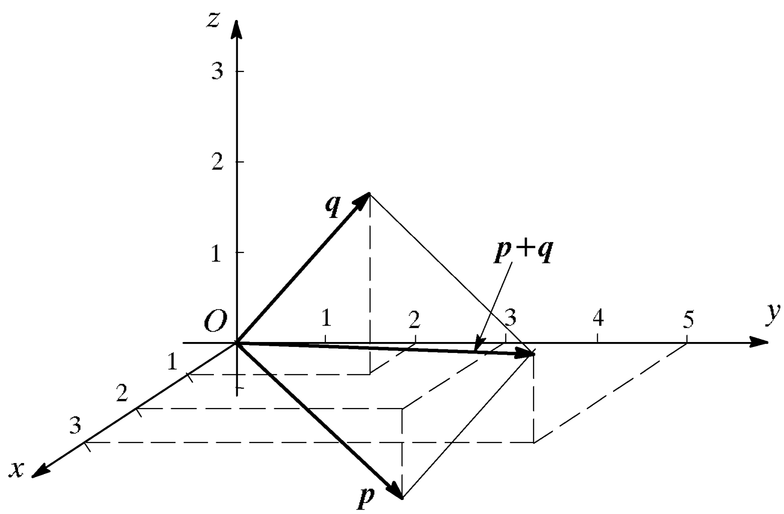
Student Solution Manual for Introduction to Linear Algebra

Geza Schay and Dennis Wortman

1.1.1. $\overrightarrow{PR} = \mathbf{r} - \mathbf{p}$, $\overrightarrow{PQ} = \mathbf{q} - \mathbf{p}$, and $\overrightarrow{QP} = \mathbf{p} - \mathbf{q}$.
$\overrightarrow{QC} = \frac{1}{2}\overrightarrow{QP} = \frac{1}{2}\mathbf{p} - \frac{1}{2}\mathbf{q}$, $\overrightarrow{PC} = \frac{1}{2}\overrightarrow{PQ} = \frac{1}{2}\mathbf{q} - \frac{1}{2}\mathbf{p}$, and $\overrightarrow{OC} = \frac{1}{2}\mathbf{r} = \frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{q}$.
1.1.3. $\mathbf{p} + \mathbf{q} = (2, 3, -1) + (1, 2, 2) = (3, 5, 1)$.



1.1.5. $\mathbf{r} = \mathbf{p} + \mathbf{q}$, $\mathbf{2r} = \mathbf{2}(\mathbf{p} + \mathbf{q}) = \mathbf{2p} + \mathbf{2q}.$

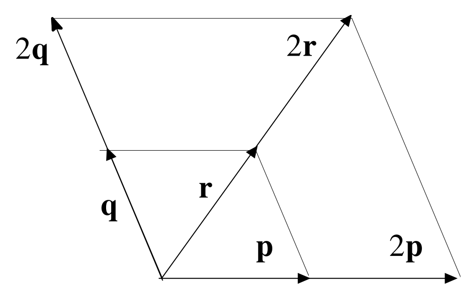


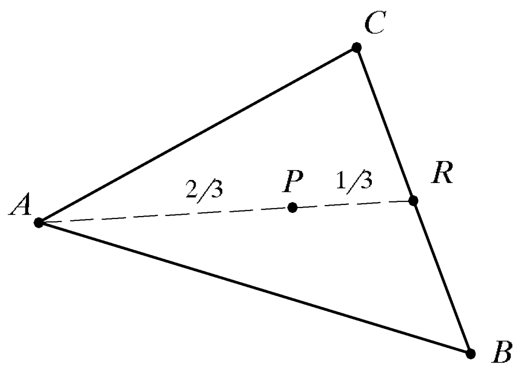
Figure 0.1.

1.1.7. $M = \sum_{i=1}^{3} m_i = 2 + 3 + 5 = 10$ and

$$\mathbf{r} = \frac{1}{M}\sum_{i=1}^{3} m_i \mathbf{r}_i = \frac{1}{10}[2(2, -1, 4) + 3(1, 5, -6) + 5(-2, -5, 4)] = (-\frac{3}{10}, -\frac{6}{5}, 1).$$

1.1.9. The center R of the side BC has position vector $\mathbf{r} = \frac{1}{2}(\mathbf{b}+\mathbf{c})$ and the vector from R to A is $\mathbf{a}-\mathbf{r}$. Thus the point P, 1/3 of the way from R to A, has position vector

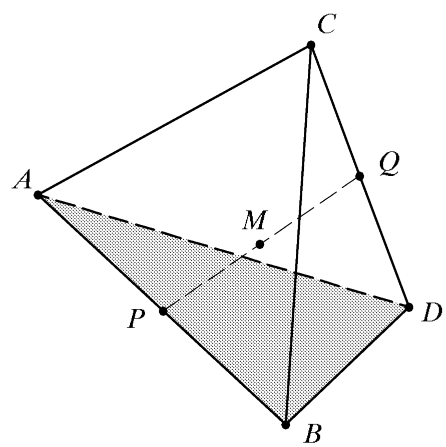
$$\mathbf{p} = \tfrac{1}{2}(\mathbf{b}+\mathbf{c}) + \tfrac{1}{3}(\mathbf{a}-\mathbf{r}) = \tfrac{1}{2}\mathbf{b} + \tfrac{1}{2}\mathbf{c} + \tfrac{1}{3}\mathbf{a} - \tfrac{1}{3}\cdot\tfrac{1}{2}(\mathbf{b}+\mathbf{c}) = \tfrac{1}{3}(\mathbf{a}+\mathbf{b}+\mathbf{c}).$$



1.1.11. Opposite edges are those with no common vertex; for example, the one from A to B, and the one from C to D. The midpoint of the former has position vector $\mathbf{p} = \frac{1}{2}(\mathbf{a}+\mathbf{b})$ and of the latter, $\mathbf{q} = \frac{1}{2}(\mathbf{c}+\mathbf{d})$. The halfway point M on the line joining those midpoints has the position vector

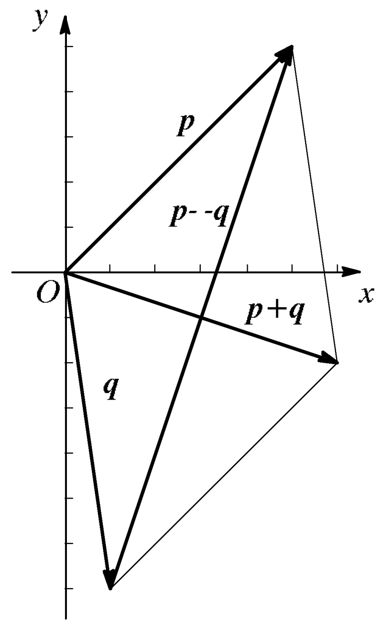
$$\begin{aligned}\mathbf{m} &= \mathbf{p} + \frac{1}{2}(\mathbf{q}-\mathbf{p}) = \frac{1}{2}(\mathbf{p}+\mathbf{q}) = \frac{1}{2}\left[\frac{1}{2}(\mathbf{a}+\mathbf{b}) + \frac{1}{2}(\mathbf{c}+\mathbf{d})\right] \\ &= \frac{1}{4}(\mathbf{a}+\mathbf{b}+\mathbf{c}+\mathbf{d}),\end{aligned}$$

which is the position vector of the centroid.



1.2.1. Let $\mathbf{p} = (5, 5)$ and $\mathbf{q} = (1, -7)$. Then
a. $\mathbf{p} + \mathbf{q} = (5 + 1, 5 - 7) = (6, -2)$, $\mathbf{p} - \mathbf{q} = (5 - 1, 5 + 7) = (4, 12)$.

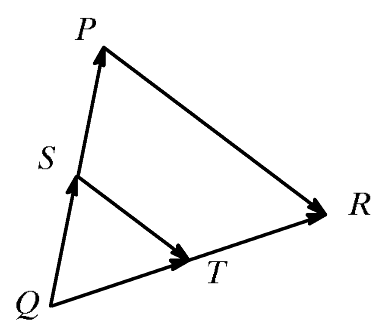
b.



c. $|\mathbf{p}| = \sqrt{5^2+5^2} = \sqrt{50},\ |\mathbf{q}| = \sqrt{1^2+(-7)^2} = \sqrt{50},$
$|\mathbf{p}+\mathbf{q}| = \sqrt{6^2+(-2)^2} = \sqrt{40},\ |\mathbf{p}-\mathbf{q}| = \sqrt{4^2+12^2} = \sqrt{160}.$
d. $|\mathbf{p}+\mathbf{q}|^2 = 40,\ |\mathbf{p}|^2 + |\mathbf{q}|^2 = 50+50 = 100$, and so $|\mathbf{p}+\mathbf{q}|^2 \neq |\mathbf{p}|^2 + |\mathbf{q}|^2$.

1.2.3. Let S be the midpoint of $\overrightarrow{QP}$ and T the midpoint of $\overrightarrow{QR}$. Then

$$\overrightarrow{ST} = \overrightarrow{QT} - \overrightarrow{QS} = \frac{1}{2}\overrightarrow{QR} - \frac{1}{2}\overrightarrow{QP} = \frac{1}{2}(\overrightarrow{QR} - \overrightarrow{QP}) = \frac{1}{2}\overrightarrow{PR}.$$

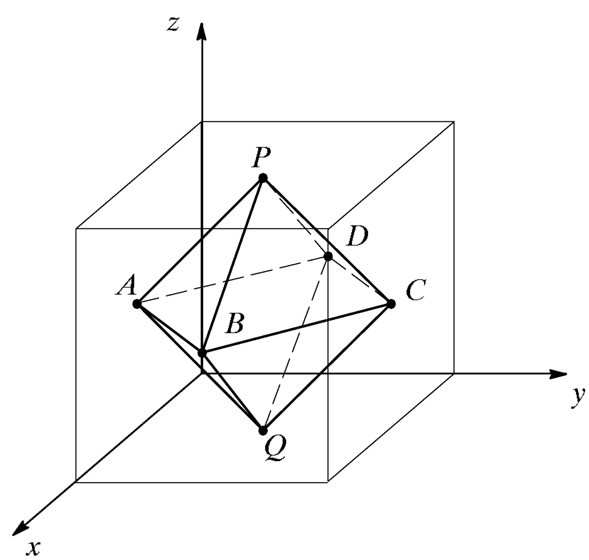


1.2.5.

$$\cos\theta = \frac{\mathbf{p}\cdot\mathbf{q}}{|\mathbf{p}||\mathbf{q}|} = \frac{1\cdot 3 + (-2)\cdot 5 + 4\cdot 2}{\sqrt{21}\cdot\sqrt{38}} = \frac{1}{\sqrt{798}} \approx 0.0354.$$

Hence $\theta \approx 1.535$ radians $\approx 87.9\,^\circ$

1.2.7.



The above diagram illustrates the regular octahedron inscribed in the unit cube, joining the centers of the sides of the cube by line segments. Using the notation from the diagram, we find four distinct types of angles formed by the sides of the octahedron, represented, for example, by $\angle PAB$, $\angle DAB$, $\angle PAQ$ and $\angle APC$. To determine those angles, first calculate the coordinates of the relevant points:

$$P(\frac{1}{2},\frac{1}{2},1),\ A(\frac{1}{2},0,\frac{1}{2}),\ B(1.\frac{1}{2},\frac{1}{2}),\ D(0,\frac{1}{2},\frac{1}{2}),\ Q(\frac{1}{2},\frac{1}{2},0),\ C(\frac{1}{2},1,\frac{1}{2}).$$

Hence

$$\overrightarrow{AP} = (0,\frac{1}{2},\frac{1}{2}) \text{ and } \overrightarrow{AB} = (\frac{1}{2},\frac{1}{2},0),$$

and thus

$$\cos\angle PAB = \frac{\overrightarrow{AP}\cdot\overrightarrow{AB}}{|\overrightarrow{AP}||\overrightarrow{AB}|} = \frac{1/4}{1/2} = \frac{1}{2};\ \therefore\ \angle PAB = 60^\circ$$

Also, $\overrightarrow{AD} = (-\frac{1}{2}, \frac{1}{2}, 0)$ and thus

$$\cos\angle DAB = \frac{\overrightarrow{AB}\cdot\overrightarrow{AD}}{|\overrightarrow{AB}||\overrightarrow{AD}|} = 0;\ \therefore\ \angle DAB = 90^\circ$$

Also $\overrightarrow{AQ} = (0, \frac{1}{2}, -\frac{1}{2})$ and

$$\cos\angle PAQ = \frac{\overrightarrow{AP}\cdot\overrightarrow{AQ}}{|\overrightarrow{AP}||\overrightarrow{AQ}|} = 0;\ \therefore\ \angle PAQ = 90^\circ$$

Lastly,

$$\overrightarrow{PA} = (0, -\frac{1}{2}, -\frac{1}{2}) \text{ and } \overrightarrow{PC} = (0, \frac{1}{2}, -\frac{1}{2}),$$

and so

$$\cos\angle APC = \frac{\overrightarrow{PA}\cdot\overrightarrow{PC}}{|\overrightarrow{PA}||\overrightarrow{PC}|} = 0;\ \therefore\ \angle APC = 90^\circ$$

Remark: we used vector methods in the above solution, but the angles could also have been determined using simple geometric properties.

1.2.9. The parallel component is

$$\mathbf{p}_1 = \frac{\mathbf{p}\cdot\mathbf{q}}{|\mathbf{q}|^2}\mathbf{q} = \frac{24-9+4}{12^2+3^2+4^2}(12,3,4) = \frac{19}{169}(12,3,4)$$

and the perpendicular component is

$$\mathbf{p}_2 = \mathbf{p} - \mathbf{p}_1 = (2,-3,1) - \frac{19}{169}(12,3,4) = (\frac{110}{169}, -\frac{564}{169}, \frac{93}{169}).$$

1.2.11. Call the end points of the given diameter A and B. Then, with the notation of Figure 1.20,

$$\overrightarrow{AP} \cdot \overrightarrow{BP} = (\mathbf{p} - \mathbf{r}) \cdot (\mathbf{p} + \mathbf{r}) = \mathbf{p} \cdot \mathbf{p} - \mathbf{r} \cdot \mathbf{r} = |\mathbf{p}|^2 - |\mathbf{r}|^2 = 0.$$

The last equality holds because $\mathbf{p}$ and $\mathbf{r}$ are both radius vectors, and so their lengths are equal.

1.2.13.

a. Note first that

$$|\mathbf{p} + \mathbf{q}|^2 = (\mathbf{p} + \mathbf{q}) \cdot (\mathbf{p} + \mathbf{q}) = \mathbf{p} \cdot \mathbf{p} + 2\mathbf{p} \cdot \mathbf{q} + \mathbf{q} \cdot \mathbf{q} \leq \mathbf{p} \cdot \mathbf{p} + 2|\mathbf{p} \cdot \mathbf{q}| + \mathbf{q} \cdot \mathbf{q}.$$

In turn, by Cauchy's inequality,

$$\mathbf{p} \cdot \mathbf{p} + 2|\mathbf{p} \cdot \mathbf{q}| + \mathbf{q} \cdot \mathbf{q} \leq \mathbf{p} \cdot \mathbf{p} + 2|\mathbf{p}||\mathbf{q}| + \mathbf{q} \cdot \mathbf{q}.$$

Since $\mathbf{p} \cdot \mathbf{p} = |\mathbf{p}|^2$ and $\mathbf{q} \cdot \mathbf{q} = |\mathbf{q}|^2$, we have

$$\mathbf{p} \cdot \mathbf{p} + 2|\mathbf{p}||\mathbf{q}| + \mathbf{q} \cdot \mathbf{q} = |\mathbf{p}|^2 + 2|\mathbf{p}||\mathbf{q}| + |\mathbf{q}|^2 = (|\mathbf{p}| + |\mathbf{q}|)^2.$$

Finally, this chain of steps implies that $|\mathbf{p} + \mathbf{q}|^2 \leq (|\mathbf{p}| + |\mathbf{q}|)^2$, from which the Triangle Inequality follows by taking square roots of both sides.

b. In Cauchy's inequality, equality occurs when $\mathbf{q} = \mathbf{0}$ or $\mathbf{p} = \lambda\mathbf{q}$, for any λ, but the first inequality above becomes an equality for $\mathbf{p} = \lambda\mathbf{q}$ only if $\lambda \geq 0$. Thus, the chain of steps above remains true with equality replacing the inequalities if and only if the vectors $\mathbf{p}$ and $\mathbf{q}$ are parallel and point in the same direction.

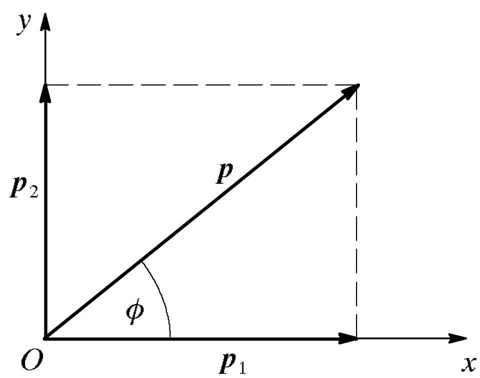
1.2.15.

a.

$$\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2 = |\mathbf{p}| \cos\phi\mathbf{i} + |\mathbf{p}| \sin\phi\mathbf{j} = (|\mathbf{p}| \cos\phi, |\mathbf{p}| \sin\phi) = |\mathbf{p}|(\cos\phi, \sin\phi).$$

b.

$$\mathbf{u}_p = \frac{\mathbf{p}}{|\mathbf{p}|} = \frac{|\mathbf{p}|(\cos\phi, \sin\phi)}{|\mathbf{p}|} = (\cos\phi, \sin\phi).$$



$\mathbf{1.2.17}$. If $\mathbf{p} = (3, -4, 12)$, then $|\mathbf{p}| = \sqrt{9 + 16 + 144} = 13$ and therefore

$$\cos\alpha_1 = \frac{\mathbf{p}\cdot\mathbf{i}}{|\mathbf{p}|} = \frac{p_1}{|\mathbf{p}|} = \frac{3}{13},\ \cos\alpha_2 = \frac{\mathbf{p}\cdot\mathbf{j}}{|\mathbf{p}|} = \frac{p_2}{|\mathbf{p}|} = -\frac{4}{13},$$

and

$$\cos\alpha_3 = \frac{\mathbf{p}\cdot\mathbf{k}}{|\mathbf{p}|} = \frac{p_3}{|\mathbf{p}|} = \frac{12}{13}.$$

$\mathbf{1.2.19}$. We have to verify that the first three parts of Theorem 1.2.2 remain valid for the product defined by $\mathbf{p}\cdot\mathbf{q} = 2p_1q_1 + p_2q_2$:

1. $\mathbf{p}\cdot\mathbf{q} = 2p_1q_1 + p_2q_2 = 2q_1p_1 + q_2p_2 = \mathbf{q}\cdot\mathbf{p}$,
2. $\mathbf{p}\cdot(\mathbf{q}+\mathbf{r}) = 2p_1(q_1+r_1) + p_2(q_2+r_2) = 2p_1q_1 + 2p_1r_1 + p_2q_2 + p_2r_2$
$= 2p_1q_1 + p_2q_2 + 2p_1r_1 + p_2r_2 = \mathbf{p}\cdot\mathbf{q} + \mathbf{p}\cdot\mathbf{r}$,
3. $c(\mathbf{p}\cdot\mathbf{q}) = c(2p_1q_1 + p_2q_2) = 2(cp_1)q_1 + (cp_2)q_2 = (c\mathbf{p})\cdot\mathbf{q}$ and
$c(\mathbf{p}\cdot\mathbf{q}) = c(2p_1q_1 + p_2q_2) = 2p_1(cq_1) + p_2(cq_2) = \mathbf{p}\cdot(c\mathbf{q})$.

If we stretch the first component of every vector by a factor of $\sqrt{2}$, then this product applied to the original vectors gives the lengths of and the angles between the altered vectors the same way as the standard dot product would when applied to the altered vectors.

$\mathbf{1.3.1}$. The vector parametric form of the equation of the line is $\mathbf{p} = \mathbf{p}_0 + t\mathbf{v} = (1, -2, 4) + t(2, 3, -5) = (1 + 2t, -2 + 3t, 4 - 5t)$, and thus the scalar parametric form is $x = 1+2t,\ y = -2+3t,\ z = 4-5t$, and the nonparametric form is $\frac{x-1}{2} = \frac{y+2}{3} = \frac{z-4}{-5}$.

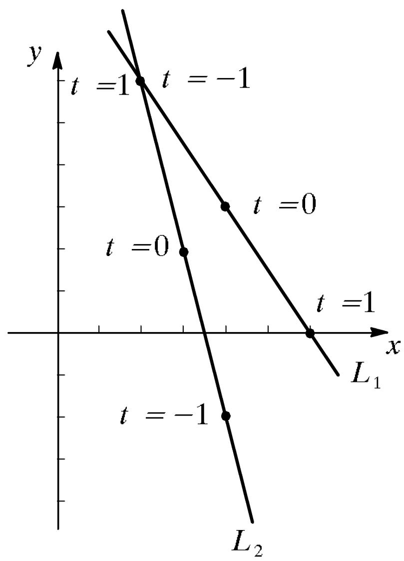
$\mathbf{1.3.3}$. A direction vector of the line through P_0 and P_1 is $\mathbf{v} = \overrightarrow{P_0P_1} =$

$(5-7, 6-(-2), -3-5) = (-2, 8, -8)$. Hence the vector parametric form of the equation of the line is $\mathbf{p} = \mathbf{p}_0 + t\mathbf{v} = (7, -2, 5) + t(-2, 8, -8) = (7-2t, -2+8t, 5-8t)$, and thus the scalar parametric form is $x = 7-2t,\ y = -2+8t,\ z = 5-8t$, and the nonparametric form is $\frac{x-7}{-2} = \frac{y+2}{8} = \frac{z-5}{-8}$.

1.3.5. A direction vector of the line through the given points P_0 and P_1 is $\mathbf{v} = \overrightarrow{P_0P_1} = (1-1, -2-(-2), -3-4) = (0, 0, -7)$. Hence the vector parametric form of the equation of the line is $\mathbf{p} = \mathbf{p}_0 + t\mathbf{v} = (1, -2, 4) + t(0, 0, -7) = (1, -2, 4-7t)$, and thus the scalar parametric form is $x = 1$, $y = -2,\ z = 4-7t$, and the nonparametric form is $x = 1,\ y = -2$.

1.3.7. As noted in the given hint, the direction vector $\mathbf{w}$ of the line is orthogonal to the normal vectors to the two planes, $\mathbf{u} = (3, -4, 3)$ and $\mathbf{v} = (0, 0, 1)$. Proceeding as in Example 1.3.7 in the text, the equation of the plane through the origin determined by the vectors $\mathbf{u}$ and $\mathbf{v}$ can be expressed as $\mathbf{p} = s\mathbf{u} + t\mathbf{v}$, where s and t are parameters, or, in scalar form, as $x = 3s,\ y = -4s,\ z = 3s+t$. Eliminating the parameters results in the equation $4x+3y = 0$ (the variable z may take any value since the parameter t is "free"), and thus $\mathbf{w} = (4, 3, 0)$. Therefore, the vector parametric form of the equation of the line is $\mathbf{p} = \mathbf{p}_0 + t\mathbf{w} = (5, 4, -8) + t(4, 3, 0)$, and thus the scalar parametric form is $x = 5+4t$, $y = 4+3t$, $z = -8$, and the nonparametric form is $\frac{x-5}{4} = \frac{y-4}{3}$, $z = -8$.

1.3.9. a.



b. The t-scales on the two lines have their zeros at different points. Or, alternatively, if t denotes time, then the two equations describe two moving points that are at the intersection at different times.

c. $\mathbf{p} = (2, 6) + s(2, -3)$ and $\mathbf{p} = (2, 6) + s(-1, 4)$. (Alternatively, any nonzero scalar multiples of the direction vectors $(2, -3)$ and $(-1, 4)$ will do in their places.)

1.3.11. The vector $\mathbf{p} = r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ represents the point A when $r = 1$ (and $s = t = 0$), the point B when $s = 1$, and the point C when $t = 1$. When $r = 0$, $\mathbf{p}$ represents a point on the line segment joining B and C (by the results of Exercise 1.3.10), and similarly for the cases $s = 0$ and $t = 0$, $\mathbf{p}$ represents a point on the other two sides of the triangle formed by A, B, and C.

When none of the variables r, s, or t is 0 or 1, then $\mathbf{p}$ represents a point in the interior of that triangle. This can be seen by first noting that since $r = 1 - s - t$, the equation $\mathbf{p} = r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ can be expressed in the form $\mathbf{p} = (1 - s - t)\mathbf{a} + s\mathbf{b} + t\mathbf{c} = \mathbf{a} + s(\mathbf{b} - \mathbf{a}) + t(\mathbf{c} - \mathbf{a})$ Now, let D be the point on side AB represented by $\mathbf{d} = \mathbf{a}+s(\mathbf{b} - \mathbf{a})$, and let E be the point on side BC represented by $\mathbf{e} = \mathbf{a}+s(\mathbf{b} - \mathbf{a}) + (1 - s)(\mathbf{c} - \mathbf{a}) = s\mathbf{b} + (1 - s)\mathbf{c}$.

Then $\mathbf{p} = \mathbf{d} + \frac{t}{1-s}(1-s)(\mathbf{c}-\mathbf{a}) = \mathbf{d} + \frac{t}{1-s}(\mathbf{e}-\mathbf{d})$, and thus P is on the line segment joining D and E, $\frac{t}{1-s}$ of the way from D towards E (note that $\frac{t}{1-s} < 1$ since $s + t < 1$).

$\mathbf{1.3.13}$. The vector $\mathbf{p}_0 = (3, -2, 1)$ must lie in the required plane, and so we may choose $\mathbf{u} = \mathbf{p}_0$. For $\mathbf{v}$ we may choose the same $\mathbf{v}$ that we have for the given line: $\mathbf{v} = (2, 1, -3)$, and for the fixed point of the plane we may choose O. Thus the parametric equation can be written as $\mathbf{p} = s(3, -2, 1) + t(2, 1, -3)$.

The nonparametric equations are obtained by eliminating s and t from $x = 3s+2t,\ y = -2s+t,\ z = s-3t$. This elimination results in $5x+11y+7z = 0$.

$\mathbf{1.3.15}$. In this case it is easier to start with the nonparametric form of the equation of the plane. Since the required plane is to be orthogonal to the line given by $\mathbf{p} = (3, -2, 1) + t(2, 1, -3)$, we may take its normal vector to be the direction vector of the line, that is, take $\mathbf{n} = (2, 1, -3)$. Next, we may take $\mathbf{p}_0 = \overrightarrow{OP}_0 = (5, 4, -8)$, and therefore the equation $\mathbf{n} \cdot \mathbf{p} = \mathbf{n} \cdot \mathbf{p}_0$ becomes $(2, 1, -3) \cdot (x, y, z) = (2, 1, -3) \cdot (5, 4, -8) = 38$, that is, $2x + y - 3z = 38$.

To write parametric equations, we need to determine two nonparallel vectors that lie in the plane. To do so, note that the points $(19, 0, 0)$ and $(0, 38, 0)$ lie in the plane and thus $\mathbf{u} = (19, 0, 0) - (5, 4, -8) = (14, -4, 8)$ and $\mathbf{v} = (0, 38, 0) - (5, 4, -8) = (-5, 34, 8)$ are nonparallel vectors in the plane. Hence $\mathbf{p} = (5, 4, -8) + s(14, -4, 8) + t(-5, 34, 8)$ is a parametric vector equation of the plane. (Note that this answer is not unique: there are infinitely many other correct answers depending on the choices for $\mathbf{p}_0$, $\mathbf{u}$ and $\mathbf{v}$.)

$\mathbf{1.3.17}$. In this case it is again easier to start with the nonparametric form of the equation of the plane. Since normal vectors to parallel planes must be parallel, we may use as the normal vector $\mathbf{n}$ to the required plane that of the given plane, namely $\mathbf{n} = (7, 1, 2)$. Also, because $P_0(5, 4, -8)$ is a point in the required plane, we may take $\mathbf{p}_0 = \overrightarrow{OP}_0 = (5, 4, -8)$, and hence the equation $\mathbf{n} \cdot \mathbf{p} = \mathbf{n} \cdot \mathbf{p}_0$ becomes $(7, 1, 2) \cdot (x, y, z) = (7, 1, 2) \cdot (5, 4, -8) = 23$, that is, $7x + y + 2z = 23$.

