,(b\mathbf{u})
\end{aligned}
$$

(vi)

$$1\mathbf{u} = 1\,(u_1,\ldots,u_n) = (1u_1,\ldots,1u_n) = (u_1,\ldots,u_n) = \mathbf{u}$$

(vii)

$$0\mathbf{u} = 0\,(u_1,\ldots,u_n) = (0u_1,\ldots,0u_n) = (0,\ldots,0) = \mathbf{0}$$

(viii)

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + \cdots + u_nv_n = v_1u_1 + \cdots + v_nu_n = \mathbf{v}\cdot\mathbf{u}$$

(ix)

$$
\begin{aligned}
(\mathbf{u}+\mathbf{v})\cdot\mathbf{w} &= (u_1+v_1)\,w_1 + \cdots + (u_n+v_n)\,w_n\\
&= u_1w_1 + v_1w_1 + \cdots + u_nw_n + v_nw_n\\
&= (u_1w_1 + \cdots + u_nw_n) + (v_1w_1 + \cdots + v_nw_n)\\
&= \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}
\end{aligned}
$$

(x)

$$
\begin{aligned}
(a\mathbf{u})\cdot\mathbf{v} = (au_1)\,v_1 + \cdots + (au_n)\,v_n &= a\,(u_1v_1) + \cdots + a\,(u_nv_n)\\
&= a\,(u_1v_1 + \cdots + u_nv_n)\\
&= a\,(\mathbf{u}\cdot\mathbf{v})
\end{aligned}
$$

(xi)

$$\mathbf{u}\cdot\mathbf{u} = u_1u_1 + \cdots + u_nu_n = u_1^2 + \cdots + u_n^2 = |\mathbf{u}|^2 \geq 0$$

(xii)

$$\mathbf{u}\cdot\mathbf{u} = u_1u_1+\cdots+u_nu_n = u_1^2+\cdots+u_n^2 = 0 \text{ if and only if } u_i = 0, i \in [1,n] \Rightarrow \mathbf{u} = \mathbf{0}$$

7. Lets consider two vectors $\mathbf{u}$ and $\mathbf{v}$, where $\mathbf{v} \neq \mathbf{0}$ and $\mathbf{u}$ is not a scalar multiple of $\mathbf{v}$. Then $\mathbf{u} + t\mathbf{v} \neq \mathbf{0}$ for every scalar t. Hence

$$|\mathbf{u}+t\mathbf{v}|^2 = (\mathbf{u}+t\mathbf{v})\cdot(\mathbf{u}+t\mathbf{v}) > 0 \text{ or } P\,(t) = t^2|\mathbf{v}|^2 + 2t\,(\mathbf{u}\cdot\mathbf{v}) + |\mathbf{u}|^2 > 0$$

implying that the discriminant of the quadratic $P\,(t)$ must be negative, that is

$$4\,(\mathbf{u}\cdot\mathbf{v})^2 - 4|\mathbf{u}|^2|\mathbf{v}|^2 < 0 \qquad \Longrightarrow \qquad |\mathbf{u}\cdot\mathbf{v}| < |\mathbf{u}||\mathbf{v}|$$

If $\mathbf{u} = c\mathbf{v}$, then

$$|\mathbf{u} \cdot \mathbf{v}| = |c\mathbf{v} \cdot \mathbf{v}| = |c\,(\mathbf{v} \cdot \mathbf{v})\,| = |c||\mathbf{v}|^2 = |c\mathbf{v}||\mathbf{v}| = |\mathbf{u}||\mathbf{v}|$$

Hence, we have proved that

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$$

Squaring both sides of this inequality and writing it out explicitly as a sum over the components of each vector then gives

$$\left(\sum_{i=1}^{n} u_i v_i\right)^2 \leq \sum_{i=1}^{n} u_i^2 \sum_{j=1}^{n} v_j^2$$

8. To prove the triangle inequality, in particular the equality case, we start with writing

$$\begin{aligned} |\mathbf{u} + \mathbf{v}|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2\mathbf{u} \cdot \mathbf{v} \\ &= |\mathbf{u}|^2 + |\mathbf{v}|^2 + 2\mathbf{u} \cdot \mathbf{v} + 2|\mathbf{u}||\mathbf{v}| - 2|\mathbf{u}||\mathbf{v}| \\ &= (|\mathbf{u}| + |\mathbf{v}|)^2 + 2\mathbf{u} \cdot \mathbf{v} - 2|\mathbf{u}||\mathbf{v}| \end{aligned}$$

Next, making use of the Cauchy-Schwartz inequality $|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$, we may further manipulate the last equality to obtain

$$|\mathbf{u} + \mathbf{v}|^2 \leq (|\mathbf{u}| + |\mathbf{v}|)^2 \qquad \Longrightarrow \qquad |\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$$

To prove the case of equality is a matter of examining the case when $\mathbf{u} \cdot \mathbf{v} - |\mathbf{u}||\mathbf{v}| = 0$, which has already been done as part of problem (7). Again, we consider the case where $\mathbf{u} = c\mathbf{v}$ with $c \geq 0$ (i.e. the case when $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent). For then

$$\mathbf{u} \cdot \mathbf{v} = (c\mathbf{v}) \cdot \mathbf{v} = c\,(\mathbf{v} \cdot \mathbf{v}) = c|\mathbf{v}|^2 = |c\mathbf{v}||\mathbf{v}| = |\mathbf{u}||\mathbf{v}|$$

Hence, equality rule for the triangle inequality is satisfied exactly when $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent and when both or either $\mathbf{u}$ and/or $\mathbf{v}$ is equal to the $\mathbf{0}$ vector.

9. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be a system of k non-zero vectors. Then for this system to be orthogonal means that the dot product of the i'th and j'th vector needs to be zero: $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, whereas the dot product of the i'th vector with itself should be a positive, nonzero scalar: $\mathbf{v}_i \cdot \mathbf{v}_i > 0$. Next, consider the condition for linear independence. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent if the only scalars $c_1, \ldots, c_k$ such that

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

are $c_1 = c_2 = \cdots = c_k = 0$. Now let us take the dot product of the i'th vector with the the equation above. Since it is assumed that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are orthogonal, we have

$$\begin{aligned} (c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i &= \mathbf{0} \cdot \mathbf{v}_i \\ (c_i\mathbf{v}_i) \cdot \mathbf{v}_i &= 0 \\ c_i|\mathbf{v}_i|^2 &= \end{aligned}$$

The only way for this to make sense is if $c_1 = 0$. By repeating the same steps for all other vectors in the orthogonal system we can conclude that in fact all $c_1, \dots, c_k$ need to equal zero and hence, that the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent by the definition of linear independence just given.

10. (a) A system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ is linearly dependent if one of the vectors is expressible as a linear combination of the others. In other words when in the expression

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

there is at least one c_i, $i \in [1, k]$ that is nonzero, say c_k. For then we can write

$$\begin{aligned}\mathbf{v}_k &= -\frac{c_1}{c_k}\mathbf{v}_1 + \cdots + \frac{c_{k-1}}{c_k}\mathbf{v}_{k-1} \\ &= \tilde{c}_1\mathbf{v}_1 + \cdots + \tilde{c}_{k-1}\mathbf{v}_{k-1}\end{aligned}$$

where $\tilde{c}_i = -c_i/c_k$, $i \in [1, k-1]$.

(b) Suppose $\mathbf{v}_k = \mathbf{0}$ and $c_1, \dots, c_k$ not all zero. In particular lets consider the case where $c_1 = 0, c_2 = 0, \dots, c_{k-1} = 0$, but $c_k \neq 0$. Then there does exist a $c_i \neq 0$, $i \in [1, k]$ (namely c_k) for which

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = 0$$

is satisfied. Hence, by the definition of linear independence, i.e. that the only scalars $c_1, \dots, c_k$ for which the above equation holds are $c_1 = c_2 = \cdots = c_k = 0$, any system of vectors which contains the $\mathbf{0}$ vector is linearly dependent.

(c) If the system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_{k+1}$ is linearly dependent then there exists a $c_i \neq 0$ $i \in [1, k+1]$ such that

$$c_1\mathbf{v}_1 + \cdots + c_{k+1}\mathbf{v}_{k+1} = \mathbf{0}$$

However, $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linear independent, and so

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

with $c_1 = c_2 = \cdots = c_k = 0$. Hence, c_{k+1} cannot be zero and $\mathbf{v}_{k+1}$ is expressible as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$:

$$\mathbf{v}_{k+1} = -\frac{c_1}{c_{k+1}}\mathbf{v}_1 - \cdots - \frac{c_k}{c_{k+1}}\mathbf{v}_k$$

(d) If $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent then

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

only holds if $c_1 = c_2 = \cdots = c_k = 0$. We can always choose h of these linearly independent vectors, where $h < k$, such that

$$c_1\mathbf{v}_1 + \cdots + c_h\mathbf{v}_h = \mathbf{0}$$

only when $c_1 = c_2 = \cdots = c_h = 0$, for if instead we could find some non-zero c_i, $i \in [1, h]$ such that the equation above would be satisfied, it would be impossible to find $k - h$ more linearly independent vectors to form the system of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, which would be contradictory to our initial assumptions.

(e) To show equality of both sets of coefficients we write

$$\begin{aligned} a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k &= b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k \\ (a_1 - b_1)\,\mathbf{v}_1 + \cdots + (a_k - b_k)\,\mathbf{v}_k &= \mathbf{0} \end{aligned}$$

According to the definition, the difference between each pair of coefficients should vanish for the system of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ to be linear independent. Hence, $a_1 = b_1, a_2 = b_2, \ldots, a_k = b_k$.

(f) Let V^n be a vector space of dimensionality n. Then an arbitrary vector $\mathbf{v}$ in V^n can be expressed uniquely as a linear combination of n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ also in V^n. These n linearly independent vectors form a basis for V^n. Furthermore, let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be an arbitrary set of $n+1$ vectors in V^n. If these $n+1$ vectors are linearly independent then the only scalars $c_1, \ldots, c_k$ such that

$$c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$$

are $c_1 = c_2 = \cdots = c_k = 0$. However, interpreting the above vector equation as a system of linear equations we see that we now end up with an underdetermined system of n equations (since each vector u_i, $i \in [1, n+1]$ only contains n entrees) in $n+1$ unknowns. Hence, there are infinitely many non-trivial solutions and so there exist numbers $c_1, \ldots, c_k$ that are not all zero, which obviously is in disagreement with the condition for linear independence. In conclusion, the $n+1$ vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are linearly dependent and we thus can confirm there do not exist $n+1$ linearly independent vectors in V^n.

(g) A set of n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is said to form a basis for V^n when every vector $\mathbf{v}$ can be expressed uniquely as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$, i.e. if

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

for unique choices of the scalars $c_1, \ldots, c_n$. Since each vector space V^n contains exactly n linearly independent vectors according to rule g) and we can associate a unique set of scalars $c_1, \ldots, c_n$ with them as per rule e), we can always form this basis for V^n.

(h) By rule i) $\mathbf{v}_1, \ldots, \mathbf{v}_k$ do not form a basis, since it is not possible to construct a vector $\mathbf{v} \in V^n$ through linearly combining only $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Hence, we may choose $\mathbf{v}_{k+1}$ such that $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ are linearly independent and we can continue repeating this process $n-k-1$ times more until we have found $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ such that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ does form a basis for V^n.

(i) According to rule j), if $k < n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent vectors in V^n, we can find $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ such that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis for V^n. Next, according to rule g) there do not exist $n+1$ linearly independent vectors in V^n, hence, the set of $n+1$ vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ are linearly dependent, so that

$$c_1\mathbf{u}_1 + \cdots + c_{k+1}\mathbf{u}_{k+1} + c_{k+2}\mathbf{v}_{k+1} + \cdots + c_{n+1}\mathbf{v}_n = \mathbf{0}$$

for scalars $c_1, \ldots, c_{n+1}$ not all equal to zero. Now let us choose $c_1 = 0, \ldots, c_{k+1} = 0$ so that we require that

$$c_{k+1}\mathbf{v}_{k+1} + \cdots + c_{n+1}\mathbf{v}_n = \mathbf{0}$$

for scalars $c_{k+1}, \ldots, c_{n+1}$ not all equal to zero. However, this would contradict the statement that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent and so $c_{k+2}, \ldots, c_{n+1}$ must in fact all equal zero. Hence, there must exist at least one non-zero c_i, where $i \in [1, k+1]$ so that

$$c_1\mathbf{u}_1 + \cdots + c_{k+1}\mathbf{u}_{k+1} = \mathbf{0}$$

from which finally we may conclude that $\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}$ are linearly dependent.

(j) First note that an orthogonal system of non-zero vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is necessarily linear independent (Problem 9, section 1.15). Let us denote these vectors as $\mathbf{v}_1 = \mathbf{u}_1, \ldots, \mathbf{v}_k = \mathbf{u}_k$. Then according to rule j) we can find $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$, where it is assumed that $k < n$, such that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ form a basis for V^n. Next, given these n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ we can construct an orthogonal system of n vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ by applying the Gramm-Schmidt process to $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Hence, we will recover $\mathbf{u}_1, \ldots, \mathbf{u}_k$ together with $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ additional vectors such that $\mathbf{u}_1, \ldots, \mathbf{u}_n$ form an orthogonal basis for V^n.

11. (a) From (1.9) it follows that for an orthogonal system of vectors in V^n it should hold that $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ when $i \neq j$ for then $\theta = \pi/2$ and $\mathbf{u}_i \cdot \mathbf{u}_j > 0$ when $i = j$ for then $\theta = 0$. Since the zero vector $\mathbf{0}$ is agreed to be orthogonal (and parallel) to all vectors, including itself it cannot be used to construct an orthogonal set of vectors. Hence, $\mathbf{u}_i \neq \mathbf{0}$ for $i \in [1, n]$. We can prove the orthogonality of $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ simply by taking the inner product $\mathbf{u}_i \cdot \mathbf{u}_j$ for $i \neq j$, substituting and

finally verifying the result equals 0:

$$\begin{aligned}
\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{v}_1 \cdot \left[\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\,\frac{\mathbf{u}_1}{|\mathbf{u}_1|^2}\right] &= \mathbf{v}_1 \cdot \left[\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2}\right] \\
&= \mathbf{v}_1 \cdot \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_1 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2} \\
&= \mathbf{v}_1 \cdot \mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\,\frac{|\mathbf{v}_1|^2}{|\mathbf{v}_1|^2} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\mathbf{u}_1 \cdot \mathbf{u}_3 &= \mathbf{v}_1 \cdot \left[\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\,\frac{\mathbf{u}_1}{|\mathbf{u}_1|^2} - (\mathbf{v}_3 \cdot \mathbf{u}_2)\,\frac{\mathbf{u}_2}{|\mathbf{u}_2|^2}\right] \\
&= \mathbf{v}_1 \cdot \mathbf{v}_3 - \mathbf{v}_3 \cdot \mathbf{v}_1 - \left[\mathbf{v}_3 \cdot \left(\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\,\frac{\mathbf{u}_1}{|\mathbf{u}_1|^2}\right)\right] \frac{\mathbf{v}_1 \cdot \left[\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\,\frac{\mathbf{u}_1}{|\mathbf{u}_1|^2}\right]}{|\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\,\frac{\mathbf{u}_1}{|\mathbf{u}_1|^2}|^2} \\
&= \mathbf{v}_1 \cdot \mathbf{v}_3 - \mathbf{v}_1 \cdot \mathbf{v}_3 - \left[\mathbf{v}_3 \cdot \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_3 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\right] \frac{\mathbf{v}_1 \cdot \mathbf{v}_2 - \mathbf{v}_2 \cdot \mathbf{v}_1}{|\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2}|^2} \\
&= 0 - \left[\mathbf{v}_3 \cdot \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_3 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2}\right] \frac{0}{|\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2}|^2} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\mathbf{u}_2 \cdot \mathbf{u}_3 &= \mathbf{u}_2 \cdot \left[\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\,\frac{\mathbf{u}_1}{|\mathbf{u}_1|^2} - (\mathbf{v}_3 \cdot \mathbf{u}_2)\,\frac{\mathbf{u}_2}{|\mathbf{u}_2|^2}\right] \\
&= \mathbf{u}_2 \cdot \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{v}_1)\,\frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{|\mathbf{v}_1|^2} - \mathbf{v}_3 \cdot \mathbf{u}_2 \\
&= -\frac{(\mathbf{v}_3 \cdot \mathbf{v}_1)}{|\mathbf{v}_1|^2}\left[\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2}\right] \cdot \mathbf{v}_1 \\
&= -\frac{(\mathbf{v}_3 \cdot \mathbf{v}_1)}{|\mathbf{v}_1|^2}\,(\mathbf{v}_2 \cdot \mathbf{v}_1 - \mathbf{v}_2 \cdot \mathbf{v}_1) \\
&= 0
\end{aligned}$$

