# Chapter 5

## Section 5.1

**5.1.1** $\|\vec{v}\| = \sqrt{7^2 + 11^2} = \sqrt{49 + 121} = \sqrt{170} \approx 13.04$

**5.1.2** $\|\vec{v}\| = \sqrt{2^2 + 3^2 + 4^2} = \sqrt{4 + 9 + 16} = \sqrt{29} \approx 5.39$

**5.1.3** $\|\vec{v}\| = \sqrt{2^2 + 3^2 + 4^2 + 5^2} = \sqrt{4 + 9 + 16 + 25} = \sqrt{54} \approx 7.35$

**5.1.4** $\theta = \arccos \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \arccos \frac{7+11}{\sqrt{2}\sqrt{170}} = \arccos \frac{18}{\sqrt{340}} \approx 0.219$ (radians)

**5.1.5** $\theta = \arccos \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \arccos \frac{2+6+12}{\sqrt{14}\sqrt{29}} \approx 0.122$ (radians)

**5.1.6** $\theta = \arccos \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \arccos \frac{2-3+8-10}{\sqrt{10}\sqrt{54}} \approx 1.700$ (radians)

**5.1.7** Use the fact that $\vec{u}\cdot\vec{v} = \|\vec{u}\|\|\vec{v}\|\cos\theta$, so that the angle is acute if $\vec{u}\cdot\vec{v} > 0$, and obtuse if $\vec{u}\cdot\vec{v} < 0$. Since $\vec{u}\cdot\vec{v} = 10 - 12 = -2$, the angle is obtuse.

**5.1.8** Since $\vec{u}\cdot\vec{v} = 4 - 24 + 20 = 0$, the two vectors enclose a right angle.

**5.1.9** Since $\vec{u}\cdot\vec{v} = 3 - 4 + 5 - 3 = 1$, the angle is acute (see Exercise 7).

**5.1.10** $\vec{u}\cdot\vec{v} = 2 + 3k + 4 = 6 + 3k$. The two vectors enclose a right angle if $\vec{u}\cdot\vec{v} = 6 + 3k = 0$, that is, if $k = -2$.

**5.1.11** a $\theta_n = \arccos \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \arccos \frac{1}{\sqrt{n}}$

$\theta_2 = \arccos \frac{1}{\sqrt{2}} = \frac{\pi}{4} (= 45^\circ)$

$\theta_3 = \arccos \frac{1}{\sqrt{3}} \approx 0.955$ (radians)

$\theta_4 = \arccos \frac{1}{2} = \frac{\pi}{3} (= 60^\circ)$

b Since $y = \arccos(x)$ is a continuous function,

$$\lim_{n\to\infty} \theta_n = \arccos\left(\lim_{n\to\infty} \frac{1}{\sqrt{n}}\right) = \arccos(0) = \frac{\pi}{2} (= 90^\circ)$$

**5.1.12** $\|\vec{v} + \vec{w}\|^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w})$ (by hint)

$= \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2(\vec{v}\cdot\vec{w})$ (by definition of length)

$\leq \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\|\|\vec{w}\|$ (by Cauchy-Schwarz)

$= (\|\vec{v}\| + \|\vec{w}\|)^2$, so that

$\|\vec{v} + \vec{w}\|^2 \leq (\|\vec{v}\| + \|\vec{w}\|)^2$

Taking square roots of both sides, we find that $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$, as claimed.

5.1.**13** Figure 5.1 shows that $\|\vec{F}_2 + \vec{F}_3\| = 2\cos\left(\frac{\theta}{2}\right)\|\vec{F}_2\| = 20\cos\left(\frac{\theta}{2}\right)$.

It is required that $\|\vec{F}_2 + \vec{F}_3\| = 16$, so that $20\cos\left(\frac{\theta}{2}\right) = 16$, or $\theta = 2\arccos(0.8) \approx 74°$.

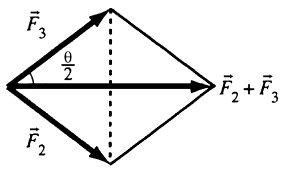



Figure 5.1: for Problem 5.1.13.

5.1.**14** The horizontal components of $\vec{F}_1$ and $\vec{F}_2$ are $-\|\vec{F}_1\|\sin\beta$ and $\|\vec{F}_2\|\sin\alpha$, respectively (the horizontal component of $\vec{F}_3$ is zero).

Since the system is at rest, the horizontal components must add up to 0, so that $-\|\vec{F}_1\|\sin\beta + \|\vec{F}_2\|\sin\alpha = 0$ or $\|\vec{F}_1\|\sin\beta = \|\vec{F}_2\|\sin\alpha$ or $\frac{\|\vec{F}_1\|}{\|\vec{F}_2\|} = \frac{\sin\alpha}{\sin\beta}$.

To find $\frac{\overline{EA}}{\overline{EB}}$, note that $\overline{EA} = \overline{ED}\tan\alpha$ and $\overline{EB} = \overline{ED}\tan\beta$ so that $\frac{\overline{EA}}{\overline{EB}} = \frac{\tan\alpha}{\tan\beta} = \frac{\sin\alpha}{\sin\beta}\cdot\frac{\cos\beta}{\cos\alpha} = \frac{\|\vec{F}_1\|}{\|\vec{F}_2\|}\frac{\cos\beta}{\cos\alpha}$. Since $\alpha$ and $\beta$ are two distinct acute angles, it follows that $\frac{\overline{EA}}{\overline{EB}} \neq \frac{\|\vec{F}_1\|}{\|\vec{F}_2\|}$, so that Leonardo was mistaken.