To write parametric equations, choose any two distinct vectors $\mathbf{a}$ and $\mathbf{b}$, other than $\mathbf{p}_0$, whose components satisfy the above equation, and let $\mathbf{u} = \mathbf{a} - \mathbf{p}_0$ and $\mathbf{v} = \mathbf{b} - \mathbf{p}_0$. Since such vectors $\mathbf{a}$ and $\mathbf{b}$ represent points of the plane, $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors lying in the plane. We need to be careful to choose the vectors $\mathbf{a}$ and $\mathbf{b}$ so that $\mathbf{u}$ and $\mathbf{v}$ are not parallel. For instance, if we choose $\mathbf{a} = (0, 23, 0)$ and $\mathbf{b} = (1, 16, 0)$, then $\mathbf{u} = (0, 23, 0) - (5, 4, -8) = (-5, 19, 8)$ and $\mathbf{v} = (1, 16, 0) - (5, 4, -8) = (-4, 12, 8)$. Hence we conclude that $\mathbf{p} = (5, 4, -8) + s(-5, 19, 8) + t(-4, 12, 8)$ is a parametric vector equation of the plane. (As a check, notice that $s = 1,\ t = 0$ give $\mathbf{p} = \mathbf{a}$, and $s = 0,\ t = 1$ give $\mathbf{p} = \mathbf{b}$.)

$\mathbf{1.3.19}$. We may choose $\mathbf{u} = \mathbf{p}_1 - \mathbf{p}_0 = (1, 6, -3) - (5, 4, -8) = (-4, 2, 5)$ and $\mathbf{v} = \mathbf{p}_2 - \mathbf{p}_0 = (7, -2, 5) - (5, 4, -8) = (2, -6, 13)$. Thus we obtain $\mathbf{p} = (5, 4, -8) + s(-4, 2, 5) + t(2, -6, 13)$ as a parametric vector equation of the plane. The corresponding scalar equations are

$$x = 5 - 4s + 2t,\ y = 4 + 2s - 6t,\ z = -8 + 5s + 13t.$$

To obtain a nonparametric equation, eliminate s and t from the last three equations as in Example 1.3.5. Thus

$$28(x - 5) + 31(y - 4) + 10(z + 8) = 0.$$

or, equivalently,

$$28x + 31y + 10z = 184.$$

$\mathbf{1.3.21}$. We can proceed as in Example 1.3.3 in the text to decompose the vector equation $(-5, 4, -1) + s(2, 1, -7) = (9, -9, -2) + t(2, -4, 5)$ into three scalar equations:

$$-5 + 2s = 9 + 2t,\ 4 + s = -9 - 4t,\ -1 - 7s = -2 + 5t.$$

Solving this system yields $s = 3$ and $t = -4$. Hence the lines intersect at the point $(1, 7, -22)$.

$\mathbf{1.3.23}$. In terms of scalar components, the equation of the given line is

$$x = 3 - 3s,\ y = -2 + 5s,\ z = 6 + 7s.$$

Substituting these equations into that of the given plane results in $3(3 - 3s) + 2(-2 + 5s) - 2(6 + 7s) = 3$, or, simplifying, $13s = -10$; the solution is $s = -\frac{10}{13}$. Hence the point of intersection is $(69/13, -76/13, 8/13)$.

$\mathbf{1.3.25}$. Rewriting the equation of the given line in components and replacing the parameter s by r yield

$$x = 3 + 7r,\ y = 2 - 5r,\ z = -4 + 4r$$

and rewriting the equation of the given plane in components gives

$$x = -3s + 2t,\ y = -2 - 3t,\ z = 1\ + 3s + 4t.$$

Combining those sets of equations and simplifying result in the system

$$7r + 3s - 2t = -3$$

$$5r - 3t = 4$$

$$4r - 3s - 4t = 5.$$

The solution of this system is $r = -6$ (and $s = 49/9$ and $t = -34/3$), so the point of intersection is $(-39, 32, -28)$.

1.3.27. Pick a point P on the plane, say, $P(0, 5, 0)$. Then $\overrightarrow{PP_0} = (3, -1, 0)$. From the given equation, $\mathbf{n} = (0, 1, -2)$. Thus, as in Example 1.3.6,

$$D = \frac{|\overrightarrow{PP_0} \cdot \mathbf{n}|}{|\mathbf{n}|} = \frac{|3 \cdot 0 + (-1) \cdot 1 + 0 \cdot (-2)|}{\sqrt{0 + 1 + 4}} = \frac{1}{\sqrt{5}}.$$

1.3.29. First find the plane through O that is parallel to the direction vectors of the lines, as in Example 1.3.7, $\mathbf{p} = s(-4, 0, 3) + t(5, 0, -2)$, or, in scalar parametric form, $x = -4s + 5t,\ y = 0,\ z = 3s - 2t$. The equation of the plane in scalar form is $y = 0$, from which we obtain a normal vector, $\mathbf{n} = (0, 1, 0)$, to the plane, and hence also to the two given lines.

The point $P(2, 1, 5)$ lies on the first line and $Q(0, -2, 3)$ lies on the second one. Thus $\overrightarrow{PQ} = (-2, -3, -2)$, and as in Example 1.3.6,

$$D = \frac{|\overrightarrow{PQ} \cdot \mathbf{n}|}{|\mathbf{n}|} = \frac{|(-2) \cdot 0 + (-3) \cdot 1 + (-2) \cdot 0|}{\sqrt{0 + 1 + 0}} = 3.$$

1.3.31. Following the hint given in Exercise 1.3.30, let Q be the point $(3, 2, -4)$ on the line L, obtained by setting $s = 0$ in the equation, so $\overrightarrow{QP_0} =$

$(1,-2,4)-(3,2,-4)=(-2,-4,8)$. The direction of the line L is $\mathbf{u}=(7,-5,4)$, and thus the component of $\overrightarrow{QP}_0$ parallel to L is given by.

$$\left(\overrightarrow{QP}_0\cdot\frac{\mathbf{u}}{|\mathbf{u}|}\right)\frac{\mathbf{u}}{|\mathbf{u}|}=\frac{38}{90}(7,-5,4)=\frac{19}{45}(7,-5,4).$$

Then the component of $\overrightarrow{QP}_0$ orthogonal to L is

$$(2,4,-8)-\frac{19}{45}(7,-5,4)=\frac{1}{45}(223,85,-284),$$

and therefore the distance from P_0 to P is ,$\left|\frac{1}{45}(223,85,-284)\right|\approx 8.24$.

1.3.33. Let P_0 denote a point on the plane S, and let $\mathbf{p}_0=\overrightarrow{OP}_0$ denote the radius vector of P_0. Then $d=\mathbf{n}\cdot\mathbf{p}_0$ and, since $|\mathbf{n}|=1$, we note that

$$f(\mathbf{q})=\mathbf{n}\cdot\mathbf{q}-d=\mathbf{n}\cdot\mathbf{q}-\mathbf{n}\cdot\mathbf{p}_0=\mathbf{n}\cdot(\mathbf{q}-\mathbf{p}_0)=|\mathbf{q}-\mathbf{p}_0|\cos\theta,$$

where θ denotes the angle between the vectors $\mathbf{n}$ and $\mathbf{q}-\mathbf{p}_0$. Thus $f(\mathbf{q})=|\mathbf{q}-\mathbf{p}_0||\cos\theta|$ is the length of the component of $\mathbf{q}-\mathbf{p}_0$ perpendicular to the plane S; that is, it is the distance between S and Q. We also note that $\cos\theta\geq 0$ if $0\leq\theta\leq\frac{\pi}{2}$, when $\mathbf{n}$ points from S towards Q, and $\cos\theta\leq 0$ if $\frac{\pi}{2}\leq\theta\leq\pi$, when $\mathbf{n}$ points from Q towards S.

1.3.35. Applying the result of Exercise 1.3.33, we have $\mathbf{q}=\overrightarrow{OP}_0=(3,4,0)$, $\mathbf{n}=\frac{(0,1,-2)}{\sqrt{0+1+4}}=\frac{(0,1,-2)}{\sqrt{5}}$, and hence $d=\frac{5}{\sqrt{5}}$, determined by dividing both sides of the equation of the given plane by $\sqrt{5}$. Therefore, $f(\mathbf{q})=\mathbf{n}\cdot\mathbf{q}-d=\frac{4}{\sqrt{5}}-\frac{5}{\sqrt{5}}=-\frac{1}{\sqrt{5}}$ and thus $D=\frac{1}{\sqrt{5}}$.

1.3.37. In the solution to Exercise 1.3.29, it was determined that $\mathbf{v}=(0,1,0)$ is orthogonal to the given lines. Thus an equation of the plane containing $\mathbf{v}$ and the line $\mathbf{p}=(2,1,5)+s(-4,0,3)$ is given by $\mathbf{p}=(2,1,5)+s(-4,0,3)+t(0,1,0)$. In terms of scalar components: $x=2-4s$, $y=1+t$, $z=5+3s$. The scalar form of the equation $\mathbf{p}=(0,-2,6)+r(5,0,-2)$ of the second line, is then $x=5r$, $y=-2$, $z=6-2r$. At the point of intersection of the plane S and that second line, we have $2-4s=5r$, $1+t=-2$, $5+3s=6-2r$, for which the solution is $s=1/7$, $t=-3$, and $r=2/7$. Thus $(10/7,-2,38/7)$ is the point of intersection of S and the second line. Therefore the equation of the normal transverse L is $\mathbf{p}=(10/7,-2,38/7)+t(0,1,0)$. It is a line that intersects both the given lines and is orthogonal to both of them.

2.1.1. First row-reduce the corresponding augmented matrix to echelon form:

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

Next apply back substitution to solve the corresponding system. The last row corresponds to the equation $-\frac{7}{2}x_3 = 1$, so $x_3 = -\frac{2}{7}$. The second row corresponds to the equation $4x_2 + \frac{7}{2}x_3 = 1$ and thus $x_2 = -\frac{7}{8}x_3 + \frac{1}{4} = \frac{1}{2}$. Finally, the first row corresponds to $2x_1 + 2x_2 - 3x_3 = 0$, from which we obtain $x_1 = -x_2 + \frac{3}{2}x_3 = -\frac{13}{14}$. Hence, in vector form the solution is

$$\mathbf{x} = \begin{bmatrix} -13/14 \\ 1/2 \\ -2/7 \end{bmatrix}.$$

2.1.3. First row-reduce the corresponding augmented matrix to echelon form:

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

Next apply back substitution to solve the corresponding system. The second row corresponds to the equation $4x_2 + \frac{7}{2}x_3 = 1$. Here x_3 is free; set $x_3 = t$. Then we obtain $4x_2 + \frac{7}{2}t = 1$; hence $x_2 = \frac{1}{4} - \frac{7}{8}t$. Finally, the first row corresponds to $2x_1 + 2(\frac{1}{4} - \frac{7}{8}t) - 3t = 0$, from which we obtain $x_1 = -\frac{1}{4} + \frac{19}{8}t$. Hence, in vector form the solution is

$$\mathbf{x} = t\begin{bmatrix} 19/8 \\ -7/8 \\ 1 \end{bmatrix} + \begin{bmatrix} -1/4 \\ 1/4 \\ 0 \end{bmatrix}.$$

2.1.5. First row-reduce to echelon form:

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

Next apply back substitution to solve the corresponding system. The last row just gives the consistent trivial equation $0 = 0$. The second row corresponds to the equation $3x_2 - 3x_3 = 0$. Here x_3 is free; set $x_3 = t$. Then we obtain $3x_2 - 3t = 0$; hence $x_2 = t$. Finally, the first row corresponds to $x_1 - x_3 = 0$, from which we obtain $x_1 = t$. Thus the solution is $x_1 = x_2 = x_3 = t$ or in vector form,

$$\mathbf{x} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

2.1.7

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

The last row of the reduced matrix corresponds to the self-contradictory equation $0 = 2$, and consequently the system is inconsistent and has no solution.

2.1.9. First row-reduce to echelon form:

$$\left[\begin{array}{rrrr|r} 1 & 4 & 9 & 2 & 0 \\ 2 & 2 & 6 & -3 & 0 \\ 2 & 7 & 16 & 3 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 - 2\mathbf{r}_1 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - 2\mathbf{r}_1 \end{array} \left[\begin{array}{rrrr|r} 1 & 4 & 9 & 2 & 0 \\ 0 & -6 & -12 & -7 & 0 \\ 0 & -1 & -2 & -1 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_3 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_2 \end{array} \left[\begin{array}{rrrr|r} 1 & 4 & 9 & 2 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & -6 & -12 & -7 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - 6\mathbf{r}_2 \end{array} \left[\begin{array}{rrrr|r} 1 & 4 & 9 & 2 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right].$$

Next apply back substitution. The last row corresponds to the equation $-x_4 = 0$ and so $x_4 = 0$. The second row corresponds to the equation $-x_2 - 2x_3 - x_4 = 0$. Here x_3 is free; set $x_3 = t$. Then we obtain $x_2 = -2t$. Finally, the first row corresponds to $x_1 + 4x_2 + 9x_3 + 2x_4 = 0$, from which we determine $x_1 = -t$. In vector form the solution is

$$\mathbf{x} = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \end{bmatrix} t.$$

2.1.11. First row-reduce to echelon form:

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

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{rrrrr|r} -1 & 2 & 2 & 3 & 3 & 0 \\ 3 & -6 & -1 & 1 & 5 & 0 \\ 4 & -8 & -3 & -2 & 1 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 + 3\mathbf{r}_1 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 + 4\mathbf{r}_1 \end{array} \left[\begin{array}{rrrrr|r} -1 & 2 & 2 & 3 & 3 & 0 \\ 0 & 0 & 5 & 10 & 14 & 0 \\ 0 & 0 & 5 & 10 & 13 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - \mathbf{r}_2 \end{array} \left[\begin{array}{rrrrr|r} -1 & 2 & 2 & 3 & 3 & 0 \\ 0 & 0 & 5 & 10 & 14 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right].$$

Next apply back substitution. The last row corresponds to the equation $-x_5 = 0$ and so $x_5 = 0$. The second row corresponds to the equation $5x_3 + 10x_4 + 14x_5 = 0$. Here x_4 is free; set $x_4 = t$. Then we obtain $x_3 = -2t$. The variable x_2 is free; set $x_2 = s$. Finally, the first row corresponds to the equation $-x_1 + 2x_2 + 2x_3 + 3x_4 + 3x_5 = 0$, and so to $-x_1 + 2s + 2(-2t) + 3t + 3 \cdot 0 = 0$. Hence $x_1 = 2s - t$. In vector form the solution is

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

2.1.13.

$$\left[\begin{array}{rrrr|r} 3 & -6 & -1 & 1 & 5 \\ -1 & 2 & 2 & 3 & 3 \\ 6 & -8 & -3 & -2 & 1 \end{array}\right] \mathbf{r}_1 \leftrightarrow \mathbf{r}_2 \left[\begin{array}{rrrr|r} -1 & 2 & 2 & 3 & 3 \\ 3 & -6 & -1 & 1 & 5 \\ 6 & -8 & -3 & -2 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrrr|r} -1 & 2 & 2 & 3 & 3 \\ 0 & 0 & 5 & 10 & 14 \\ 0 & 4 & 9 & 16 & 19 \end{array}\right] \mathbf{r}_2 \leftrightarrow \mathbf{r}_3 \left[\begin{array}{rrrr|r} -1 & 2 & 2 & 3 & 3 \\ 0 & 4 & 9 & 16 & 19 \\ 0 & 0 & 5 & 10 & 14 \end{array}\right].$$

The last matrix is in echelon form and the forward elimination is finished. The fourth column has no pivot and so x_4 is free and we set $x_4 = t$. Then the last row corresponds to the equation $5x_3 + 10t = 14$, which gives $x_3 = \frac{14}{5} - 2t$. The second row yields $4x_2 + 9\left(\frac{14}{5} - 2t\right) + 16t = 19$, whence $x_2 = \frac{1}{2}t - \frac{31}{20}$. Finally, the first row gives $x_1 = -3 + 2\left(\frac{1}{2}t - \frac{31}{20}\right) + 2\left(\frac{14}{5} - 2t\right) + 3t = -\frac{1}{2}$. In vector form the solution is

$$\mathbf{x} = \begin{bmatrix} -1/2 \\ -31/20 \\ 14/5 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1/2 \\ -2 \\ 1 \end{bmatrix} t.$$

$\mathbf{2.1.15}$. The line of intersection of the first two planes is obtained from the reduction

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 3 & 6 & -1 & 8 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 0 & -1 & 2 \end{array}\right].$$

The last matrix represents the same planes as Equations (2.25) in the text.

The line of intersection of the last two planes is obtained from the reduction

$$\left[\begin{array}{ccc|c} 3 & 6 & -1 & 8 \\ 1 & 2 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 3 & 6 & -1 & 8 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 0 & -4 & 8 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 0 & -1 & 2 \end{array}\right] \to$$
$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 0 & -1 & 2 \end{array}\right],$$

In the last step, we added the second row to the first row. Again, the last matrix represents the same planes as Equations (2.25) in the text.

The line of intersection of the first and last planes is obtained from the reduction

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

the same planes as before.

2.2.**1**.

$$\begin{bmatrix} p_1 & * \\ 0 & p_2 \end{bmatrix}, \begin{bmatrix} p_1 & * \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & p_1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2.2.**3**. Following the given hint, reduce the augmented matrix $[A|\mathbf{b}]$ to echelon form:

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & b_1 \\ -2 & 3 & -1 & b_2 \\ -6 & 6 & 0 & b_3 \end{array}\right] \begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 + 2\mathbf{r}_1 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 + 6\mathbf{r}_1 \end{array} \left[\begin{array}{ccc|l} 1 & 0 & -1 & b_1 \\ 0 & 3 & -3 & b_2 + 2b_1 \\ 0 & 6 & -6 & b_3 + 6b_1 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - 2\mathbf{r}_2 \end{array} \left[\begin{array}{ccc|l} 1 & 0 & -1 & b_1 \\ 0 & 3 & -3 & b_2 + 2b_1 \\ 0 & 0 & 0 & b_3 + 6b_1 - 2\,(b_2 + 2b_1) \end{array}\right].$$

Hence the condition for consistency is $b_3 + 6b_1 - 2(b_2 + 2b_1) = 0$, or equivalently, $2b_1 - 2b_2 + b_3 = 0$.

2.2.**5**. Reduce the augmented matrix $[A|\mathbf{b}]$ to echelon form:

$$\left[\begin{array}{ccc|c} 1 & 2 & -6 & b_1 \\ -2 & -4 & 12 & b_2 \end{array}\right] \begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 + 2\mathbf{r}_1 \end{array} \left[\begin{array}{ccc|l} 1 & 2 & -6 & b_1 \\ 0 & 0 & 0 & b_2 + 2b_1 \end{array}\right].$$

Hence the condition for consistency is $b_2 + 2b_1 = 0$.

2.2.**7**. The rank of a matrix is the number r of non-zero rows in an echelon matrix obtained by the forward elimination phase of the Gaussian elimination algorithm. The same row operations that reduce the matrix A to an echelon matrix U will also row-reduce the augmented matrix $[A|\mathbf{b}]$ to an augmented matrix $[U|\mathbf{c}]$, with the same echelon matrix U. Thus the system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $c_{r+1} = c_{r+2} = \cdots = 0$; that is, if and only if all rows in $[U|\mathbf{c}]$ below the $r\mathrm{th}$ row are zero rows. In this case $[U|\mathbf{c}]$ is an echelon matrix too, and the number of non-zero rows in it is the rank of $[A|\mathbf{b}]$. Hence the system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if the rank of $[A|\mathbf{b}]$ equals the number r of non-zero rows in the matrix U, which is, by definition, the rank of A.

2.3.**1**..

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

2.3.**3**. To apply the method of Gauss-Jordan elimination for solving the given system of equations, continue row-reducing the echelon matrix obtained in the solution of Exercise 2.1.5:

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2/3 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Next, apply back substitution to the last matrix, which is in row-reduced echelon form. The third row corresponds to the equation $0 = 0$, and so the system is consistent. The second row corresponds to the equation $x_2 - x_3 = 0$; x_3 is free, so set $x_3 = t$, and thus also $x_2 = t$. Finally, the first row corresponds to the equation $x_1 - x_3 = 0$, and hence also $x_1 = t$.

2.3.5. To apply the method of Gauss-Jordan elimination for solving the given system of equations, continue row-reducing the echelon matrix obtained in the solution of Exercise 2.1.11:

$$\left[\begin{array}{ccccc|c} -1 & 2 & 2 & 3 & 3 & 0 \\ 0 & 0 & 5 & 10 & 14 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow -\mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2/5 \\ \mathbf{r}_3 \leftarrow -\mathbf{r}_3 \end{array} \left[\begin{array}{ccccc|c} 1 & -2 & -2 & -3 & -3 & 0 \\ 0 & 0 & 1 & 2 & 14/5 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 + 3\mathbf{r}_3 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 - (14/5)\,\mathbf{r}_3 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{ccccc|c} 1 & -2 & -2 & -3 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 + 2\mathbf{r}_2 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{ccccc|c} 1 & -2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right].$$

The pivots correspond to the variables $x_1,\ x_3$ and x_5; the variables x_2 and x_4 are free: set $x_2 = s$ and $x_4 = t$. Then, by back substitution, we obtain $x_5 = 0$ from the third row, $x_3 = -2t$ from the second, and $x_1 = 2s - t$ from the first.

2.3.7. First represent the system as an augmented matrix and row-reduce it to echelon form:

$$\left[\begin{array}{rrr|r} 2 & 3 & -1 & 4 \\ 3 & 5 & 2 & 1 \end{array}\right] \begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 - (3/2)\,\mathbf{r}_1 \end{array} \left[\begin{array}{rrr|r} 2 & 3 & -1 & 4 \\ 0 & 1/2 & 7/2 & -5 \end{array}\right].$$

The second row corresponds to the equation $\frac{1}{2}x_2 + \frac{7}{2}x_3 = -5$, and the first row to the equation $2x_1 + 3x_2 - x_3 = 4$. We can get a particular solution by setting the free variable $x_3 = 0$. Then $\frac{1}{2}x_2 = -5$, and so $x_2 = -10$, and $2x_1 - 30 = 4$, and thus $x_1 = 17$. Hence $\mathbf{x}_b = \begin{bmatrix} 17 \\ -10 \\ 0 \end{bmatrix}$ is a particular solution of $A\mathbf{x} = \mathbf{b}$. Similarly, setting $x_3 = 1$ gives $\frac{1}{2}x_2 + \frac{7}{2} = -5$, or $x_2 = -17$, and $2x_1 - 51 - 1 = 4$, or $x_1 = 28$. Thus $\mathbf{x}_b' = \begin{bmatrix} 28 \\ -17 \\ 1 \end{bmatrix}$ is another particular solution of $A\mathbf{x} = \mathbf{b}$.

For the homogeneous equation $A\mathbf{x} = \mathbf{0}$ we can use the same echelon matrix as above, except that the entries of its last column must be replaced by zeros. The general solution is obtained by setting $x_3 = t$, and solving for the other variables as $\frac{1}{2}x_2 + \frac{7}{2} = 0,\ x_2 = -7t,\ 2x_1 - 21t - t = 0,\ x_1 = 11t$. Thus the general solution of $A\mathbf{x} = \mathbf{0}$ is $\mathbf{v} = t\begin{bmatrix} 11 \\ -7 \\ 1 \end{bmatrix}$.

Hence the general solution of the inhomogeneous equation $A\mathbf{x} = \mathbf{b}$ can be written either as

$$\mathbf{x} = \begin{bmatrix} 17 \\ -10 \\ 0 \end{bmatrix} + t\begin{bmatrix} 11 \\ -7 \\ 1 \end{bmatrix} \text{ or as } \mathbf{x} = \begin{bmatrix} 28 \\ -17 \\ 1 \end{bmatrix} + s\begin{bmatrix} 11 \\ -7 \\ 1 \end{bmatrix}$$

The two equations represent the same line, as can be seen by setting $s = t - 1$.

2.4.**1** .

$$C = 2A + 3B = \begin{bmatrix} 13 & -6 \\ 8 & 2 \end{bmatrix} \text{ and } D = 4A - 3B = \begin{bmatrix} -1 & 24 \\ -2 & -14 \end{bmatrix}.$$

2.4.**3**.

$$AB = 2 \text{ and } BA = \begin{bmatrix} 3 & -6 & 9 \\ 2 & -4 & 6 \\ 1 & -2 & 3 \end{bmatrix}.$$

2.4.**5**.

$AB = \begin{bmatrix} 17 & -17 \\ 2 & -17 \\ 3 & -26 \end{bmatrix}$, while BA does not exist since B is 3×2 and A is 3×3.

2.4.**7**.

$AB = [(3-4+3+8),\ (-4-4-9-20)] = [10,\ -37]$, and BA is undefined.

2.4.**9**.

$AB = [-1\ \ -8],\ (AB)C = [-1,\ -8]\begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix} = [-25,\ 3]$, and $BC = \begin{bmatrix} -9 & -9 \\ 8 & -6 \end{bmatrix},\ A(BC) = [1,\ -2]\begin{bmatrix} -9 & -9 \\ 8 & -6 \end{bmatrix} = [-25,\ 3]$.

2.4.**11**. There are many possible answers; for example,

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

2.4.**13**. a. Each of the terms in the sum $\mathbf{a}_1\mathbf{b}^1+\mathbf{a}_2\mathbf{b}^2+\cdots+\mathbf{a}_P\mathbf{b}^P$ is the outer product of an $m \times 1$ column vector and a $1 \times n$ row vector and is therefore of the same mxn size as AB. Now the ith element of any column $\mathbf{a}_k$ is the element in the ith row kth column of A, that is, a_{ik}. Similarly, the jth element of the row $\mathbf{b}^k$ is b_{kj}. Thus $(\mathbf{a}_k\mathbf{b}^k)_{ij} = a_{ik}b_{kj}$. Summing over k from 1 to p, we get

$$\left(\sum_{k=1}^{p} \mathbf{a}_k\mathbf{b}^k\right)_{ij} = \sum_{k=1}^{p} (\mathbf{a}_k\mathbf{b}^k)_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} = (AB)_{ij}$$

and so

$$AB = \sum_{k=1}^{p} \mathbf{a}_k\mathbf{b}^k.$$

b. Let us expand the jth sum on the right of the equation given in part (b) of the exercise:

$$\sum_{i=1}^{p} \mathbf{a}_i b_{ij} = \sum_{i=1}^{p} \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} b_{ij} = \begin{bmatrix} \sum_{i=1}^{p} a_{1i} b_{ij} \\ \sum_{i=1}^{p} a_{2i} b_{ij} \\ \vdots \\ \sum_{i=1}^{p} a_{mi} b_{ij} \end{bmatrix}.$$

The column vector on the right is exactly the $j\text{th}$ column of AB, and so AB is the row of such columns for $j = 1, \ldots, n$.
c. Let us expand the $i\text{th}$ sum on the right of the equation given in part (c) of the exercise:

$$\sum_{j=1}^{p} a_{ij} \mathbf{b}^j = \sum_{j=1}^{p} a_{ij} (b_{j1}\ b_{j2}\ \ldots\ b_{jn}) = \left(\sum_{j=1}^{p} a_{ij} b_{j1}\ \sum_{j=1}^{p} a_{ij} b_{j2}\ \ldots\ \sum_{j=1}^{p} a_{ij} b_{jn} \right).$$

The row vector on the right is exactly the $i\text{th}$ row of AB, and so AB is the column of such rows for $i = 1, \ldots, m$.

2.4.**15**. There are many possible answers; for example,
$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ or $A = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$.

2.4.**17**. $(M^3)_{24} = (M \cdot M^2)_{24}$ can be computed by summing the products of the elements of the second row of M with the corresponding elements in the fourth column of M^2. Thus, from Equations (2.104) and (2.105), we obtain $(M^3)_{24} = 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 1 + 0 \cdot 2 + 0 \cdot 0 = 0$. Hence there are no three-leg flights between B and D.

2.4.**19**. On the one hand, using the blocks:

$$\begin{aligned} & \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix} \\ = & \begin{bmatrix} 1 \cdot 0 - 2 \cdot 0 & 1 \cdot 0 - 2 \cdot 0 \\ 3 \cdot 0 + 4 \cdot 0 & 3 \cdot 0 + 4 \cdot 0 \end{bmatrix} + \begin{bmatrix} 1 \cdot 3 + 0 \cdot 1 & 1 \cdot 2 + 0 \cdot (-1) \\ 0 \cdot 3 + 1 \cdot 1 & 0 \cdot 2 + 1 \cdot (-1) \end{bmatrix} \\ = & \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}, \end{aligned}$$

and on the other hand, without the blocks:

$$
\begin{bmatrix} 1 & -2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 2 \\ 1 & -1 \end{bmatrix}
$$

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

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

2.4.**21**. Expand the given matrix using Corollary 2.4.3 and then apply the result of Exercise 2.4.20.