Showing that $\mathbf{u}_h$ is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_h$ and $\mathbf{v}_h$ is a linear combination of $\mathbf{u}_1, \dots, \mathbf{u}_h$ is trivially done by substituting and gives

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2} \\
&= a\mathbf{v}_1 + \mathbf{v}_2 \\
\mathbf{u}_3 &= \mathbf{v}_3 - (\mathbf{v}_1 \cdot \mathbf{v}_3)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2} - \left[\mathbf{v}_2 \cdot \mathbf{v}_3 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\,\frac{\mathbf{v}_1 \cdot \mathbf{v}_3}{|\mathbf{v}_1|^2}\right] \frac{\mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2}}{|\mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\,\frac{\mathbf{v}_1}{|\mathbf{v}_1|^2}|^2} \\
&= b\mathbf{v}_1 + c\mathbf{v}_2 + \mathbf{v}_3
\end{aligned}$$

where the scalars a, b and c are given by

$$
\begin{aligned}
a &= -\frac{(\mathbf{v}_1 \cdot \mathbf{v}_2)}{|\mathbf{v}_1|^2} \\
b &= \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1|^2} \frac{\mathbf{v}_2 \cdot \mathbf{v}_3 - (\mathbf{v}_1 \cdot \mathbf{v}_2)(\mathbf{v}_1 \cdot \mathbf{v}_3)/|\mathbf{v}_1|^2}{|\mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\mathbf{v}_1/|\mathbf{v}_1|^2|^2} - \frac{\mathbf{v}_1 \cdot \mathbf{v}_3}{|\mathbf{v}_1|^2} \\
c &= -\frac{\mathbf{v}_2 \cdot \mathbf{v}_3 - (\mathbf{v}_1 \cdot \mathbf{v}_2)(\mathbf{v}_1 \cdot \mathbf{v}_3)/|\mathbf{v}_1|^2}{|\mathbf{v}_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2)\mathbf{v}_1/|\mathbf{v}_1|^2|^2}
\end{aligned}
$$

and

$$
\mathbf{v}_1 = \mathbf{u}_1 \qquad \mathbf{v}_2 = -a\mathbf{u}_1 + \mathbf{u}_2 \qquad \mathbf{v}_3 = (ac - b)\,\mathbf{u}_1 - c\mathbf{u}_2 + \mathbf{u}_3
$$

(b)

$$
\mathbf{u}_1 = \mathbf{i} \qquad \mathbf{u}_2 = \mathbf{j} + \mathbf{k} \qquad \mathbf{u}_3 = \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}
$$

12.

$$
\begin{aligned}
\mathbf{u} + \mathbf{v} &= (u_1^*\mathbf{e}_1^* + \cdots + u_n^*\mathbf{e}_n^*) + (v_1^*\mathbf{e}_1^* + \cdots + v_n^*\mathbf{e}_n^*) \\
&= u_1^*\mathbf{e}_1^* + v_1^*\mathbf{e}_1^* + \cdots + u_n^*\mathbf{e}_n^* + v_n^*\mathbf{e}_n^* \\
&= (u_1^* + v_1^*)\,\mathbf{e}_1^* + \cdots + (u_n^* + v_n^*)\,\mathbf{e}_n^* \\
h\mathbf{v} &= h\,(v_1^*\mathbf{e}_1^* + \cdots + v_n^*\mathbf{e}_n^*) \\
&= v_1^*\,(h\mathbf{e}_1^*) + \cdots + v_n^*\,(h\mathbf{e}_n^*) \\
&= (hv_1^*)\,\mathbf{e}_1^* + \cdots + (hv_n^*)\,\mathbf{e}_n^* \\
\mathbf{u} \cdot \mathbf{v} &= (u_1^*\mathbf{e}_1^* + \cdots + u_n^*\mathbf{e}_n^*) \cdot (v_1^*\mathbf{e}_1^* + \cdots + v_n^*\mathbf{e}_n^*) \\
&= u_1^*v_1^*\,(\mathbf{e}_1^* \cdot \mathbf{e}_1^*) + \cdots + u_n^*v_n^*\,(\mathbf{e}_n^* \cdot \mathbf{e}_n^*) \\
&= u_1^*v_1^* + \cdots + u_n^*v_n^*
\end{aligned}
$$

13. (a) Let $\mathbf{v}$ be an arbitrary vector in V^n and $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$ and $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be two different orthonormal bases of V^n, so that

$$
\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n = v_1^*\mathbf{e}_1^* + \cdots + v_n^*\mathbf{e}_n^*
$$

Taking an inner product with $\mathbf{e}_i^*$ for $i = 1, ..., n$ then gives

$$
\mathbf{v} \cdot \mathbf{e}_i^* = v_1\,(\mathbf{e}_1 \cdot \mathbf{e}_i^*) + \cdots + v_n\,(\mathbf{e}_n \cdot \mathbf{e}_i^*) = \sum_j^n v_j\,(\mathbf{e}_j \cdot \mathbf{e}_i^*) = \sum_j^n \mathbf{e}_i^* \cdot (v_j\mathbf{e}_j) = v_i^*
$$

We may associate each inner product $\mathbf{e}_j \cdot \mathbf{e}_i^*$ with an entree a_{ij} of the matrix $\mathbf{A}$ for then we can write $v_i^* = \mathbf{a}_i \cdot \mathrm{col}\,(v_i, \ldots, v_n)$, where $\mathbf{a}_i$ denotes the i'th row of the matrix $\mathbf{A}$. Hence, we have

$$
\begin{aligned}
\mathrm{col}\,(v_1^*, \ldots, v_n^*) &= \begin{bmatrix} \mathbf{e}_1^* \cdot \mathbf{e}_1 & \cdots & \mathbf{e}_1^* \cdot \mathbf{e}_n \\ \vdots & & \vdots \\ \mathbf{e}_n^* \cdot \mathbf{e}_1 & \cdots & \mathbf{e}_n^* \cdot \mathbf{e}_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\
&= \mathbf{A}\mathrm{col}\,(v_1, \ldots, v_n)
\end{aligned}
$$

(b) The jth column of $\mathbf{A}$ is given by $\mathbf{e}_1^* \cdot \mathbf{e}_j, \ldots, \mathbf{e}_n^* \cdot \mathbf{e}_j = \mathbf{e}_j \cdot \mathbf{e}_1^*, \ldots, \mathbf{e}_j \cdot \mathbf{e}_n^*$ so that

$$e_{1j} = \mathbf{e}_j \cdot \mathbf{e}_1^*, \ldots, e_{nj} = \mathbf{e}_j \cdot \mathbf{e}_n^*$$

from which follows that the jth column of $\mathbf{A}$ gives the components of $\mathbf{e}_j$ with respect to the basis $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$. Similarly, the ith column of $\mathbf{A}$ is given by $\mathbf{e}_i^* \cdot \mathbf{e}_1, \ldots, \mathbf{e}_i^* \cdot \mathbf{e}_n$ so that

$$e_{1i}^* = \mathbf{e}_i^* \cdot \mathbf{e}_1, \ldots, e_{ni}^* = \mathbf{e}_i^* \cdot \mathbf{e}_n$$

from which follows that the ith row of $\mathbf{A}$ gives the components of $\mathbf{e}_i^*$ with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

(c) If $\mathbf{A}$ is in fact an orthogonal matrix then it should hold that $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$. Writing out the matrix product explicitly

$$\mathbf{A}\mathbf{A}^\top = \begin{bmatrix} \mathbf{e}_1^* \cdot \mathbf{e}_1 & \cdots & \mathbf{e}_1^* \cdot \mathbf{e}_n \\ \vdots & & \vdots \\ \mathbf{e}_n^* \cdot \mathbf{e}_1 & \cdots & \mathbf{e}_n^* \cdot \mathbf{e}_n \end{bmatrix} \begin{bmatrix} \mathbf{e}_1^* \cdot \mathbf{e}_1 & \cdots & \mathbf{e}_n^* \cdot \mathbf{e}_1 \\ \vdots & & \vdots \\ \mathbf{e}_1^* \cdot \mathbf{e}_n & \cdots & \mathbf{e}_n^* \cdot \mathbf{e}_n \end{bmatrix}$$

we see that the product of the ith row with the jth column is given by

$$\begin{aligned}
\mathbf{a}_i \cdot \mathrm{col}\left(\mathbf{a}_j^\top\right) &= \left(\mathbf{e}_i^* \cdot \mathbf{e}_1\right)\left(\mathbf{e}_j^* \cdot \mathbf{e}_1\right) + \cdots + \left(\mathbf{e}_i^* \cdot \mathbf{e}_n\right)\left(\mathbf{e}_j^* \cdot \mathbf{e}_n\right) \\
&= \left(\mathbf{e}_i^* \cdot \mathbf{e}_1\right)\left(\mathbf{e}_1 \cdot \mathbf{e}_j^*\right) + \cdots + \left(\mathbf{e}_i^* \cdot \mathbf{e}_n\right)\left(\mathbf{e}_n \cdot \mathbf{e}_j^*\right) \\
&= \mathbf{e}_i^* \cdot \left[\mathbf{e}_1\left(\mathbf{e}_1 \cdot \mathbf{e}_j^*\right) + \cdots + \mathbf{e}_n\left(\mathbf{e}_n \cdot \mathbf{e}_j^*\right)\right] \\
&= \mathbf{e}_i^* \cdot \left[\mathbf{e}_1\mathbf{e}_1 + \cdots + \mathbf{e}_n\mathbf{e}_n\right] \cdot \mathbf{e}_j^* \\
&= \mathbf{e}_i^* \mathbf{I} \mathrm{col}\left(\mathbf{e}_j^*\right) \\
&= \mathbf{e}_i^* \cdot \mathbf{e}_j^* \\
&= \delta_{jk}
\end{aligned}$$

In the equality above it is assumed that $\mathbf{e}_1\mathbf{e}_1 + \cdots + \mathbf{e}_n\mathbf{e}_n = \mathbf{I}$, although we haven't proven this yet. Lets consider the matrix product $\left(\mathbf{e}_1\mathbf{e}_1 + \cdots + \mathbf{e}_n\mathbf{e}_n\right)\mathbf{v}$ for an arbitrary column vector $\mathbf{v} = \mathrm{col}\left(v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n\right)$. We then get

$$\begin{aligned}
\left(\mathbf{e}_1\mathbf{e}_1 + \cdots + \mathbf{e}_n\mathbf{e}_n\right)\mathbf{v} = \left(\sum_i^n \mathbf{e}_i\mathbf{e}_i\right)\left(\sum_j^n v_j\mathbf{e}_j\right) &= \sum_j^n v_j \sum_i^n \mathbf{e}_i\left(\mathbf{e}_i \cdot \mathbf{e}_j\right) \\
&= \sum_j^n v_j \sum_i^n \mathbf{e}_i\delta_{ij} \\
&= \sum_j^n v_j\mathbf{e}_j \\
&= \mathbf{v}
\end{aligned}$$

from which we may conclude that indeed $\mathbf{e}_1\mathbf{e}_1 + \cdots + \mathbf{e}_n\mathbf{e}_n = \mathbf{I}$.

14. (a) Let the point $\tilde{x}_i = x_i - h_i$, $i = 1, \ldots, n$ be the shifted coordinates in the orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ with new origin O^*. Then following Problem 13, the coordinates in the new orthonormal basis $\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*$ are given by the matrix product

$$\operatorname{col}(x_1^*, \ldots, x_n^*) = \mathbf{A}\operatorname{col}(\tilde{x}_i, \ldots, \tilde{x}_n) = \mathbf{A}\operatorname{col}(x_1 - h_1, \ldots, x_n - h_n)$$

where the matrix $\mathbf{A}$ is the same matrix that was found for Problem 13.

(b) This follows readily from the 1st part of (a). In the case the origin is changed each original coordinate x_i will be shifted by a possible different amount h_i, $i = 1, \ldots, n$. Hence, for an arbitrary vector $\overrightarrow{O^*P}$ we have

$$\overrightarrow{O^*P} = x_1^*\mathbf{e}_1 + \cdots + x_n^*\mathbf{e}_n = (x_1 - h_1)\,\mathbf{e}_1 + \cdots + (x_n - h_n)\,\mathbf{e}_n$$

Note that since there is no change of basis, we use the same basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ on each side of the equation. Dotting both sides with $\mathbf{e}_i$ for $i = 1, \ldots, n$ then gives

$$x_i^* = x_i - h_i, \qquad i = 1, \ldots, n$$

Section 1.16

1. (a)

$$T(1,0) = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad T(0,1) = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$
$$T(2,-1) = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad T(-1,1) = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

(b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

from which follows that the kernel is the $\mathbf{0}$ vector alone. Hence, T is one to one. All $\mathbf{x}$ such that $T(\mathbf{x}) = (2,3)$ are found by solving the equation

$$T(\mathbf{x}) = \operatorname{col}(2,3) = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

from which follows $T(\mathbf{x}) = \operatorname{col}(2,3)$ only for $\mathbf{x} = \operatorname{col}(1,0)$.

(c) The range of T is all of V^2 and so T maps V^2 onto V^2.

2. (a)

$$T(1,0) = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \qquad T(0,1) = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$
$$T(1,-1) = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} \qquad T(-1,-1) = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -5 \\ -10 \end{bmatrix}$$

(b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

from which follows that the kernel contains all vectors of the form $\mathbf{x} = t(-3, 2)$ for $-\infty < t < \infty$. Hence, T is not one to one. All $\mathbf{x}$ such that $T(\mathbf{x}) = (2, 4)$ are found by solving the equation

$$T(\mathbf{x}) = \mathrm{col}(2, 4) = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

from which follows $T(\mathbf{x}) = \mathrm{col}(2, 4)$ for all vectors $\mathbf{x} = (1, 0) + t(-3, 2)$ for $-\infty < t < \infty$.

(c) The range of T is the set of all vectors $\mathbf{y}$ for which $T(x) = \mathbf{y}$ for at least one vector $\mathbf{x}$. Now

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} x_1 + \begin{bmatrix} 3 \\ 6 \end{bmatrix} x_2 = (2x_1 + 3x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{y}$$

I.e. the two columns of $\mathbf{B}$ are linearly dependent. Hence, the range of T is given by all vectors $t(1, 2)$, where $t = 2x_1 + 3x_2$ and so T does not map V^2 onto V^2.

3. (a) The number of columns of $\mathbf{F}$ gives $n = 3$, the number of row of $\mathbf{F}$ gives $m = 2$.

(b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

from which follows that the kernel contains all vectors of the form $\mathbf{x} = t(0, -2, 1)$ for $-\infty < t < \infty$. Hence, T is not one to one.