2.4.**23**. In order that all relevant matrices be conformable, the second matrix should be partitioned as

$$
\left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 2 & 0 & -3 & 1 \\ \hline 0 & 0 & 2 & 3 \\ 0 & 0 & 7 & 4 \end{array}\right].
$$

Then, applying the result of Exercise 2.4.21, we obtain:

$$
\left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 \end{array}\right]
\left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 2 & 0 & -3 & 1 \\ \hline 0 & 0 & 2 & 3 \\ 0 & 0 & 7 & 4 \end{array}\right]
= \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix},
$$

where

$$
X_{11} = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -3 & -2 \\ 11 & -6 \\ -1 & 2 \end{bmatrix},
$$

$$
X_{12} = \begin{bmatrix} 1 & -2 \\ 3 & 4 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 7 & 4 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ -2 & 8 \\ -1 & 0 \end{bmatrix},
$$

$$X_{21} = \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 0 \end{bmatrix},$$

$$X_{22} = \begin{bmatrix} 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 7 & 4 \end{bmatrix} = \begin{bmatrix} 3 & -1 \end{bmatrix}.$$

Thus

$$\left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 \end{array}\right] \left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 2 & 0 & -3 & 1 \\ \hline 0 & 0 & 2 & 3 \\ 0 & 0 & 7 & 4 \end{array}\right]$$
$$= \left[\begin{array}{cc|cc} -3 & -2 & 9 & 1 \\ 11 & -6 & -2 & 8 \\ -1 & 2 & -1 & 0 \\ \hline -2 & 0 & 3 & -1 \end{array}\right].$$

2.5.**1**. To find A^{-1}, form the augmented matrix $[A|I]$ and row-reduce it to the form $[I|A^{-1}]$, if possible.

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

$$\to \left[\begin{array}{cc|cc} 2 & 0 & 1/7 & 3/7 \\ 0 & 1 & 2/7 & -1/7 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 1/14 & 3/14 \\ 0 & 1 & 2/7 & -1/7 \end{array}\right].$$

Hence

$$A^{-1} = \frac{1}{14} \begin{bmatrix} 1 & 3 \\ 4 & -2 \end{bmatrix}.$$

2.5.**3**. To find A^{-1}, form the augmented matrix $[A|I]$ and row-reduce it to the form $[I|A^{-1}]$, if possible.

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

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1/2 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{ccc|ccc} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 4 & -1 & 1 & 0 & 1 & 0 \\ 3 & 2 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 - 4\mathbf{r}_1 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - 3\mathbf{r}_1 \end{array} \left[\begin{array}{ccc|ccc} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 0 & -7 & -9 & -2 & 1 & 0 \\ 0 & -5/2 & -19/2 & -3/2 & 0 & 1 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow -\mathbf{r}_2/7 \\ \mathbf{r}_3 \leftarrow -2\mathbf{r}_3/5 \end{array} \left[\begin{array}{ccc|ccc} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 0 & 1 & 9/7 & 2/7 & -1/7 & 0 \\ 0 & 1 & 19/5 & 3/5 & 0 & -2/5 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 - 3\mathbf{r}_2/2 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - \mathbf{r}_2 \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 4/7 & 1/14 & 3/14 & 0 \\ 0 & 1 & 9/7 & 2/7 & -1/7 & 0 \\ 0 & 1 & 88/35 & 11/35 & 1/7 & -2/5 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \rightarrow 35\mathbf{r}_3/88 \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 4/7 & 1/14 & 3/14 & 0 \\ 0 & 1 & 9/7 & 2/7 & -1/7 & 0 \\ 0 & 0 & 1 & 1/8 & 5/88 & -7/44 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 - 4\mathbf{r}_3/7 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 - 9\mathbf{r}_3/7 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 2/11 & 1/11 \\ 0 & 1 & 0 & 1/8 & -19/88 & 9/44 \\ 0 & 0 & 1 & 1/8 & 5/88 & -7/44 \end{array}\right].$$

Hence

$$A^{-1} = \frac{1}{88}\begin{bmatrix} 0 & 16 & 8 \\ 11 & -19 & 18 \\ 11 & 5 & -14 \end{bmatrix}.$$

2.5.**5**. To find A^{-1}, form the augmented matrix $[A|I]$ and row-reduce it to the form $[I|A^{-1}]$, if possible.

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

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1/2 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 \end{array} \left[\begin{array}{rrr|rrr} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 4 & -1 & 3 & 0 & 1 & 0 \\ 3 & 2 & 5 & 0 & 0 & 1 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 - 4\mathbf{r}_1 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - 3\mathbf{r}_1 \end{array} \left[\begin{array}{rrr|rrr} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 0 & -7 & -7 & -2 & 1 & 0 \\ 0 & -5/2 & -5/2 & -3/2 & 0 & 1 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow -\mathbf{r}_2/7 \\ \mathbf{r}_3 \leftarrow -2\mathbf{r}_3/5 \end{array} \left[\begin{array}{rrr|rrr} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 0 & 1 & 1 & 2/7 & -1/7 & 0 \\ 0 & 1 & 1 & 3/5 & 0 & -2/5 \end{array}\right]$$

$$\begin{array}{l} \mathbf{r}_1 \leftarrow \mathbf{r}_1 \\ \mathbf{r}_2 \leftarrow \mathbf{r}_2 \\ \mathbf{r}_3 \leftarrow \mathbf{r}_3 - \mathbf{r}_2 \end{array} \left[\begin{array}{rrr|rrr} 1 & 3/2 & 5/2 & 1/2 & 0 & 0 \\ 0 & 1 & 1 & 2/7 & -1/7 & 0 \\ 0 & 0 & 0 & 11/35 & 1/7 & -2/5 \end{array}\right]$$

Since the last row corresponds to inconsistent equations, the matrix A cannot be reduced to the identity matrix I, and therefore A is not invertible.

2.5.**7**. For example, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ are invertible, but $A + B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ is not.

2.5.**9**. a. The matrix equation $XA = I$ can be spelled out as

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 4 & -1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We can write the corresponding system in augmented matrix form, and reduce it as follows:

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

(Notice the appearance of A^T in the augmented matrix. This is due to the fact that the equation $XA = I$ is equivalent to $A^T X^T = I$, and it is the latter equation with the X^T on the right in the product that is directly translated to the augmented matrix form.)

The unknowns x_{13} and x_{23} are free. Choosing $x_{13} = s$ and $x_{23} = t$, we get the systems of equations

$$\begin{aligned} 2x_{11} + \ 4x_{12} + \ 2s \ &= 1 \\ x_{12} + \ 3s \ &= 1/2 \end{aligned} \quad \text{and} \quad \begin{aligned} 2x_{21} + \ 4x_{22} + \ 2t \ &= 0 \\ x_{22} + \ 3t \ &= 1, \end{aligned}$$

from which $x_{12} = \frac{1}{2} - 3s$, $x_{11} = -\frac{1}{2} + 5s,\ x_{22} = 1 - 3t$, and $x_{21} = -2 + 5t$. Thus

$$X = \begin{bmatrix} -\frac{1}{2} + 5s & \frac{1}{2} - 3s & s \\ -2 + 5t & 1 - 3t & t \end{bmatrix}$$

is a solution for any $s,\ t$, and every left inverse of A must be of this form.

b. $AY = I$ can be written as

$$\begin{bmatrix} 2 & -1 \\ 4 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence the entries of the first column of Y must satisfy

$$\begin{aligned} 2y_{11} - \quad y_{12} \ &= 1 \\ 4y_{11} - \quad y_{21} \ &= 0 \\ 2y_{11} + \ 2y_{21} \ &= 0. \end{aligned}$$

From the last two equations we get $y_{11} = y_{21} = 0$ and, substituting these values into the first equation above, we obtain the contradiction $0 = 1$. Thus there is no solution matrix Y.

2.5.**11**. For the given square matrix A, we know that $AX = I$ and $YA = I$. If we multiply the first of these equations by Y from the left, we get $Y(AX) = YI$, which is equivalent to $(YA)X = Y$. Substituting $YA = I$ into the latter equation, we obtain $X = Y$.

2.5.**13**. Writing $\mathbf{e}^i$ for the ith row of I, we must have

$$P = PI = P \begin{bmatrix} \mathbf{e}^1 \\ \mathbf{e}^2 \\ \mathbf{e}^3 \end{bmatrix} = \begin{bmatrix} c\mathbf{e}^1 \\ \mathbf{e}^2 \\ \mathbf{e}^3 \end{bmatrix} = \begin{bmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

But then P also produces the desired multiplication of the first row by c for any $3 \times n$ matrix A:

$$PA = \begin{bmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \mathbf{a}^3 \end{bmatrix} = \begin{bmatrix} c\mathbf{a}^1 \\ \mathbf{a}^2 \\ \mathbf{a}^3 \end{bmatrix}.$$

2.5.**15**. Since $AA^{-1} = I$, and since $c \cdot \frac{1}{c} = 1$, a reasonable guess is $(cA)^{-1} = \frac{1}{c}A^{-1}$. To prove this, we need only show that $(cA)(\frac{1}{c}A^{-1}) = I$. If we denote the ikth element of A by a_{ik} and the kjth element of A^{-1} by b_{kj}, then the ikth element of cA *is* ca_{ik}, and the kjth element of $\frac{1}{c}A^{-1}$ *is* $\frac{1}{c}b_{kj}$, and hence the ijth element of $(cA)(\frac{1}{c}A^{-1})$ is

$$\sum_{k=1}^{m}(ca_{ik})(\frac{1}{c}b_{kj}) = \sum_{k=1}^{m} a_{ik}b_{kj} = (AA^{-1})_{ij} = I_{ij},$$

which is what we wanted to show.

2.5.**17**.

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, P_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, P_5 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, P_6 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$P_i^{-1} = P_i$ for $i = 1, 2, 3, 4$, and $P_5^{-1} = P_6$ and $P_6^{-1} = P_5$.

Furthermore, for instance

$$P_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \mathbf{a}^3 \end{bmatrix} = \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^3 \\ \mathbf{a}^2 \end{bmatrix},$$

which shows that P_2 permutes the second row of A with the third one. Since $P_2I = P_2$, the second and third rows of P_2 are those of I permuted. Analogous results hold for each P_i.

If we apply any P_i to an $n \times 3$ matrix B from the right, then it permutes the *columns* $\mathbf{b}_i$ of B the same way as P_i has the columns of I permuted. For instance,

$$BP_2 = [\mathbf{b}_1 \;\; \mathbf{b}_2 \;\; \mathbf{b}_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = [\mathbf{b}_1 \;\; \mathbf{b}_3 \;\; \mathbf{b}_2].$$

2.5.**19**. The matrix $(A^{-1})^{-1}$ is defined as the solution X of the equation $A^{-1}X = I$. Now, multiplying both sides by A and simplifying, we get $A\left(A^{-1}X\right) = \left(AA^{-1}\right)X = IX = X$ on the left, and $AI = A$ on the right. Thus, $X = A$. Also, earlier we had $X = (A^{-1})^{-1}$. Hence $(A^{-1})^{-1} = A$.

3.1.**1**. The set of all polynomials of degree two and the zero polynomial is not a vector space. For example, it is not closed under addition: the sum of $x^2 + x - 1$ and $-x^2$ is $x - 1$, which is neither a polynomial of degree two nor the zero polynomial.

3.1.**3**. The set of all solutions (x, y) of the equation $2x + 3y = 1$ is not a vector space. For example, $(-1, 1)$ and $(-4, 3)$ are solutions, but their sum $(-1, 1) + (-4, 3) = (-5, 4)$ is not a solution since $2(-5) + 3(4) = 2 \neq 1$. Thus the set is not closed under addition.

3.1.**5**. The set of all twice differentiable functions f for which $f''(x) + 2f(x) = 1$ holds is not a vector space. For example, $f(x) = 1/2$ is a solution, but $f(x) = 2\,(1/2) = 1$ is not a solution. Thus the set is not closed under multiplication by all scalars.

3.1.**7**. This set is not a vector space, because addition is not commutative: For example, let $(p_1, p_2) = (1, 2)$ and $(q_1, q_2) = (1, 3)$. Then, by the given addition rule, we have $(1, 2) + (1, 3) = (1 + 3, 2 + 1) = (4, 3)$, but $(1, 3) + (1, 2) = (1 + 2, 3 + 1) = (3, 4)$.

3.1.**9**. This set is not a vector space, because Axiom 7 fails: Let $a = b = 1$, and $p_2 \neq 0$. Then $(a + b)\mathbf{p} = (a + b)(p_1, p_2) = 2(p_1, p_2) = (2p_1, 2p_2)$, but $a\mathbf{p} + b\mathbf{p} = (p_1, p_2) + (p_1, p_2) = (2p_1, 0)$.

3.1.**11**. Prove the last three parts of Theorem 3.1.1.

Proof of Part 6:

For any real number c we have $-c = (-1)c$, and so

$(-c)\mathbf{p} = ((-1)c)\mathbf{p}$

$= (-1)(c\mathbf{p})$ by Axiom 6 of Definition 3.1.1, and

$= -(c\mathbf{p})$ by Part 5 of Theorem 3.1.1.

Also, $((-1)c)\mathbf{p} = (c(-1))\mathbf{p}$ by the commutativity of multiplication of numbers,

$= c((-1)\mathbf{p})$ by Axiom 6 of Definition 3.1.1, and

$= c(-\mathbf{p})$ by Part 5 of Theorem 3.1.1.

Proof of Part 7:

$c(\mathbf{p} - \mathbf{q}) = c[\mathbf{p} + (-\mathbf{q})]$ by Definition 3.1.3,

$= c\mathbf{p} + c(-\mathbf{q})$ by Axiom 8 of Definition 3.1.1,

$= c\mathbf{p} + (-(c\mathbf{q}))$ by Part 6 above, and

$= c\mathbf{p} - c\mathbf{q}$ by Definition 3.1.3.

Proof of Part 8:

$(c - d)\mathbf{p} = [c + (-d)]\mathbf{p}$ by the definition of subtraction of numbers,

$= c\mathbf{p} + (-d)\mathbf{p}$ by Axiom 7 of Definition 3.1.1,

$= c\mathbf{p} + (-(d\mathbf{p}))$ by Part 6 above, and

$= c\mathbf{p} - d\mathbf{p}$ by Definition 3.1.3.

3.1.13. If $\mathbf{p} + \mathbf{q} = \mathbf{p} + \mathbf{r}$, then adding $-\mathbf{p}$ to both sides we get $-\mathbf{p} + (\mathbf{p} + \mathbf{q}) = -\mathbf{p} + (\mathbf{p} + \mathbf{r})$. By applying the associative rule for addition, we may change this to $(-\mathbf{p} + \mathbf{p}) + \mathbf{q} = (-\mathbf{p} + \mathbf{p}) + \mathbf{r}$. By Axioms 1 and 4 of Definition 3.1.1 we can replace the $-\mathbf{p} + \mathbf{p}$ terms by $\mathbf{0}$, and then Axioms 1 and 3 give $\mathbf{q} = \mathbf{r}$.

3.2.1. This U is a subspace of $\mathbb{R}^3$:

a. $\mathbf{0} \in U$, and so U is nonempty.

b. U is closed under addition: Let $\mathbf{u}, \mathbf{v} \in U$. Then $u_1 = u_2 = u_3$ and $v_1 = v_2 = v_3$. Hence $u_1 + v_1 = u_2 + v_2 = u_3 + v_3$, which shows that $\mathbf{u} + \mathbf{v} \in U$.

c. U is closed under multiplication by scalars: If $\mathbf{u} \in U$ and $c \in \mathbb{R}$, then $cu_1 = cu_2 = cu_3$, which shows that $c\mathbf{u} \in U$.

3.2.3. This U is not a subspace of $\mathbb{R}^3$:

Take for instance $\mathbf{u} = [1, -1, 1]^T$ and $\mathbf{v} = [1, 1, 1]^T$. Then $\mathbf{u}, \mathbf{v} \in U$, but $\mathbf{u} + \mathbf{v} = [2, 0, 2]^T \notin U$, because $|2| \neq |0|$. Thus U is not closed under addition.

$\mathbf{3.2.5}$. This U is not a subspace of $\mathbb{R}^3$:

Take for instance $\mathbf{u} = [1,1,1]^T$ and $\mathbf{v} = [1,2,0]^T$. Then $\mathbf{u}, \mathbf{v} \in U$, but $\mathbf{u} + \mathbf{v} = [2,3,1]^T \notin U$, because $2 \neq 3$ and neither is the third component of the sum 0. Thus U is not closed under addition.

$\mathbf{3.2.7}$. This U is not a subspace of $\mathbb{R}^n$:

Take for instance $\mathbf{u} = \mathbf{e}_1 = [1, 0, \cdots, 0]^T$ and $c = -1$. Then $\mathbf{u} \in U$, but $c\mathbf{u} = -\mathbf{e}_1 = [-1, 0, \cdots, 0]^T \notin U$. Thus U is not closed under multiplication by scalars.

$\mathbf{3.2.9}$. If U and V are subspaces of a vector space X, then $U \cap V$ is a subspace of X :

a. $\mathbf{0} \in U \cap V$, and so $U \cap V$ is nonempty.

b. $U \cap V$ is closed under addition: Let $\mathbf{u},\ \mathbf{v} \in U \cap V$. Then we have both $\mathbf{u}, \mathbf{v} \in U$ and $\mathbf{u}, \mathbf{v} \in V$. Since both U and V are vector spaces, they are closed under addition. that is, $\mathbf{u} + \mathbf{v} \in U$ and $\mathbf{u} + \mathbf{v} \in V$. Thus, by the definition of intersection, $\mathbf{u} + \mathbf{v} \in U \cap V$.

c. $U \cap V$ is closed under multiplication by scalars: If $\mathbf{u} \in U \cap V$ and $c \in \mathbb{R}$, then we have both $\mathbf{u} \in U$ and $\mathbf{u} \in V$. Since both U and V are vector spaces, they are closed under multiplication by scalars: that is, $c\mathbf{u} \in U$ and $c\mathbf{u} \in V$. Thus, by the definition of intersection, $c\mathbf{u} \in U \cap V$.

$\mathbf{3.2.11}$. No. $\mathbf{0} \notin \overline{U}$ and so $\overline{U}$ is not a vector space.

$\mathbf{3.3.1}$. Let us prove the "if" part first: So, assume that one of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ is a linear combination of the other two, say

$$\mathbf{a}_3 = s_1\mathbf{a}_1 + s_2\mathbf{a}_2$$

for some coefficients s_1 and s_2. (The proof would be similar if $\mathbf{a}_1$ or $\mathbf{a}_2$ were given as a linear combination of the other two vectors.) Then we can write equivalently

$$s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + (-1)\mathbf{a}_3 = \mathbf{0}.$$

This equation shows that there is a nontrivial linear combination of $\mathbf{a}_1,\ \mathbf{a}_2$ and $\mathbf{a}_3$ that equals the zero vector. But the existence of such a linear combination is precisely the definition of the dependence of the given vectors.

To prove the "only if" part, assume that the given vectors are dependent, that is, that there exist coefficients $s_1,\ s_2$ and s_3, not all zero, such that

$$s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + s_3\mathbf{a}_3 = \mathbf{0}.$$

Assume further, without any loss of generality, that $s_3 \neq 0$. Then we can solve the above equation for $\mathbf{a}_3$ as

$$\mathbf{a}_3 = \left(-\frac{s_1}{s_3}\right)\mathbf{a}_1 + \left(-\frac{s_2}{s_3}\right)\mathbf{a}_2.$$

This equation exhibits $\mathbf{a}_3$ as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$; just what we had to show.

3.3.3. $\mathbf{b} = (7, 32, 16, -3)^T$, $\mathbf{a}_1 = (4, 7, 2, 1)^T$, $\mathbf{a}_2 = (4, 0, -3, 2)^T$, $\mathbf{a}_3 = (1, 6, 3, -1)^T$.

We have to solve

$$\begin{bmatrix} 4 & 4 & 1 \\ 7 & 0 & 6 \\ 2 & -3 & 3 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 32 \\ 16 \\ -3 \end{bmatrix}.$$

We solve this system in augmented matrix form by row reduction as follows:

$$\left[\begin{array}{rrr|r} 4 & 4 & 1 & 7 \\ 7 & 0 & 6 & 32 \\ 2 & -3 & 3 & 16 \\ 1 & 2 & -1 & -3 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 7 & 0 & 6 & 32 \\ 2 & -3 & 3 & 16 \\ 4 & 4 & 1 & 7 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 0 & -14 & 13 & 53 \\ 0 & -7 & 5 & 22 \\ 0 & -4 & 5 & 19 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 0 & -7 & 5 & 22 \\ 0 & -14 & 13 & 53 \\ 0 & -4 & 5 & 19 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 0 & -7 & 5 & 22 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 15/7 & 45/7 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 0 & -7 & 5 & 22 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Then, by back substitution, we obtain $s_3 = 3$, $-7s_2 + 5 \cdot 3 = 22$, and so $s_2 = -1$, $s_1 - 2 - 3 = -3$, and so $s_1 = 2$. Hence $\mathbf{b} = 2\mathbf{a}_1 - \mathbf{a}_2 + 3\mathbf{a}_3$.

3.3.5. We have to solve

$$\begin{bmatrix} 4 & 4 & 0 \\ 2 & -3 & 5 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 16 \\ -3 \end{bmatrix}.$$

We solve this system in augmented matrix form by row reduction as:

$$\left[\begin{array}{rrr|r} 4 & 4 & 0 & 7 \\ 2 & -3 & 5 & 16 \\ 1 & 2 & -1 & -3 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 2 & -3 & 5 & 16 \\ 4 & 4 & 0 & 7 \end{array}\right]$$

$$\to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 0 & -7 & 7 & 22 \\ 0 & -4 & 4 & 19 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 2 & -1 & -3 \\ 0 & -7 & 3 & 22 \\ 0 & 0 & 0 & 45/7 \end{array}\right].$$

The last row of the reduced matrix yields the self-contradictory equation $0 = 45/7$, and so the vector $\mathbf{b}$ cannot be written as a linear combination of the given $\mathbf{a}_i$ vectors.

$\mathbf{3.3.7}$. These four vectors are not independent since, by the result of Exercise 3.3.4, the equation $s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + s_3\mathbf{a}_3 + s_4\mathbf{b} = \mathbf{0}$ has the nontrivial solution $s_1 = 2,\ s_2 = -1,\ s_3 = 3,\ s_4 = -1$.

$\mathbf{3.3.9}$. The three vectors $\mathbf{a}_1 = (4,2,1)^T,\ \mathbf{a}_2 = (4,-3,2)^T,\ \mathbf{a}_3 = (0,5,-1)^T$ from Exercise 3. 3. 5 are not independent: We have to solve

$$\begin{bmatrix} 4 & 4 & 0 \\ 2 & -3 & 5 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

By row-reduction:

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

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

Hence s_3 is free, and there are nontrivial solutions. Consequently the $\mathbf{a}_i$ vectors are dependent.

This result also follows from the equivalence of Parts 4 and 6 of Theorem 2.5.5, which implies that if $A\mathbf{x} = \mathbf{b}$ has no solution for some $\mathbf{b}$, as happens for the $\mathbf{b}$ of Exercise 3.3.5, then $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions.

3.3.11. To test the vectors $\mathbf{a}_1 = (1,0,0,1)^T$, $\mathbf{a}_2 = (0,0,1,1)^T$, $\mathbf{a}_3 = (1,1,0,0)^T$, $\mathbf{a}_4 = (1,0,1,1)^T$ for independence, we solve $A\mathbf{s} = \mathbf{0}$ by row reduction:

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 \end{array}\right]$$

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

Back substitution yields $\mathbf{s} = \mathbf{0}$, and so the columns of A are independent.

3.3.13.

a. The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ of any vector space X are linearly independent if and only if the equation $s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \cdots + s_n\mathbf{a}_n = \mathbf{0}$ has only the trivial solution $s = \mathbf{0}$.

b. The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ of any vector space X are linearly independent if and only if none of them can be written as a linear combination of the others.

c. The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ of any vector space X are linearly independent if and only if the solution of the equation $s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \cdots + s_n\mathbf{a}_n = \mathbf{b}$ is unique for any $\mathbf{b} \in X$ for which a solution exists.

3.3.15. Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ be any vectors in $\mathbb{R}^3$, with $n > 3$. Then, to test them for dependence, we solve $s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \cdots + s_n\mathbf{a}_n = \mathbf{0}$. We can do this by reducing the augmented matrix $[A|\mathbf{0}]$ to echelon form, where A is the $3 \times n$ matrix that has the given vectors for its columns. Since A is $3 \times n$, we cannot have more than three pivots. On the other hand, since there are n columns. with $n > 3$, there must exist at least one pivot-free column and a corresponding free variable. Thus there are nontrivial solutions, and consequently the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are dependent.

3.3.17. Let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ span $\mathbb{R}^3$. Let A be the 3×3 matrix that has the given vectors as columns. Then, by the definition of spanning, $A\mathbf{s} = \mathbf{b}$ has a solution for every vector $\mathbf{b}$ in $\mathbb{R}^3$. By the equivalence of Parts 4 and 6 of Theorem 2.5.5, this implies that the equation $A\mathbf{s} = \mathbf{0}$ has only the trivial solution. Thus $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are independent.

3.3.19. Let A be the $n \times m$ matrix with columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. If these vectors are independent, then, by the definition of linear independence, the equation $A\mathbf{s} = \mathbf{0}$ has only the trivial solution. Thus, if A is row-reduced to the echelon form U, then the corresponding equation $U\mathbf{s} = \mathbf{0}$ cannot have any free variables, that is, the number of rows n must be at least the number of columns m, and there cannot be any free columns. In other words, we must have $m \leq n$ and $r = m$. The definition of independence also requires $0 < m$.

Conversely, if $0 < m \leq n$ and $r = m$, then $U\mathbf{s} = \mathbf{0}$ has no free variables, and so $A\mathbf{s} = \mathbf{0}$ has only the trivial solution. Thus the columns of A are independent.

5

3.4.1. First we apply Theorems 3.3.1 and 3.3.3:

$$\begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & 4 \\ 2 & -1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 \\ 0 & -7 & 1 \\ 0 & -7 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 \\ 0 & -7 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus a basis for Col(A) is the set

$$\left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} \right\},$$

and a basis for Row(A) is

$$\left\{ \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -7 \\ 1 \end{bmatrix} \right\}.$$

To find a basis for Null(A) we need to solve $Ax = \mathbf{0}$. The row reduction of $[A|\mathbf{0}]$ would result in the same echelon matrix as above, just augmented with a zero column. Thus the variable x_3 is free. Set $x_3 = s$; then the second row of the reduced matrix gives $x_2 = s/7$, and the first row leads to $x_1 = -10s/7$. Hence

$$\mathbf{x} = \frac{s}{7}\begin{bmatrix} -10 \\ 1 \\ 7 \end{bmatrix}$$

is the general form of a vector in Null(A), and so the one-element set

$$\left\{\begin{bmatrix} -10 \\ 1 \\ 7 \end{bmatrix}\right\}$$

is a basis for Null(A).

3.4.3. As in Exercise 3.3.1 above,

$$\begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & 4 \\ 2 & -1 & 3 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 \\ 0 & -7 & 1 \\ 0 & -7 & 1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus a basis for Col(A) is the set

$$\left\{\begin{bmatrix} 1 \\ 3 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 3 \\ 1 \end{bmatrix}\right\},$$

and a basis for Row(A) is

$$\left\{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}\right\}.$$

For Null(A): $x_3 = 0,\ x_2 = 0,\ x_1 = 0$. Hence Null$(A) = \{\mathbf{0}\}$, which has no basis or, expressed differently, its basis is the empty set.

3.4.5. A simple example is given by

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

Clearly, the first two rows being the same, they span a one-dimensional space, which does not contain the third row. Thus the first two rows transposed do not form a basis for Row(A), since Row(A) must contain the transpose of every row. On the other hand, the transposed first two rows of

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

do form a basis for Row(A), since the third row of A can be obtained as the difference of the first two rows of E.