(c) The range of T is the set of all vectors $\mathbf{y}$ of the form

$$T(\mathbf{x}) = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_2 + \begin{bmatrix} 2 \\ 4 \end{bmatrix} x_3 = x_1 \begin{bmatrix} 3 \\ 0 \end{bmatrix} + (x_2 + 2x_3) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{y}$$

I.e. the second and third column of $\mathbf{F}$ are linearly dependent. Hence, the range of T is V^2 and T maps V^3 onto V^2.

4. (a) The number of columns of $\mathbf{G}$ gives $n = 3$, the number of rows of $\mathbf{G}$ gives $m = 2$.

(b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 8 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

from which follows that the kernel contains all vectors of the form $\mathbf{x} = t_1(-4, 1, 0) + t_2(-3, 0, 1)$ for $-\infty < t_1, t_2 < \infty$. Hence, T is not one to one.

(c) The range of T is the set of all vectors $\mathbf{y}$ of the form

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 8 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 4 \\ 8 \end{bmatrix} x_2 + \begin{bmatrix} 3 \\ 6 \end{bmatrix} x_3 = (x_1 + 4x_2 + 3x_3) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{y}$$

I.e. all three columns of $\mathbf{G}$ are linearly dependent. Hence, the range of T are all vectors $t(1,2)$ where $t = x_1 + 4x_2 + 3x_3$ and T does not map V^3 onto V^2.

5. (a) The number of columns of $\mathbf{H}$ gives $n = 2$, the number of rows of $\mathbf{H}$ gives $m = 3$.

   (b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \\ -6 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

   from which follows that the kernel contains all vectors of the form $\mathbf{x} = t(1,-2)$ for $-\infty < t < \infty$. Hence, T is not one to one.

   (c) The range of T is the set of all vectors $\mathbf{y}$ of the form

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 1 \\ 4 & 2 \\ -6 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -6 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} x_2 = (2x_1 + x_2) \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = \mathbf{y}$$

   I.e. the two columns of $\mathbf{H}$ are linearly dependent. Hence, the range of T are all vectors $t(1,2,-3)$ where $t = 2x_1 + x_2$ and T does not map V^2 onto V^3.

6. (a) The number of columns of $\mathbf{J}$ gives $n = 2$, the number of rows of $\mathbf{H}$ gives $m = 3$.

   (b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

   from which follows that the kernel is the $\mathbf{0}$ vector alone. Hence, T is one to one.

   (c) The range of T is the set of all vectors $\mathbf{y}$ of the form

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} x_2 = \mathbf{y}$$

   I.e. the range of T are all vectors $t_1(2,1,1) + t_2(1,2,2)$ where $t_1 = x_1$ and $t_2 = x_2$ and T does not map V^2 onto V^3.

7. (a) The number of columns of $\mathbf{K}$ gives $n = 3$, the number of rows of $\mathbf{K}$ gives $m = 3$.

(b) The kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \\ 5 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

from which follows that the kernel contains all vectors of the form $\mathbf{x} = t(-1, 1, 1)$ for $-\infty < t < \infty$. Hence, T is not one to one.

(c) The range of T is the set of all vectors $\mathbf{y}$ of the form

$$T(\mathbf{x}) = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \\ 5 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} x_2 + \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} x_3 = \mathbf{y}$$

Referring back to part (b) of the problem, it was found that the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \\ 5 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

where $c_1 = x_1$, $c_2 = x_2$, $c_3 = x_3$ and $\mathbf{v}_i$ for $1 \le i \le 3$ is the ith column of $\mathbf{K}$, has infinitely many solutions of the form $\mathbf{x} = t(-1, 1, 1)$. Hence, the three columns of $\mathbf{K}$ are linearly dependent. Choosing $t = 1$, we can thus write $\mathbf{v}_3 = \mathbf{v}_1 - \mathbf{v}_2$ and so

$$\begin{aligned} \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} x_2 + \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} x_3 &= \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} x_2 + \left(\begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right) x_3 \\ &= (x_1 + x_3) \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} + (x_2 - x_3) \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \\ &= \mathbf{y} \end{aligned}$$

Hence, the range of T are all vectors $t_1(3, 1, 5) + t_2(1, 0, 2)$ where $t_1 = x_1 + x_3$ and $t_2 = x_2 - x_3$ and T does not map V^3 onto V^3.

8. (a) The number of columns of $\mathbf{L}$ gives $n = 3$, the number of rows of $\mathbf{L}$ gives $m = 3$.

(b) the kernel of T may be found by solving the equation

$$T(\mathbf{x}) = \mathbf{0} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

from which follows that the kernel is the $\mathbf{0}$ vector alone. Hence, T is one to one.

(c) The range of T is the set of all vectors $\mathbf{y}$ of the form

$$T(\mathbf{x}) = \begin{bmatrix}1\\0\\0\end{bmatrix} x_1 + \begin{bmatrix}0\\1\\1\end{bmatrix} x_2 + \begin{bmatrix}1\\0\\1\end{bmatrix} x_3 = \mathbf{y}$$

Since all columns of $\mathbf{L}$ are linearly independent the range of T are all vectors $t_1(1,0,0) + t_2(0,1,1) + t_3(1,0,1)$ where $t_1 = x_1$, $t_2 = x_2$ and $t_3 = x_3$ and T maps V^3 onto V^3.

9. The linear mapping $\mathbf{y} = T(\mathbf{x}) = \mathbf{Cx}$ is explicitly given by

$$\mathbf{y} = T(\mathbf{x}) = \mathbf{Cx} = \begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix} \begin{bmatrix}x_1\\x_2\end{bmatrix}$$

from which follows

$$\mathbf{x} = (x_1, x_2) \qquad \mathbf{y} = \left(\frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2, \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2\right)$$

The angle between the two vectors $\mathbf{x}$ and $\mathbf{y}$ is given by $\theta = \cos^{-1}((\mathbf{x}\cdot\mathbf{y})/(|\mathbf{x}||\mathbf{y}|))$. Now $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$ and

$$|\mathbf{y}| = \sqrt{\left(\frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2\right)^2 + \left(\frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2\right)^2} = \sqrt{x_1^2 + x_2^2}$$

Hence, $|\mathbf{x}| = |\mathbf{y}|$. Next

$$\mathbf{x}\cdot\mathbf{y} = (x_1, x_2) \cdot \left(\frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2, \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2\right) = \frac{1}{\sqrt{2}}\left(x_1^2 + x_2^2\right)$$

From this it follows that

$$\theta = \cos^{-1}\left(\frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|}\right) = \cos^{-1}\left(\frac{1}{\sqrt{2}}\frac{x_1^2 + x_2^2}{x_1^2 + x_2^2}\right) = \cos^{-1}\frac{1}{\sqrt{2}} = \frac{\pi}{4}$$

The geometric interpretation of this is that the vector $\mathbf{y}$ is the vector $\mathbf{x}$ rotated by $\pi/4$. The linear mapping T thus may be interpreted as a rotation operator for vectors in V^2.

10. Applying the linear mapping T with equation $\mathbf{y} = \mathbf{Dx}$ to an arbitrary vector $\mathbf{x}$ in V^2 gives

$$\mathbf{y} = T(\mathbf{x}) = \mathbf{Dx} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = (x_1, -x_2)$$

Hence, the vector $\mathbf{y}$ is the reflection of the vector $\mathbf{x}$ about the x-axis.

11. (a) Applying the linear mapping T with equation $\mathbf{y} = \mathbf{E}\mathbf{x}$ to an arbitrary vector $\mathbf{x}$ in V^2 gives

$$\mathbf{y} = T(\mathbf{x}) = \mathbf{E}\mathbf{x} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = (-x_1, -x_2)$$

Hence, the vector $\mathbf{y}$ is the reflection of the vector $\mathbf{x}$ about the origin.

(b) Applying the linear mapping T with equation $\mathbf{y} = \mathbf{M}\mathbf{x}$ to an arbitrary vector $\mathbf{x}$ in V^3 gives

$$\mathbf{y} = T(\mathbf{x}) = \mathbf{M}\mathbf{x} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = 2(x_1, x_2, x_3)$$

Hence, the vector $\mathbf{y}$ is the vector $\mathbf{x}$ stretched by a ratio of 2 to 1.

12. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be eigenvectors of the linear mapping T having matrix $\mathbf{A}$, associated with the distinct eigenvalues $\lambda_1, \dots, \lambda_k$ so that $\mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$. It is then stated that $\mathbf{v}_1, \dots, \mathbf{v}_k$ are linearly independent. We start by verifying that this holds for $k = 1$. By the definition of linear independence $c_1\mathbf{v}_1 = \mathbf{0}$ only when c_1. Multiplying each side of the equation by $\mathbf{A}$ from the left gives

$$c_1\mathbf{A}\mathbf{v}_1 = \mathbf{0} \qquad \Longrightarrow \qquad c_1\lambda_1\mathbf{v}_1$$

Since $\mathbf{v}_1$ cannot be the zero vector (as it is considered to be an eigenvector) and λ_1 can be any scalar including zero, we conclude at once that $c_1 = 0$ is the only option. For $k = m < n$ we have

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

and thus we need to prove that $c_1, \dots, c_k$ are zero for $\mathbf{v}_1, \dots, \mathbf{v}_k$ to be linearly independent. To this end we again multiply the equation above by $\mathbf{A}$ from the left to obtain

$$c_1\mathbf{A}\mathbf{v}_1 + \cdots + c_k\mathbf{A}\mathbf{v}_k = \mathbf{0} \qquad \Longrightarrow \qquad c_1\lambda_1\mathbf{v}_1 + \cdots + c_k\lambda_k\mathbf{v}_k = \mathbf{0}$$

Multiplying the same equation by λ_1 and subtracting it from the first then gives

$$\begin{aligned} c_1\lambda_1\mathbf{v}_1 + \cdots + c_k\lambda_k\mathbf{v}_k - (c_1\lambda_1\mathbf{v}_1 + \cdots + c_k\lambda_1\mathbf{v}_k) = \mathbf{0} \\ 0\mathbf{v}_1 + c_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \cdots + c_k(\lambda_k - \lambda_1)\mathbf{v}_k = \end{aligned}$$

Since each λ_i is distinct and $\mathbf{v}_i \neq \mathbf{0}$ for $1 \leq i \leq k$ since it is an eigenvector, we may conclude that $c_2 = \cdots = c_k = 0$, which by the argument previously given automatically implies that $c_1 = 0$. Next, consider the case for $k = m + 1$. The condition for linear independence is then to state that the only scalars $c_1, \dots, c_{m+1}$ such that

$$c_1\mathbf{v}_1 + \cdots + c_{m+1}\mathbf{v}_{m+1} = \mathbf{0}$$

are $c_1 = c_2 = \cdots = c_{m+1} = 0$. Multiplying the equation above by $\mathbf{A}$ from the left as before and subtracting from it the product of λ_{m+1} with the same equation gives

$$\begin{aligned} c_1\lambda_1\mathbf{v}_1 + \cdots + c_{m+1}\lambda_{m+1}\mathbf{v}_{m+1} - (c_1\lambda_{m+1}\mathbf{v}_1 + \cdots + c_{m+1}\lambda_{m+1}\mathbf{v}_{m+1}) &= \mathbf{0} \\ c_1(\lambda_1 - \lambda_{m+1})\mathbf{v}_1 + \cdots + c_m(\lambda_m - \lambda_{m+1})\mathbf{v}_m + 0\mathbf{v}_{m+1} &= \end{aligned}$$

Again, since each λ_i is distinct and $\mathbf{v}_i \neq \mathbf{0}$ for $1 \leq i \leq m$ since it is an eigenvector, we may conclude that $c_1 = \cdots = c_m = 0$. Hence, the linear independence equation reduces to

$$c_{m+1}\mathbf{v}_{m+1} = \mathbf{0}$$

which by the same reasoning as for the case $k = 1$ (i.e. that $\mathbf{v}_{m+1}$ cannot be the zero vector, since it is assumed to be an eigenvector) implies $c_{m+1} = 0$.

13. (a) If the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly dependent vectors of V^n, then there exists at least one non-zero scalar c_i for $1 \leq i \leq k$ such that

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Now consider the linear mapping $T(\mathbf{v}_i) = \mathbf{w}_i$ for $1 \leq i \leq k$ which maps each vector $\mathbf{v}_i$ of V^n to a vector $\mathbf{w}_i$ of V^m. Note that since the mapping is into, but not necessarily one to one there may exist multiple $\mathbf{v}_i$'s of V^n that map to the same vector $\mathbf{w}_i$ in V^m. This will be relevant for the last part of the proof. Furthermore, since the mapping is linear we can write $T(\mathbf{v}_i) = \mathbf{A}\mathbf{v}_i = \mathbf{w}_i$ for some $m \times n$ matrix $\mathbf{A}$. Applying the linear transformation T to both sides of the equation for linear dependence written down earlier and assuming there exists at least one nonzero c_i then gives

$$\begin{aligned} T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) &= T(\mathbf{0}) \\ \mathbf{A}(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) &= \mathbf{A}\mathbf{0} \\ c_1\mathbf{A}\mathbf{v}_1 + \cdots + c_k\mathbf{A}\mathbf{v}_k &= \mathbf{0} \\ c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k &= \end{aligned}$$

Since the scalars $c_1, \ldots, c_k$ are the same as before the mapping and hence, at least one of them is nonzero, we may conclude that the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_k$ are linear dependent vectors of V^m. The converse is not necessarily true, since as mentioned earlier, the mapping T may not be one to one in which case the matrix $\mathbf{A}$ is singular. In other words, applying the inverse mapping T^{-1} to the condition for linear dependence of the set $\mathbf{w}_1, \ldots, \mathbf{w}_k$ of V^m:

$$\begin{aligned} T^{-1}(c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k) &= T^{-1}(\mathbf{0}) \\ \mathbf{A}^{-1}(c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k) &= \mathbf{A}^{-1}\mathbf{0} \end{aligned}$$

will only make sense if the inverse matrix $\mathbf{A}^{-1}$ exists.

(b) Let $\mathbf{A}$ be a nonsingular $n \times n$ matrix so that its inverse matrix $\mathbf{A}^{-1}$ exists. Now a set of vectors $\mathbf{w}_1, \ldots, \mathbf{w}_k = \mathbf{A}\mathbf{v}_1, \ldots, \mathbf{A}\mathbf{v}_k$ of V^n is linear independent if the only scalars $c_1, \ldots, c_k$ such that

$$c_1\mathbf{A}\mathbf{v}_1 + \cdots + c_k\mathbf{A}\mathbf{v}_k = \mathbf{0}$$

are $c_1 = c_2 = \cdots = c_k = 0$. Since $\mathbf{A}^{-1}$ exists we may apply the inverse transformation to the equation above to obtain

$$\begin{aligned} T^{-1}\left(c_1\mathbf{A}\mathbf{v}_1 + \cdots + c_k\mathbf{A}\mathbf{v}_k\right) &= T^{-1}\left(\mathbf{0}\right) \\ \mathbf{A}^{-1}\left(c_1\mathbf{A}\mathbf{v}_1 + \cdots + c_k\mathbf{A}\mathbf{v}_k\right) &= \mathbf{A}^{-1}\mathbf{0} \\ c_1\mathbf{A}^{-1}\mathbf{A}\mathbf{v}_1 + \cdots + c_k\mathbf{A}^{-1}\mathbf{A}\mathbf{v}_k &= \mathbf{0} \\ c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k &= \end{aligned}$$

Note that since the scalars $c_1, \ldots, c_k$ are the same as before and so $c_1 = c_2 = \cdots = c_k = 0$ still, we may conclude that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of V^n are linearly dependent also.

14. (a) Let us start by re-writing the vector $\mathbf{u} = p_1\mathbf{u}_1 + \cdots + p_n\mathbf{u}_n$ as the matrix equation

$$\mathbf{u} = p_1\mathbf{u}_1 + \cdots + p_n\mathbf{u}_n = \begin{bmatrix}\mathbf{u}_1 & \ldots & \mathbf{u}_n\end{bmatrix}\begin{bmatrix}p_1 \\ \vdots \\ p_n\end{bmatrix} = \mathbf{U}\mathrm{col}\left(p_1, \ldots, p_n\right)$$

where the $n \times n$ matrix $\mathbf{U}$ comprises of the n column vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Similarly, for $\mathbf{v} = q_1\mathbf{v}_1, \ldots, q_m\mathbf{v}_m$ we get

$$\mathbf{v} = q_1\mathbf{v}_1 + \cdots + q_m\mathbf{v}_m = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_m\end{bmatrix}\begin{bmatrix}q_1 \\ \vdots \\ q_m\end{bmatrix} = \mathbf{V}\mathrm{col}\left(q_1, \ldots, q_m\right)$$

where the $m \times m$ matrix $\mathbf{V}$ contains the m column vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Next, let T be a linear mapping of V^n into V^m so that T assigns to each vector $\mathbf{u}$ of V^n a vector $\mathbf{v}$ of V^m. Since the mapping T is linear we can write $T\left(\mathbf{u}\right) = \mathbf{A}\mathbf{u} = \mathbf{v}$ or

$$T\left(\mathbf{u}\right) = \mathbf{A}\left(p_1\mathbf{u}_1 + \cdots + p_n\mathbf{u}_n\right) = \mathbf{A}\mathbf{U}\mathrm{col}\left(p_1, \ldots, p_n\right) = \mathbf{V}\mathrm{col}\left(q_1, \ldots, q_m\right)$$

for some $m \times n$ matrix $\mathbf{A}$. Provided that $\mathbf{V}$ is invertible (which it is because all column vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent so that $\det\left(\mathbf{V}\right) \neq 0$), we thus may write

$$\mathbf{V}^{-1}\mathbf{A}\mathbf{U}\mathrm{col}\left(p_1, \ldots, p_n\right) = \mathbf{B}\left(p_1, \ldots, p_n\right) = \mathrm{col}\left(q_1, \ldots, q_m\right)$$

(b) Similar to part (a), the vector $\mathbf{u}$ may be re-written as the matrix equation

$$\begin{aligned}
\mathbf{u} = p_1\mathbf{u}_1 + \cdots + p_n\mathbf{u}_n &= p_1\left(u_{11}\mathbf{e}_1 + \cdots + u_{n1}\mathbf{e}_n\right) + \cdots + p_n\left(u_{1n}\mathbf{e}_1 + \cdots + u_{nn}\mathbf{e}_n\right) \\
&= p_1\begin{bmatrix}\mathbf{e}_1 & \ldots & \mathbf{e}_n\end{bmatrix}\begin{bmatrix}u_{11}\\ \vdots \\ u_{n1}\end{bmatrix} + \cdots + p_n\begin{bmatrix}\mathbf{e}_1 & \ldots & \mathbf{e}_n\end{bmatrix}\begin{bmatrix}u_{1n}\\ \vdots \\ u_{nn}\end{bmatrix} \\
&= \begin{bmatrix}\mathbf{e}_1 & \ldots & \mathbf{e}_n\end{bmatrix}\left(p_1\begin{bmatrix}u_{11}\\ \vdots \\ u_{n1}\end{bmatrix} + \cdots + p_n\begin{bmatrix}u_{1n}\\ \vdots \\ u_{nn}\end{bmatrix}\right) \\
&= \begin{bmatrix}\mathbf{e}_1 & \ldots & \mathbf{e}_n\end{bmatrix}\left(p_1\mathbf{u}_1 + \cdots + p_n\mathbf{u}_n\right) \\
&= \begin{bmatrix}\mathbf{e}_1 & \ldots & \mathbf{e}_n\end{bmatrix}\begin{bmatrix}\mathbf{u}_1 & \ldots & \mathbf{u}_n\end{bmatrix}\begin{bmatrix}p_1\\ \vdots \\ p_n\end{bmatrix} \\
&= \mathbf{E}^{(n)}\mathbf{U}\mathrm{col}\left(p_1, \ldots, p_n\right) \\
&= \mathbf{U}\mathrm{col}\left(p_1, \ldots, p_n\right)
\end{aligned}$$

where the last step follows from the fact that $\mathbf{E}^{(n)} = \mathbf{I}$ and as for part (a), the $n \times n$ matrix $\mathbf{U}$ comprises of the column vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Going through the same steps for $\mathbf{v}$ we find that

$$\mathbf{v} = \mathbf{E}^{(m)}\mathbf{V}\mathrm{col}\left(q_1, \ldots, q_m\right) = \mathbf{V}\mathrm{col}\left(q_1, \ldots, q_m\right)$$

Hence, we again find that

$$T\left(\mathbf{u}\right) = \mathbf{A}\mathbf{u} = \mathbf{A}\mathbf{U}\mathrm{col}\left(p_1, \ldots, p_n\right) = \mathbf{V}\mathrm{col}\left(q_1, \ldots, q_m\right)$$

so that $\mathbf{B} = \mathbf{V}^{-1}\mathbf{A}\mathbf{U}$.

(c) If only one basis is used so that $\mathbf{v}_1 = \mathbf{u}_1, \ldots, \mathbf{v}_n = \mathbf{u}_n$, then it should be obvious that $\mathbf{V} = \mathbf{U}$ and so

$$\mathbf{B} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$$

which shows that the $n \times n$ matrix $\mathbf{B}$ is similar to the $n \times n$ matrix $\mathbf{A}$. If the basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is orthonormal, it is both orthogonal and normalized and so $\mathbf{U}^\top\mathbf{U} = \mathbf{I}$, since $\mathbf{u}_i \cdot \mathrm{col}\left(\mathbf{u}_j\right) = 1$ only if $i = j$ and $\mathbf{u}_i \cdot \mathrm{col}\left(\mathbf{u}_j\right) = 0$ whenever $i \neq j$. I.e. the $n \times n$ matrix $\mathbf{U}$ is orthogonal. Hence, by (1.94) the matrix $\mathbf{B}$ is said to be orthogonally congruent to matrix $\mathbf{A}$.

Section 1.17

1. (a) *Finding the rank by finding the maximum number of linearly independent row vectors.* By employing Gaussian elimination we find

$$\begin{bmatrix}3 & 1 & 2\\ 2 & 5 & 0\\ 5 & -7 & 6\end{bmatrix} \to \begin{bmatrix}1 & 1/3 & 2/3\\ 0 & 13/3 & -4/3\\ 3 & -23/3 & 14/3\end{bmatrix} \to \begin{bmatrix}1 & 1/3 & 2/3\\ 0 & 1 & -4/13\\ 0 & 1 & -4/13\end{bmatrix} \to \begin{bmatrix}1 & 1/3 & 2/3\\ 0 & 1 & -4/13\\ 0 & 0 & 0\end{bmatrix}$$

Hence $\mathbf{u}_1 = (3, 1, 2)$ and $\mathbf{u}_2 = (0, 13, -4)$ form a basis and the matrix has dimension $2 = r =$ rank of the matrix.

Finding the rank by finding the maximum number of linearly independent column vectors. To find a basis for the range of the matrix we employ Gaussian elimination to the matrix transpose to find

$$\begin{bmatrix} 3 & 2 & 5 \\ 1 & 5 & -7 \\ 2 & 0 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -7 \\ 0 & -13 & 26 \\ 0 & -10 & 20 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, $\mathbf{z}_1 = \operatorname{col}(1, 5, -7)$ and $\mathbf{z}_2 = \operatorname{col}(0, 1, -2)$ are a basis for the range and so we conclude that again, the dimension of the range $= 2 = r$ is the rank of the matrix.

(b) *Determinant definition of rank.* We take the determinant of the matrix (i.e. the 3×3 minor to get

$$\begin{vmatrix} 1 & 2 & 2 \\ 3 & 5 & 4 \\ -1 & 4 & 2 \end{vmatrix} = (1)(10 - 16) - (2)(6 + 4) + (2)(12 + 5) = 8 \neq 0$$

Hence, we find the rank $r = 3$.

Finding the rank by finding the maximum number of linearly independent row vectors. By employing Gaussian elimination we find

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

from which follows that the 3 vectors making up the columns of the matrix are linearly independent and so we again verify the rank $r = 3$.

(c) *Finding the rank by finding the maximum number of linearly independent row vectors.* By employing Gaussian elimination we find

$$\begin{bmatrix} 0 & 1 & 3 & 4 \\ 1 & 0 & 0 & 5 \\ 3 & 0 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 0 & 2 & -12 \\ 0 & 1 & 3 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 22 \\ 0 & 0 & 1 & -6 \end{bmatrix}$$

From which follows that the vectors $\mathbf{u}_1 = (0, 1, 3)$, $\mathbf{u}_2 = (1, 0, 0)$ and $\mathbf{u}_3 = (3, 0, 2)$ form a basis and so the matrix has dimension $3 = r =$ rank of the matrix.

Finding the rank by finding the maximum number of linearly independent column vectors. To find a basis for the range of the matrix we employ Gaussian elimination to the matrix transpose to find

$$\begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 0 \\ 3 & 0 & 2 \\ 4 & 5 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 5 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, $\mathbf{z}_1 = \text{col}\,(1, 0, 0)$, $\mathbf{z}_2 = \text{col}\,(0, 1, 3)$ and $\mathbf{z}_3 = \text{col}\,(0, 0, 1)$ are a basis for the range and so we conclude that again, the dimension of the range $= 3 = r$ is the rank of the matrix.

(d) *Finding the rank by finding the maximum number of linearly independent row vectors.* By employing Gaussian elimination we find

$$\begin{bmatrix} 2 & 1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 & 3 \\ 5 & 3 & -1 & 2 & 3 \\ 0 & -1 & 2 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 1 & 3 \\ 2 & 1 & 0 & 1 & 2 \\ 5 & 3 & -1 & 2 & 3 \\ 0 & -1 & 2 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 1 & 3 \\ 0 & 1 & -2 & -1 & -4 \\ 0 & 3 & -6 & -3 & -12 \\ 0 & -1 & 2 & 1 & 4 \end{bmatrix}$$
$$\to \begin{bmatrix} 1 & 0 & 1 & 1 & 3 \\ 0 & 1 & -2 & -1 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

From which follows that the vectors $\mathbf{u}_1 = (1, 0, 1, 1, 3)$ and $\mathbf{u}_2 = (0, 1, -2, -1, -4)$ form a basis and so the matrix has dimension $2 = r =$ rank of the matrix.

Finding the rank by finding the maximum number of linearly independent column vectors. To find a basis for the range of the matrix we employ Gaussian elimination to the matrix transpose to find

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

Hence, $\mathbf{z}_1 = \text{col}\,(1, 0, 3, -1)$ and $\mathbf{z}_2 = \text{col}\,(0, 1, -1, 2)$ are a basis for the range and so we conclude that again, the dimension of the range $= 2 = r$ is the rank of the matrix.

2. (a) Part (a) of problem 1 showed that $r = 2$. Thus the kernel of the matrix is represented as the set of linear combinations of just $n - r = 3 - 2 = 1$ vector $\mathbf{u}$. Hence, the nullity $h = 1$ in this case. To determine the basis for the kernel we use the result from part (a) and consider the corresponding homogeneous simultaneous linear equations

$$\begin{aligned} 3x_1 + \quad x_2 + 2x_3 &= 0 \\ 13x_2 - 4x_3 &= 0 \\ x_3 &= k \end{aligned}$$

where $-\infty \leq k \leq \infty$ is an arbitrary scalar. From this follows that $x_1 = -(10/13)k$, $x_2 = (4/13)k$ and $x_3 = k$, and so the basis for the kernel is given by

$$\mathbf{x} = k\text{col}\,(10, -4, -13) = k\mathbf{u}$$

(b) Employing Gaussian elimination as for part (b) of problem 1 gives

$$x_1 = 1 \qquad x_2 = 1 \qquad x_3 = 1$$

and so $r = n = 3$. Hence, the nullity $h = n - r = 0$.

(c) Employing Gaussian elimination as for part (c) of problem 1 results in the corresponding homogeneous simultaneous linear equations

$$\begin{aligned} x_1 \qquad\qquad + 5x_4 &= 0 \\ x_2 + 3x_3 + 4x_4 &= 0 \\ x_3 - 6x_4 &= 0 \end{aligned}$$

Hence, there are more unknowns than equations and the solutions are given by linear expressions for $r = 3$ of the unknowns in terms of the single remaining unknown $x_4 = k$:

$$x_1 = -5k \qquad x_2 = -22k \qquad x_3 = 6k$$

From this it follows that the nullity $h = n - r = 1$ and the basis for the kernel is given by

$$\mathbf{x} = k\mathrm{col}\,(-5, -22, 6, 1) = k\mathbf{u}$$

(d) Employing Gaussian elimination as for part (d) of problem 1 results in the corresponding homogeneous simultaneous linear equations

$$\begin{aligned} x_1 \quad + \quad x_3 + x_4 + 3x_5 &= 0 \\ x_2 - 2x_3 - x_4 - 4x_5 &= 0 \end{aligned}$$

Next, we can express $r = 2$ of the unknowns in terms of the remaining $h = n - r = 3$ unknowns:

$$\begin{aligned} x_1 &= -x_3 - x_4 - 3x_5 \\ x_2 &= 2x_3 + x_4 + 4x_5 \end{aligned}$$

Choosing $x_3 = k_1$, $x_4 = k_2$ and $x_5 = k_3$, where k_1, k_2 and k_3 are arbitrary scalars, the basis for the kernel is then given by

$$\begin{aligned} \mathbf{x} &= k_1\mathrm{col}\,(-1, 2, 1, 0, 0) + k_2\mathrm{col}\,(-1, 1, 0, 1, 0) + k_3\mathrm{col}\,(-3, 4, 0, 0, 1) \\ &= k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + k_3\mathbf{u}_3 \end{aligned}$$

3. (a) Forming a matrix from the four vectors and employing Gaussian elimination gives

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & -1 & 2 \\ 1 & 10 & 4 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 2 \\ 0 & -7 & -2 \\ 0 & 7 & 2 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 7 & 2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, $\mathbf{u}_1 = (1, 3, 2)$, $\mathbf{u}_2 = (0, 1, 1)$ and $\mathbf{u}_3 = (0, 0, -5)$ form a basis and W has dimension $3 = r =$ rank of the matrix.

(b) Forming a matrix out of the three vectors and employing Gaussian elimination gives

$$\begin{bmatrix} 3 & 6 & 2 \\ 1 & 3 & 1 \\ 5 & 1 & 7 \end{bmatrix} \to \begin{bmatrix} 3 & 6 & 2 \\ 0 & -3 & -1 \\ 0 & -14 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}$$

and so the only solution is the trivial solution $x_1 = x_2 = x_3 = 0$. Hence, the three vectors $\mathbf{u}_1 = (3, 6, 2)$, $\mathbf{u}_2 = (1, 3, 1)$ and $\mathbf{u}_3 = (5, 1, 7)$ are linear independent and form a basis for W, where W has dimension $3 = r =$ rank of the matrix.

(c) Forming a matrix out of the four vectors and employing Gaussian elimination gives

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

Hence, $\mathbf{u}_1 = (1, 2, 2, 3)$ and $\mathbf{u}_2 = (0, 1, 1, 2)$ form a basis and W has dimension $2 = r =$ rank of the matrix.