3.4.7. Since $\mathbf{a}$ is the third column of the given matrix A, it is in Col(A). It can be expressed as a linear combination of the first two columns $\mathbf{a}_1$ and $\mathbf{a}_2$ by Gaussian elimination as follows:

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

Hence $s_2 = -1/7,\ s_1 - 3/7 = 1$, and so s_1 /7. Thus $\mathbf{a} = (10\mathbf{a}_1 - \mathbf{a}_2)/7$.

For $\mathbf{b}$, similarly:

$$\left[\begin{array}{rr|r} 1 & 3 & -10 \\ 3 & 2 & 1 \\ 2 & -1 & 7 \end{array}\right] \to \left[\begin{array}{rr|r} 1 & 3 & -10 \\ 0 & -7 & 31 \\ 0 & -7 & 27 \end{array}\right] \to \left[\begin{array}{rr|r} 1 & 3 & -10 \\ 0 & -7 & 31 \\ 0 & 0 & -4 \end{array}\right].$$

The last row shows that this is the augmented matrix of an inconsistent system, and so $\mathbf{b}$ is not in Col(A).

Since $\mathbf{c} = -(\mathbf{a} + \mathbf{b})$ and $\mathbf{a}$ is in Col(A), but $\mathbf{b}$ is not, $\mathbf{c}$ cannot be in

Col(A) either. (If it were in Col(A), then $\mathbf{b} = -(\mathbf{a}+\mathbf{c})$ would have to be too, because Col(A), being a subspace, is closed under vector addition and under multiplication by scalars.)

For $\mathbf{d}$, row-reduce again:

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

Hence $s_2 = 6/7,\ s_1 + 18/7 = 5$, and $s_1 = 17/7$. Thus we can write $\mathbf{d}$ as $\mathbf{d} = (17\mathbf{a}_1 + 6\mathbf{a}_2)/7$.

$\mathbf{3.4.9}$. First we show that any vector $\mathbf{x} \in$ Null(B) is also in Null(AB). Now $\mathbf{x} \in$ Null(B) means that $\mathbf{x}$ is a solution of $B\mathbf{x} = \mathbf{0}$. Multiplying both sides by A, we get $AB\mathbf{x} = \mathbf{0}$. Thus $\mathbf{x}$ is also in Null(AB).

Conversely, if $\mathbf{x} \in$ Null(AB), then $AB\mathbf{x} = \mathbf{0}$, and multiplying both sides by A^{-1} (which exists by assumption), we get $A^{-1}AB\mathbf{x} = B\mathbf{x} = \mathbf{0}$. Thus $\mathbf{x}$ is also in Null(B).

Taken together, the two arguments above show that Null(B) and Null(AB) contain exactly the same vectors $\mathbf{x}$. This means that Null(B) = Null(AB).

$\mathbf{3.4.11}$. Let $A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ be the list of the given spanning vectors, and $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ the list of independent vectors that we use for comparison. By the Exchange Theorem, the number of independent vectors must not exceed the number of spanning vectors. More formally: the n here is the k of the Exchange Theorem, and the m here is the n there. Thus the conclusion $k \le n$ of the theorem becomes $n \le m$.

$\mathbf{3.4.13}$. Let A be the $n \times m$ matrix with columns $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$, with $m > n$. These vectors are dependent if and only if the equation $A\mathbf{s} = \mathbf{0}$ has nontrivial solutions. Now, if A is row-reduced to the echelon form U, then the corresponding equation $U\mathbf{s} = \mathbf{0}$ must have free variables, because the number of pivots cannot exceed the number n of rows, and the number of columns m is greater than n. So there are free columns and, correspondingly, nontrivial solutions.

For an alternative proof, assume that $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$, with $n < m$, is a list of independent vectors in $\mathbb{R}^n$. For comparison use the standard vectors as the spanning list, that is, let $A = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ in the Exchange

Theorem.. Then the conclusion $k \le n$ of the theorem becomes $m \le n$. This inequality contradicts the previous one, and so the vectors of B cannot be independent.

3.5.1. We reduce A to echelon form:

$$\begin{bmatrix} 1 & 1 & 1 & 2 & 0 \\ 3 & 3 & 0 & 6 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & -3 & 0 & 0 \end{bmatrix}.$$

Hence a basis for Row(A) is given by the vectors $\mathbf{b}_1 = (1,1,1,2,0)^T$ and $\mathbf{b}_2 = (0,0,-3,0,0)^T$. Thus $\dim(\text{Row}(A)) = 2$. Consequently, by Theorem 3.4.2, $\dim(\text{Col}(A)) = 2$, too, and, by Corollary 3.4.2 and Theorem 3.4.4, $\dim(\text{Null}(A)) = 3$, and $\dim(\text{Left-null}(A)) = 0$.

To find a basis for Null(A), solve $A\mathbf{x} = \mathbf{0}$ by setting $x_2 = s$, $x_4 = t$, and $x_5 = u$. Then $x_3 = 0$ and $x_1 = -s - 2t$. Hence

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

is the general form of a vector in Null(A), and so

$$\mathbf{c}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{c}_2 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{c}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for Null(A). Thus, to get the desired decomposition of $\mathbf{x} = (-2,0,1,4,1)^T$, we must solve $(-2,0,1,4,1)^T = s_1\mathbf{b}_1 + s_2\mathbf{b}_2 + t_1\mathbf{c}_1 + t_2\mathbf{c}_2 + t_3\mathbf{c}_3$. By row-reduction:

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

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & -2 & 0 & -2 \\ 0 & 0 & 2 & 2 & 0 & 2 \\ 0 & -3 & 1 & 2 & 0 & 3 \\ 0 & 0 & 2 & 5 & 0 & 8 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \rightarrow$$

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & -2 & 0 & -2 \\ 0 & -3 & 1 & 2 & 0 & 3 \\ 0 & 0 & 2 & 2 & 0 & 2 \\ 0 & 0 & 2 & 5 & 0 & 8 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \rightarrow$$

$$\left[\begin{array}{ccccc|c} 1 & 0 & -1 & -2 & 0 & -2 \\ 0 & -3 & 1 & 2 & 0 & 3 \\ 0 & 0 & 2 & 2 & 0 & 2 \\ 0 & 0 & 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

From here, by back substitution, we get $t_3 = 1,\ t_2 = 2,\ t_1 = -1$, $3s_2 = t_1 + 2t_2 - 3 = -1 + 4 - 3 = 0$, and $s_1 = t_1 + 2t_2 - 2 = 1$. Thus

$$\mathbf{x}_0 = -\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix},$$

and

$$\mathbf{x}_R = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}.$$

Indeed,

$$\mathbf{x}_0 + \mathbf{x}_R = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 4 \\ 1 \end{bmatrix}$$

and

$$\mathbf{x}_0 \cdot \mathbf{x}_R = \begin{bmatrix} -3 & -1 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = 0.$$

3.5.3.

a. $\dim(\text{Row}(A)) = \dim(\text{Col}(A)) = 1$, and, by Corollary 3.4.2 and Theorem 3.4.4, $\dim(\text{Null}(A)) = 4$, and $\dim(\text{Left-null}(A)) = 0$.

b. This A is an echelon matrix, and so $\mathbf{b}_1 = (3,3,0,4,4)^T$ forms a basis for its row space, and we must have $\mathbf{x}_R = s\mathbf{b}_1$ with an unknown value for s. Let $\mathbf{x} = (1,1,1,1,1)^T$. We use the shortcut mentioned at the end of Example 3.4.2 to solve $\mathbf{x} = s\mathbf{b}_1 + \mathbf{x}_0$ for s by left-multiplying it with $\mathbf{b}_1^T$. We obtain, because of the orthogonality of the row space to the nullspace, $\mathbf{b}_1^T\mathbf{x} = s\mathbf{b}_1^T\mathbf{b}_1$, which becomes $14 = 50s$. Thus $s = 7/25$ and $\mathbf{x}_R = \frac{7}{25}(3,3,0,4,4)^T$. From here we get $\mathbf{x}_0 = \mathbf{x} - \mathbf{x}_R = \frac{1}{25}(4,4,25,-3,-3)^T$.

3.5.5. This matrix needs no reduction: it is already in echelon form. Thus $\mathbf{b}_1 = (0,2,0,0,4)^T$ and $\mathbf{b}_2 = (0,0,0,2,2)^T$ form a basis for $\text{Row}(A)$. Hence $\dim(\text{Row}(A)) = 2$. Consequently, by Theorem 3.4.2, $\dim(\text{Col}(A)) = 2$, too, and, by Corollary 3.4.2 and Theorem 3.4.4, $\dim(\text{Null}(A)) = 3$, and $\dim(\text{Left-null}(A)) = 0$.

To find a basis for $\text{Null}(A)$ we solve $A\mathbf{x} = \mathbf{0}$ by setting $x_1 = s,\ x_3 = t$, and $x_5 = u$. Then $2x_4 + 2u = 0, 2x_2 + 4u = 0$, and so $x_4 = -u$ and $x_2 = -2u$. Hence $\text{Null}(A)$ consists of the vectors

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

and so

$$\mathbf{c}_1 = \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix},\ \mathbf{c}_3 = \begin{bmatrix}0\\-2\\0\\-1\\0\end{bmatrix}$$

form a basis for Null(A). To decompose $\mathbf{x} = (1,2,3,4,5)^T$ into the sum of an $\mathbf{x}_0 \in$ Null(A) and an $\mathbf{x}_R \in$ Row(A), solve $\mathbf{x} = s_1\mathbf{b}_1 + s_2\mathbf{b}_2 + t_1\mathbf{c}_1 + t_2\mathbf{c}_2 + t_3\mathbf{c}_3$ by row-reduction:

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

$$\to \left[\begin{array}{ccccc|c} 2 & 0 & 0 & 0 & -2 & 2\\ 0 & 2 & 0 & 0 & -1 & 4\\ 0 & 0 & 1 & 0 & 0 & 1\\ 0 & 0 & 0 & 1 & 0 & 3\\ 0 & 0 & 0 & 0 & 6 & -3 \end{array}\right]$$

Thus, $t_3 = -1/2,\ t_2 = 3,\ t_1 = 1,\ s_2 = 7/4,\ s_1 = 1/2,$

$$\mathbf{x}_0 = \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} + 3\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix} - \frac{1}{2}\begin{bmatrix}0\\-2\\0\\-1\\0\end{bmatrix} = \frac{1}{2}\begin{bmatrix}2\\2\\6\\1\\-1\end{bmatrix}$$

and

$$\mathbf{x}_R = \frac{1}{2}\begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 4 \end{bmatrix} + \frac{7}{4}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 \\ 2 \\ 0 \\ 7 \\ 11 \end{bmatrix}.$$

3.5.7. Null(A) is defined as the solution set of $A\mathbf{x} = \mathbf{0}$. Writing $\mathbf{a}^i$ for the ith row of A, we have, equivalently, $\mathbf{a}^i\mathbf{x} = 0$ for all i. If $\mathbf{x}$ is also in the row space of A, then by the definition of Row(A), $\mathbf{x}^T$ is a linear combination of the rows of A, that is, $\mathbf{x}^T = x_{Ai}\mathbf{a}^i$, for some coefficients x_{Ai}. Multiplying both sides of $\mathbf{a}^i\mathbf{x} = 0$ by x_{Ai}, and summing over i, we get $\sum x_{Ai}\mathbf{a}^i\mathbf{x} = \mathbf{x}^T\mathbf{x} = 0$. The last equality can be written as $|\mathbf{x}|^2 = 0$, which is true only for $\mathbf{x} = \mathbf{0}$. Thus $\mathbf{0}$ is the only vector lying in both Null(A) and Row(A).

3.5.9. From Exercise 3.2.9 we know that $U \cap V$ is a subspace of X. Since, by the result of Exercise 3.5.8, $U + V$ too is a subspace of X, all we need to prove is that $U \cap V$ is a *subset* of $U + V$.

Let $\mathbf{x}$ be any element of $U \cap V$. Then $\mathbf{x}$ is an element of U and can therefore be written as $\mathbf{u} + \mathbf{v}$ with $\mathbf{u} = \mathbf{x} \in U$ and $\mathbf{v} = \mathbf{0} \in V$. Thus $\mathbf{x} \in U + V$, and so $U \cap V$ is a subset, and consequently a subspace, of $U + V$.

3.5.11. If both A and B have n columns, then clearly each of the three nullspaces in the problem are subspaces of $\mathbb{R}^n$. Now, by definition,

$$\text{Null}\begin{bmatrix} A \\ B \end{bmatrix} = \left\{\mathbf{x} \in \mathbb{R}^n \middle| \begin{bmatrix} A \\ B \end{bmatrix}\mathbf{x} = \mathbf{0}\right\}.$$

Here the condition $\begin{bmatrix} A \\ B \end{bmatrix}\mathbf{x} = \mathbf{0}$ is equivalent to $\begin{bmatrix} A\mathbf{x} \\ B\mathbf{x} \end{bmatrix} = \mathbf{0}$, and therefore to the simultaneous occurrence of $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$. Thus a vector $\mathbf{x} \in \mathbb{R}^n$ is in Null$\begin{bmatrix} A \\ B \end{bmatrix}$ if and only if it is both in Null(A) and in Null(B), that is, in Null$(A) \cap$ Null(B).

3.5.13. For any subspace U of an inner product space X, we have defined $U^\perp = \{\mathbf{v} \in X | \mathbf{u} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{u} \in U\}$. Hence

a. $\mathbf{0} \in U^\perp$, because $\mathbf{u} \cdot \mathbf{0} = 0$ for all $\mathbf{u} \in U$, and so $U^\perp$ is nonempty.

b. $U^\perp$ is closed under addition: Let $\mathbf{v}, \mathbf{w} \in U^\perp$. Then $\mathbf{u} \cdot \mathbf{v} = 0$ and

$\mathbf{u} \cdot \mathbf{w} = 0$ for all $\mathbf{u} \in U$. Hence $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = 0$, for all $\mathbf{u} \in U$, which shows that $\mathbf{v} + \mathbf{w} \in U^{\perp}$.

c. $U^{\perp}$ is closed under multiplication by scalars: If $\mathbf{v} \in U^{\perp}$ and $c \in \mathbb{R}$, then $\mathbf{u} \cdot \mathbf{v} = 0$ for all $\mathbf{u} \in U$, and therefore $\mathbf{u} \cdot (c\mathbf{v}) = 0$ for all $\mathbf{u} \in U$, which shows that $c\mathbf{v} \in U^{\perp}$. Thus $U^{\perp}$ is a subspace of X.

$\mathbf{3.5.15}$. The exercises cited in the hint show that every elementary row operation on a matrix A is equivalent to multiplication by an elementary matrix E, and that every such E is invertible. Thus, if A and B are row-equivalent matrices, then there exist elementary matrices E_i such that $A = E_1E_2 \cdots E_kB$. Since each of the E_i matrices is invertible, so too is their product $R = E_1E_2 \cdots E_k$. Consequently, by the result of Exercise 3.5.14, B and $A = RB$ have the same rank.

$\mathbf{3.5.17}$. Assuming B has n columns, they are linearly dependent, if and only if there exist n constants t_i, not all zero, such that $\sum_{i=1}^{n} t_i\mathbf{b}_i = \mathbf{0}$. Multiplying both sides of this equation by A, we get $\sum_{i=1}^{n} t_iA\mathbf{b}_i = \mathbf{0}$, which shows that the vectors $A\mathbf{b}_i$, that is, the columns of AB, are dependent with the same coefficients as the $\mathbf{b}_i$ vectors.

Conversely, if the matrix A is not invertible, then we cannot conclude that the equation $\sum_{i=1}^{n} t_iA\mathbf{b}_i = \mathbf{0}$ implies $\sum_{i=1}^{n} t_i\mathbf{b}_i = \mathbf{0}$. Thus we cannot expect the converse of the original statement to be true, and should look for a counterexample.

Indeed, if

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

then

$$AB = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Here the columns of AB are linearly dependent, but the columns of B are linearly independent.

$\mathbf{3.5.19}$. The first three columns of A form an echelon matrix, and so they are independent and form the basis $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ for U. The last three columns of A can be reduced to an echelon matrix with three pivots, and so $\{\mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5\}$ is a basis for V.

To find a basis for $U \cap V$, we must find all vectors $\mathbf{p}$ that can be written both as $s_1\mathbf{a}_1+s_2\mathbf{a}_2+s_3\mathbf{a}_3$ and as $t_3\mathbf{a}_3+t_4\mathbf{a}_4+t_5\mathbf{a}_5$, that is, find all solutions of the equation $s_1\mathbf{a}_1+s_2\mathbf{a}_2+s_3\mathbf{a}_3 = t_3\mathbf{a}_3+t_4\mathbf{a}_4+t_5\mathbf{a}_5$. Writing $x_1 = s_1,\ x_2 = s_2,\ x_3 = s_3 - t_3,\ x_4 = -t_4$, and $x_5 = -t_5$, the equation we wish to solve becomes $A\mathbf{x} = \mathbf{0}$. Since A is an echelon matrix, we need no row-reduction, and back substitution gives $x_5 = t,\ x_4 = -t,\ x_3 = 0,\ x_2 = -t/2,\ x_1 = 0$. Equivalently, $s_1 = 0,\ s_2 = -t/2,\ s_3 - t_3 = 0,\ t_4 = t$, and $t_5 = -t$. Thus $U \cap V = \{\mathbf{p}|\mathbf{p} = -t\mathbf{a}_2/2 + s_3\mathbf{a}_3\} = \{\mathbf{p}|\mathbf{p} = s_3\mathbf{a}_3 + t(\mathbf{a}_5 - \mathbf{a}_4)\}$. Hence we observe that $\{\mathbf{a}_2, \mathbf{a}_3\}$ is a basis for $U \cap V$.

Since A is an echelon matrix, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$ is a basis for $U + V = \mathbb{R}^4$.

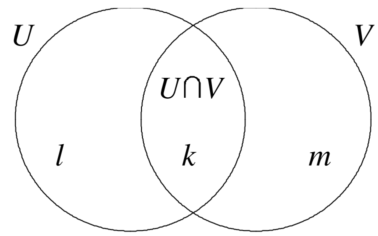
To find a basis for $U^\perp$ we need to solve $\mathbf{x}^T(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = 0$, that is, find a basis for the left nullspace of $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$. This computation results in $\{\mathbf{e}_4\}$ as a basis for $U^\perp$.

To find a basis for $V^\perp$ we need to solve $\mathbf{x}^T(\mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5) = 0$, that is, find a basis for the left nullspace of $(\mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5)$. This computation results in $\{(1, 0, -1, 1)^T\}$ as a basis for $V^\perp$.

3.5.21. Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be a basis for U and $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ a basis for V. Then $\dim U = m$ and $\dim V = n$. Furthermore, any $\mathbf{u} \in U$ and $\mathbf{v} \in V$ can be written as $\mathbf{u} = \sum_{i=1}^m a_i\mathbf{u}_i$ and $\mathbf{v} = \sum_{j=1}^n b_j\mathbf{v}_j$. Now $U \oplus V$ consists of all sums of vectors from U and from V, that is, of all sums $\sum_{i=1}^m a_i\mathbf{u}_i + \sum_{j=1}^n b_j\mathbf{v}_j$. Thus the combined set $C = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ spans $U \oplus V$. It is also an independent set, for the following reason: If $\mathbf{u} + \mathbf{v} = \sum_{i=1}^m a_i\mathbf{u}_i + \sum_{j=1}^n b_j\mathbf{v}_j = \mathbf{0}$, then $\mathbf{u} = -\mathbf{v}$, and so $\mathbf{u}$ is also in V. Since, by the definition of direct sum, the only vector that is in both U and V is the zero vector, we have $\mathbf{u} = \mathbf{v} = \mathbf{0}$. Thus both $\sum_{i=1}^m a_i\mathbf{u}_i = \mathbf{0}$ and $\sum_{j=1}^m b_j\mathbf{v}_j = \mathbf{0}$. Since the $\mathbf{u}_i$ vectors form a basis for U, and the $\mathbf{v}_j$ vectors a basis for V, the last two equations imply that all coefficients a_i and b_j must be zero. Hence C is an independent set, too, and therefore a basis for $U \oplus V$. This shows that $\dim(U \oplus V) = m + n$ and so that $\dim(U \oplus V) = \dim U + \dim V$.

3.5.23. Let $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ be a basis for $U \cap V$. Extend $\mathcal{B}$ to a basis $\mathcal{B}_1 = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{u}_1, \ldots, \mathbf{u}_l\}$ for U and to a basis
$\mathcal{B}_2 = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{v}_1, \ldots, \mathbf{v}_m\}$ for V. (These subspaces, together with the symbol for the number of basis vectors in each intersection, are illustrated schematically in a Venn diagram below.) Then
$\mathcal{B}_1 \cup \mathcal{B}_2 = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{u}_1, \ldots, \mathbf{u}_l, \mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a basis for $U + V$. Hence $\dim(U \cap V) = k$, $\dim(U + V) = k + l + m$, $\dim U = k + l$, and $\dim V =$

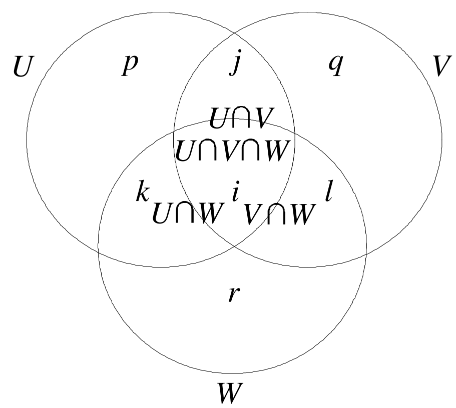
$k + m$. Now, $k + l + m = (k + l) + (k + m) - k$, and this proves the given formula.



3.5.25

$\dim(U + V + W) = \dim(U) + \dim(V) + \dim(W)$
$- \dim(U \cap V) - \dim(U \cap W) - \dim(V \cap W) + \dim(U \cap V \cap W).$

(The relevant subspaces, together with the symbol for the number of basis vectors in each intersection, are illustrated schematically in the Venn diagram below.)



Proof: Consider a basis $\mathcal{B}$ for $U \cap V \cap W$. Say, it has i vectors. Extend $\mathcal{B}$ to a basis for $U \cap V$, so that $\dim(U \cap V) = i + j$. Next, extend $\mathcal{B}$ to a basis for $U \cap W$, so that $\dim(U \cap W) = i + k$, and to a basis for $V \cap W$, so that $\dim(V \cap W) = i + l$. Since $U \cap V$ and $U \cap W$ are subspaces of U, the union of their bases can be extended to a basis for U, and so $\dim(U) = i + j + k + p$, for some p. Similarly, $\dim(V) = i + j + l + q$ and $\dim(W) = i + k + l + r$ for some q and r. By the result of Exercise 3.5.20, $U + V + W$ is generated by the union of U , V and W, and so also by the union of their bases. Thus

$\dim(U+V+W) = i+j+k+l+p+q+r$. On the other hand, the other dimension formulas above yield $\dim(U)+\dim(V)+\dim(W)-\dim(U\cap V)-\dim(U\cap W)-\dim(V\cap W)+\dim(U\cap V\cap W)$

$= (i+j+k+p)+(i+j+l+q)+(i+k+l+r)-(i+j)-(i+k)-(i+l)+i$

$= i+j+k+l+p+q+r$, the same as above.

3.5.27. Let $\mathcal{A}$ be a basis for $U\cap V$, $\mathcal{B}$ a basis for $U\cap V^{\perp}$, $\mathcal{C}$ a basis for $U^{\perp}\cap V$, and $\mathcal{D}$ a basis for $U^{\perp}\cap V^{\perp}$. Then, clearly, $\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}\cup\mathcal{D}$ is a basis for $\mathbb{R}^n$, $\mathcal{C}\cup\mathcal{D}$ for the subspace $U^{\perp}$, $\mathcal{B}\cup\mathcal{D}$ for $V^{\perp}$, and $\mathcal{B}\cup\mathcal{C}\cup\mathcal{D}$ for $(U\cap V)^{\perp}$. However, by the result of Exercise 3.5.20, the set $(\,(\mathcal{B}\cup\mathcal{D})\cup(\mathcal{C}\cup\mathcal{D})\,)$ $=\mathcal{B}\cup\mathcal{C}\cup\mathcal{D}$ generates $U^{\perp}+V^{\perp}$, and so we must have $(U\cap V)^{\perp}=U^{\perp}+V^{\perp}$.

3.5.29. $U\subset V$ means that every $\mathbf{u}\in U$ is also in V. Thus if $\overline{\mathbf{v}}\in V^{\perp}$, then $\overline{\mathbf{v}}^T\mathbf{u}=0$ for every $\mathbf{u}\in U$. On the other hand, the set $U^{\perp}$ was defined exactly as the set of those vectors that are orthogonal to every $\mathbf{u}\in U$. Hence $\overline{\mathbf{v}}\in V^{\perp}$ implies that $\overline{\mathbf{v}}\in U^{\perp}$, and this means that $V^{\perp}\subset U^{\perp}$ holds.

3.5.31. We want to find a basis for Left-null(B), as in Example 3.5.6, because the vectors of such a basis are orthogonal to the columns of B, and the transposes of those basis vectors can therefore be taken as the rows of the desired matrix A. Since Left-null(B) = Null(B^T) we solve $B^T\mathbf{x}=\mathbf{0}$ by reducing B^T as follows:

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

Thus x_3 is free and $x_1=x_2=x_4=0$, and so $A=[0,0,1,0]$ is a matrix such that Null(A) = Col(B).

3.5.33. For any A as stated, the equation $A\mathbf{x}=\mathbf{b}$ is consistent if and only if $\mathbf{b}\in$ Col(A). For any $\mathbf{b}\in$ Col(A) write the above equation as $A\mathbf{x}=I\mathbf{b}$ and reduce the latter, by elementary row operations, until A is in an echelon form $\begin{bmatrix} U \\ O \end{bmatrix}$ with U having no zero rows. On the right-hand side denote the result of this reduction of the matrix I by $\begin{bmatrix} L \\ M \end{bmatrix}$ Thus we get the equations $U\mathbf{x}=L\mathbf{b}$ and $\mathbf{0}=M\mathbf{b}$. The last equation shows that the rows of M must be orthogonal to any vector in the column space of A, and so their transposes are in the left null- space of A. Furthermore, the matrix $\begin{bmatrix} L \\ M \end{bmatrix}$ has full rank, since it is obtained from I by elementary row operations, which are invertible. Consequently the rows of M are independent. On the other hand, since the dimension of the left nullspace of A is $m-r$ and M has $m-r$ independent rows, their transposes span the left nullspace of A.

3.5.35. The construction of Exercise 3.5.33 applied to the transpose A^T of A in place of A yields a basis for the nullspace of A.

3.5.37. By the algorithm of Exercise 3.5.36 we do the following reduction

$$\left[\begin{array}{ccc|cccc} 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ -2 & 3 & -1 & 0 & 1 & 0 & 0 \\ 3 & -3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & -2 & 0 & 0 & 0 & 1 \\ \hline 5 & -6 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \rightarrow$$

$$\left[\begin{array}{ccc|cccc} 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 3 & -3 & 2 & 1 & 0 & 0 \\ 0 & -3 & 3 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 & 1 \\ \hline 0 & -6 & 6 & -5 & 0 & 0 & 0 \end{array}\right] \rightarrow$$

$$\left[\begin{array}{ccc|cccc} 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 3 & -3 & 2 & 1 & 0 & 0 \\ 0 & 0 & 3 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & -1 & 2 & 0 & 0 \end{array}\right].$$

Thus $\mathbf{x} = (-1, 2, 0, 0)^T$ is a particular solution, and the general solution is

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

where s and t are arbitrary parameters.

3.5.39. Applying the algorithm of Exercise 3.5.38, we have:

$$\left[\begin{array}{rr|rr} 1 & -2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rr|rr} 1 & -2 & 1 & 0 \\ 0 & 10 & -3 & 1 \\ \hline 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{rr|rr} 1 & -2 & 1 & 0 \\ 0 & 10 & -3 & 1 \\ \hline 0 & 0 & 4/10 & 2/10 \\ 0 & 0 & -3/10 & 1/10 \end{array}\right].$$

Thus

$$A^{-1} = \frac{1}{10}\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix}.$$

3.5.41. Let $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ be a list of spanning vectors. By the definition of dimension, any basis of an n-dimensional vector space X consists of n vectors. Thus, let $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ be the list of the vectors of a basis and use this list as the list of independent vectors in the Exchange Theorem. According to the theorem, the number of independent vectors cannot exceed the number of spanning vectors, and so we must have $n \leq m$.