(d) Forming a matrix out of the four vectors and employing Gaussian elimination gives

$$\begin{bmatrix} 3 & 1 & 5 & 0 & 1 \\ 2 & 0 & 4 & 2 & 0 \\ 1 & -1 & 3 & 4 & -1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 3 & 1 & 5 & 0 & 1 \\ 2 & 0 & 4 & 2 & 0 \\ 1 & -1 & 3 & 4 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 5 & 0 & 1 \\ 0 & -2 & 4 & 2 & 0 \\ 0 & -2 & 3 & 4 & -1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 4 & 2 & 0 \\ 0 & -2 & 5 & 0 & 1 \\ 0 & -2 & 3 & 4 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & -1 & 0 \\ 0 & -2 & 5 & 0 & 1 \\ 0 & -2 & 3 & 4 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & -1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence, $\mathbf{u}_1 = (1, 1, 0, 0, 0)$, $\mathbf{u}_2 = (0, 1, -2, -1, 0)$ and $\mathbf{u}_3 = (0, 0, 1, -2, 1)$ form a basis and W has dimension $3 = r =$ rank of the matrix.

4. By definition, the dimension k denotes the number of linearly independent vectors in W. Also by definition, a basis of W is a set of linearly independent vectors that when linearly combined allows one to express every vector $\mathbf{w}$ of W in unique fashion. Hence, since there exist no more than k linearly independent vectors in W it follows immediately that the basis cannot contain more than k vectors. Next, let us suppose that there exists a basis for W that has $l < k$ linearly independent vectors. This would imply that every vector $\mathbf{w}$ of W can be expressed as a linear combination of the set $\mathbf{w}_1, \ldots, \mathbf{w}_l$. However, since the dimension of W is k, this would imply that the remaining $k - l$ linearly independent vectors $\mathbf{w}_{l+1}, \ldots, \mathbf{w}_k$ of W can also be expressed as a linear combination of $\mathbf{w}_1, \ldots, \mathbf{w}_l$, i.e.

$$\mathbf{w}_i = w_1\mathbf{w}_1 + \ldots + w_l\mathbf{w}_l \qquad \text{for} \quad l + 1 \le i \le k$$

This contradicts the very fact that $\mathbf{w}_{l+1}, \ldots, \mathbf{w}_k$ are linearly independent however and so we conclude that the basis $\mathbf{w}_1, \ldots, \mathbf{w}_l$ can and even must be extended to the basis $\mathbf{w}_1, \ldots, \mathbf{w}_k$ in order to exhaust/span W. This concludes the prove that every basis of the subspace W of dimension k contains exactly k (linearly independent) vectors.

5. To show that $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ are linearly independent it should hold that

$$w_{r+1}\mathbf{u}_{r+1} + \cdots + w_n\mathbf{u}_n = \mathbf{0}$$

only if the arbitrary constant $w_{r+1} = \cdots = w_n = 0$, which indeed is the case. This is easily seen by writing out $\mathbf{u}_r, \ldots, \mathbf{u}_n$ explicitly:

$$w_{r+1}\begin{bmatrix} d_{1,r+1} \\ \vdots \\ d_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + w_n \begin{bmatrix} d_{1,n} \\ \vdots \\ d_{r,n} \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \mathbf{0}$$

Since there will be at least one entree in each column which is equal to one, the only way for the $\mathbf{0}$ vector to be written as a linear combination of $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ is when all arbitrary constants $w_{r+1} = \cdots = w_n = 0$. Hence, this concludes the prove.

6. Let a $k \times n$ matrix $\mathbf{A}$ be given and W be a subspace, consisting of all linear combinations of the rows $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of $\mathbf{A}$. What remains to be proven is that the rank r of $\mathbf{A}$ is the largest integer such that some $r \times r$ minor of $\mathbf{A}$ is nonzero. As a first step, we use Gaussian elimination on $\mathbf{A}$ to reduce it to row echelon form. Applying the two steps (I: interchanging rows and II: adding multiples of one row to the other rows) leaves the row space W of $\mathbf{A}$ unchanged. Step I needs no further prove. Regarding step II, we replace $\mathbf{v}_1, \ldots, \mathbf{v}_k$ by the linear combinations $\mathbf{v}_1 + c_1\mathbf{v}_h, \ldots, \mathbf{v}_k + c_k\mathbf{v}_h$ where $c_h = 0$ (here we assume a multiple of the row vector $\mathbf{v}_h$ is added to all the other row vectors). Every linear combination of these new vectors is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and hence, will also be in W. Repeatedly applying these two operations will eventually transform $\mathbf{A}$ to row echelon form with nonzero, linearly independent row vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$, where $r \leq k$. If $r < k$, then the last $p = k - r$ rows of $\mathbf{A}$ in row echelon form will be $\mathbf{0}$. Since the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ form a basis for W and Gaussian elimination has left the row space of $\mathbf{A}$ unchanged (the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ can be expressed as a linear combination of the original vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$), the rank r of $\mathbf{A}$ (the same r as for $\mathbf{u}_r$) has not changed either. Now that we have established that Gaussian elimination has left the rank of $\mathbf{A}$ unchanged, let us first consider the

case $r < k$ so that the matrix $\mathbf{A}$ in row echelon form looks like

$$\bar{\mathbf{A}} = \begin{bmatrix} 1 & \bar{a}_{1,2} & \bar{a}_{1,3} & \dots & \bar{a}_{1,r+1} & \dots & \bar{a}_{1,n} \\ 0 & 1 & \bar{a}_{2,3} & \dots & \bar{a}_{2,r+1} & \dots & \bar{a}_{2,n} \\ \vdots & & \ddots & & \vdots & & \vdots \\ 0 & \dots & 0 & 1 & \bar{a}_{r,r+1} & \dots & \bar{a}_{r,n} \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

As should be apparent immediately from looking at $\mathbf{A}$ in row echelon form, any $q \times q$ (where $r < q \leq k$ and $k \leq n$) determinant formed from a $q \times q$ sub matrix $\bar{\mathbf{A}}^{(q,q)}$ of $\bar{\mathbf{A}}$ will give a 0 result, since it has at least one row i consisting entirely of entrees that are 0. Expanding the determinant about row i gives

$$\det\left(\bar{\mathbf{A}}^{(q,q)}\right) = \sum_{j=1}^{q} (-1)^{i+j}\, \bar{a}_{i,j} \det\left(\bar{\mathbf{A}}^{(q,q)}\,(i|j)\right) = \sum_{j=1}^{q} (-1)^{i+j}\, 0 \det\left(\bar{\mathbf{A}}^{(q,q)}\,(i|j)\right) = 0$$

Here $\det\left(\bar{\mathbf{A}}^{(q,q)}\,(i|j)\right)$ denotes the $(q-1) \times (q-1)$ minor obtained from $\bar{\mathbf{A}}^{(q,q)}$ by striking out the i'th row and j'th column. In the case $r = k$ there will exist at least one $r \times r$ minor of $\bar{\mathbf{A}}$ that is nonzero. Again, forming an $r \times r$ sub matrix $\bar{\mathbf{A}}^{(r,r)}$ of $\bar{\mathbf{A}}$ and expanding about row i where $i \leq r$ gives

$$\det\left(\bar{\mathbf{A}}^{(r,r)}\right) = \sum_{j=1}^{r} (-1)^{i+j}\, \bar{a}_{i,j} \det\left(\bar{\mathbf{A}}^{(r,r)}\,(i|j)\right) \neq 0$$

Regardless of what sub matrix $\bar{\mathbf{A}}^{(r,r)}$ will be chosen, there will exist at least one $\bar{a}_{i,j}$ that is nonzero. In addition, since the r row vectors of the row echelon form of $\mathbf{A}$ are linear independent, so will any combination of r of its column vectors, and hence, the $(r-1) \times (r-1)$ minor $\det\left(\bar{\mathbf{A}}^{(r,r)}\,(i|j)\right)$ will always be nonzero. Hence, this concludes the prove that the evaluation of rank by nonzero minors is correct.

7. For problem 6 it was proved that the rank r of a matrix $\mathbf{A}$ is the largest integer r such that some $r \times r$ minor of $\mathbf{A}$ is nonzero. As for the previous problem, let us denote this minor by $\det(\mathbf{A}^{(r,r)})$. According to Rule I of section 1.4 $\det(\mathbf{A}) = \det(\mathbf{A}^{\top})$. Applying this rule to our minor then gives

$$\det\left(\mathbf{A}^{(r,r)}\right) = \det\left[\left(\mathbf{A}^{(r,r)}\right)^{\top}\right]$$

which proves at once that the row rank of $\mathbf{A}$ equals its column rank.

8. Let $a\mathbf{x}_i$ and $b\mathbf{x}_j$ be two eigenvectors with the same eigenvalue λ. We thus have $a\mathbf{A}\mathbf{x}_i = a\lambda\mathbf{x}_i$ and $b\mathbf{A}\mathbf{x}_j = b\lambda\mathbf{x}_j$. Adding both equations gives

$$a\mathbf{A}\mathbf{x}_i + b\mathbf{A}\mathbf{x}_j = a\lambda\mathbf{x}_i + b\lambda\mathbf{x}_j \qquad \Longrightarrow \qquad \mathbf{A}\left(a\mathbf{x}_i + b\mathbf{x}_j\right) = \lambda\left(a\mathbf{x}_i + b\mathbf{x}_j\right)$$

Hence, the linear combination $a\mathbf{x}_i + b\mathbf{x}_j$ is also a valid eigenvector for λ and hence, will also be in W. In conclusion, the set of vectors $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$ together with the $\mathbf{0}$ vector is indeed a subspace W of V^n.

9. Applying Gaussian elimination to the augmented matrix $\begin{bmatrix}\mathbf{A} & \mathbf{y}\end{bmatrix}$ associated with the equation $\mathbf{Ax} = \mathbf{y}$ can lead to three different cases. First, consider the case where Gaussian elimination will produce a row echelon matrix where the $k\times(k+1)$ augmented matrix $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ and the $k \times k$ coefficient matrix $\bar{\mathbf{A}}$ have the same rank k:

$$\left[\begin{array}{ccccc|c}
1 & \bar{a}_{12} & \dots & \dots & \bar{a}_{1,k} & \bar{y}_1 \\
0 & \ddots & & & \vdots & \vdots \\
\vdots & & \ddots & & \vdots & \vdots \\
0 & \dots & 0 & 1 & \bar{a}_{k-1,k} & \bar{y}_{k-1} \\
0 & \dots & \dots & 0 & 1 & \bar{y}_k
\end{array}\right]$$

Here, both $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ and $\bar{\mathbf{A}}$ have rank k and hence, the equation $\mathbf{Ax} = \mathbf{y}$ has a unique solution. Next, consider the case where the rank of the $m \times (n+1)$ augmented matrix $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ is bigger than the rank k of the $m \times n$ coefficient matrix $\bar{\mathbf{A}}$:

$$\left[\begin{array}{ccccccc|c}
1 & \bar{a}_{12} & \dots & \dots & \bar{a}_{1,n} & \dots & \bar{a}_{1,n} & \bar{y}_1 \\
0 & \ddots & & & \vdots & & \vdots & \vdots \\
\vdots & & \ddots & & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 1 & \bar{a}_{k,k+1} & \dots & \bar{a}_{k,n} & \bar{y}_k \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \bar{y}_{k+1} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & \bar{y}_m
\end{array}\right]$$

Here, the rank m of $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ is larger than the rank k of $\bar{\mathbf{A}}$ and hence, the equation $\mathbf{Ax} = \mathbf{y}$ has no solution (i.e. is inconsistent). Lastly, consider the case where again, the rank of the $m \times (n+1)$ augmented matrix $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ is equal to the rank of the $m \times n$ coefficient matrix $\bar{\mathbf{A}}$, but $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ and $\bar{\mathbf{A}}$ contain one or more rows with zeros:

$$\left[\begin{array}{ccccccc|c}
1 & \bar{a}_{12} & \dots & \dots & \bar{a}_{1,n} & \dots & \bar{a}_{1,n} & \bar{y}_1 \\
0 & \ddots & & & \vdots & & \vdots & \vdots \\
\vdots & & \ddots & & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 1 & \bar{a}_{k,k+1} & \dots & \bar{a}_{k,n} & \bar{y}_k \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0
\end{array}\right]$$

Here, the rank k of $\begin{bmatrix}\bar{\mathbf{A}} & \bar{\mathbf{y}}\end{bmatrix}$ is equal to the rank $\bar{\mathbf{A}}$ and the equation $\mathbf{Ax} = \mathbf{y}$ has infinitely many solutions.

Section 1.18

1. (a) An arbitrary polynomial of degree at most 2 is of the general form

$$P_2(x) = a_0 + a_1 x + a_2 x^2$$

The 0 polynomial exists. Now let $p(x)$ and $q(x)$ be two arbitrary (but different) polynomials of degree at most 2. It is then simple to see that rules I through VII are all satisfied. For rule I we have

$$\begin{aligned} p(x) + q(x) &= \left(p_0 + p_1 x + p_2 x^2\right) + \left(q_0 + q_1 x + q_2 x^2\right) \\ &= (p_0 + q_0) + (p_1 + q_1)\, x + (p_2 + q_2)\, x^2 \\ &= (q_0 + p_0) + (q_1 + p_1)\, x + (q_2 + p_2)\, x^2 \\ &= \left(q_0 + q_1 x + q_2 x^2\right) + \left(p_0 + p_1 x + p_2 x^2\right) \\ &= q(x) + p(x) \end{aligned}$$

Let $p(x)$ and $q(x)$ be as before and let $r(x)$ be a third arbitrary polynomial of at most degree 2. For rule II we thus have

$$\begin{aligned} (p(x) + q(x)) + r(x) &= \left[\left(p_0 + p_1 x + p_2 x^2\right) + \left(q_0 + q_1 x + q_2 x^2\right)\right] + \left(r_0 + r_1 x + r_2 x^2\right) \\ &= (p_0 + q_0 + r_0) + (p_1 + q_1 + r_1)\, x + (p_2 + q_2 + r_2)\, x^2 \\ &= [p_0 + (q_0 + r_0)] + [p_1 + (q_1 + r_1)]\, x + [p_2 + (q_2 + r_2)]\, x^2 \\ &= \left(p_0 + p_1 x + p_2 x^2\right) + \left[\left(q_0 + q_1 x + q_2 x^2\right) + \left(r_0 + r_1 x + r_2 x^2\right)\right] \\ &= p(x) + (q(x) + r(x)) \end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned} h(p(x) + q(x)) &= h\left[\left(p_0 + p_1 x + p_2 x^2\right) + \left(q_0 + q_1 x + q_2 x^2\right)\right] \\ &= h\left(p_0 + p_1 x + p_2 x^2\right) + h\left(q_0 + q_1 x + q_2 x^2\right) \\ &= hp(x) + hq(x) \end{aligned}$$

Let a and b be two arbitrary scalars, then for rule IV we have

$$\begin{aligned} (a + b)\, p(x) &= (a + b)\left(p_0 + p_1 x + p_2 x^2\right) \\ &= a\left(p_0 + p_1 x + p_2 x^2\right) + b\left(p_0 + p_1 x + p_2 x^2\right) \\ &= ap(x) + bp(x) \end{aligned}$$

Again, let a and b be two arbitrary scalars. For rule V we then have

$$\begin{aligned} (ab)\, p(x) &= (ab)\left(p_0 + p_1 x + p_2 x^2\right) \\ &= a\left(bp_0 + bp_1 x + bp_2 x^2\right) \\ &= a\left[b\left(p_0 + p_1 x + p_2 x^2\right)\right] \\ &= a(bp(x)) \end{aligned}$$

For rule VI we have

$$1p\left(x\right)=1\left(p_0+p_1x+p_2x^2\right)=1p_0+1p_1x+1p_2x^2=p_0+p_1x+p_2x^2=p\left(x\right)$$

And lastly, for rule VII we have

$$0p\left(x\right)=0\left(p_0+p_1x+p_2x^2\right)=0p_0+0p_1x+0p_2x^2=0+0+0=0$$

Hence, the polynomial $P_2(x)$ forms a vector space V. Consider the equation

$$p_0 1+p_1x+p_2x^2=0$$

That is, let the polynomial $p_0+p_1x+p_2x^2$ coincide with the 0 polynomial and hence, have the value 0 for any x. Now since a polynomial of degree 2 has at most 2 roots, the only way to guarantee the polynomial is 0 for any x is to have $p_0=p_1=p_2=0$. Therefore, 1, x and x^2 are linearly independent and form a basis for V and V is a 3-dimensional subspace of the vector space of all polynomials.