3.5.43. Assume that $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m)$, with $m > n$, is a list of independent vectors in X. For comparison use the vectors $\mathbf{a}_i$ of any basis as the spanning list, that is, let $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ in the Exchange Theorem. Then the conclusion $k \leq n$ of the theorem becomes $m \leq n$. This inequality contradicts the assumption of $m > n$, and so the vectors of B cannot be independent.

3.6.1. a. By Corollary 3.6.1,

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

b. Compute S^{-1}:

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

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

Thus

$$S^{-1} = \begin{bmatrix} 1/5 & 2/5 \\ -2/5 & 1/5 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

and

$$\mathbf{x}_A = S^{-1}\mathbf{x} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 13 \\ -1 \end{bmatrix}.$$

$\mathbf{3.6.3}$. a. The matrix A is the same as in Exercise 3. 6. 1, and so

$$A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \text{ and } A^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$$

Now

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

and

$$S = A^{-1}B = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 7 & 3 \\ -4 & -1 \end{bmatrix}.$$

b. By Theorem 3.6.1,

$$\mathbf{x}_A = S\mathbf{x}_B = \frac{1}{5}\begin{bmatrix} 7 & 3 \\ -4 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

Thus $\mathbf{x} = 3\mathbf{b}_1 - 2\mathbf{b}_2 = 3\mathbf{a}_1 - 2\mathbf{a}_2$. (It is only a coincidence that the coordinates of this $\mathbf{x}$ are the same in both bases.)

c. From Part (a) above,

$$S^{-1} = \begin{bmatrix} -1 & -3 \\ 4 & 7 \end{bmatrix}.$$

Hence

$$\mathbf{x}_B = S^{-1}\mathbf{x}_A = \begin{bmatrix} -1 & -3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -14 \\ 36 \end{bmatrix},$$

and so

$$\mathbf{x} = 2\mathbf{a}_1 + 4\mathbf{a}_2 = -14\mathbf{b}_1 + 36\mathbf{b}_2.$$

$\mathbf{3.6.5}$. From the given data

$$\begin{aligned}\mathbf{x}_A &= \begin{bmatrix} x_2 - x_3 \\ x_3 - x_1 \\ x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 0x_1 & +1x_2 & -1x_3 \\ -1x_1 & +0x_2 & +1x_3 \\ 1x_1 & +1x_2 & +0x_3 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.\end{aligned}$$

By Corollary 3.6.1, with A in place of B, we have $\mathbf{x}_A = A^{-1}\mathbf{x}$, and so

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

Hence

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

The new basis vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are the columns of this matrix (including the factor 1/2).

$\mathbf{3.6.7}$. a. From the statement of the problem,

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

By Corollary 3.6.1, with A in place of B, we have $S = A$.

b. From A we compute

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

and

$$\mathbf{x}_A = A^{-1}\mathbf{x} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}.$$

c.

$$M_A = A^{-1}MA \; = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

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

$\mathbf{3.6.9}$. a. We need to solve $AS = B$. As in Example 3.6.3, we reduce the augmented matrix $[A|B]$:

$$\left[\begin{array}{cc|cc} 1 & 0 & 2 & 3 \\ 2 & 4 & 0 & -2 \\ 3 & -1 & 7 & 11 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 2 & 3 \\ 1 & 2 & 0 & -1 \\ 3 & -1 & 7 & 11 \end{array}\right] \to$$

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

Thus

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

b. The transition matrix in the reverse direction is S^{-1}, which happens to equal S.

$\mathbf{3.6.11}$. a. As in Example 3.6.4, in the space

$$\mathcal{P}_3 = \{\mathbf{p} = P : P(x) = p_0 + p_1 x + p_2 x^2 + p_3 x^3; p_0, p_1, p_2, p_3 \in \mathbb{R}\}$$

we choose the basis A to consist of the monomials, that is, $\mathbf{a}_i = x^i$ for $i = 0, \ldots, 3$, and the basis B to consist of the first four Legendre polynomials, that is, $\mathbf{b}_i = L_i$ for $i = 0, \ldots, 3$. Then, according to Theorem 3.6.2, the columns of the matrix S are given by the coordinates of the $\mathbf{b}_i$ vectors relative to A. These can be read off the definitions of the Legendre polynomials, to give (in ascending order of degrees)

$$S = \begin{bmatrix} 1 & 0 & -1/2 & 0 \\ 0 & 1 & 0 & -3/2 \\ 0 & 0 & 3/2 & 0 \\ 0 & 0 & 0 & 5/2 \end{bmatrix}.$$

b. The coordinate vector of any $\mathbf{p}$ relative to A is $\mathbf{p}_A = (p_0, p_1, p_2, p_3)^T$ and its coordinate vector relative to B is given by $\mathbf{p}_B = S^{-1}\mathbf{p}_A$. Thus multiplication of $\mathbf{p}_A$ by

$$S^{-1} = \begin{bmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 3/5 \\ 0 & 0 & 2/3 & 0 \\ 0 & 0 & 0 & 2/5 \end{bmatrix}$$

will give the coordinates of any $\mathbf{p}$ relative to B. For the polynomial $P(x) = 1 - 2x + 3x^2 - 4x^3$ we have

$$\begin{bmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 3/5 \\ 0 & 0 & 2/3 & 0 \\ 0 & 0 & 0 & 2/5 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ -22/5 \\ 2 \\ -8/5 \end{bmatrix}.$$

Thus $P(x) = 2L_0(x) - 22L_1(x)/5 + 2L_2(x) - 8L_3(x)/5$.

3.6.13. For any $\mathbf{x} \in \mathbb{R}^n$, let $\mathbf{y} = M\mathbf{x}$ be the corresponding vector of $\mathbb{R}^m$. We can rewrite this equation, by using $\mathbf{x} = A\mathbf{x}_A$ and $\mathbf{y} = B\mathbf{y}_B$, as $A\mathbf{x}_A = MB\mathbf{y}_B$. Hence $\mathbf{x}_A = A^{-1}MB\mathbf{y}_B$ and so $M_{A,B} = A^{-1}MB$.

3.6.15. The second one of the given matrices is the 2×2 identity matrix I. Now, for any invertible matrix S, $SIS^{-1} = I$, and so there is no matrix other than I itself that is similar to the identity matrix.

3.6.17. a. B is similar to A if there exists an invertible matrix S such that $B = S^{-1}AS$, or, equivalently, $SB = AS$. Taking the transpose of both sides, we get $B^TS^T = S^TA^T$ and from this $B^T = S^TA^T(S^T)^{-1}$. Thus B^T is similar to A^T with $(S^T)^{-1}$ in place of S if B is similar to A.

b. If $B = S^{-1}AS$, then $B^k = S^{-1}ASS^{-1}AS \cdots S^{-1}AS = S^{-1}A^kS$.

c. For any invertible matrices A and S, we have $(S^{-1}AS)^{-1} = SA^{-1}S^{-1}$, and so, if $B = S^{-1}AS$, then B^{-1} exists and equals $SA^{-1}S^{-1}$, and B^{-1} is similar to A^{-1} with S^{-1} in place of S.

3.6.19. An $n \times n$ matrix B is similar to an $n \times n$ matrix A if there exists an invertible $n \times n$ matrix S such that $B = S^{-1}AS$, or, equivalently, $SB = AS$. Furthermore, by the result of Exercise 3.4.9, $\mathrm{Null}(B) = \mathrm{Null}(SB)$.

To see that $\mathrm{nullity}(A) = \mathrm{nullity}(AS)$, consider the definitions of these nullspaces: Null(AS) is the solution set of the equation $AS\mathbf{x} = \mathbf{0}$, and $\mathrm{Null}(A)$ that of $A\mathbf{y} = \mathbf{0}$. Since S is invertible, the equation $\mathbf{y} = S\mathbf{x}$ establishes a one-to-one linear correspondence between the elements of Null(AS) and $\mathrm{Null}(A)$. In particular, any basis of Null(AS) is mapped by S onto a basis of $\mathrm{Null}(A)$. Thus the dimensions of the two nullspaces must be equal.

Coupling the statements in the two paragraphs above, we find that $\mathrm{nullity}(A) = \mathrm{nullity}(B)$. Thus, from this equation and Corollary 3.5.1 we get the desired result: $\mathrm{rank}(A) = \mathrm{rank}(B)$.

4.1.1. a. We just prove Corollary 4.1.1, because Lemma 4.1.1 is a special case of it with $n = 2$.

Assume first that Equation 4.4 holds for some $n \geq 2$ and for all vectors and all scalars in it. Then, choosing $c_1 = c_2 = 1$ and $c_i = 0$ for all $i > 2$, we get Equation 4.1. On the other hand, choosing $c_1 = c$ and $c_i = 0$ for all $i > 1$, we get Equation 4.2. Thus, if T preserves all linear combinations with $n \geq 2$ terms, then it is linear.

Assume conversely that T is linear, that is, that Equations 4.1 and 4.2 hold for all choices of x_1, x_2, and c. Then, for any $n \geq 2$,

$$T\left(\sum_{i=1}^{n} c_i\mathbf{x}_i\right) = T\left(c_1\mathbf{x}_1 + \sum_{i=2}^{n} c_i\mathbf{x}_i\right) = T(c_1\mathbf{x}_1) + T\left(\sum_{i=2}^{n} c_i\mathbf{x}_i\right)$$

$$= c_1T(\mathbf{x}_1) + T\left(\sum_{i=2}^{n} c_i\mathbf{x}_i\right).$$

We can similarly reduce the remaining sum by repeated applications of Equations 4.1 and 4.2 until we get

$$T\left(\sum_{i=1}^{n} c_i\mathbf{x}_i\right) = c_1T(\mathbf{x}_1) + c_2T(\mathbf{x}_2) + \cdots + c_nT(\mathbf{x}_n) = \sum_{i=1}^{n} c_iT(\mathbf{x}_i).$$

b. For any $\mathbf{x}$ we have $T(\mathbf{0}) = T(\mathbf{x} - \mathbf{x}) = T(\mathbf{x} + (-1)\mathbf{x}) = T(\mathbf{x}) + T((-1)\mathbf{x}) = T(\mathbf{x}) + (-1)T(\mathbf{x}) = T(\mathbf{x}) - T(\mathbf{x}) = \mathbf{0}$.

4.1.3. a. It is certainly true that if a mapping T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is linear, then it preserves straight lines: If T is any such mapping, then, applying T to $\mathbf{x} = \mathbf{x}_0 + t\mathbf{a}$ with any $\mathbf{x}, \mathbf{x}_0, \mathbf{a} \in \mathbb{R}^n$ and scalar t, we must have $\mathbf{y} = T(\mathbf{x}) = T(\mathbf{x}_0 + t\mathbf{a}) = T(\mathbf{x}_0) + tT(\mathbf{a})$. Thus the line given by $\mathbf{x} = \mathbf{x}_0 + t\mathbf{a}$ is mapped into the line given by $\mathbf{y} = \mathbf{y}_0 + t\mathbf{b}$ with $y_0 = T(\mathbf{x}_0)$ and $\mathbf{b} = T(\mathbf{a})$.

b. The converse statement is not true in general, that is, not every transformation that preserves straight lines is linear. For example a shift within $\mathbb{R}^n$, defined by $\mathbf{y} = \mathbf{S}(\mathbf{x}) = \mathbf{x} + \mathbf{c}$ with fixed nonzero $\mathbf{c}$, preserves straight lines but is nonlinear, since $\mathbf{S}(\mathbf{0}) = \mathbf{0} + \mathbf{c} = \mathbf{c}$ and not $\mathbf{0}$ as required for linear transformations (see Exercise 4.1.1).

Let us try to find out what additional conditions will make such transformations linear. So, let us assume that T maps every $\mathbf{x} = \mathbf{x}_0 + t\mathbf{a}$ to some $\mathbf{y} = \mathbf{y}_0 + t\mathbf{b}$ for every scalar t. Setting $t = 0$ yields $y_0 = T(\mathbf{x}_0)$, and so we must have $T(\mathbf{x}_0 + t\mathbf{a}) = T(\mathbf{x}_0) + t\mathbf{b}$. For T to be linear we must also have $\mathbf{b} = T(\mathbf{a})$. These two conditions are also sufficient, that is, if T maps every $\mathbf{x} = \mathbf{x}_0 + t\mathbf{a}$ to $\mathbf{y} = \mathbf{y}_0 + t\mathbf{b}$ with $y_0 = T(\mathbf{x}_0)$ and $\mathbf{b} = T(\mathbf{a})$, then T is linear. Indeed, in that case, $T(\mathbf{x}_0 + t\mathbf{a}) = T(\mathbf{x}_0) + tT(\mathbf{a})$ with $t = 1$ shows the additivity of T. Setting $\mathbf{x}_0 = \mathbf{a} = \mathbf{0}$, from here we obtain $T(\mathbf{0}) = 2T(\mathbf{0})$, and so $T(\mathbf{0}) = \mathbf{0}$. Hence $T(t\mathbf{a}) = T(\mathbf{0} + t\mathbf{a}) = T(\mathbf{0}) + tT(\mathbf{a}) = \mathbf{0} + tT(\mathbf{a}) = tT(\mathbf{a})$, which shows the homogeneity of T.

4.1.5. We can write

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

and so

$$[T] = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ 3 & 2 \end{bmatrix}.$$

4.1.7. In the formula $T(\mathbf{x}) = (\mathbf{a}^T\mathbf{x})\mathbf{b}$ the product on the right is understood as the scalar $\mathbf{a}^T\mathbf{x}$ multiplying the vector $\mathbf{b}$, that is, multiplying each component of $\mathbf{b}$. Thus the ith component of $T(\mathbf{x})$ is

$$[T(\mathbf{x})]_i = \left(\sum_{j=1}^{n} a_f x_j\right) b_i = \sum_{j=1}^{n}(b_i a_j)x_j,$$

and so $[T]_{ij} = b_i a_j$.

Alternatively, we may write $T(\mathbf{x})$ as the matrix product $\mathbf{b}(\mathbf{a}^T\mathbf{x})$, in which $\mathbf{a}^T\mathbf{x}$ is considered to be a 1×1 matrix. Thus, by the associativity of matrix multiplication, $T(\mathbf{x}) = (\mathbf{b}\mathbf{a}^T)\mathbf{x}$. Hence

$$[T] = \mathbf{b}\mathbf{a}^T = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} (a_1, a_2, \ldots, a_n) = \begin{bmatrix} b_1a_1 & b_1a_2 & \cdots & b_1a_n \\ b_2a_1 & b_2a_2 & \cdots & b_2a_n \\ \vdots & & & \\ b_ma_1 & b_ma_2 & \cdots & b_ma_n \end{bmatrix}.$$

This matrix is called the tensor product of $\mathbf{b}$ and $\mathbf{a}$ and also the outer product of $\mathbf{b}$ and $\mathbf{a}^T$.

4.1.9. $\mathbf{a}_1 = (1,1,1)^T$, $\mathbf{a}_2 = (1,-1,-1)^T$ and $\mathbf{a}_3 = (1,1,0)^T$ are the columns of the basis-matrix

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

Similarly, the vectors $\mathbf{b}_1 = (1,1)^T$, $\mathbf{b}_2 = (1,-1)^T$ and $\mathbf{b}_3 = (1,0)^T$ may be combined into the (nonbasis) matrix

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

By the definition of the matrix $[T]$,

$$[T]A = B$$

and so

$$[T] = BA^{-1}.$$

Now A^{-1} can be obtained in the usual way:

$$
\begin{aligned}
&\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \\
\rightarrow &\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 & 1 \end{array}\right] \\
\rightarrow &\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 \end{array}\right] \\
\rightarrow &\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{array}\right] \\
\rightarrow &\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & -1/2 & 1 \\ 0 & 1 & 0 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{array}\right].
\end{aligned}
$$

Thus,

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

and

$$
[T] = BA^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

Hence

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

4.1.11. A rotation by -45° takes the $y = x$ line into the x-axis. Then stretch by the matrix S of Example 4.1.11 and rotate back by 45° :

$$[T] = \left(\frac{\sqrt{2}}{2}\right)^2 \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

It is easy to check that this matrix doubles the vector $(1,1)^T$ and leaves the vector $(1,-1)^T$ unchanged, as it should.

4.1.13. The matrix $[T]$ given by

$$[T] = \frac{1}{a^2+b^2} \begin{bmatrix} b^2-a^2 & -2ab \\ -2ab & a^2-b^2 \end{bmatrix}$$

represents a reflection across the line $ax + by = 0$ (see Exercise 4.1.12). If θ denotes the inclination of that line, with $0 \leq \theta \leq \pi$, then $\cos 2\theta = (b^2-a^2)/(a^2+b^2)$ and $\sin 2\theta = (-2ab)/(a^2+b^2)$. Thus we can rewrite $[T]$ as

$$[T] = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

Similarly, the matrix $[S]$ given by

$$[S] = \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix}$$

represents a reflection across a line $c\mathbf{x} + d\mathbf{y} = 0$ whose inclination is ϕ. The matrix that represents the result of a reflection across the line $ax + by = 0$ followed by a reflection across the line $cx + dy = 0$ is given by

$$\begin{aligned} [S][T] &= \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos 2(\phi-\theta) & -\sin 2(\phi-\theta) \\ \sin 2(\phi-\theta) & \cos 2(\phi-\theta) \end{bmatrix}. \end{aligned}$$

The resultant matrix corresponds to a rotation about the origin through the angle $2(\phi-\theta)$, which is twice the angle between the lines, measured from $ax + by = 0$ to $cx + dy = 0$.

4.1.15. A linear transformation T from $\mathbb{R}^2$ to $\mathbb{R}^2$ will map a line $\mathbf{p} = \mathbf{p}_1 + t\mathbf{v}_1$ onto the line $T(\mathbf{p}) = T(\mathbf{p}_1 + t\mathbf{v}_1) = T(\mathbf{p}_1) + tT(\mathbf{v}_1)$. Similarly T

maps the line $\mathbf{p} = \mathbf{p}_2 + t\mathbf{v}_2$ onto the line $T(\mathbf{p})$ $=T(\mathbf{p}_2) + tT(\mathbf{v}_2)$ Thus T has the property that it maps perpendicular lines onto perpendicular lines if and only if it maps orthogonal vectors onto orthogonal vectors. The zero transformation satisfies these conditions trivially. Suppose that $T \neq \mathbf{0}$, and that the matrix of T relative to the standard basis for $\mathbb{R}^2$ is given by

$$[T] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then $T(\mathbf{e}_1) = (a, c)^T$, and $T(\mathbf{e}_2) = (b, d)^T$. Since $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal, their images are orthogonal; that is, $(a, c)^T \cdot (b, d)^T = 0$, or, (1) $ab + cd = 0$. Also, the vectors $\mathbf{e}_1 - \mathbf{e}_2$ and $\mathbf{e}_1 + \mathbf{e}_2$ are orthogonal, and thus so are their images; that is, $(a - b, c - d)^T \cdot (a + b, c + d)^T = 0$, or, (2) $a^2 + c^2 = b^2 + d^2$. The assumption that $T \neq \mathbf{0}$, together with Equations (1) and (2) above, imply that $(a, c)^T \neq \mathbf{0}$ and $(b, d)^T \neq \mathbf{0}$. In addition, we may write $(a, c)^T = k(\cos\theta, \sin\theta)^T$ and $(b, d)^T = k(\cos\phi, \sin\phi)^T$, where $k = \sqrt{a^2 + c^2} = \sqrt{b^2 + d^2} > 0$, and $\phi = \theta \neq \pi/2$. There are two cases to consider:

(1) $\phi = \theta + \pi/2$.

In this case, $\sin\phi = \cos\theta$ and $\cos\phi = -\sin\theta$, and so

$$[T] = k \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Hence T is a rotation through an angle θ, followed by a dilation or contraction by a factor k.

(2) $\phi = \theta - \pi/2$.

In this case, $\sin\phi = -\cos\theta$ and $\cos\phi = \sin\theta$, and so

$$[T] = k \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

Hence T is a reflection across the line through the origin with inclination $\theta/2$, followed by a dilation or contraction by a factor k.

Since a *pure* dilation or contraction is the special case of (1) when $\theta = 0$, and the zero transformation is obtained when $k = 0$, we can summarize the possibilities as follows. A linear transformation T preserves orthogonality if and only if it is a rotation, a reflection, a dilation or a contraction, or a composition of such transformations.

4.1.17. For the transformation T of Exercise 4.1.6, we have

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

and thus

$$[T] = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

We may apply Corollary 4.1.2 to determine the matrix $T_{A,B}$. To do so, we need to compute B^{-1}, which we find to be

$$B^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

Then matrix multiplication gives

$$\begin{aligned} T_{A,B} &= B^{-1}[T]A \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}. \end{aligned}$$

4.1.19. To determine the matrix $T_{A,B}$ that represents the integration map T relative to the ordered bases $A = (1, x, \ldots, x^n)$ and $B = (1, x, \ldots, x^{n+1})$, we can proceed as in Example 4.1.13 in the text. Using the notation $\mathbf{a}_j = x^{j-1}$ and $\mathbf{b}_i = x^{i-1}$, we have

$$T(\mathbf{a}_j) = \frac{1}{j}\mathbf{b}_{j+1} \text{ for } j = 1,\ 2,\ \ldots,\ n.$$

Thus

$$T(\mathbf{a}_1) = 0\mathbf{b}_1 + 1\mathbf{b}_2 + \cdots + 0\mathbf{b}_{n+1},$$

$$T(\mathbf{a}_2) = 0\mathbf{b}_1 + 0\mathbf{b}_2 + (1/2)\mathbf{b}_3 + \cdots + 0\mathbf{b}_{n+1},$$

$$T(\mathbf{a}_3) = 0\mathbf{b}_1 + 0\mathbf{b}_2 + 0\mathbf{b}_3 + (1/3)\mathbf{b}_4 + \cdots + 0\mathbf{b}_{n+1},$$

etc.

Therefore

$$T_{A,B} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1/2 & 0 & \cdots & 0 \\ 0 & 0 & 1/3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1/(n+1) \end{bmatrix}.$$

4.2.1. a. W must be closed under addition: Let $\mathbf{x}_1, \mathbf{x}_2 \in W$. Then $T(\mathbf{x}_1) + T(\mathbf{x}_2)$ is well-defined, but if $\mathbf{x}_1 + \mathbf{x}_2$ were not in the domain W of T, then $T(\mathbf{x}_1 + \mathbf{x}_2)$ would not be defined, and $T(\mathbf{x}_1) + T(\mathbf{x}_2) = T(\mathbf{x}_1 + \mathbf{x}_2)$ could not hold, in contradiction to our assumption.

b. W must be closed under multiplication by scalars: Let $\mathbf{x}_1 \in W$, and let c be any scalar. Then $cT(\mathbf{x}_1)$ is well-defined, but if $c\mathbf{x}_1$ were not in the domain W of T, then $T(c\mathbf{x}_1)$ would not be defined, and $cT(\mathbf{x}_1) = T(c\mathbf{x}_1)$ could not hold, in contradiction to our assumption.

4.2.3. a. T being one-to-one implies $\mathrm{Ker}(T) = \{\mathbf{0}\}$: By definition, T is one-to-one if $T(\mathbf{x}) = T(\mathbf{y})$ implies $\mathbf{x} = \mathbf{y}$. Choose $\mathbf{y} = \mathbf{0}$. By the result of Exercise 4.1.1b, $T(\mathbf{0}) = \mathbf{0}$, and so $T(\mathbf{x}) = \mathbf{0}$ implies $T(\mathbf{x}) = T(\mathbf{0})$, which, by the one-to-oneness assumption, implies $\mathbf{x} = \mathbf{0}$. Thus $\mathrm{Ker}(T) = \{\mathbf{x} : T(\mathbf{x}) = \mathbf{0}\} = \{\mathbf{0}\}$.

b. $\mathrm{Ker}(T) = \{\mathbf{0}\}$ implies T is one-to-one: If $\mathrm{Ker}(T) = \{\mathbf{0}\}$, then $T(\mathbf{x} - \mathbf{y}) = \mathbf{0}$ implies $\mathbf{x}-\mathbf{y} = \mathbf{0}$. But, by the linearity of T, the equation $T(\mathbf{x}-\mathbf{y}) = \mathbf{0}$ is equivalent to $T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{0}$, that is, to $T(\mathbf{x}) = T(\mathbf{y})$. Also, $\mathbf{x} - \mathbf{y} = \mathbf{0}$ is equivalent to $\mathbf{x} = \mathbf{y}$, and so, if $\mathrm{Ker}(T) = \{\mathbf{0}\}$, then $T(\mathbf{x}) = T(\mathbf{y})$ implies $\mathbf{x} = \mathbf{y}$. This is precisely the definition of T being one-to-one.

4.2.5. If $N = M^T$, then Row(N) = Col(M) and Col(N) = $\mathrm{Row}(M)$. Thus we can use the same bases as in Example 4.2.5, just switched, and so

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

is a basis matrix for Row(N), and

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

is a basis matrix for Col (N).

If $\mathbf{N}$: $\mathrm{Row}(N) \to$ Col(N) is the isomorphism given by $\mathbf{y} = N\mathbf{x}$, then this equation can be written, in terms of the coordinate vectors, as

$$B\mathbf{y}_B = NA\mathbf{x}_A,$$

that is, as

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} y_{B1} \\ y_{B2} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_{A1} \\ x_{A2} \end{bmatrix} \\ &= \begin{bmatrix} 2 & 3 \\ 3 & 6 \\ 2 & 3 \\ 5 & 9 \end{bmatrix} \begin{bmatrix} x_{A1} \\ x_{A2} \end{bmatrix}. \end{aligned}$$

This equation can be reduced to

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{B1} \\ y_{B2} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{A1} \\ x_{A2} \end{bmatrix}.$$

Hence

$$N_{A,B} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}.$$

4.2.7. a. $\mathrm{Ker}(T)$ is not empty, since $\mathbf{0}$ is always a member of it.

b. It is closed under vector addition, since, if $T(\mathbf{x}) = \mathbf{0}$ and $T(\mathbf{y}) = \mathbf{0}$, then, by the linearity of T, $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) = \mathbf{0}$.

c. It is closed under multiplication by scalars: If $T(\mathbf{x}) = \mathbf{0}$ and c is any scalar, then, by the linearity of T, we have $T(c\mathbf{x}) = cT(\mathbf{x}) = \mathbf{0}$.

4.2.9. Let us first test the $\mathbf{b}_i$ vectors for independence: Assume $\sum_{i=1}^{n} c_i\mathbf{b}_i = \mathbf{0}$. Then, by the linearity of T, we may rewrite this as

$$\sum_{i=1}^{n} c_i T(\mathbf{a}_i) = T(\sum_{i=1}^{n} c_i\mathbf{a}_i) = \mathbf{0}.$$

Also, as for any linear transformation, $T(\mathbf{0}) = \mathbf{0}$. Thus, because T is one-to-one, we must have $\sum_{i=1}^{n} c_i\mathbf{a}_i = \mathbf{0}$. Since the $\mathbf{a}_i$ vectors form a basis for U, they are independent, and so, $c_i = 0$ for each i. Hence the $\mathbf{b}_i$ vectors are also independent, and consequently, by the result of Exercise 3.4.40, they form a basis for V.

All that remains to show is that T is a mapping *onto* V. But this is now easy: Let $\mathbf{b} \in V$. Since the $\mathbf{b}_i$ vectors form a basis for V, we can write $\mathbf{b}$ as a linear combination of these, say, $\mathbf{b} = \sum_{i=1}^{n} c_i\mathbf{b}_i$. But then $\mathbf{b} = \sum_{i=1}^{n} c_i T(\mathbf{a}_i) = T(\sum_{i=1}^{n} c_i\mathbf{a}_i)$, which shows that any $\mathbf{b} \in V$ is the image of some $\mathbf{a} \in U$ under T.

4.2.11. Consider the ordered basis $A = B = (1, x, \ldots, x^n)$ of $\mathcal{P}_n$, that is, take the vectors $\mathbf{a}_i = x^{i-1}$, $i = 1, 2, \ldots, n+1$, for a basis. Then, as in Example 4.1.13, $\mathbf{D}\mathbf{a}_i = (i-1)x^{i-2}$ for $i = 2, \ldots, n+1$,.and $\mathbf{D}\mathbf{a}_1 = 0$. Hence, $\mathbf{X} \circ \mathbf{D}\mathbf{a}_i = (i-1)x^{i-1} = (i-1)\mathbf{a}_i$ for $i = 1, 2, \ldots, n+1$, and, similarly, $\mathbf{D} \circ \mathbf{X}\mathbf{a}_i = ix^{i-1} = i\mathbf{a}_i$ for all i.

Thus, $T\mathbf{a}_i = (\mathbf{X} \circ \mathbf{D} - \mathbf{D} \circ \mathbf{X})\mathbf{a}_i = -\mathbf{a}_i$. Consequently, we find that $T_{A,B} = -I$, $\mathrm{Range}(T) = \mathcal{P}_n$, $\mathrm{Ker}(T) = \{\mathbf{0}\}$, and T is both one-to-one and onto.

4.3.1. As in Example 4.3.1, we have

$$\begin{aligned}
\overline{R}(30^\circ) &= T(1,-2)R(30^\circ)T(-1,2)\\
&= \frac{1}{2}\begin{bmatrix} 1 & 0 & 1\\ 0 & 1 & -2\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} \sqrt{3} & -1 & 0\\ 1 & \sqrt{3} & 0\\ 0 & 0 & 2\end{bmatrix}\begin{bmatrix} 1 & 0 & -1\\ 0 & 1 & 2\\ 0 & 0 & 1\end{bmatrix}\\
&= \begin{bmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & -\frac{1}{2}\sqrt{3}\\ \frac{1}{2} & \frac{1}{2}\sqrt{3} & \sqrt{3}-\frac{5}{2}\\ 0 & 0 & 1\end{bmatrix}\\
&= \frac{1}{2}\begin{bmatrix} 1 & 0 & 1\\ 0 & 1 & -2\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} \sqrt{3} & -1 & -\sqrt{3}-2\\ 1 & \sqrt{3} & -1+2\sqrt{3}\\ 0 & 0 & 2\end{bmatrix}\\
&= \frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 & -\sqrt{3}\\ 1 & \sqrt{3} & 2\sqrt{3}-5\\ 0 & 0 & 2\end{bmatrix}.
\end{aligned}$$

4.3.3.

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

This matrix represents the inverse of the mapping of Exercise 4.3.2, that is, it maps the unit square onto the given rectangle without any rotation. Indeed, $L^{-1}(0,0,1)^T = (1,-2,1)^T,\ L^{-1}(1,0,1)^T = (4,-2,1)^T,\ L^{-1}(0,1,1)^T = (1,2,1)^T,\ L^{-1}(1,1,1)^T = (4,2,1)^T$.

4.3.5. First we use the matrix

$$R = \frac{1}{pp_{12}}\begin{bmatrix} pp_2 & -pp_1 & 0\\ p_1p_3 & p_2p_3 & -p_{12}^2\\ p_1p_{12} & p_2p_{12} & p_3p_{12}\end{bmatrix}$$

resulting from Exercise 4.3.4, to rotate the vector $\mathbf{p}$ into the z-axis. Here $p = |\mathbf{p}|$ and $p_{12} = \sqrt{p_1^2 + p_2^2}$. Next, we rotate by the angle θ about the z-axis, using the matrix of Equation (4.107) and, finally, we rotate back by R^{-1}. Thus

$$R_\theta = \frac{1}{p^2 p_{12}^2}\begin{bmatrix} pp_2 & p_1p_3 & p_1p_{12} \\ -pp_1 & p_2p_3 & -p_{12}^2 \\ 0 & p_2p_{12} & p_3p_{12} \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} pp_2 & -pp_1 & 0 \\ p_1p_3 & p_2p_3 & -p_{12}^2 \\ p_1p_{12} & p_2p_{12} & p_3p_{12} \end{bmatrix}$$

4.3.7. From the solution of Exercise 4.3.5, with $p_1 = p_2 = 1,\ p_3 = 0$ and $p = p_{12} = \sqrt{2}$, we obtain

$$\overline{R}_\theta = \begin{bmatrix} 2\cos\theta + 2 & -2\cos\theta + 2 & 2\sqrt{2}\sin\theta & 0 \\ -2\cos\theta + 2 & 2\cos\theta + 2 & -2\sqrt{2}\sin\theta & 0 \\ -2\sqrt{2}\sin\theta & 2\sqrt{2}\sin\theta & 4\cos\theta & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

4.3.9. Let $\mathbf{u} = \frac{1}{\sqrt{5}}(2,-1,0)^T$ and $\mathbf{v} = (0,0,1)^T$. Thus

$$P_V(\mathbf{u},\mathbf{v}) = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

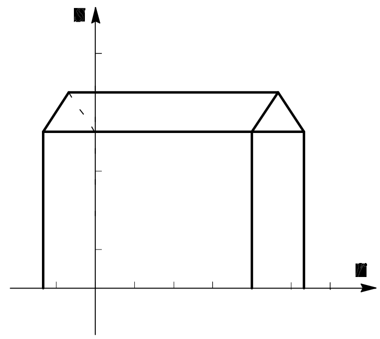
$$P_V\begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -3/\sqrt{5} \\ 4 \end{bmatrix},\ P_V\begin{bmatrix} 6 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 9/\sqrt{5} \\ 4 \end{bmatrix},$$

$$P_V\begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -3/\sqrt{5} \\ 0 \end{bmatrix},\ P_V\begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 9/\sqrt{5} \\ 0 \end{bmatrix},\ P_V\begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix},$$

$$P_V\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ P_V\begin{bmatrix} 6 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 12/\sqrt{5} \\ 4 \end{bmatrix},$$

$$P_V\begin{bmatrix} 0 \\ 3/2 \\ 5 \end{bmatrix} = \begin{bmatrix} -3\sqrt{5}/10 \\ 5 \end{bmatrix},\ P_V\begin{bmatrix} 6 \\ 3/2 \\ 5 \end{bmatrix} = \begin{bmatrix} 21\sqrt{5}/10 \\ 5 \end{bmatrix}.$$

Hence we get the following picture:



5.1.1. To find a matrix A with independent columns with a given column space, is just another way of saying that we want to find a basis for the given space.

In the present case the x-axis is one-dimensional and is spanned by any non-zero multiple of the vector $\mathbf{i}$. Thus $A = (1, 0, 0)^T$ is a matrix with independent columns whose column space is the x-axis.

5.1.3. Since the given column space is a plane, it is two-dimensional and we just have to find two independent vectors whose components satisfy the condition $x = y$. Clearly the columns of the matrix

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

will do.

Alternatively, using the formalism of Chapter 2, we may write the equation $x = y$ in the form $1x - 1y + 0z = 0$, whose augmented matrix is the echelon matrix $B = [1, -1,\ 0|\ 0]$. Columns 2 and 3 are free, and so we set y and z equal to parameters s and t. Then we get the general solution as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

from which we see that the columns of the matrix A above form a basis for Null(B), that is for the $x = y$ plane.

$\mathbf{5.1.5}$. In one direction the proof is easy: If $\mathbf{x}$ is a solution of $A\mathbf{x} = \mathbf{p}$, then left-multiplying both sides by A^T shows that $\mathbf{x}$ is also a solution of $A^TA\mathbf{x} = A^T\mathbf{p}$.

In the other direction, the assumption that $\mathbf{p}$ is in Col(A) must be used. This condition means that there exists a vector $\mathbf{p}_A$ such that $\mathbf{p} = A\mathbf{p}_A$. Substituting this expression for $\mathbf{p}$ into $A^TA\mathbf{x} = A^T\mathbf{p}$, we get $A^TA\mathbf{x} = A^TA\mathbf{p}_A$. By Lemma 5.1.3 the matrix A^TA is invertible for any A with independent columns, and so the last equation yields $\mathbf{x} = \mathbf{p}_A$. Thus if $\mathbf{x}$ satisfies $A^TA\mathbf{x} = A^T\mathbf{p}$, then it also satisfies $\mathbf{p} = A\mathbf{p}_A = A\mathbf{x}$.

$\mathbf{5.1.7}$. Since such a matrix has to map any vector $(x, y, z)^T$ to $(x, y, 0)^T$, it must be the matrix

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

$\mathbf{5.1.9}$. The solution of Exercise 5.1.3 gives a matrix A with independent columns, whose column space is the plane $x = y$. Thus, by Theorem 5.1.1, a matrix of the projection onto this plane is given by $P = A(A^TA)^{-1}A^T$ Now

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

and

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

Hence

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

Using this P, we obtain the projections of the given vectors as

$$\mathbf{q}_1 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix},$$

$$\mathbf{q}_2 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 3/2 \\ 3 \end{bmatrix},$$

and

$$\mathbf{q}_3 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ -2 \end{bmatrix}.$$

$\mathbf{5.1.11}$.

$$\left(A(A^TA)^{-1}A^T\right)^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T,$$

because the A^TA in the middle can be cancelled with one of the $(A^TA)^{-1}$ factors. Also,

$$(A(A^TA)^{-1}A^T)^T = A^{TT}((A^TA)^{-1})^TA^T = A((A^TA)^T)^{-1}A^T = A(A^TA)^{-1}A^T.$$

$\mathbf{5.1.13}$. This "proof" makes use of P^{-1}, which does not exist for any projection matrix other than I. Indeed, if P^{-1} exists, then multiplying both sides of the idempotency relation $P^2 = P$ by it results in $P = I$.

$\mathbf{5.1.15}$. We have Null(A) $=$ Left-null(A^T) $=$ Col$(A^T)^\perp$, and so, by the result of Exercise 5.1.12, with A^T in place of A, the matrix of the projection onto Null(A) is $P = I - A^T(AA^T)^{-1}A$.

$\mathbf{5.1.17}$. The normal system can be written as

$$\begin{bmatrix} \Sigma x_i^2 & \Sigma x_i \\ \Sigma x_i & m \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \Sigma x_i y_i \\ \Sigma y_i \end{bmatrix}.$$

The second row corresponds to the equation

$$a \sum x_i + mb = \sum y_i.$$

Division by m gives $y = ax + b$ with $x = \frac{1}{m}\sum x_i$ and $y = \frac{1}{m}\sum y_i$. The latter are the coordinates of the centroid of the given points.

$\mathbf{5.1.19}$. The third equation of the normal system is

$$a \sum x_i + b \sum y_i + mc = \sum z_i.$$

Division by m gives $z = ax + by + c$ with $x = \frac{1}{m}\sum x_i,\ y = \frac{1}{m}\sum y_i$, and $z = \frac{1}{m}\sum z_i$. The latter are the coordinates of the centroid of the given points.

5.2.1. Normalizing the given vectors to $\mathbf{a}_1 = \frac{1}{3}(2,1,2)^T$ and $\mathbf{a}_2 = \frac{1}{\sqrt{2}}(1,0,-1)^T$, we can apply Equation 5.49, since these vectors are obviously orthogonal. Now $\mathbf{a}_1 \cdot \mathbf{x} = \frac{1}{3}(2,1,2)^T(2,3,4) = \frac{15}{3}$ and $\mathbf{a}_2 \cdot \mathbf{x} = \frac{1}{\sqrt{2}}(1,0,-1)^T(2,3,4) = -\frac{2}{\sqrt{2}}$. Thus the required projection is

$$\mathbf{x}_C = \frac{15}{9}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + \frac{-2}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 7 \\ 5 \\ 13 \end{bmatrix}.$$

5.2.3. a. The subspace U is spanned by the orthonormal vectors $\mathbf{q}_1,\ \mathbf{q}_2$, and $\mathbf{q}_3$ given by the Equations 5.56. Thus we may use Equation 5.49 with these vectors in place of the $\mathbf{a}_i$ there. Now

$$\mathbf{q}_1 \cdot \mathbf{x} = \frac{1}{\sqrt{6}}(2,0,-1,1)^T(1,2,3,-1) = \frac{-2}{\sqrt{6}},$$

$$\mathbf{q}_2 \cdot \mathbf{x} = \frac{1}{\sqrt{6}}(0,2,1,1)^T(1,2,3,-1) = \sqrt{6},$$

and

$$\mathbf{q}_3 \cdot \mathbf{x} = \frac{1}{\sqrt{102}}(-3,-5,2,8)^T(1,2,3,-1) = \frac{-15}{\sqrt{102}}.$$

Thus, substituting into Equation 5.49, we get

$$\mathbf{x}_U = \frac{-2}{6}\begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 1 \\ 1 \end{bmatrix} - \frac{15}{102}\begin{bmatrix} -3 \\ -5 \\ 2 \\ 8 \end{bmatrix} = \frac{1}{102}\begin{bmatrix} -23 \\ 279 \\ 106 \\ -52 \end{bmatrix}.$$

b. The projection $\mathbf{x}_{U^\perp}$ of $\mathbf{x}$ into $U^\perp$ can be obtained simply by subtracting

$\mathbf{x}_U$ from $\mathbf{x}$:

$$\mathbf{x}_{U^\perp} = \mathbf{x} - \mathbf{x}_U = \begin{bmatrix} 1 \\ 2 \\ 3 \\ -1 \end{bmatrix} - \frac{1}{102}\begin{bmatrix} -23 \\ 279 \\ 106 \\ -52 \end{bmatrix} = \frac{25}{102}\begin{bmatrix} 5 \\ -3 \\ 8 \\ -2 \end{bmatrix}.$$

c. Since $U^\perp$ is one-dimensional, the vector $\mathbf{x}_{U^\perp}$ forms a basis for it.

d. To obtain an appropriate $\mathbf{q}_4$, we just have to normalize $\mathbf{x}_{U^\perp}$. Actually, we can ignore the factor 25/102, and normalize without it. $5^2+3^2+8^2+2^2 = 102$, and so $\mathbf{q}_4 = \frac{1}{\sqrt{102}}(5,-3,8,-2)^T$.

5. 2. 5. Apply first $|Q\mathbf{x}| = |\mathbf{x}|$ to $\mathbf{x} = \mathbf{e}_i$. Now

$$Q\mathbf{e}_1 = (\mathbf{q}_1\ \mathbf{q}_2\ \cdots\ \mathbf{q}_n)\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 1\mathbf{q}_1 + 0\mathbf{q}_2 + \cdots + 0\mathbf{q}_n = \mathbf{q}_1$$

and similarly $Q\mathbf{e}_i = \mathbf{q}_i$ for all i. Thus $|Q\mathbf{e}_i| = |\mathbf{e}_i| = 1$ implies $|\mathbf{q}_i| = 1$.

Next, apply $|Q\mathbf{x}| = |\mathbf{x}|$ to $\mathbf{x} = \mathbf{e}_i + \mathbf{e}_j$ with $i \neq j$:

$$\begin{aligned} |\mathbf{q}_i + \mathbf{q}_j|^2 &= |Q\mathbf{e}_i + Q\mathbf{e}_j|^2 = |Q(\mathbf{e}_i + \mathbf{e}_j)|^2 = |\mathbf{e}_i + \mathbf{e}_j|^2 \\ &= |\mathbf{e}_i|^2 + 2\mathbf{e}_i^T\mathbf{e}_j + |\mathbf{e}_j|^2 = 1 + 0 + 1 = 2. \end{aligned}$$

On the other hand,

$$|\mathbf{q}_i + \mathbf{q}_j|^2 = |\mathbf{q}_i|^2 + 2\mathbf{q}_i^T\mathbf{q}_j + |\mathbf{q}_j|^2 = 2 + 2\mathbf{q}_i^T\mathbf{q}_j.$$

Comparing this result with the previous equation, we see that $\mathbf{q}_i^T\mathbf{q}_j = 0$ must hold. Thus the columns of Q are orthonormal, which means that Q is an orthogonal matrix.

$\mathbf{5.2.7}$. The vectors $\mathbf{q}_1 = \frac{1}{3}(-1,2,2)^T$ and $\mathbf{q}_2 = \frac{1}{3}(2,-1,2)^T$ are orthonormal. Thus, we need to find a vector $\mathbf{x}$ of length 1 such that $\mathbf{q}_1^T\mathbf{x} = 0$ and $\mathbf{q}_2^T\mathbf{x} = 0$ hold. With $\mathbf{x} = (x,y,z)^T$, the above equations reduce to $-x+2y+2z = 0$ and $2x-y+2z = 0$. Subtracting the first equation from the second one, we get $3x-3y = 0$ or $x = y$. Letting $x = y = 2$, we obtain

$z = -1$. Thus $\mathbf{x} = (2, 2, -1)^T$ is a vector orthogonal to $\mathbf{q}_1$ and $\mathbf{q}_2$. Its length is 3, and so $\mathbf{q}_3 = \frac{1}{3}(2, 2, -1)^T$ is a unit vector in the same direction. Thus $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ form an orthonormal basis for $\mathbb{R}^3$.

5.2.9. By Exercise 2.4.13a, $\sum_{i=1}^n \mathbf{q}_i \mathbf{q}_i^T = QQ^T$ and, by Theorem 5.2.2, $Q^T = Q^{-1}$. Consequently, $\sum_{i=1}^n \mathbf{q}_i \mathbf{q}_i^T = QQ^{-1} = I$.

6.1.1.

$$\begin{aligned}
\begin{vmatrix} 2 & -3 & 2 \\ 1 & 4 & 0 \\ 0 & 1 & -5 \end{vmatrix} &= (-1)\begin{vmatrix} 1 & 4 & 0 \\ 2 & -3 & 2 \\ 0 & 1 & -5 \end{vmatrix} = \\
(-1)\begin{vmatrix} 1 & 4 & 0 \\ 0 & -11 & 2 \\ 0 & 1 & -5 \end{vmatrix} &= (-1)^2\begin{vmatrix} 1 & 4 & 0 \\ 0 & 1 & -5 \\ 0 & -11 & 2 \end{vmatrix} \\
&= (-1)^2\begin{vmatrix} 1 & 4 & 0 \\ 0 & 1 & -5 \\ 0 & 0 & -53 \end{vmatrix} = -53.
\end{aligned}$$

6.1.3.

$$\begin{aligned}
\begin{vmatrix} -1 & 1 & 2 & 3 \\ 2 & 0 & -5 & 0 \\ 0 & 0 & 0 & -1 \\ 3 & -3 & 1 & 4 \end{vmatrix} &= \begin{vmatrix} -1 & 1 & 2 & 3 \\ 0 & 2 & -1 & 6 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 7 & 13 \end{vmatrix} \\
&= (-1)\begin{vmatrix} -1 & 1 & 2 & 3 \\ 0 & 2 & -1 & 6 \\ 0 & 0 & 7 & 13 \\ 0 & 0 & 0 & -1 \end{vmatrix} = -14.
\end{aligned}$$

6.1.5. Axiom 2 does not apply, because it relates two distinct columns of a matrix to each other, and when $n = 1$, there is only one column, since $A = [a_{11}]$.

Axiom 1 becomes $\det(s\mathbf{a}_1 + t\mathbf{a}_1') = s\det(\mathbf{a}_1) + t\det(\mathbf{a}_1')$. Choosing $s = a_{11}$, $t = 0$, and $\mathbf{a}_1 = \mathbf{e}_1 = [1]$, we can thus write $\det(A) = \det(a_{11}) = \det(a_{11}\mathbf{e}_1) = a_{11}\det(\mathbf{e}_1)$, which, by Axiom 3, equals a_{11}.

6.1.7. This result could be proved much the same way as Theorem 6.1.7 was in the text. Alternatively, it follows from Theorems 6.1.6 and 6.1.7, since

A^T is upper triangular:

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ 0 & a_{22} & a_{32} \\ 0 & 0 & a_{33} \end{bmatrix}.$$

6.1.9. If A invertible, then $AA^{-1} = I$. Hence, by Theorem 6.1.9 and Axiom 3 of Definition 6.1.1, we have $|A||A^{-1}| = |AA^{-1}| = |I| = 1$. Dividing by $|A|$, we get $|A^{-1}| = 1/|A|$.

6.1.11. If AB is invertible, then $\det(AB) = \det(A)\det(B) \neq 0$, and so neither $\det(A)$ nor $\det(B)$ is zero, which implies that A and B are invertible. Conversely, if A and B are invertible, then neither $\det(A)$ nor $\det(B)$ is zero, and so $\det(AB) = \det(A)\det(B) \neq 0$, which implies that AB is invertible.

6.1.13. By the definition of an orthogonal matrix, $Q^TQ = I$. Then Theorems 6.1.9 and 6.1.6 give $|Q^TQ| = |Q^T||Q| = |Q|^2$. On the other hand, by Axiom 3 of Definition 6.1.1, we have $|Q^TQ| = |I| = 1$, and so $|Q|^2 = 1$, from which we get the desired result: $\det(Q) = \pm 1$.

6.1.15. The formula for the Vandermonde determinant of order $n > 3$ can be expressed as

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (a_j - a_i).$$

The notation on the right-hand side represents the product of all factors of the form $a_j - a_i$, where i and j are subscripts from 1 to n, with $i < j$.

To prove the formula, we will show how to express it for the case of an order n determinant in terms of the corresponding one of order $n-1$. Starting with the Vandermonde determinant of order n, subtracting the first row from each of the succeeding rows results in the equivalent determinant

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ 0 & a_3 - a_1 & a_3^2 - a_1^2 & \cdots & a_3^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix},$$

which is equal to

$$\begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ a_3 - a_1 & a_3^2 - a_1^2 & \cdots & a_3^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

Each of the terms in the first column is a factor of all of the terms in its row and thus may be factored out of the determinant, resulting in the equivalent expression

$$\prod_{j=2}^{n}(a_j - a_1) \begin{vmatrix} 1 & a_2+a_1 & a_2^2+a_2a_1 + a_1^2 & \cdots & a_2^{n-1}+\cdots + a_1^{n-1} \\ 1 & a_3+a_1 & a_3^2+a_3a_1 + a_1^2 & \cdots & a_3^{n-1}+\cdots + a_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n+a_1 & a_n^2+a_na_1 + a_1^2 & \cdots & a_n^{n-1}+\cdots + a_1^{n-1} \end{vmatrix}.$$

The last determinant can be simplified by multiplying the first column by a_1 and subtracting the product from the second column, then multiplying the first column by a_1^2 and subtracting the product from the third column, etc., resulting in the equivalent expression

$$\prod_{j=2}^{n}(a_j - a_1) \begin{vmatrix} 1 & a_2 & a_2^2+a_2a_1 & \cdots & a_2^{n-2}+\cdots + a_1^{n-3}a_2 \\ 1 & a_3 & a_3^2+a_3a_1 & \cdots & a_3^{n-2}+\cdots + a_1^{n-3}a_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2+a_na_1 & \cdots & a_n^{n-2}+\cdots + a_1^{n-3}a_n \end{vmatrix}.$$

The last expression can be further simplified by multiplying the second column by a_1 and subtracting the product from the third column, then multiplying the second column by a_1^2 and subtracting the product from the fourth column, etc., resulting in the equivalent expression

$$\prod_{j=2}^{n}(a_j - a_1)\begin{vmatrix} 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} + \cdots + a_1^{n-4}a_2^2 \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-2} + \cdots + a_1^{n-4}a_3^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-2} + \cdots + a_1^{n-4}a_n^2 \end{vmatrix}.$$

Continuing to simplify in a similar fashion leads to the following expression, which is equal to the original Vandermonde determinant of order n :

$$\prod_{j=2}^{n}(a_j - a_1)\begin{vmatrix} 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-2} \end{vmatrix}.$$

Note that the determinant in this expression is another Vandermonde determinant, but of order $n-1$. Evaluating it in the same way indicated above, results in the expression

$$\prod_{j=2}^{n}(a_j - a_1)\prod_{k=3}^{n}(a_k - a_2)\begin{vmatrix} 1 & a_3 & a_3^2 & \cdots & a_3^{n-3} \\ 1 & a_4 & a_4^2 & \cdots & a_4^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-3} \end{vmatrix},$$

and then similarly evaluating the resulting Vandermonde determinants of lower order leads eventually to the desired formula. (A more rigorous proof would employ mathematical induction. The argument above provides the crucial inductive step in such a proof.)

6.2.1. Expanding along the first column, we get

$$\begin{aligned} \begin{vmatrix} 2 & -3 & 2 \\ 1 & 4 & 0 \\ 0 & 1 & -5 \end{vmatrix} &= 2\begin{vmatrix} 4 & 0 \\ 1 & -5 \end{vmatrix} - (-3)\begin{vmatrix} 1 & 0 \\ 0 & -5 \end{vmatrix} + 2\begin{vmatrix} 1 & 4 \\ 0 & 1 \end{vmatrix} \\ &= 2\,(-20+0) + 3\,(-5+0) + 2\,(1-0) = -53. \end{aligned}$$

6.2.3. Expansion along the third row, and then along the second row, gives:

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

$$-2\begin{vmatrix} 1 & 2 \\ -3 & 1 \end{vmatrix} - (-5)\begin{vmatrix} -1 & 1 \\ 3 & -3 \end{vmatrix} = -2\,(1+6) + 5\,(3-3) = -14.$$

6.2.5. Expanding the determinant in the numerator of Cramer's rule along the column $\mathbf{b}$, we obtain

$$x_i = \frac{1}{|A|}\sum_j b_j A_{ji}.$$

Substituting this expression for x_i into $A\mathbf{x}$ results in

$$(A\mathbf{x})_k = \sum_i a_{ki}x_i = \frac{1}{|A|}\sum_i \sum_j b_j a_{ki} A_{ji} = \frac{1}{|A|}\sum_j (b_j \sum_i a_{ki} A_{ji}).$$

By Lemma 6.2.1, the inside sum on the right is 0 if $j \neq k$, and, by Theorem 6.2.1, it is $|A|$ if $j = k$. Thus the sum over j has only one nonzero term, which equals $b_k|A|$. Therefore $(A\mathbf{x})_k = b_k$ and $A\mathbf{x} = \mathbf{b}$, as we had to show.

6.2.7. Using Cramer's Rule and the expansion of Theorem 6.1.3, we get

$$\begin{aligned} x_1 &= \begin{vmatrix} 1 & 1 & 2 \\ 2 & 0 & 4 \\ 3 & 3 & 1 \end{vmatrix} \div \begin{vmatrix} 1 & 1 & 2 \\ 2 & 0 & 4 \\ 0 & 3 & 1 \end{vmatrix} \\ &= \frac{(0+12+12)-(12+2+0)}{(0+12+0)-(12+2+0)} = \frac{10}{-2} = -5, \end{aligned}$$

and similarly, $x_2 = 0,\ x_3 = 3.$

6.2.9. $A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$ implies $A_{11} = 1,\ A_{12} = -2,\ A_{21} = -3,\ A_{22} = 1,$

and so $\text{adj}(A) = \begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix}$. Hence

$$A^{-1} = \frac{1}{|A|}\,\text{adj}(A) = \frac{1}{-5}\begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix}.$$

6.2.11. From Theorem 6.2.3, $\text{adj}(A) = |A|A^{-1}$. Apply the result $\det(cA) = c^n \det(A)$ of Exercise 6.1.6 here, with $c = |A|$, and A^{-1} in place of A, to get $\det(\text{adj}(A)) = \det(|A|A^{-1}) = |A|^n \det(A^{-1})$. Now, by Corollary 6.1.2, $\det(A^{-1}) = 1/|A|$, and so

$$\det(\text{adj}(A)) = (\det(A))^{n-1}.$$

6.2.13. The area A of the triangle with vertices $(a_1, a_2),\ (b_1, b_2),\ (c_1, c_2)$ is half that of the parallelogram spanned by the vectors $\mathbf{b} - \mathbf{a}$ and $\mathbf{c} - \mathbf{a}$, where $\mathbf{a} = (a_1, a_2),\ \mathbf{b} = (b_1, b_2)$, and $\mathbf{c} = (c_1, c_2)$. Thus, by Theorem 6.2.4,

$$\begin{aligned}
\pm A &= \frac{1}{2}|\mathbf{b} - \mathbf{a},\ \mathbf{c} - \mathbf{a}| = \frac{1}{2}\begin{vmatrix} b_1 - a_1 & c_1 - a_1 \\ b_2 - a_2 & c_2 - a_2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} b_1 - a_1 & b_2 - a_2 \\ c_1 - a_1 & c_2 - a_2 \end{vmatrix} \\
&= \frac{1}{2}\begin{vmatrix} 0 & 0 & 1 \\ b_1 - a_1 & b_2 - a_2 & 0 \\ c_1 - a_1 & c_2 - a_2 & 0 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} a_1 & a_2 & 1 \\ b_1 - a_1 & b_2 - a_2 & 0 \\ c_1 - a_1 & c_2 - a_2 & 0 \end{vmatrix} \\
&= \frac{1}{2}\begin{vmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{vmatrix}.
\end{aligned}$$

6.2.15. If we expand the determinant along the first row, then we obtain a linear combination of the elements of that row, which can be rearranged into the form $y = ax^2 + bx + c$, since the coefficient of y is a nonzero determinant (see Exercise 6.1.14). Note that, by the result in Exercise 6.2.13, the coefficient of x^2 is also a nonzero determinant if the points $(x_1, y_1),\ (x_2, y_2)$, and (x_3, y_3) are noncollinear. If they are collinear, then we get an equation of a straight line, which may be considered to be a degenerate parabola. If we substitute the coordinates of any of the given points for x, y, then two rows become equal, which makes the determinant vanish. Thus the given points lie on the parabola.