(b) A generic formulation for all polynomials containing no terms of odd degree is given by

$$P\left(x\right)=p_0+p_2x^2+\cdots+p_{2n}x^{2n}\qquad\text{where } n=2,3,\ldots$$

The 0 polynomial exists. Now let $p(x)$ and $q(x)$ be two arbitrary polynomials containing no terms of odd degree. For rule I we have

$$\begin{aligned}
p\left(x\right)+q\left(x\right)&=\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)+\left(q_0+q_2x^2+\cdots+q_{2n}x^{2n}\right)\\
&=\left(p_0+q_0\right)+\left(p_2+q_2\right)x^2+\cdots+\left(p_{2n}+q_{2n}\right)x^{2n}\\
&=\left(q_0+p_0\right)+\left(q_2+p_2\right)x^2+\cdots+\left(q_{2n}+p_{2n}\right)x^{2n}\\
&=\left(q_0+q_2x^2+\cdots+q_{2n}x^{2n}\right)+\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)\\
&=q\left(x\right)+p\left(x\right)
\end{aligned}$$

Let $p(x)$ and $q(x)$ be as before and let $r(x)$ be a third arbitrary polynomial containing no terms of odd degree. For rule II we have

$$\begin{aligned}
\left(p\left(x\right)+q\left(x\right)\right)+r\left(x\right)&=\left[\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)+\left(q_0+q_2x^2+\cdots+q_{2n}x^{2n}\right)\right]\\
&\quad+\left(r_0+r_2x^2+\cdots+r_{2n}x^{2n}\right)\\
&=\left(p_0+q_0+r_0\right)+\left(p_2+q_2+r_2\right)x^2+\cdots+\left(p_{2n}+q_{2n}+r_{2n}\right)x^{2n}\\
&=\left[p_0+\left(q_0+r_0\right)\right]+\left[p_2+\left(q_2+r_2\right)\right]x^2+\cdots+\left[p_{2n}+\left(q_{2n}+r_{2n}\right)\right]x^{2n}\\
&=\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)\\
&\quad+\left[\left(q_0+q_2x^2+\cdots+q_{2n}x^{2n}\right)+\left(r_0+r_2x^2+\cdots+r_{2n}x^{2n}\right)\right]\\
&=p\left(x\right)+\left(q\left(x\right)+r\left(x\right)\right)
\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned} h\left(p\left(x\right)+q\left(x\right)\right) &= h\left[\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)+\left(q_0+q_2x^2+\cdots+q_{2n}^{2n}\right)\right] \\ &= h\left(p_0+p_2x^2+\cdots+p_{2n}^{2n}\right)+h\left(q_0+q_2x^2+\cdots+q_{2n}^{2n}\right) \\ &= hp\left(x\right)+hq\left(x\right) \end{aligned}$$

Let a and b be two arbitrary scalars, then for rule IV we have

$$\begin{aligned} \left(a+b\right)p\left(x\right) &= \left(a+b\right)\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right) \\ &= a\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)+b\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right) \\ &= ap\left(x\right)+bp\left(x\right) \end{aligned}$$

Again, let a and b be two arbitrary scalars. For rule V we then have

$$\begin{aligned} \left(ab\right)p\left(x\right) &= \left(ab\right)\left(p_0+p_2x^2+\cdots+p_{2n}^{2n}\right) \\ &= a\left(bp_0+bp_2x^2+\cdots+bp_{2n}x^{2n}\right) \\ &= a\left[b\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right)\right] \\ &= a\left(bp\left(x\right)\right) \end{aligned}$$

For rule VI we have

$$\begin{aligned} 1p\left(x\right)=1\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right) &= 1p_0+1p_2x^2+\cdots+1p_{2n}x^{2n} \\ &= p_0+p_2x^2+\cdots+p_{2n}x^{2n} \\ &= p\left(x\right) \end{aligned}$$

And lastly, for rule VII we have

$$\begin{aligned} 0p\left(x\right)=0\left(p_0+p_2x^2+\cdots+p_{2n}x^{2n}\right) &= 0p_0+0p_2x^2+\cdots+0p_{2n}x^{2n} \\ &= 0+0+0 \\ &= 0 \end{aligned}$$

Hence, the polynomial $P(x)$ containing no terms of odd degree forms a vector space V. Consider the equation

$$p_0 1+p_2x^2+\cdots+p_{2n}x^{2n}=0$$

That is, let the polynomial $p_0 1+p_2x^2+\cdots+p_{2n}x^{2n}$ coincide with the 0 polynomial and hence, have the value 0 for any x. Now since a polynomial of degree $2n$ where $0\le n\le k$ has at most $2k$ roots, the only way to guarantee the polynomial is 0 for any x is to have $p_0=p_2=\cdots=p_{2n}=0$. Therefore, 1, x^2, ..., x^{2n} are linearly independent and form a basis for V and V is an infinite-dimensional subspace of the vector space of all polynomials.

(c) A generic formulation for all trigonometric polynomials is given by

$$P\left(x\right)=a_0+a_1\cos x+b_1\sin x+\cdots+a_n\cos nx+b_n\sin nx$$

The 0 polynomial exists. Now let $p(x)$ and $q(x)$ be two arbitrary trigonometric polynomials. For rule I we have

$$\begin{aligned}
p\left(x\right)+q\left(x\right)&=\left(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\right)\\
&\quad+\left(a_0^{(q)}+a_1^{(q)}\cos x+b_1^{(q)}\sin x+\cdots+a_n^{(q)}\cos nx+b_n^{(q)}\sin nx\right)\\
&=\left(a_0^{(p)}+a_0^{(q)}\right)+\left(a_1^{(p)}+a_1^{(q)}\right)\cos x+\left(b_1^{(p)}+b_1^{(q)}\right)\sin x+\dots\\
&\quad+\left(a_n^{(p)}+a_n^{(q)}\right)\cos nx+\left(b_n^{(p)}+b_n^{(q)}\right)\sin nx\\
&=\left(a_0^{(q)}+a_0^{(p)}\right)+\left(a_1^{(q)}+a_1^{(p)}\right)\cos x+\left(b_1^{(q)}+b_1^{(p)}\right)\sin x+\dots\\
&\quad+\left(a_n^{(q)}+a_n^{(p)}\right)\cos nx+\left(b_n^{(q)}+b_n^{(p)}\right)\sin nx\\
&=\left(a_0^{(q)}+a_1^{(q)}\cos x+b_1^{(q)}\sin x+\cdots+a_n^{(q)}\cos nx+b_n^{(q)}\sin nx\right)\\
&\quad+\left(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\right)\\
&=q\left(x\right)+p\left(x\right)
\end{aligned}$$

Let $p(x)$ and $q(x)$ be as before and let $r(x)$ be a third arbitrary trigonometric polynomial. For rule II we have

$$
\begin{aligned}
(p(x)+q(x))+r(x) &= \Big[\Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&\quad + \Big(a_0^{(q)}+a_1^{(q)}\cos x+b_1^{(q)}\sin x+\cdots+a_n^{(q)}\cos nx+b_n^{(q)}\sin nx\Big)\Big] \\
&\quad + \Big(a_0^{(r)}+a_1^{(r)}\cos x+b_1^{(r)}\sin x+\cdots+a_n^{(r)}\cos nx+b_n^{(r)}\sin nx\Big) \\
&= \Big(a_0^{(p)}+a_0^{(q)}+a_0^{(r)}\Big) \\
&\quad + \Big[\Big(a_1^{(p)}+a_1^{(q)}+a_1^{(r)}\Big)\cos x+\Big(b_1^{(p)}+b_1^{(q)}+b_1^{(r)}\Big)\sin x\Big]+\dots \\
&\quad + \big[\big(a_n^{(p)}+a_n^{(q)}+a_n^{(r)}\big)\cos nx+\big(b_n^{(p)}+b_n^{(q)}+b_n^{(r)}\big)\sin nx\big] \\
&= \Big[a_0^{(p)}+\Big(a_0^{(q)}+a_0^{(r)}\Big)\Big] \\
&\quad + \Big[\Big(a_1^{(p)}+\Big(a_1^{(q)}+a_1^{(r)}\Big)\Big)\cos x+\Big(b_1^{(p)}+\Big(b_1^{(q)}+b_1^{(r)}\Big)\Big)\sin x\Big]+\dots \\
&\quad + \big[\big(a_n^{(p)}+\big(a_n^{(q)}+a_n^{(r)}\big)\big)\cos nx+\big(b_n^{(p)}+\big(b_n^{(q)}+b_n^{(r)}\big)\big)\sin nx\big] \\
&= \Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&\quad + \Big[\Big(a_0^{(q)}+a_1^{(q)}\cos x+b_1^{(q)}\sin x+\cdots+a_n^{(q)}\cos nx+b_n^{(q)}\sin nx\Big) \\
&\quad + \Big(a_0^{(r)}+a_1^{(r)}\cos x+b_1^{(r)}\sin x+\cdots+a_n^{(r)}\cos nx+b_n^{(r)}\sin nx\Big)\Big] \\
&= p(x)+(q(x)+r(x))
\end{aligned}
$$

Let h be an arbitrary scalar, then for rule III we have

$$
\begin{aligned}
h(p(x)+q(x)) &= h\Big[\Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&\quad + \Big(a_0^{(q)}+a_1^{(q)}\cos x+b_1^{(q)}\sin x+\cdots+a_n^{(q)}\cos nx+b_n^{(q)}\sin nx\Big)\Big] \\
&= h\Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&\quad + h\Big(a_0^{(q)}+a_1^{(q)}\cos x+b_1^{(q)}\sin x+\cdots+a_n^{(q)}\cos nx+b_n^{(q)}\sin nx\Big) \\
&= hp(x)+hq(x)
\end{aligned}
$$

Let α and β be two arbitrary scalars, then for rule IV we have

$$
\begin{aligned}
(\alpha+\beta)\,p(x) &= (\alpha+\beta)\Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&= \alpha\Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&\quad + \beta\Big(a_0^{(p)}+a_1^{(p)}\cos x+b_1^{(p)}\sin x+\cdots+a_n^{(p)}\cos nx+b_n^{(p)}\sin nx\Big) \\
&= \alpha p(x)+\beta p(x)
\end{aligned}
$$

Again, let α and β be two arbitrary scalars. For rule V we then have

$$\begin{aligned}
(\alpha\beta)\, p\,(x) &= (\alpha\beta) \left(a_0^{(p)} + a_1^{(p)} \cos x + b_1^{(p)} \sin x + \cdots + a_n^{(p)} \cos nx + b_n^{(p)} \sin nx \right) \\
&= \alpha \left(\beta a_0^{(p)} + \beta a_1^{(p)} \cos x + \beta b_1^{(p)} \sin x + \cdots + \beta a_n^{(p)} \cos nx + \beta b_n^{(p)} \sin nx \right) \\
&= \alpha \left[\beta \left(a_0^{(p)} + a_1^{(p)} \cos x + b_1^{(p)} \sin x + \cdots + a_n^{(p)} \cos nx + b_n^{(p)} \sin nx \right) \right] \\
&= \alpha \left(\beta p\,(x) \right)
\end{aligned}$$

For rule VI we have

$$\begin{aligned}
1p\,(x) &= 1 \left(a_0^{(p)} + a_1^{(p)} \cos x + b_1^{(p)} \sin x + \cdots + a_n^{(p)} \cos nx + b_n^{(p)} \sin nx \right) \\
&= 1a_0^{(p)} + 1a_1^{(p)} \cos x + 1b_1^{(p)} \sin x + \cdots + 1a_n^{(p)} \cos nx + 1b_n^{(p)} \sin nx \\
&= a_0^{(p)} + a_1^{(p)} \cos x + b_1^{(p)} \sin x + \cdots + a_n^{(p)} \cos nx + b_n^{(p)} \sin nx \\
&= p\,(x)
\end{aligned}$$

And lastly, for rule VII we have

$$\begin{aligned}
0p\,(x) &= 0 \left(a_0^{(p)} + a_1^{(p)} \cos x + b_1^{(p)} \sin x + \cdots + a_n^{(p)} \cos nx + b_n^{(p)} \sin nx \right) \\
&= 0a_0^{(p)} + 0a_1^{(p)} \cos x + 0b_1^{(p)} \sin x + \cdots + 0a_n^{(p)} \cos nx + 0b_n^{(p)} \sin nx \\
&= 0 + 0 + 0 + \cdots + 0 + 0 \\
&= 0
\end{aligned}$$

Hence, all trigonometric polynomials $P(x)$ form a vector space V. Consider the equation

$$a_0 1 + a_1 \cos x + b_1 \sin x + \cdots + a_n \cos nx + b_n \sin nx = 0$$

That is, let the trigonometric polynomial above coincide with the 0 polynomial and hence, have the value 0 for any x. Now since $\cos nx = 0$, but $\sin nx = \pm 1$ for $\pi(n + 1/2)$ and $\cos nx = \pm 1$, but $\sin nx = 0$ for πn, where $-\infty < n < \infty$, the only way to guarantee the polynomial is 0 for any x is to have $a_0 = a_1 = b_1 = \cdots = a_n = b_n = 0$. Therefore, 1, $\cos x$, $\sin x$, ..., $\cos nx$, $\sin nx$ are linearly independent and form a basis for V and V is an n-dimensional vector space.

(d) Next, we consider all functions of the form

$$F\,(x) = ae^{x} + be^{-x}$$

Choosing $a = -1$ and $b = 1$, we can verify that $F(0) = 0$. Now let $f(x)$ and $g(x)$

be two arbitrary functions of the form above. For rule I we have

$$\begin{aligned}
f(x)+g(x) &= \left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)+\left(a^{(g)}e^{x}+b^{(g)}e^{-x}\right)\\
&= \left(a^{(f)}+a^{(g)}\right)e^{x}+\left(b^{(f)}+b^{(g)}\right)e^{-x}\\
&= \left(a^{(g)}+a^{(f)}\right)e^{x}+\left(b^{(g)}+b^{(f)}\right)e^{-x}\\
&= \left(a^{(g)}e^{x}+b^{(g)}e^{-x}\right)+\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)\\
&= g(x)+f(x)
\end{aligned}$$

Let $f(x)$ and $g(x)$ be as before and let $h(x)$ be a third arbitrary function of the given form. For rule II we have

$$\begin{aligned}
(f(x)+g(x))+h(x) &= \left[\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)+\left(a^{(g)}e^{x}+b^{(g)}e^{-x}\right)\right]+\left(a^{(h)}e^{x}+b^{(h)}e^{-x}\right)\\
&= \left(a^{(f)}+a^{(g)}+a^{(h)}\right)e^{x}+\left(b^{(f)}+b^{(g)}+b^{(h)}\right)e^{-x}\\
&= \left[a^{(f)}+\left(a^{(g)}+a^{(h)}\right)\right]e^{x}+\left[b^{(f)}+\left(b^{(g)}+b^{(h)}\right)\right]e^{-x}\\
&= \left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)+\left[\left(a^{(g)}e^{x}+b^{(g)}e^{-x}\right)+\left(a^{(h)}e^{x}+b^{(h)}e^{-x}\right)\right]\\
&= f(x)+(g(x)+h(x))
\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned}
h(f(x)+g(x)) &= h\left[\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)+\left(a^{(g)}e^{x}+b^{(g)}e^{-x}\right)\right]\\
&= h\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)+h\left(a^{(g)}e^{x}+b^{(g)}e^{-x}\right)\\
&= hf(x)+hg(x)
\end{aligned}$$

Let α and β be two arbitrary scalars, then for rule IV we have

$$\begin{aligned}
(\alpha+\beta)f(x) &= (\alpha+\beta)\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)\\
&= \alpha\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)+\beta\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)\\
&= \alpha f(x)+\beta f(x)
\end{aligned}$$

Again, let α and β be two arbitrary scalars. For rule V we then have

$$\begin{aligned}
(\alpha\beta)f(x) &= (\alpha\beta)\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)\\
&= \alpha\left(\beta a^{(f)}e^{x}+\beta b^{(f)}e^{-x}\right)\\
&= \alpha\left[\beta\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)\right]\\
&= \alpha(\beta f(x))
\end{aligned}$$

For rule VI we have

$$1f(x)=1\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)=1a^{(f)}e^{x}+1b^{(f)}e^{-x}=a^{(f)}e^{x}+b^{(f)}e^{-x}=f(x)$$

And lastly, for rule VII we have

$$0f(x)=0\left(a^{(f)}e^{x}+b^{(f)}e^{-x}\right)=0a^{(f)}e^{x}+0b^{(f)}e^{-x}=0+0=0$$

Hence, all functions of the form $f(x) = ae^x + be^{-x}$ form a vector space V. Consider the equation

$$ae^x + be^{-x} = 0$$

That is, let the function $f(x)$ coincide with the 0 function and hence, have the value 0 for any x. Let us rewrite the equation above as

$$e^x = \tilde{a}e^{-x}$$

where $\tilde{a} = -b/a \neq 0$ and $-\infty < x < \infty$. Now consider the special cases $x = 0$ and $x = 1$ which give the two equations

$$\begin{aligned} x = 0 \quad &\Longrightarrow \quad a + b = 0 \\ x = 1 \quad &\Longrightarrow \quad ae + be^{-1} = 0 \end{aligned}$$

Subtracting e times the first equation from the second gives $b(e^{-1} - e) = 0$. Hence, $b = 0$. Substituting $b = 0$ in the first (or second) equation then also gives $a = 0$. Therefore, e^x and e^{-x} are linearly independent and form a basis for V and V is a 2-dimensional vector space.

(e) A generic formulation for all 3×3 diagonal matrices is given by

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

By choosing $a_{11} = a_{22} = a_{33} = 0$ we can verify the $\mathbf{0}$ exists. Let $\mathbf{A}$ and $\mathbf{B}$ be two arbitrary 3×3 diagonal matrices. For rule I we have

$$\begin{aligned} \mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} &= \begin{bmatrix} a_{11} + b_{11} & 0 & 0 \\ 0 & a_{22} + b_{22} & 0 \\ 0 & 0 & a_{33} + b_{33} \end{bmatrix} \\ &= \begin{bmatrix} b_{11} + a_{11} & 0 & 0 \\ 0 & b_{22} + a_{22} & 0 \\ 0 & 0 & b_{33} + a_{33} \end{bmatrix} \\ &= \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} + \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\ &= \mathbf{B} + \mathbf{A} \end{aligned}$$

Let $\mathbf{A}$ and $\mathbf{B}$ be as before and let $\mathbf{C}$ be a third arbitrary 3×3 diagonal matrix.

For rule II we have

$$
\begin{aligned}
(\mathbf{A}+\mathbf{B})+\mathbf{C} &= \left(\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}\right) + \begin{bmatrix} c_{11} & 0 & 0 \\ 0 & c_{22} & 0 \\ 0 & 0 & c_{33} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}+b_{11}+c_{11} & 0 & 0 \\ 0 & a_{22}+b_{22}+c_{22} & 0 \\ 0 & 0 & a_{33}+b_{33}+c_{33} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}+(b_{11}+c_{11}) & 0 & 0 \\ 0 & a_{22}+(b_{22}+c_{22}) & 0 \\ 0 & 0 & a_{33}+(b_{33}+c_{33}) \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} + \left(\begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} + \begin{bmatrix} c_{11} & 0 & 0 \\ 0 & c_{22} & 0 \\ 0 & 0 & c_{33} \end{bmatrix}\right) \\
&= \mathbf{A}+(\mathbf{B}+\mathbf{C})
\end{aligned}
$$

Let h be an arbitrary scalar, then for rule III we have

$$
\begin{aligned}
h\,(\mathbf{A}+\mathbf{B}) &= h\left(\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} + \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}\right) \\
&= h\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} + h\begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} \\
&= h\mathbf{A}+h\mathbf{B}
\end{aligned}
$$

Let α and β be two arbitrary scalars, then for rule IV we have

$$
\begin{aligned}
(\alpha+\beta)\,\mathbf{A} = (\alpha+\beta)\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} &= \alpha\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} + \beta\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\
&= \alpha\mathbf{A}+\beta\mathbf{A}
\end{aligned}
$$

Again, let α and β be two arbitrary scalars. For rule V we then have

$$
\begin{aligned}
(\alpha\beta)\,\mathbf{A} = (\alpha\beta)\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} &= \alpha\begin{bmatrix} \beta a_{11} & 0 & 0 \\ 0 & \beta a_{22} & 0 \\ 0 & 0 & \beta a_{33} \end{bmatrix} \\
&= \alpha\left(\beta\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}\right) \\
&= \alpha\,(\beta\mathbf{A})
\end{aligned}
$$

For rule VI we have

$$1\mathbf{A} = 1\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} = \begin{bmatrix} 1a_{11} & 0 & 0 \\ 0 & 1a_{22} & 0 \\ 0 & 0 & 1a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} = \mathbf{A}$$

And lastly, for rule VII we have

$$0\mathbf{A} = 0\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} = \begin{bmatrix} 0a_{11} & 0 & 0 \\ 0 & 0a_{22} & 0 \\ 0 & 0 & 0a_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0}$$

Hence, all 3×3 diagonal matrices form a vector space V. Consider the equation

$$\begin{aligned} \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} &= \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \\ &= a_{11}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + a_{22}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + a_{33}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \mathbf{0} \end{aligned}$$

The only way for the equation above to make sense is to choose $a_{11} = a_{22} = a_{33} = 0$. Therefore, the three matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

form a basis for V and V is a 3-dimensional vector space.

(f) A generic formulation for all 4×4 symmetric matrices is given by

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{bmatrix}$$

By choosing all 16 coefficients as 0 we can verify the $\mathbf{0}$ matrix exists. Proving the addition and multiplication rules is redundant as this would be the same as for the 3×3 diagonal matrix case. Hence, we will move on directly to determining a basis for the vector space V formed of all 4×4 symmetric matrices. Consider the equation

$$
\begin{aligned}
\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{bmatrix} &= \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & 0 & 0 \\ a_{12} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & a_{13} & 0 \\ 0 & 0 & 0 & 0 \\ a_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 & 0 & 0 & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{14} & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & a_{23} & 0 \\ 0 & a_{23} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & a_{24} & 0 & 0 \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & a_{34} & 0 \end{bmatrix} \\
&= a_{11} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + a_{33} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
&\quad + a_{44} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + a_{13} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
&\quad + a_{14} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} + a_{23} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + a_{24} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\
&\quad + a_{34} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\
&= \mathbf{0}
\end{aligned}
$$

The only way for the equation above to make sense is to choose $a_{11} = a_{22} = a_{33} =$

$a_{44} = a_{12} = a_{13} = a_{14} = a_{23} = a_{24} = a_{34} = 0$. Therefore, the ten matrices

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

form a basis for V and V is a 10-dimensional vector space.

(g) The general solution to the second order equation $y'' + y = 0$ is given by

$$y = F(x) = a\cos x + b\sin x \quad \text{for } -\infty < x < \infty$$

Choosing $a = b = 1$ and $x = -\pi/4$, we can verify that $F(-pi/4) = 0$. Now let $f(x)$ and $g(x)$ be two arbitrary function of the form above. For rule I we have

$$\begin{aligned}
f(x) + g(x) &= \left(a^{(f)}\cos x + b^{(f)}\sin x\right) + \left(a^{(g)}\cos x + b^{(g)}\sin x\right)\\
&= \left(a^{(f)} + a^{(g)}\right)\cos x + \left(b^{(f)} + b^{(g)}\right)\sin x\\
&= \left(a^{(g)} + a^{(f)}\right)\cos x + \left(b^{(g)} + b^{(f)}\right)\sin x\\
&= \left(a^{(g)}\cos x + b^{(g)}\sin x\right) + \left(a^{(f)}\cos x + b^{(f)}\sin x\right)\\
&= g(x) + f(x)
\end{aligned}$$

Let $f(x)$ and $g(x)$ be as before and let $h(x)$ be a third arbitrary function of the given form. For rule II we have

$$\begin{aligned}
(f(x) + g(x)) + h(x) &= \left[\left(a^{(f)}\cos x + b^{(f)}\sin x\right) + \left(a^{(g)}\cos x + b^{(g)}\sin x\right)\right]\\
&\quad + \left(a^{(h)}\cos x + b^{(h)}\sin x\right)\\
&= \left(a^{(f)} + a^{(g)} + a^{(h)}\right)\cos x + \left(b^{(f)} + b^{(g)} + b^{(h)}\right)\sin x\\
&= \left[a^{(f)} + \left(a^{(g)} + a^{(h)}\right)\right]\cos x + \left[b^{(f)} + \left(b^{(g)} + b^{(h)}\right)\right]\sin x\\
&= \left(a^{(f)}\cos x + b^{(f)}\sin x\right)\\
&\quad + \left[\left(a^{(g)}\cos x + b^{(g)}\sin x\right) + \left(a^{(h)}\cos x + b^{(h)}\sin x\right)\right]\\
&= f(x) + (g(x) + h(x))
\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned}
h(f(x) + g(x)) &= h\left[\left(a^{(f)}\cos x + b^{(f)}\sin x\right) + \left(a^{(g)}\cos x + b^{(g)}\sin x\right)\right]\\
&= h\left(a^{(f)}\cos x + b^{(f)}\sin x\right) + h\left(a^{(g)}\cos x + b^{(g)}\sin x\right)\\
&= hf(x) + hg(x)
\end{aligned}$$

Let α and β be two arbitrary scalars, then for rule IV we have

$$\begin{aligned}(\alpha+\beta) f(x) &= (\alpha+\beta)\left(a^{(f)}\cos x + b^{(f)}\sin x\right)\\ &= \alpha\left(a^{(f)}\cos x + b^{(f)}\sin x\right) + \beta\left(a^{(f)}\cos x + b^{(f)}\sin x\right)\\ &= \alpha f(x) + \beta g(x)\end{aligned}$$

Again, let α and β be two arbitrary scalars. For rule V we then have

$$\begin{aligned}(\alpha\beta) f(x) &= (\alpha\beta)\left(a^{(f)}\cos x + b^{(f)}\sin x\right)\\ &= \alpha\left(\beta a^{(f)}\cos x + \beta b^{(f)}\sin x\right)\\ &= \alpha\left[\beta\left(a^{(f)}\cos x + b^{(f)}\sin x\right)\right]\\ &= \alpha(\beta f(x))\end{aligned}$$

For rule VI we have

$$\begin{aligned}1f(x) = 1\left(a^{(f)}\cos x + b^{(f)}\sin x\right) = 1a^{(f)}\cos x + 1b^{(f)}\sin x &= a^{(f)}\cos x + b^{(f)}\sin x\\ &= f(x)\end{aligned}$$

And lastly, for rule VII we have

$$0f(x) = 0\left(a^{(f)}\cos x + b^{(f)}\sin x\right) = 0a^{(f)}\cos x + 0b^{(f)}\sin x = 0 + 0 = 0$$

Hence, all functions of the form $f(x) = a^{(f)}\cos x + b^{(f)}\sin x$ form a vector space V. Consider the equation

$$a\cos x + b\sin x = 0$$

That is, let the function $f(x)$ coincide with the 0 function and hence, have the value 0 for any x. In particular, let us consider the special cases $x = 0$ and $x = \pi/2$:

$$\begin{aligned}a\cos 0 + b\sin 0 = 0 \quad &\Longrightarrow \quad a = 0\\ a\cos\pi/2 + b\sin\pi/2 = 0 \quad &\Longrightarrow \quad b = 0\end{aligned}$$

Since we require $a\cos x + b\sin x = 0$ for any x, thus also in particular when $x = 0$ or $x = \pi/2$, we conclude that the only way to guarantee this is to choose $a = b = 0$. Therefore, $\cos x$ and $\sin x$ are linearly independent and form a basis for V and V is a 2-dimensional vector space.

(h) The general solution to the third order equation $y''' - y' = 0$ is given by

$$y = F(x) = a + be^{x} + ce^{-x} \quad \text{for } -\infty < x < \infty$$

Choosing $a = 1$, $b = 1$, $c = -2$ and $x = 0$, we can verify that $F(0) = 0$. Now let $f(x)$ and $g(x)$ be two arbitrary functions of the form above. For rule I we have

$$\begin{aligned}
f(x) + g(x) &= \left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) + \left(a^{(g)} + b^{(g)}e^{x} + c^{(g)}e^{-x}\right) \\
&= \left(a^{(f)} + a^{(g)}\right) + \left(b^{(f)} + b^{(g)}\right)e^{x} + \left(c^{(f)} + c^{(g)}\right)e^{-x} \\
&= \left(a^{(g)} + a^{(f)}\right) + \left(b^{(g)} + b^{(f)}\right)e^{x} + \left(c^{(g)} + c^{(f)}\right)e^{-x} \\
&= \left(a^{(g)} + b^{(g)}e^{x} + c^{(g)}e^{-x}\right) + \left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) \\
&= g(x) + f(x)
\end{aligned}$$

Let $f(x)$ and $g(x)$ be as before and let $h(x)$ be a third arbitrary function of the given form. For rule II we have

$$\begin{aligned}
(f(x) + g(x)) + h(x) &= \left[\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) + \left(a^{(g)} + b^{(g)}e^{x} + c^{(g)}e^{-x}\right)\right] \\
&\quad + \left(a^{(h)} + b^{(h)}e^{x} + c^{(h)}e^{-x}\right) \\
&= \left(a^{(f)} + a^{(g)} + a^{(h)}\right) + \left(b^{(f)} + b^{(g)} + b^{(h)}\right)e^{x} \\
&\quad + \left(c^{(f)} + c^{(g)} + c^{(h)}\right)e^{-x} \\
&= \left[a^{(f)} + \left(a^{(g)} + a^{(h)}\right)\right] + \left[b^{(f)} + \left(b^{(g)} + b^{(h)}\right)\right]e^{x} \\
&\quad + \left[c^{(f)} + \left(c^{(g)} + c^{(h)}\right)\right]e^{-x} \\
&= \left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) \\
&\quad + \left[\left(a^{(g)} + b^{(g)}e^{x} + c^{(g)}e^{-x}\right) + \left(a^{(h)} + b^{(h)}e^{x} + c^{(h)}e^{-x}\right)\right] \\
&= f(x) + (g(x) + h(x))
\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned}
h(f(x) + g(x)) &= h\left[\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) + \left(a^{(g)} + b^{(g)}e^{x} + c^{(g)}e^{-x}\right)\right] \\
&= h\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) + h\left(a^{(g)} + b^{(g)}e^{x} + c^{(g)}e^{-x}\right) \\
&= hf(x) + hg(x)
\end{aligned}$$

Let α and β be two arbitrary scalars, then for rule IV we have

$$\begin{aligned}
(\alpha + \beta) f(x) &= (\alpha + \beta)\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) \\
&= \alpha\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) + \beta\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) \\
&= \alpha f(x) + \beta f(x)
\end{aligned}$$

Again, let α and β be two arbitrary scalars. For rule V we then have

$$\begin{aligned}
(\alpha\beta) f(x) &= (\alpha\beta)\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) \\
&= \alpha\left(\beta a^{(f)} + \beta b^{(f)}e^{x} + \beta c^{(f)}e^{-x}\right) \\
&= \alpha\left[\beta\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right)\right] \\
&= \alpha(\beta f(x))
\end{aligned}$$