6.2.17. Applying the formula in Exercise 6.2.16, the equation of the circle can be expressed in the form

$$\begin{vmatrix} x^2+y^2 & x & y & 1 \\ 0 & 0 & 0 & 1 \\ 5 & 2 & -1 & 1 \\ 16 & 4 & 0 & 1 \end{vmatrix} = 0.$$

Expanding along the second row yields

$$\begin{vmatrix} x^2+y^2 & x & y \\ 5 & 2 & -1 \\ 16 & 4 & 0 \end{vmatrix} = 0.$$

Then expanding along the top row results in

$$\begin{vmatrix} 2 & -1 \\ 4 & 0 \end{vmatrix}(x^2+y^2) - \begin{vmatrix} 5 & -1 \\ 16 & 0 \end{vmatrix} x + \begin{vmatrix} 5 & 2 \\ 16 & 4 \end{vmatrix} y = 0.$$

From here, evaluating the coefficient determinants, we obtain

$$4(x^2+y^2) - 16x - 12y = 0, \text{ or, } x^2+y^2-4x-3y = 0.$$

Completing the square of the terms in x and y results in

$$(x-2)^2 + (y-3/2)^2 = 25/4,$$

and so the circle has center $(2, 3/2)$ and radius $5/2$.

6.2.19. If we expand the determinant along the first row, then we obtain a linear combination of the elements of that row. That combination is a linear equation in x, y, and z, which is the equation of a plane. If we substitute the coordinates of any of the given points for x, y, and z, then two rows become equal, which makes the determinant vanish. Thus the given points lie on the plane.

6.3.1. If $\mathbf{u} = (1, -1, 0)^T$ and $\mathbf{v} = (1, 2, 0)^T$, then

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -1 & 0 \\ 1 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 3\mathbf{k}.$$

The triangle spanned by $\mathbf{u}$ and $\mathbf{v}$ has an altitude, parallel to the x-axis, of length 1 and a corresponding base of length 3. Its area is 3/2, and thus the area of the parallelogram spanned by the two vectors is 3, which is equal to $|\mathbf{u} \times \mathbf{v}| = |3\mathbf{k}|$.

$\mathbf{6.3.3}$. First, we will verify Statement 11 of Theorem 6.3.1:

If $\mathbf{u} = (1, -1, 0)^T,\ \mathbf{v} = (1, 2, 0)^T$, and $\mathbf{w} = (1, 0, 3)^T$, then

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 0 \\ 1 & 0 & 3 \end{vmatrix} = 6\mathbf{i} - 3\mathbf{j} - 2\mathbf{k}$$

$$\mathbf{w} \times \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 3 \\ 1 & -1 & 0 \end{vmatrix} = 3\mathbf{i} + 3\mathbf{j} - \mathbf{k}$$

and, as shown in Exercise 6.3.1,

$$\mathbf{u} \times \mathbf{v} = 3\mathbf{k}.$$

Thus

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (1, -1, 0)^T \cdot (6, -3, -2)^T = 6 + 3 + 0 = 9,$$

$$\mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = (1, 2, 0)^T \cdot (3, 3, -1)^T = 3 + 6 + 0 = 9,$$

$$\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = (1, 0, 3)^T \cdot (0, 0, 3)^T = 0 + 0 + 9 = 9,$$

and

$$\det(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \begin{vmatrix} 1 & 1 & 1 \\ -1 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 9.$$

Next we will verify Statement 12:

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (0, 0, 3)^T \times (1, 0, 3)^T = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 0 & 3 \\ 1 & 0 & 3 \end{vmatrix} = 3\mathbf{j},$$

and

$$\begin{aligned}&(\mathbf{u}\cdot\mathbf{w})\mathbf{v}-(\mathbf{v}\cdot\mathbf{w})\mathbf{u}\\ =\ &[(1,-1,0)^T\cdot(1,0,3)^T](1,2,0)^T-[(1,2,0)^T\cdot(1,0,3)](1,-1,0)^T\\ =\ &(1,2,0)^T-(1,-1,0)^T=(0,3,0)^T=3\mathbf{j}.\end{aligned}$$

Finally, we will verify Statement 13:

$$\begin{aligned}\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) &= (1,-1,0)^T\times(6,-3,-2)^T=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 1 & -1 & 0\\ 6 & -3 & -2\end{vmatrix}\\ &= 2\mathbf{i}+2\mathbf{j}+3\mathbf{k}=(2,2,3)^T,\end{aligned}$$

and

$$\begin{aligned}&(\mathbf{u}\cdot\mathbf{w})\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\mathbf{w}\\ =\ &[(1,-1,0)^T\cdot(1,0,3)^T](1,2,0)^T-[(1,-1,0)^T\cdot(1,2,0)^T](1,0,3)^T\\ =\ &(1,2,0)^T+(1,0,3)^T=(2,2,3)^T.\end{aligned}$$

6.3.5. Geometrically, in the case of nonzero vectors, $\mathbf{v}\times\mathbf{w}$ is a vector whose magnitude is equal to the area of the parallelogram spanned by the vectors $\mathbf{v}$ and $\mathbf{w}$ and whose direction is orthogonal to the plane of that parallelogram. Thus, if $\mathbf{n}=(\mathbf{v}\times\mathbf{w})/|\mathbf{v}\times\mathbf{w}|$ denotes the unit vector in the direction of $\mathbf{v}\times\mathbf{w}$, then $|\mathbf{u}\cdot\mathbf{n}|$ is equal to the length of the orthogonal projection of $\mathbf{u}$ onto the direction of $\mathbf{v}\times\mathbf{w}$, which is the height of the parallelepiped spanned by the vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. Therefore, the volume V of that parallelepiped is given by

$$\begin{aligned}V &= \text{height}\times\text{area of base } = |\mathbf{u}\cdot\mathbf{n}|\,|\mathbf{v}\times\mathbf{w}|\\ &= \left|\mathbf{u}\cdot\frac{\mathbf{v}\times\mathbf{w}}{|\mathbf{v}\times\mathbf{w}|}\right||\mathbf{v}\times\mathbf{w}|=|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})|.\end{aligned}$$

If either $\mathbf{v}$ or $\mathbf{w}$ is $\mathbf{0}$, then the parallelepiped spanned by the three vectors degenerates into one of zero volume, and the scalar triple product is also zero in this case.

6.3.7. If we expand the given expression by applying Statement 11 in Theorem 6.3.1, with $\mathbf{u}=\mathbf{a}\times\mathbf{b},\ \mathbf{v}=\mathbf{c}$, and $\mathbf{w}=\mathbf{d}$, then we obtain

$$(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})=\mathbf{c}\cdot(\mathbf{d}\times(\mathbf{a}\times\mathbf{b}))$$

Evaluating the triple cross product in the expression on the right side of the above equation with the aid of Statement 12 in Theorem 6.3.1 results in

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= \mathbf{c}\cdot[(\mathbf{d}\cdot\mathbf{b})\mathbf{a}-(\mathbf{d}\cdot\mathbf{a})\mathbf{b}] \\ &= (\mathbf{d}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{a})-(\mathbf{d}\cdot\mathbf{a})(\mathbf{c}\cdot\mathbf{b}) \\ &= (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}).\end{aligned}$$

6.3.9. a. 1. T is additive: By Part 6 of Theorem 6.3.1, we have $T(\mathbf{x}+\mathbf{y}) = \mathbf{a}\times(\mathbf{x}+\mathbf{y}) = \mathbf{a}\times\mathbf{x}+\mathbf{a}\times\mathbf{y} = T(\mathbf{x})+T(\mathbf{y})$.

2. T is homogeneous: By Part 1 of Theorem 6.3.1 we have $T(c\mathbf{x}) = \mathbf{a}\times(c\mathbf{x}) = c(\mathbf{a}\times\mathbf{x}) = cT(\mathbf{x})$.

b. The columns of $[T]$ are given by the action of T on the standard vectors. Thus

$\mathbf{t}_1 = T(\mathbf{i}) = \mathbf{a}\times\mathbf{i} = (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\times\mathbf{i} = -a_2\mathbf{k}+a_3\mathbf{j} = (0, a_3, -a_2)^T$,
$\mathbf{t}_2 = T(\mathbf{j}) = \mathbf{a}\times\mathbf{j} = (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\times\mathbf{j} = a_1\mathbf{k}-a_3\mathbf{i} = (-a_3, 0, a_1)^T$,
$\mathbf{t}_3 = T(\mathbf{k}) = \mathbf{a}\times\mathbf{k} = (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\times\mathbf{k} = -a_1\mathbf{j}+a_2\mathbf{i} = (a_2, -a_1, 0)^T$.

Hence

$$[T] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}.$$

c. Solve $[T]\mathbf{x} = \mathbf{0}$ by row reduction, assuming without loss of generality $a_3 \neq 0$:

$$\left[\begin{array}{ccc|c} 0 & -a_3 & a_2 & 0 \\ a_3 & 0 & -a_1 & 0 \\ -a_2 & a_1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} a_3 & 0 & -a_1 & 0 \\ 0 & -a_3 & a_2 & 0 \\ -a_2 & a_1 & 0 & 0 \end{array}\right] \to$$
$$\left[\begin{array}{ccc|c} a_3 & 0 & -a_1 & 0 \\ 0 & -a_3 & a_2 & 0 \\ 0 & a_1 & -a_1a_2/a_3 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} a_3 & 0 & -a_1 & 0 \\ 0 & -a_3 & a_2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Now x_3 is free, and so, set $x_3 = t$. Then $-a_3x_2 + a_2t = 0$ gives $x_2 = a_2t/a_3$ and $a_3x_2 - a_1t = 0$ gives $x_1 = a_1t/a_3$. If we set $t = ca_3$, then we obtain from the above $\mathbf{x} = c\mathbf{a}$. Thus the nullspace of the matrix $[T]$ is the line of $\mathbf{a}$. We could have found this result alternatively by observing that $[T]\mathbf{x} = \mathbf{0}$ is equivalent to $T(\mathbf{x}) = \mathbf{a}\times\mathbf{x} = \mathbf{0}$, and the cross product is zero if and only if its factors are parallel.

d. Since, for any matrix, rank + nullity = number of columns, in the present case we have $\operatorname{rank}([T]) = 3 - 1 = 2$. Alternatively, the rank is the number of non-zero rows in the echelon form, which is clearly 2 in the matrix above, because of the assumption $a_3 \neq 0$. Since $\mathbf{a}$ was assumed to be nonzero, at least one of its components has to be nonzero, and we end up with two nonzero rows in the echelon form, regardless of which component is nonzero.

7.1.1. The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 3 \\ 3 & 2-\lambda \end{vmatrix} = 0,$$

or equivalently, $\lambda^2 - 4\lambda - 5 = (\lambda + 1)(\lambda - 5) = 0$. The eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 5$.

To find the eigenvectors corresponding to $\lambda_1 = -1$, we need to solve the equation

$$(A - \lambda_1 I)\mathbf{s} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(-1, 1)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_1 = -1$ form a one-dimensional subspace of $\mathbb{R}^2$ with basis vector $\mathbf{s}_1 = (-1, 1)^T$

To find the eigenvectors corresponding to $\lambda_2 = 5$, we need to solve the equation

$$(A - \lambda_2 I)\mathbf{s} = \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(1, 1)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_2 = 5$ form a one-dimensional subspace of $\mathbb{R}^2$ with basis vector $\mathbf{s}_2 = (1, 1)^T$.

7.1.3. The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 0-\lambda & 0 \\ 0 & 0-\lambda \end{vmatrix} = 0,$$

or equivalently, $\lambda^2 = 0$. The only solution is $\lambda = 0$. Thus 0 is the only eigenvalue and it has multiplicity 2. To find the corresponding eigenvectors, we need to solve $(A - \lambda I)\mathbf{s} = \mathbf{0}$. Clearly, every $\mathbf{s} \in \mathbb{R}^2$ is a solution, that is, an eigenvector. In other words, $\lambda = 0$ is the sole eigenvalue with the corresponding eigenspace being the whole of $\mathbb{R}^2$.

7.1.5. The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} = 0,$$

or equivalently, $(2-\lambda)^3 - (2-\lambda) = 0$. Factoring on the left, we may write the equation in the form $(2-\lambda)[(2-\lambda)^2 - 1] = 0$, or $(2-\lambda)(1-\lambda)(3-\lambda) = 0$. Thus the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$.

To find the eigenvectors corresponding to $\lambda_1 = 1$, we need to solve the equation

$$(A - \lambda_1 I)\mathbf{s} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(-1, 0, 1)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_1 = 1$ form a one-dimensional subspace of $\mathbb{R}^3$ with basis vector $\mathbf{s}_1 = (-1, 0, 1)^T$

Next, to find the eigenvectors corresponding to $\lambda_2 = 2$, we need to solve the equation

$$(A - \lambda_2 I)\mathbf{s} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(0, 1, 0)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_2 = 2$ form a one-dimensional subspace of $\mathbb{R}^3$ with basis vector $\mathbf{s}_2 = (0, 1, 0)^T$.

Finally, to find the eigenvectors corresponding to $\lambda_3 = 3$, we need to solve the equation

$$(A - \lambda_3 I)\mathbf{s} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(1, 0, 1)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_3 = 3$ form a one-dimensional subspace of $\mathbb{R}^3$ with basis vector $\mathbf{s}_3 = (1, 0, 1)^T$.

7.1.7. The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 0 & -1 \\ -2 & 3-\lambda & -1 \\ -6 & 6 & -\lambda \end{vmatrix} = 0.$$

We can reduce this equation to $-(1-\lambda)(3-\lambda)\lambda + 12 - 6(3-\lambda) + 6(1-\lambda) = 0$; or, equivalently, to $\lambda(1 - \lambda)(3 - \lambda) = 0$. Thus the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 3$.

To find the eigenvectors corresponding to $\lambda_1 = 0$, we need to solve the equation

$$(A - \lambda_1 I)\mathbf{s} = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 3 & -1 \\ -6 & 6 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(1, 1, 1)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_1 = 0$ form a one-dimensional subspace of $\mathbb{R}^3$ with basis vector $\mathbf{s}_1 = (1, 1, 1)^T$

Next, to find the eigenvectors corresponding to $\lambda_2 = 1$, we need to solve the equation

$$(A - \lambda_2 I)\mathbf{s} = \begin{bmatrix} 0 & 0 & -1 \\ -2 & 2 & -1 \\ -6 & 6 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(1, 1, 0)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_2 = 1$ form a one-dimensional subspace of $\mathbb{R}^3$ with basis vector $\mathbf{s}_2 = (1, 1, 0)^T$

Finally, to find the eigenvectors corresponding to $\lambda_3 = 3$, we need to solve the equation

$$(A - \lambda_3 I)\mathbf{s} = \begin{bmatrix} -2 & 0 & -1 \\ -2 & 0 & -1 \\ -6 & 6 & -3 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The solutions of this equation are of the form $\mathbf{s} = s(-1, 0, 2)^T$. Thus the eigenvectors belonging to the eigenvalue $\lambda_3 = 3$ form a one-dimensional subspace of $\mathbb{R}^3$ with basis vector $\mathbf{s}_3 = (-1, 0, 2)^T$.

7.1.9. If the scalar $c \neq 0$, then we have

$$|cA - \lambda I| = \left| c(A - \frac{\lambda}{c} I) \right| = c^n \left| A - \frac{\lambda}{c} I \right|.$$

It follows from this relation that λ is an eigenvalue of the matrix cA with $c \neq 0$, if and only if λ/c is an eigenvalue of A, and that the eigenvectors corresponding to those related eigenvalues will be the same. On the other hand, if $c = 0$, then $cA = O$, and its only eigenvalue is $\lambda = 0$ with eigenspace $\mathbb{R}^n$. (See Exercise 7.1.3).

7.1.11. First suppose that the matrix A is singular. Then $A\mathbf{s} = \mathbf{0}$ has a nontrivial solution $\mathbf{s}$, and $A\mathbf{s} = 0\mathbf{s}$ for all such solutions. This equation shows that 0 is an eigenvalue of A, with all solutions $\mathbf{s}$ of $A\mathbf{s} = \mathbf{0}$ as its eigenvectors.

In the other direction, suppose that 0 is an eigenvalue of A. Then for all corresponding nonzero eigenvectors $\mathbf{s}$ we have $A\mathbf{s} = 0\mathbf{s} = \mathbf{0}$. By Part 6 of Theorem 2.5.5, it follows that A is singular.

7.1.13. Suppose that $\mathbf{s}$ is an eigenvector of a nonsingular matrix A, belonging to the eigenvalue λ. Then $A\mathbf{s} = \lambda\mathbf{s}$. Since A is nonsingular, it is invertible, and $\lambda \neq 0$ (see Exercise 7.1.11). Therefore, $A^{-1}(A\mathbf{s}) = A^{-1}(\lambda\mathbf{s})$, and thus $\mathbf{s} = \lambda A^{-1}\mathbf{s}$, or $A^{-1}\mathbf{s} = \frac{1}{\lambda}\mathbf{s}$. Hence $\mathbf{s}$ is an eigenvector of A^{-1} belonging to the eigenvalue λ^{-1}.

7.1.15. Apply the matrix A to the vector $\mathbf{u}$. That gives $A\mathbf{u} = \mathbf{u}\mathbf{u}^T\mathbf{u} = \mathbf{u}$, because $\mathbf{u}^T\mathbf{u} = 1$. This shows that $\mathbf{u}$ is an eigenvector of A belonging to the eigenvalue 1. If $\mathbf{v}$ is any vector orthogonal to $\mathbf{u}$, then $A\mathbf{v} = \mathbf{u}\mathbf{u}^T\mathbf{v} = \mathbf{0}$, because $\mathbf{u}^T\mathbf{v} = 0$. Thus any such vector $\mathbf{v}$ is an eigenvector of A belonging to the eigenvalue 0. So the eigenvalues are 1 and (for $n > 1$ only) 0. The corresponding eigenspaces are the line of $\mathbf{u}$ and its orthogonal complement, respectively.

Now, any vector $\mathbf{x} \in \mathbb{R}^n$ can be decomposed uniquely as the sum of its projection $c\mathbf{u}$ onto $\mathbf{u}$ and a vector $\mathbf{v}$ orthogonal to $\mathbf{u}$. Thus $A\mathbf{x} = A(c\mathbf{u} + \mathbf{v}) = c\mathbf{u}$. This equation shows that A represents the projection onto the line of $\mathbf{u}$.

$\mathbf{7.1.17}$. According to the definition, $\mathbf{s}^T$ is a left eigenvector of a matrix A belonging to the eigenvalue λ if and only if the equation $\mathbf{s}^T A = \lambda \mathbf{s}^T$ holds. This last equation is true if and only if $(\mathbf{s}^T A)^T = (\lambda \mathbf{s}^T)^T$, or, $A^T \mathbf{s} = \lambda \mathbf{s}$, and thus if and only if $\mathbf{s}$ is an eigenvector of A^T belonging to the eigenvalue λ.

$\mathbf{7.1.19}$. Suppose that $\mathbf{s}$ is an eigenvector of a matrix A, belonging to two eigenvalues λ_1 and $\lambda_2 \neq \lambda_1$. Then $A\mathbf{s} = \lambda_1 \mathbf{s} = \lambda_2 \mathbf{s}$. Hence $(\lambda_1 - \lambda_2)\,\mathbf{s} = \mathbf{0}$ and so, since $(\lambda_1 - \lambda_2) \neq 0$, we must have $\mathbf{s} = \mathbf{0}$.

$\mathbf{7.2.1}$. By Exercise 7.1.11, if A is invertible, then all of its eigenvalues are different from zero, even if it is not diagonalizable.

Conversely, if A is diagonalizable, say as $A = S\Lambda S^{-1}$, and $\lambda_i \neq 0$ for all i, then

$$\Lambda^{-1} = \begin{bmatrix} \lambda_1^{-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-1} \end{bmatrix}$$

exists, and $X = S\Lambda^{-1}S^{-1}$ is A^{-1}. Indeed, $XA = S\Lambda^{-1}S^{-1}S\Lambda S^{-1} = S\Lambda^{-1}\Lambda S^{-1} = SS^{-1} = I$, and similarly $AX = I$ as well.

Furthermore, for a diagonalizable and invertible A and any positive integer

k,

$$\Lambda^{-k} = \left(\Lambda^{-1}\right)^k = \begin{bmatrix} \left(\lambda_1^{-1}\right)^k & 0 & \cdots & 0 \\ 0 & \left(\lambda_2^{-1}\right)^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \left(\lambda_n^{-1}\right)^k \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1^{-k} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-k} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-k} \end{bmatrix}$$

and

$$A^{-k} = \left(A^{-1}\right)^k = \left(S\Lambda^{-1}S^{-1}\right)^k = S\Lambda^{-k}S^{-1}.$$

7.2.3. By Theorem 7.3.2, any symmetric matrix can be diagonalized with an orthogonal matrix S, so that $A = S\Lambda S^{-1}$. Furthermore, if all eigenvalues of A are nonnegative, then writing

$$\Lambda^{1/2} = \begin{bmatrix} \lambda_1^{1/2} & 0 & \cdots & 0 \\ 0 & \lambda_2^{1/2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{1/2} \end{bmatrix},$$

we can compute $\sqrt{A}$ as $\sqrt{A} = S\Lambda^{1/2}S^{-1}$, because then

$$\left(\sqrt{A}\right)^2 = \left(S\Lambda^{1/2}S^{-1}\right)^2 = S\left(\Lambda^{1/2}\right)^2 S^{-1} = S\Lambda S^{-1} = A.$$

Furthermore, the orthogonality of S implies $S^T = S^{-1}$, and so $(\sqrt{A})^T = (S^{-1})^T(\Lambda^{1/2})^T S^T = S\Lambda^{1/2}S^{-1} = \sqrt{A}$, that is, $\sqrt{A}$ is also symmetric.

7.2.5. A can be diagonalized with an orthogonal matrix as $A = S\Lambda S^{-1}$. The eigenvalues and eigenvectors of A were computed in Exercise 7.1.5. The former make up the diagonal elements of Λ, and the latter, after normalization, the columns of S. Thus $A^{100} = S\Lambda^{100}S^{-1}$ can be evaluated as

$$\begin{aligned} A^{100} &= \frac{1}{2}\begin{bmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1^{100} & 0 & 0 \\ 0 & 2^{100} & 0 \\ 0 & 0 & 3^{100} \end{bmatrix}\begin{bmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 3^{100}+1 & 0 & 3^{100}-1 \\ 0 & 2^{101} & 0 \\ 3^{100}-1 & 0 & 3^{100}+1 \end{bmatrix}. \end{aligned}$$

7.2.7. If $B = SAS^{-1}$, then

$$\begin{aligned} |B - \lambda I| &= |S^{-1}AS - \lambda I| = |S^{-1}AS - S^{-1}\lambda IS| \\ &= |S^{-1}(A - \lambda I)S| = |S^{-1}||A - \lambda I||S| = |A - \lambda I|, \end{aligned}$$

since $|S^{-1}| = 1/|S|$.

7.2.9. If $S^{-1}AS = \Lambda$, then $A = S\Lambda S^{-1}$. The orthogonality of S implies $S^T = S^{-1}$, and so

$$A^T = (S\Lambda S^{-1})^T = (S^{-1})^T\Lambda^T S^T = S\Lambda S^{-1} = A.$$

7.2.11. In matrix form, the given equations become

$$\mathbf{x}(0) = \begin{bmatrix} 1000 \\ 1000 \end{bmatrix}$$

and

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

where

$$A = \begin{bmatrix} 0.8 & 0.4 \\ -0.8 & 2.0 \end{bmatrix},$$

and so

$$\mathbf{x}(n) = A^n\mathbf{x}(0)$$

for any positive integer n. We diagonalize A to compute A^n here.

The characteristic equation for this A is

$$|A - \lambda I| = \begin{vmatrix} 0.8-\lambda & 0.4 \\ -0.8 & 2.0-\lambda \end{vmatrix} = (0.8-\lambda)(2-\lambda) + 0.32 = 0.$$

The solutions are $\lambda_1 = 1.6$ and $\lambda_2 = 1.2$. The corresponding eigenvectors can be found by substituting these eigenvalues into $(A - \lambda I)\mathbf{s} = \mathbf{0}$:

$$(A - \lambda_1 I)\mathbf{s}_1 = \begin{bmatrix} 0.8 - 1.6 & 0.4 \\ -0.8 & 2.0 - 1.6 \end{bmatrix} \mathbf{s}_1 = \begin{bmatrix} -0.8 & 0.4 \\ -0.8 & 0.4 \end{bmatrix} \mathbf{s}_1 = \mathbf{0}.$$

A solution is $\mathbf{s}_1 = (1, 2)^T$.

For the other eigenvalue we have the equation

$$(A \ - \lambda_2 I)\mathbf{s}_2 = \begin{bmatrix} 0.8 - 1.2 & 0.4 \\ -0.8 & 2.0 - 1.2 \end{bmatrix} \mathbf{s}_2 = \begin{bmatrix} -0.4 & 0.4 \\ -0.8 & 0.8 \end{bmatrix} \mathbf{s}_2 = \mathbf{0}.$$

A solution is $\mathbf{s}_2 = (1, 1)^T$. Thus

$$S = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} 1.6 & 0 \\ 0 & 1.2 \end{bmatrix}$$

and

$$S^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}.$$

According to Corollary 3.6.1, the coordinate vectors of each $\mathbf{x}(n)$, for $n = 0, 1, \ldots$, relative to the basis S, are given by

$$\mathbf{x}_S(0) = S^{-1}\mathbf{x}(0) = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1000 \\ 1000 \end{bmatrix} = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$$

and

$$\begin{aligned} \mathbf{x}_S(n) &= S^{-1}\mathbf{x}(n) = S^{-1}A^n\mathbf{x}(0) = S^{-1}A^nS\mathbf{x}_S(0) = \Lambda^n\mathbf{x}_S(0) \\ &= \begin{bmatrix} 1.6^n & 0 \\ 0 & 1.2^n \end{bmatrix} \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = 1.2^n \begin{bmatrix} 0 \\ 1000 \end{bmatrix}. \end{aligned}$$

Hence the solution in the standard basis is given by

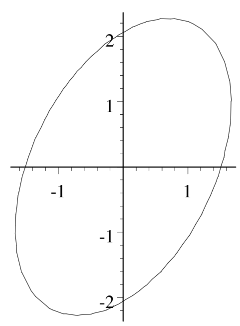
$$\mathbf{x}(n) = S\mathbf{x}_S(n) = 1.2^n \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = 1.2^n \begin{bmatrix} 1000 \\ 1000 \end{bmatrix}.$$

As the last equation shows, the proportion of predators to preys does not change over time if they were evenly split in the beginning, and the numbers of both increase by 20% each year, approaching infinity.

7.3.1. The matrix of the given quadratic form is

$$A = \frac{1}{30}\begin{bmatrix} 13 & -4 \\ -4 & 7 \end{bmatrix}.$$

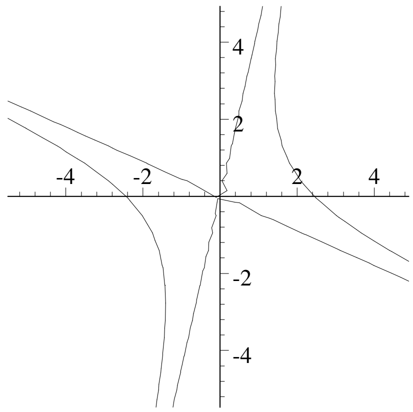
The eigenvalues and corresponding unit eigenvectors of this matrix are $\lambda_1 = 1/6,\ \lambda_2 = 1/2,\ \mathbf{s}_1 = \frac{1}{\sqrt{5}}(1,2)^T,\ \mathbf{s}_2 = \frac{1}{\sqrt{5}}(-2,1)^T$ Hence, the given equation represents the ellipse below. It is centered at the origin, its major axis has half-length $1/\sqrt{\lambda_1} = \sqrt{6}$ and points in the direction of the eigenvector $\mathbf{s}_1$, and its minor axis has half-length $1/\sqrt{\lambda_2} = \sqrt{2}$ and points in the direction of the eigenvector $\mathbf{s}_2$.