For rule VI we have

$$\begin{aligned}1f(x) = 1\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) &= 1a^{(f)} + 1b^{(f)}e^{x} + 1c^{(f)}e^{-x} \\ &= a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x} \\ &= f(x)\end{aligned}$$

And lastly, for rule VII we have

$$0f(x) = 0\left(a^{(f)} + b^{(f)}e^{x} + c^{(f)}e^{-x}\right) = 0a^{(f)} + 0b^{(f)}e^{x} + 0c^{(f)}e^{-x} = 0 + 0 + 0 = 0$$

Hence, all functions of the form $f(x) = a^{(f)} = b^{(f)}e^{x} + c^{(f)}e^{-x}$ form a vector space V. Consider the equation

$$a1 + be^{x} + ce^{-x} = 0$$

That is, let the function $f(x)$ coincide with the 0 function and hence, have the value 0 for any x. Now let us choose the special cases $x = -1$, $x = 0$ and $x = 1$, which give the three equations

$$\begin{aligned}x &= 0 & &\Longrightarrow & a + b + c &= 0 \\ x &= 1 & &\Longrightarrow & a + be + ce^{-1} &= 0 \\ x &= -1 & &\Longrightarrow & a + be^{-1} + ce &= 0\end{aligned}$$

Subtracting the second equation from the first and the third equation from the first gives

$$\begin{aligned}b(1-e) + c\left(1 - e^{-1}\right) &= 0 \\ b\left(1 - e^{-1}\right) + c(1-e) &= 0\end{aligned}$$

Again, subtracting the second equation from the first gives

$$(b - c)\left[(1 - e) - \left(1 - e^{-1}\right)\right] = 0$$

Now since the term $(1-e) - (1-e^{-1}) \neq 0$ it must be that $b = c$. Substituting for c in the first and second equations above gives

$$\begin{aligned}a + 2b &= 0 \\ a + b\left(e + e^{-1}\right) &= 0\end{aligned}$$

Once more, subtracting the second equation from the first gives

$$b\left[2 - \left(e + e^{-1}\right)\right] = 0$$

Since the term $2 - (e + e^{-1}) \neq 0$ we find $b = 0$ and so $c = 0$. Substituting for b and c in any of the three equations we also find $a = 0$. In conclusion; to guarantee that the function $a1 + be^{x} + ce^{-x} = 0$ for any x we require $a = b = c = 0$ and hence, the three functions 1, e^{x} and e^{-x} are linearly independent and form a basis for the vector space V and V is a 3-dimensional vector space.

(i) Let us first note that if $f(x)$ and $g(x)$ are two function which are defined and continuous for $0 \leq x \leq 1$ then their point wise sum $(f+g)(x) = f(x) + g(x)$ will also be defined and continuous on $0 \leq x \leq 1$. Similarly, if a is a scalar and $f(x)$ is defined and continuous for $0 \leq x \leq 1$ then the product $(af)(x) = af(x)$ will also be defined and continuous for $0 \leq x \leq 1$. Having established these definitions, we can now prove all seven rules that need to be satisfied to show that all functions $f(x)$ defined and continuous for $0 \leq x \leq 1$ form a vector space V. For rule I we have

$$f(x) + g(x) = (f+g)(x) = (g+f)(x) = g(x) + f(x)$$

Let $f(x)$ and $g(x)$ be as before and let $h(x)$ be a third arbitrary function defined and continuous for $0 \leq x \leq 1$. For rule II we have

$$\begin{aligned}(f(x) + g(x)) + h(x) = ((f+g)+h)(x) &= (f+(g+h))(x) \\ &= f(x) + (g(x) + h(x))\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned}h(f(x) + g(x)) = (h(f+g))(x) = (hf+hg)(x) &= (hf)(x) + (hg)(x) \\ &= hf(x) + hg(x)\end{aligned}$$

Let a and b be two arbitrary scalars, then for rule IV we have

$$\begin{aligned}(a+b)f(x) = ((a+b)f)(x) = (af+bf)(x) &= (af)(x) + (bf)(x) \\ &= af(x) + bf(x)\end{aligned}$$

Again, let a and b be two arbitrary scalars. For rule V we then have

$$(ab)f(x) = ((ab)f)(x) = (a(bf))(x) = a(bf(x))$$

For rule VI we have

$$1f(x) = (1f)(x) = (f)(x) = f(x)$$

And lastly, for rule VII we have

$$0f(x) = (0f)(x) = (0)(x) = 0$$

Hence, all functions $f(x)$ defined and continuous for $0 \leq x \leq 1$ form an infinite-dimensional vector space V.

(j) Let us first note that if $f(x)$ and $g(x)$ have continuous derivatives $f'(x)$ and $g'(x)$ for $0 \leq x \leq 1$ then the point wise sum $(f+g)'(x) = f'(x) + g'(x)$ will also be defined and continuous for $0 \leq x \leq 1$. Similarly, if a is a scalar and $f'(x)$ is defined and continuous for $0 \leq x \leq 1$ then the product $(af(x))' = af'(x)$ will also be defined and continuous for $0 \leq x \leq 1$. Having established these definitions, we can now prove all seven rules that need to be satisfied to show that all functions $f(x)$ which are defined and have a continuous derivative $f'(x)$ for $0 \leq x \leq 1$ form a vector space V. Note that we re-use the solutions of the preceding problem (part (i)) in order to shorten the proves for the current problem (e.g. $f(x) + g(x) = g(x) + f(x)$). For rule I we have

$$f'(x) + g'(x) = (f(x) + g(x))' = (g(x) + f(x))' = g'(x) + f'(x)$$

Let $f(x)$ and $g(x)$ be as before and let $h(x)$ be a third arbitrary function defined and continuous for $0 \leq x \leq 1$. For rule II we have

$$\begin{aligned}(f'(x) + g'(x)) + h'(x) = ((f(x) + g(x)) + h(x))' &= (f(x) + (g(x) + h(x)))' \\ &= f'(x) + (g'(x) + h'(x))\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned}h(f'(x) + g'(x)) = (h(f(x) + g(x)))' = (hf(x) + hg(x))' &= (hf(x))' + (hg(x))' \\ &= hf'(x) + hg'(x)\end{aligned}$$

Let a and b be two arbitrary scalars, then for rule IV we have

$$\begin{aligned}(a+b) f'(x) = ((a+b) f(x))' = (af(x) + bf(x))' &= (af(x))' + (bf(x))' \\ &= af'(x) + bf'(x)\end{aligned}$$

Again, let a and b be two arbitrary scalars. For rule V we then have

$$(ab) f'(x) = ((ab) f(x))' = (a(bf(x)))' = a(bf'(x))$$

For rule VI we have

$$1f'(x) = (1f(x))' = (f(x))' = f'(x)$$

And lastly, for rule VII we have

$$0f'(x) = (0f(x))' = (0)' = 0$$

Hence, all functions $f(x)$ which are defined and have a continuous derivative for $0 \leq x \leq 1$ form an infinite-dimensional vector space V.

(k) To demonstrate that all infinite sequences $x_1, x_2, \ldots, x_n, \ldots$ forms a vector space V we will prove each of the seven rules. Let $y_1, y_2, \ldots, y_n, \ldots$ be a second infinite sequence. For rule I we have

$$\begin{aligned}(x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots) &= x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots \\ &= y_1 + x_1, y_2 + x_2, \ldots, y_n + x_n, \ldots \\ &= (y_1, y_2, \ldots, y_n, \ldots) + (x_1, x_2, \ldots, x_n, \ldots)\end{aligned}$$

Let $x_1, x_2, \ldots, x_n, \ldots$ and $y_1+x_1, y_2+x_2, \ldots, y_n$ be as before and let $z_1, z_2, \ldots, z_n, \ldots$ be a third arbitrary infinite sequence. For rule II we have

$$\begin{aligned}
\left((x_k)_{k=1}^{k=\infty} + (y_k)_{k=1}^{k=\infty}\right) + (z_k)_{k=1}^{k=\infty} &= ((x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots)) \\
&\quad + (z_1, z_2, \ldots, z_n, \ldots) \\
&= x_1 + y_1 + z_1, x_2 + y_2 + z_2, \ldots, x_n + y_n + z_n, \ldots \\
&= x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), \ldots, x_n + (y_n + z_n), \ldots \\
&= (x_1, x_2, \ldots, x_n, \ldots) \\
&\quad + ((y_1, y_2, \ldots, y_n, \ldots) + (z_1, z_2, \ldots, z_n, \ldots)) \\
&= (x_k)_{k=1}^{k=\infty} + \left((y_k)_{k=1}^{k=\infty} + (z_k)_{k=1}^{k=\infty}\right)
\end{aligned}$$

Let h be an arbitrary scalar, then for rule III we have

$$\begin{aligned}
h\left((x_k)_{k=1}^{k=\infty} + (y_k)_{k=1}^{k=\infty}\right) &= h\left((x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots)\right) \\
&= (hx_1, hx_2, \ldots, hx_n, \ldots) + (hy_1, hy_2, \ldots, hy_n, \ldots) \\
&= h(x_1, x_2, \ldots, x_n, \ldots) + h(y_1, y_2, \ldots, y_n, \ldots) \\
&= h\left((x_k)_{k=1}^{k=\infty}\right) + h\left((y_k)_{k=1}^{k=\infty}\right)
\end{aligned}$$

Let a and b be two arbitrary scalars, then for rule IV we have

$$\begin{aligned}
(a+b)(x_k)_{k=1}^{k=\infty} &= (a+b)(x_1, x_2, \ldots, x_n, \ldots) \\
&= (ax_1, ax_2, \ldots, ax_n, \ldots) + (bx_1, bx_2, \ldots, bx_n, \ldots) \\
&= a(x_1, x_2, \ldots, x_n, \ldots) + b(x_1, x_2, \ldots, x_n, \ldots) \\
&= a(x_k)_{k=1}^{k=\infty} + b(x_k)_{k=1}^{k=\infty}
\end{aligned}$$

Again, let a and b be two arbitrary scalars. For rule V we then have

$$\begin{aligned}
(ab)(x_k)_{k=1}^{k=\infty} = (ab)(x_1, x_2, \ldots, x_n, \ldots) &= a(bx_1, bx_2, \ldots, bx_n, \ldots) \\
&= a\left[b(x_1, x_2, \ldots, x_n, \ldots)\right] \\
&= a\left(b(x_k)_{k=1}^{k=\infty}\right)
\end{aligned}$$

For rule VI we have

$$\begin{aligned}
1(x_k)_{k=1}^{k=\infty} = 1(x_1, x_2, \ldots, x_n, \ldots) = 1x_1, 1x_2, \ldots, 1x_n, \ldots &= x_1, x_2, \ldots, x_n, \ldots \\
&= (x_k)_{k=1}^{k=\infty}
\end{aligned}$$

And lastly, for rule VII we have

$$0(x_k)_{k=1}^{k=\infty} = 0(x_1, x_2, \ldots, x_n, \ldots) = 0x_1, 0x_2, \ldots, 0x_n, \cdots = 0, 0, \ldots, 0, \ldots$$

Hence, all infinite sequences form an infinite-dimensional vector space V.

(l) Let us start by noting that if $(x_k)_{k=1}^{k=\infty}$ and $(y_k)_{k=1}^{k=\infty}$ are two sequences that both converge then the sum $(x_k)_{k=1}^{k=\infty} + (y_k)_{k=1}^{k=\infty}$ converges also. Similarly, if a is a scalar and $(x_k)_{k=1}^{k=\infty}$ converges then the product $a(x_k)_{k=1}^{k=\infty}$ will converge as well. These facts, together with the proofs for the seven rules given for part (k) of this problem are enough to show that all convergent sequences form an infinite-dimensional vector space V.

2. Let $P_0(x) = 1$, $P_1(x) = 1 + x$, $P_2(x) = 1 + x + x^2$ and $P_3(x) = 1 + x + x^3$ and c_1, c_2, c_3 and c_4 be four arbitrary scalars and let us consider the equation

$$\begin{aligned} c_1 P_0(x) + c_2 P_1(x) + c_3 P_2(x) + c_4 P_3(x) &= 0 \\ c_1 + c_2(1+x) + c_3\left(1+x+x^2\right) + c_4\left(1+x+x^3\right) &= \\ (c_1+c_2+c_3) + (c_2+c_3+c_4)x + c_3x^2 + c_4x^3 &= \end{aligned}$$

Clearly we require $c_3 = c_4 = 0$ for this equation to make sense for any x and from that fact it also follows that $c_2 = 0$ and $c_1 = 0$. Hence, the four polynomials $P_0(x)$, $P_1(x)$, $P_2(x)$ and $P_3(x)$ are linearly independent and thus form a basis for the vector space of Example 2.

3. Let us consider the equation

$$c_1\begin{bmatrix}1 & 2\\1 & -3\end{bmatrix} + c_2\begin{bmatrix}0 & 1\\2 & -1\end{bmatrix} + c_3\begin{bmatrix}5 & 2\\7 & 1\end{bmatrix} + c_4\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

which corresponds to the system of four linear equations

$$\begin{aligned} c_1 + \quad + 5c_3 \quad &= 0 \\ 2c_1 + c_2 + 2c_3 + c_4 &= 0 \\ c_1 + 2c_2 + 7c_3 + c_4 &= 0 \\ -3c_1 - c_2 + c_3 \quad &= 0 \end{aligned}$$

Subtracting the third equation from the second gives

$$c_1 - c_2 - 5c_3 = 0$$

Adding the resulting equation to the first equation gives

$$2c_1 - c_2 = 0 \qquad \Longrightarrow \qquad c_2 = 2c_1$$

Substituting for c_2 in the fourth equation gives

$$-5c_1 + c_3 = 0 \qquad \Longrightarrow \qquad c_3 = 5c_1$$

Substituting for c_3 in the first equation then gives

$$26c_1 = 0$$

which implies $c_1 = 0$ and $c_3 = 0$. From the fourth equation it then follows that $c_2 = 0$ and finally this implies that $c_4 = 0$ also. Hence, the four matrices are linearly independent and thus form a basis for the vector space of Example 3.

4. Let us start by rewriting $\cos 2x$ and $\sin 2x$ as

$$\begin{aligned}
\cos 2x &= \frac{1}{2}\left(\cos 2x + \mathrm{i}\sin 2x\right) + \frac{1}{2}\left(\cos 2x - \mathrm{i}\sin 2x\right) \\
&= \frac{1}{2}\left(e^{2\mathrm{i}x} + e^{-2\mathrm{i}x}\right) \\
\sin 2x &= \frac{1}{2}\left(\cos 2x + \mathrm{i}\sin 2x\right) - \frac{1}{2}\left(\cos 2x - \mathrm{i}\sin 2x\right) \\
&= \frac{1}{2}\left(e^{2\mathrm{i}x} - e^{-2\mathrm{i}x}\right)
\end{aligned}$$

Next, let us consider the equation

$$a\cos 2x + b\sin 2x = \frac{a+b}{2}e^{2\mathrm{i}x} + \frac{a-b}{2}e^{-2\mathrm{i}x} = 0$$

Now let us choose the special case $x = 0$, so that the equation above becomes

$$\frac{a+b}{2} + \frac{a-b}{2} = 0 \qquad \Longrightarrow \qquad a = 0$$

And the special case $x = \pi/4$ which gives

$$\left(\frac{a+b}{2} - \frac{a-b}{2}\right)\mathrm{i} = 0 \qquad \Longrightarrow \qquad b = 0$$

Hence, in order to guarantee that we will get a zero result for any x we need $a = b = 0$ and so $\cos 2x$ and $\sin 2x$ are linearly independent and form a basis for the vector space of Example 6.