7.3.3. The matrix of the given quadratic form is

$$A = \frac{1}{12}\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}.$$

The eigenvalues and corresponding unit eigenvectors of this matrix are $\lambda_1 = 1/4,\ \lambda_2 = -1/6,\ \mathbf{s}_1 = \frac{1}{\sqrt{5}}(2,1)^T,\ \mathbf{s}_2 = \frac{1}{\sqrt{5}}(-1,2)^T$. Hence, the given equation represents a hyperbola centered at the origin, whose axis has half-length $1/\sqrt{\lambda_1} = 2$ and points in the direction of the eigenvector $\mathbf{s}_1$, and so its vertices are at $x = \pm\frac{1}{\sqrt{5}}(4,2)^T$.



7.3.5. a.

$$\begin{aligned}(\nabla(\mathbf{x}^T A\mathbf{x}))_k &= \frac{\partial}{\partial x_k}\sum_i\sum_j a_{ij}x_ix_j = \sum_i\sum_j a_{ij}(\frac{\partial x_i}{\partial x_k}x_j + x_i\frac{\partial x_j}{\partial x_k}) \\ &= \sum_i\sum_j a_{ij}(\delta_{ik}x_j + x_i\delta_{jk}) = \sum_j a_{kj}x_j + \sum_i a_{ik}x_i \\ &= (A\mathbf{x})_k + (\mathbf{x}^T A)_k = 2(A\mathbf{x})^T_k.\end{aligned}$$

(The last equality above follows from the symmetry of A.) Hence $\nabla(\mathbf{x}^T A\mathbf{x}) = 2(A\mathbf{x})^T$.

b. Since $\mathbf{x}^T\mathbf{x} = \mathbf{x}^T I\mathbf{x}$, Part (a) with $A = I$ shows that $\nabla(\mathbf{x}^T\mathbf{x}) = 2\mathbf{x}^T$. Thus Lagrange's equation $\nabla(\mathbf{x}^T A\mathbf{x}) = \lambda\nabla(\mathbf{x}^T\mathbf{x})$ becomes $2(A\mathbf{x})^T = 2\lambda\mathbf{x}^T$, from which we see that the Lagrange multiplier is an eigenvalue of A.

7.3.7. The construction in the proof of Theorem 7.3.2 remains unchanged , except that if A is not symmetric, then Equation 7.107 does not hold and Equation 7.108 becomes

$$S_1^{-1}AS_1 = \left[\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right],$$

where the stars denote appropriate, possibly nonzero entries. Similarly, the 0 entries in the first two rows of the matrix in Equation 7.110 may become nonzero; and so on, until in Equation 7.112 all the entries above the diagonal may be different from zero.

7.4.1.

a)

$\mathbf{x}^H = (2, -2i)$ and $\mathbf{y}^H = (-5i, 4 - i)$,

b)

$|\mathbf{x}|^2 = \mathbf{x}^H\mathbf{x} = (2, -2i)\begin{bmatrix} 2 \\ 2i \end{bmatrix} = 4 - 4i^2 = 8$ and $|\mathbf{x}| = \sqrt{8}$,

$|\mathbf{y}|^2 = \mathbf{y}^H\mathbf{y} = (-5i, 4-i)\begin{bmatrix} 5i \\ 4+i \end{bmatrix} = 25+16+1 = 42$ and $|\mathbf{y}| = \sqrt{42}$,

c)

$\mathbf{x}^H\mathbf{y} = (2, -2i)\begin{bmatrix} 5i \\ 4+i \end{bmatrix} = 10i - 8i + 2 = 2 + 2i$,

$\mathbf{y}^H\mathbf{x} = (-5i, 4 - i)\begin{bmatrix} 2 \\ 2i \end{bmatrix} = -10i + 8i + 2 = 2 - 2i$.

7.4.3.

a)

$\mathbf{x}^H = (2e^{-i\pi/4}, -2i)$ and $\mathbf{y}^H = (e^{-i\pi/4}, e^{i\pi/4})$,

b)

$|\mathbf{x}|^2 = (2e^{-i\pi/4}, -2i)\begin{bmatrix} 2e^{i\pi/4} \\ 2i \end{bmatrix} = 4 - 4i^2 = 8$ and $|\mathbf{x}| = \sqrt{8}$,

c)

$\mathbf{x}^H\mathbf{y} = (2e^{-i\pi/4}, -2i)\begin{bmatrix} e^{i\pi/4} \\ e^{-i\pi/4} \end{bmatrix} = 2 - 2ie^{-i\pi/4} = 2 - \sqrt{2} - \sqrt{2}i$,

$\mathbf{y}^H\mathbf{x} = \overline{\mathbf{x}^H\mathbf{y}} = 2 - \sqrt{2} + \sqrt{2}i$.

7.4.5.

a) Let $\mathbf{u}_2 = \begin{bmatrix} x_1 + iy_1 \\ x_2 + iy_2 \end{bmatrix}$. Then $\mathbf{u}_1^H\mathbf{u}_2 = 0$ becomes $x_1 + iy_1 - ix_2 + y_2 = 0$, which can be written in components as $x_1 + y_2 = 0$ and $y_1 - x_2 = 0$. We may choose $x_1 = y_2 = 0$ and $y_1 = x_2 = 1$ Thus $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} i \\ 1 \end{bmatrix}$ is a unit vector orthogonal to $\mathbf{u}_1$.

b) The matrix

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$

is unitary, and so

$$U^{-1} = U^H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}.$$

Just as in the real case, the coordinate vector $\mathbf{x}_U$ is given by $\mathbf{x}_U = U^{-1}\mathbf{x}$, which in the present case becomes

$$\mathbf{x}_U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x_1 - ix_2 \\ x_2 - ix_1 \end{bmatrix}.$$

c) From the last equation, with $x_1 = 2 + 4i$ and $x_2 = 1 - 2i$, we get

$$\mathbf{x}_U = \frac{1}{\sqrt{2}}\begin{bmatrix} 2 + 4i - i\,(1 - 2i) \\ 1 - 2i - i\,(2 + 4i) \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 3i \\ 5 - 4i \end{bmatrix}.$$

Hence

$$\mathbf{x} = \frac{3i}{\sqrt{2}}\mathbf{u}_1 + \frac{5 - 4i}{\sqrt{2}}\mathbf{u}_2.$$

7.4.7. $A^{HH} = (a_{jk})^{HH} = (\overline{a_{kj}})^H = (\overline{\overline{a_{kj}}})^T = (a_{kj})^T = (a_{jk}) = A.$

7.4.9. For any complex number z, $z^H = z$ if and only if z is real. Applying this observation and Theorem 7.4.1 to $z = \mathbf{x}^H A\mathbf{x}$, and using $(\mathbf{x}^H A\mathbf{x})^H = \mathbf{x}^H A^H \mathbf{x}^{HH} = \mathbf{x}^H A^H \mathbf{x}$, we find that $\mathbf{x}^H A^H \mathbf{x} = \mathbf{x}^H A\mathbf{x}$ for every $\mathbf{x}$, if and only if $\mathbf{x}^H A\mathbf{x}$ is real for every $\mathbf{x}$. Thus if $A^H = A$, then $\mathbf{x}^H A\mathbf{x}$ is real for every $\mathbf{x}$.

Conversely, assume that $\mathbf{x}^H A\mathbf{x}$ is real for every $\mathbf{x}$. Write $\mathbf{x}+\mathbf{y}$ in place of $\mathbf{x}$. Then $(\mathbf{x}+\mathbf{y})^H A^H(\mathbf{x}+\mathbf{y}) = (\mathbf{x}+\mathbf{y})^H A(\mathbf{x}+\mathbf{y})$ for all $\mathbf{x}$, $\mathbf{y}$, is equivalent to $\mathbf{x}^H A^H \mathbf{y} + \mathbf{y}^H A^H \mathbf{x} = \mathbf{x}^H A\mathbf{y} + \mathbf{y}^H A\mathbf{x}$ for all $\mathbf{x}, \mathbf{y}$. Choose $\mathbf{x} = \mathbf{e}_j$ and $\mathbf{y} = \mathbf{e}_k$. Then the last equation becomes

$$(A^H)_{jk} + (A^H)_{kj} = a_{jk} + a_{kj}.$$

Next, choosing $\mathbf{x} = \mathbf{e}_j$ and $\mathbf{y} = i\mathbf{e}_k$, we get

$$i(A^H)_{jk} - i(A^H)_{kj} = ia_{jk} - ia_{kj}.$$

Multiplying the last equation by i and subtracting it from the previous one, we obtain

$$(A^H)_{jk} = a_{jk}.$$

Since this equation holds for all jk, we have $A^H = A$, as was to be shown.

7.4.11. If $A^H = A$, then

$$\begin{aligned} U^H &= \left(e^{itA}\right)^H = \left(\sum_{k=0}^{\infty} \frac{(iAt)^k}{k!}\right)^H = \sum_{k=0}^{\infty} \left(\frac{(iAt)^k}{k!}\right)^H \\ &= \sum_{k=0}^{\infty} \left(\frac{\left((iAt)^H\right)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{(-iAt)^k}{k!} = e^{-At} = U^{-1}. \end{aligned}$$

7.4.13. Suppose A is Hermitian and $\mathbf{s}_1$ and $\mathbf{s}_2$ are nonzero eigenvectors belonging to the distinct eigenvalues λ_1 and λ_2 respectively. Then

$$A\mathbf{s}_1 = \lambda_1\mathbf{s}_1 \text{ and } A\mathbf{s}_2 = \lambda_2\mathbf{s}_2.$$

Multiplying these equations by $\mathbf{s}_2^H$ and $\mathbf{s}_1^H$, we get

$$\mathbf{s}_2^H A\mathbf{s}_1 = \lambda_1\mathbf{s}_2^H\mathbf{s}_1 \text{ and } \mathbf{s}_1^H A\mathbf{s}_2 = \lambda_2\mathbf{s}_1^H\mathbf{s}_2.$$

Taking the Hermitian conjugate of the last equation and using $A^H = A$, we can change it to

$$\mathbf{s}_2^H A\mathbf{s}_1 = \lambda_2\mathbf{s}_2^H\mathbf{s}_1.$$

Thus

$$(\lambda_2 - \lambda_1)\mathbf{s}_2^H\mathbf{s}_1 = 0,$$

and, since $\lambda_2 - \lambda_1 \neq 0$, we must have $\mathbf{s}_2^H\mathbf{s}_1 = 0$.

7.4.15. The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} = 0,$$

or equivalently, $(1-\lambda)^2+1=0$. Hence, the eigenvalues are $\lambda_1 = 1+i$ and $\lambda_2 = 1-i$.

To find the eigenvectors corresponding to $\lambda_1 = 1+i$, we need to solve the equation

$$(A-\lambda_1 I)\mathbf{s} = \begin{bmatrix} -i & 1 \\ -1 & -\,i \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The general solution is $\mathbf{s}_1 = t(1,i)^T$, where t is an arbitrary complex parameter.

To find the eigenvectors corresponding to $\lambda_2 = 1-i$, we need to solve the equation

$$(A-\lambda_2 I)\mathbf{s} = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The general solution is $\mathbf{s}_2 = u(i,1)^T$, where u is an arbitrary complex parameter.

8.1.1.

$$\begin{aligned}
E_{32}E_{31}E_{21} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -4 & 1 & 1 \end{bmatrix}
\end{aligned}$$

The last matrix is L^{-1}, and it too is lower diagonal. Its -4 entry does not, however, come from a single l_{ij}, but it is the combination $-(l_{21}+l_{31})$ of two coefficients of the forward elimination algorithm. While the order of the matrix multiplications in Equation 8.17 has kept the off-diagonal entries separate, the reverse order used here mixes them up.

8.1.3. Let $A = A^T$ and $A = LDU$. Since D is diagonal, we then have

$D^T = D$ and $U^T D L^T = LDU$. As we know, L is invertible, and therefore L^T must be invertible, too. Thus, multiplying the last equation by L^{-1} from the left and by $(L^T)^{-1}$ from the right, we obtain $L^{-1}U^T D = DU(L^T)^{-1}$. Notice that the product on the left is lower diagonal, and the one on the right is upper diagonal. This is possible only if both are diagonal, which, for invertible D, implies that $L^{-1}U^T = C$ must be diagonal. (If D is singular, that is, it has some zeros on its main diagonal, then the columns of C corresponding to the nonzero entries of D must still contain only diagonal nonzero elements, and the other columns, being arbitrary, can be chosen to contain only diagonal nonzero elements.) Since, by Corollary 8.1.1, the diagonal elements of L^{-1} and U^T are all 1, the diagonal elements of their product must be l's too. Thus $C = I$ and $U^T = L$ must hold.

8.1.5. In the forward phase of Gaussian elimination, assuming no row exchanges are needed, to get a 0 in place of a_{21}, we compute $l_{21} = a_{21}/a_{11}$ and subtract $l_{21}a_{1j}$ from each element a_{2j} of the second row for $j = 2, 3, \ldots, n$. Thus again, as in the $n \times n$ case, we need n long operations for the reduction of the second row of A.

Next, we do the same for each of the other rows below the first row. Thus to get all the $m-1$ zeros in the first column requires $(m-1)n$ long operations.

Now we do the same for the $(m-1) \times (n-1)$ submatrix below and to the right of a_{11}. For this we need $(m-2)(n-1)$ long operations.

Continuing in this manner, we find that the total number of long operations needed for the reduction of A, in case $m \leq n$, is

$$\sum_{k=1}^{m}((m-1)-k)(n-k) = \sum_{k=1}^{m-1}[(m-1)n - (m+n-1)k + k^2] =$$

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

If $m > n$, then the reduction must stop when $k = n$. So the total number of long operations is then

$$\sum_{k=1}^{n}((m-1)-k)(n-k) = \sum_{k=1}^{n}[(m-1)n - (m+n-1)k + k^2] =$$

$$(m-1)n^2 - (m+n-1)\frac{n(n+1)}{2} + \frac{n(n+1)(2n+1)}{6} = \frac{mn^2}{2} - \frac{n^3}{6}.$$

Since in both of the foregoing cases the reduction of A produces L and U as well, the formulas above give also the numbers of the long operations needed for the LU factorization of A.

8.1.7. By Equation 8.34, the forward elimination phase requires approximately $n^3/3$ long operations.

Assuming that none of the n pivots is zero, we need approximately $n^2/2$ divisions to change them to l's, because we have approximately that many possibly nonzero elements in the echelon matrix, which must each be divided by a pivot. Then we need $n-1$ multiplications and subtractions on the right side to produce zeros in the last column above the last pivot. (The zeros need not be computed, only the corresponding numbers on the right must be.) Similarly, we need $n-2$ multiplications and subtractions on the right side to produce zeros in the next to last column, and so on. Hence the total number of such operations is $n(n-1)/2 = n^2/2$.

For large n the number n^2 of second phase operations is negligible compared to the number $n^3/3$ in the first phase, and so the latter provides the approximate necessary total.

8.1.9. The reduction to upper triangular form requires the same operations as the forward elimination phase of Gaussian elimination, which, by Equation 8.34, uses approximately $n^3/3$ long operations when n is large. To multiply the diagonal entries we need $n-1$ multiplications, and this number can be neglected next to $n^3/3$ when n is large.

8.2.1. The first step of Gaussian Elimination would produce

$$\left[\begin{array}{lr|r} 0.002 & 1 & 4 \\ 0 & -3001 & -11998 \end{array}\right]$$

and our machine would round the second row to give

$$\left[\begin{array}{lr|r} 0.002 & 1 & 4 \\ 0 & -3000 & -12000 \end{array}\right]$$

The machine would then solve this by back substitution and obtain $x_2 = 4$ and $x_1 = 0$ This solution is, however, wrong. The correct solution is $x_2 =$

$\frac{11998}{3001} = 3.9980\ldots$ and $x_1 = \frac{6}{6.002} = 0.9996\ldots$. Thus, while the machine's answer for x_2 is close enough, for x_1 it is way off.

The reason for the discrepancy is this: In the first step of the back substitution, the machine rounded $x_2 = 3.998\ldots$ to 4. This, in itself, is certainly all right, but in the next step, the machine had to divide x_2 by 0.002 in solving for x_1. Here the small roundoff error, hidden in taking x_2 as 4, got magnified five hundredfold.

8.2.3. The scale factors are $s_1 = 1$ and $s_2 = 6$, and the ratios $r_1 = 0.0002$ and $r_2 = 1$. Since $r_2 > r_1$, we put the second row on top:

$$[A|b] = \left[\begin{array}{cc|c} 6 & -1 & 2 \\ 0.002 & 1 & 4 \end{array}\right]$$

Now we subtract $0.0002/6$ times the first row from the second, to get

$$\left[\begin{array}{cc|c} 6 & -1 & 2 \\ 0 & 1.00033\ldots & 3.99933\ldots \end{array}\right].$$

The machine of Example 8.2.1 would round the above matrix to

$$\left[\begin{array}{cc|c} 6 & -1 & 2 \\ 0 & 1 & 4 \end{array}\right]$$

and solve the system from here as $x_2 = 4$ and $x_1 = 1$.

The correct values are, as we have seen in the solution of Example 8.2.1, $x_2 = \frac{11998}{3001} = 3.998\ldots$ and $x_1 = \frac{6}{6.002} = 0.9996\ldots$. Thus, this method has produced excellent approximations. The reason for this success is that here we did not have to divide the roundoff error by a small pivot in the course of the back substitution.

8.2.5. The scale factors are $s_1 = 5,\ s_2 = 11$ and $s_3 = 9$, and the ratios $r_1 = 1/5,\ r_2 = 4/11$ and $r_3 = 5/9$. Since $r_3 > r_2 > r_1$, we put the third row on top, and then proceed with the row reduction:

$$\left[\begin{array}{ccc|c} 5 & 8 & 9 & 1 \\ 1 & 2 & 5 & 1 \\ 4 & -7 & 11 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 5 & 8 & 9 & 1 \\ 0 & 2 & 16 & 4 \\ 0 & -67 & 19 & -4 \end{array}\right] \to$$

The new scale factors are $s_2 = 16$ and $s_3 = 67$, and the ratios $r_2 = 2/16$ and

$r_3 = 67/67 = 1$. Thus we switch the last two rows:

$$\left[\begin{array}{rrr|r} 5 & 8 & 9 & 1 \\ 0 & -67 & 19 & -4 \\ 0 & 2 & 16 & 4 \end{array}\right] \to \left[\begin{array}{rrr|r} 5 & 8 & 9 & 1 \\ 0 & -67 & 19 & -4 \\ 0 & 0 & 1110 & 260 \end{array}\right].$$

Back substitution now gives $x_3 = 26/111,\ x_2 = 14/111$, and $x_1 = -47/111$.

8.3.1. a. Setting

$$\mathbf{x}_0 = \mathbf{x}_0' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

we get

$$\mathbf{x}_1 = A\mathbf{x}_0' = \begin{bmatrix} 4.0000 \\ 3.0000 \\ 1.0000 \end{bmatrix} \text{ and } \mathbf{x}_1' = \begin{bmatrix} 1 \\ 0.7500 \\ 0.2500 \end{bmatrix},$$

$$\mathbf{x}_2 = A\mathbf{x}_1' = \begin{bmatrix} 3.2500 \\ 2.2500 \\ 0.2500 \end{bmatrix} \text{ and } \mathbf{x}_2' = \begin{bmatrix} 1 \\ 0.6923 \\ 0.0769 \end{bmatrix},$$

$$\mathbf{x}_3 = A\mathbf{x}_2' = \begin{bmatrix} 3.0769 \\ 2.0769 \\ 0.0769 \end{bmatrix} \text{ and } \mathbf{x}_3' = \begin{bmatrix} 1 \\ 0.6750 \\ 0.0250 \end{bmatrix},$$

$$\mathbf{x}_4 = A\mathbf{x}_3' = \begin{bmatrix} 3.0250 \\ 2.0250 \\ 0.0250 \end{bmatrix} \text{ and } \mathbf{x}_4' = \begin{bmatrix} 1 \\ 0.6694 \\ 0.0083 \end{bmatrix},$$

$$\mathbf{x}_5 = A\mathbf{x}_4' = \begin{bmatrix} 3.0083 \\ 2.0083 \\ 0.0083 \end{bmatrix} \text{ and } \mathbf{x}_5' = \begin{bmatrix} 1 \\ 0.6676 \\ 0.0028 \end{bmatrix}.$$

Thus $\lambda_1 \approx 3$ and $\mathbf{s}_1 \approx (3, 2, 0)^T$.

If we start with

$$\mathbf{x}_0 = \mathbf{x}_0' = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix},$$

then we get

$$\mathbf{x}_1 = A\mathbf{x}_0' = \begin{bmatrix} 4.0000 \\ -3.0000 \\ 1.0000 \end{bmatrix} \text{ and } \mathbf{x}_1' = \begin{bmatrix} 1 \\ -0.7500 \\ 0.2500 \end{bmatrix},$$

$$\mathbf{x}_2 = A\mathbf{x}_1' = \begin{bmatrix} 3.2500 \\ -2.2500 \\ 0.2500 \end{bmatrix} \text{ and } \mathbf{x}_2' = \begin{bmatrix} 1 \\ -0.6923 \\ 0.0769 \end{bmatrix},$$

$$\mathbf{x}_3 = A\mathbf{x}_2' = \begin{bmatrix} 3.0769 \\ -2.0769 \\ 0.0769 \end{bmatrix} \text{ and } \mathbf{x}_3' = \begin{bmatrix} 1 \\ -0.6750 \\ 0.0250 \end{bmatrix},$$

$$\mathbf{x}_4 = A\mathbf{x}_3' = \begin{bmatrix} 3.0250 \\ -2.0250 \\ 0.0250 \end{bmatrix} \text{ and } \mathbf{x}_4' = \begin{bmatrix} 1 \\ -0.6694 \\ 0.0083 \end{bmatrix},$$

$$\mathbf{x}_5 = A\mathbf{x}_4' = \begin{bmatrix} 3.0083 \\ -2.0083 \\ 0.0083 \end{bmatrix} \text{ and } \mathbf{x}_5' = \begin{bmatrix} 1 \\ -0.6676 \\ 0.0028 \end{bmatrix}.$$

Thus $\lambda_1 \approx 3$ and $\mathbf{s}_2 \approx (3, -2, 0)^T$.

b. As we can see from the computations above, if a dominant eigenvalue has geometric multiplicity greater than 1, then different initial vectors may lead to different eigenvectors. Clearly, any eigenvector belonging to such a dominant eigenvalue can be so obtained by an appropriate choice of the initial vector, since if we started with an eigenvector, the method would stay with it. It is only this observation that needs to be added to Theorem 8.3.1.

8.3.3. The eigenvalues of a matrix A are the solutions λ of its characteristic equation $|A - \lambda I| = 0$. Hence if c is not an eigenvalue of A, then $|A - cI| \neq 0$. By Theorem 6.1.8, this inequality implies that $B = A - cI$ is nonsingular.

Alternatively, by Exercises 7.1.10 and 11, the eigenvalues of $B = A - cI$ are the $\lambda - c$ values, which are nonzero if $c \neq \lambda$, and a matrix with only nonzero eigenvalues is nonsingular.

A2.1. Equation A.22 follows from Equations A.20, A.17, and A.18, since $|\cos\phi + i\sin\phi|^2 = \cos^2\phi + \sin^2\phi = 1$ for any ϕ.

A2.3. If $P(z) = \sum_{k=0}^{n} a_k z^k$ is a polynomial with real coefficients a_k, then, by Equations A.12 and A.13, generalized to arbitrary finite sums and products, it follows that

$$\overline{P(z)} = \sum_{k=0}^{n} \overline{a_k z^k} = \sum_{k=0}^{n} \overline{a}_k \overline{z^k} = \sum_{k=0}^{n} a_k \overline{z}^k = P(\overline{z}).$$

Thus $P(z_0) = 0$ is equivalent to $P(\overline{z}_0) = \overline{P(z_0)} = \overline{0} = 0$.

$\mathbf{A2.5}$. Suppose that $\sum_{n=0}^{\infty} z_n$ converges absolutely; that is, $\sum_{n=0}^{\infty} |z_n|$ converges. Then since $|x_n| \leq |z_n|$ and $|y_n| \leq |z_n|$ for all indices n, it follows from the Comparison Test that both $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$ converge absolutely and thus both converge by the Absolute Convergence Test for real series. The result of Exercise A.2.4 implies that $\sum_{n=0}^{\infty} z_n$ also converges.

$\mathbf{A2.7}$. In order to use Formula A.33 to find the roots of a given complex number, it is first necessary to express the number in polar exponential form, that is, in the form of Equation A.29.

a. Since $|i| = 1$ and the principal argument of i is $\pi/2$, $i = e^{i((\pi/2)+2k\pi)}$ and thus $i^{1/2} = e^{i((\pi/4)+k\pi)}$. Distinct roots are obtained for $k = 0$ and 1:
$k = 0$ yields

$$z_1 = e^{i\pi/4} = \frac{\sqrt{2}}{2}(1+i),$$

and $k = 1$ yields

$$z_2 = e^{i(5\pi/4)} = -\frac{\sqrt{2}}{2}(1+i).$$

b. Since $|1+i| = \sqrt{2}$ and the principal argument of $1+i$ is $\pi/4$,

$$1 + i = \sqrt{2}e^{i((\pi/4)+2k\pi)},$$

and thus

$$(1+i)^{1/2} = \sqrt[4]{2}e^{i((\pi/8)+k\pi)}.$$

Distinct roots are obtained for $k = 0$ and 1:
$k = 0$ yields

$$z_1 = \sqrt[4]{2}e^{i\pi/8} = \sqrt[4]{2}(\cos(\pi/8) + i\sin(\pi/8))$$

and $k = 1$ yields

$$z_2 = \sqrt[4]{2}e^{i(5\pi/8)} = \sqrt[4]{2}(\cos(5\pi/8) + i\sin(5\pi/8))$$

$$= -\sqrt[4]{2}(\cos(\pi/8) + i\sin(\pi/8)).$$

c. Since $|1| = 1$ and the principal argument of 1 is 0, $1 = e^{i2\pi k}$ and thus $1^{1/3} = e^{i(2\pi k/3)}$. Distinct roots are obtained for $k = 0, 1$, and 2:
$k = 0$ yields

$$z_1 = e^0 = 1,$$

$k = 1$ yields

$$z_2 = e^{i(2\pi/3)} = \frac{1}{2}(-1 + \sqrt{3}i),$$

and $k = 2$ yields

$$z_3 = e^{i(4\pi/3)} = \frac{1}{2}(-1 - \sqrt{3}i).$$

d. Since $|-1| = 1$ and the principal argument of -1 is π,

$$-1 = e^{i(\pi + 2k\pi)},$$

and thus

$$(-1)^{1/3} = e^{i\pi(2k+1)/3}.$$

Distinct roots are obtained for $k = 0, 1$, and 2:
$k = 0$ yields

$$z_1 = e^{i\pi/3} = \frac{1}{2}(1 + \sqrt{3}i),$$

$k = 1$ yields

$$z_2 = e^{i\pi} = -1,$$

and $k = 2$ yields

$$z_3 = e^{i5\pi/3} = \frac{1}{2}(1 - \sqrt{3}i).$$

e. From part (a),

$$i = e^{i((\pi/2) + 2k\pi)},$$

and thus

$$i^{1/4} = e^{i((\pi/8) + (k\pi/4))}.$$

Distinct roots are obtained for $k = 0, 1, 2$, and 3:
$k = 0$ yields

$$z_1 = e^{i\pi/8} = \cos(\pi/8) + i\sin(\pi/8),$$

$k = 1$ yields

$$z_2 = e^{i5\pi/8} = \cos(5\pi/8) + i\sin(5\pi/8),$$

$k = 2$ yields

$$z_3 = e^{i9\pi/8} = \cos(9\pi/8) + i\sin(9\pi/8) = -[\cos(\pi/8) + i\sin(\pi/8)],$$

and $k = 3$ yields

$$z_4 = e^{i13\pi/8} = \cos(13\pi/8) + i\sin(13\pi/8) = -[\cos(5\pi/8) + i\sin(5\pi/8)].$$

http://www.springer.com/978-0-8176-8324-5

A Concise Introduction to Linear Algebra
Schay, G.
2012, XII, 330p. 50 illus.., Hardcover
ISBN: 978-0-8176-8324-5
A product of Birkhäuser Basel