5.1.**15** The subspace consists of all vectors $\vec{x}$ in $\mathbb{R}^4$ such that

$$\vec{x}\cdot\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = x_1 + 2x_2 + 3x_3 + 4x_4 = 0.$$

These are vectors of the form $\begin{bmatrix} -2r & -3s & -4t \\ r & & \\ & s & \\ & & t \end{bmatrix} = r\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -4 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

The three vectors to the right form a basis.

5.1.**16** You may be able to find the solutions by educated guessing. Here is the systematic approach: we first find all vectors $\vec{x}$ that are orthogonal to $\vec{v}_1, \vec{v}_2$, and $\vec{v}_3$, then we identify the unit vectors among them.

Finding the vectors $\vec{x}$ with $\vec{x}\cdot\vec{v}_1 = \vec{x}\cdot\vec{v}_2 = \vec{x}\cdot\vec{v}_3 = 0$ amounts to solving the system

$$\begin{bmatrix} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 - x_3 - x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 0 \end{bmatrix}$$

(we can omit all the coefficients $\frac{1}{2}$).

The solutions are of the form $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} t \\ -t \\ -t \\ t \end{bmatrix}$.

Since $\|\vec{x}\| = 2|t|$, we have a unit vector if $t = \frac{1}{2}$ or $t = -\frac{1}{2}$. Thus there are two possible choices for $\vec{v}_4$:

$$\begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \text{ and } \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

5.1.**17** The orthogonal complement $W^\perp$ of $W$ consists of the vectors $\vec{x}$ in $\mathbb{R}^4$ such that

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = 0 \text{ and } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = 0.$$

Finding these vectors amounts to solving the system $\begin{bmatrix} x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 = 0 \end{bmatrix}$.

The solutions are of the form

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

The two vectors to the right form a basis of $W^\perp$.

5.1.**18** a $\|\vec{x}\|^2 = 1 + \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots = \frac{1}{1-\frac{1}{4}} = \frac{4}{3}$ (use the formula for a geometric series, with $a = \frac{1}{4}$), so that $\|\vec{x}\| = \frac{2}{\sqrt{3}} \approx 1.155$.

b If we let $\vec{u} = (1, 0, 0, \ldots)$ and $\vec{v} = \left(1, \frac{1}{2}, \frac{1}{4}, \cdots\right)$, then

$\theta = \arccos \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \arccos \frac{1}{\frac{2}{\sqrt{3}}} = \arccos \frac{\sqrt{3}}{2} = \frac{\pi}{6} (= 30°)$.

c $\vec{x} = \left(1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, \cdots, \frac{1}{\sqrt{n}}, \cdots\right)$ does the job, since the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \cdots$ diverges (a fact discussed in introductory calculus classes).

d If we let $\vec{v} = (1, 0, 0, \ldots), \vec{x} = \left(1, \frac{1}{2}, \frac{1}{4}, \cdots\right)$ and $\vec{u} = \frac{\vec{x}}{\|\vec{x}\|} = \frac{\sqrt{3}}{2}\left(1, \frac{1}{2}, \frac{1}{4}, \cdots\right)$ then

$\text{proj}_L \vec{v} = (\vec{u} \cdot \vec{v})\vec{u} = \frac{3}{4}\left(1, \frac{1}{2}, \frac{1}{4}, \cdots\right)$.

5.1.**19** See Figure 5.2.

5.1.**20** On the line $L$ spanned by $\vec{x}$ we want to find the vector $m\vec{x}$ closest to $\vec{y}$ (that is, we want $\|m\vec{x} - \vec{y}\|$ to be minimal). We want $m\vec{x} - \vec{y}$ to be perpendicular to $L$ (that is, to $\vec{x}$), which means that $\vec{x} \cdot (m\vec{x} - \vec{y}) = 0$ or $m(\vec{x} \cdot \vec{x}) - \vec{x} \cdot \vec{y} = 0$ or $m = \frac{\vec{x}\cdot\vec{y}}{\vec{x}\cdot\vec{x}} \approx \frac{4182.9}{198.53^2} \approx 0.106$.

Recall that the correlation coefficient $r$ is $r = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\|\vec{y}\|}$, so that $m = \frac{\|\vec{y}\|}{\|\vec{x}\|} r$. See Figure 5.3.

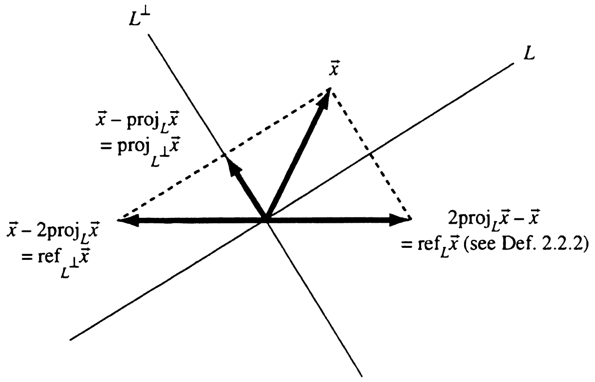



Figure 5.2: for Problem 5.1.19.

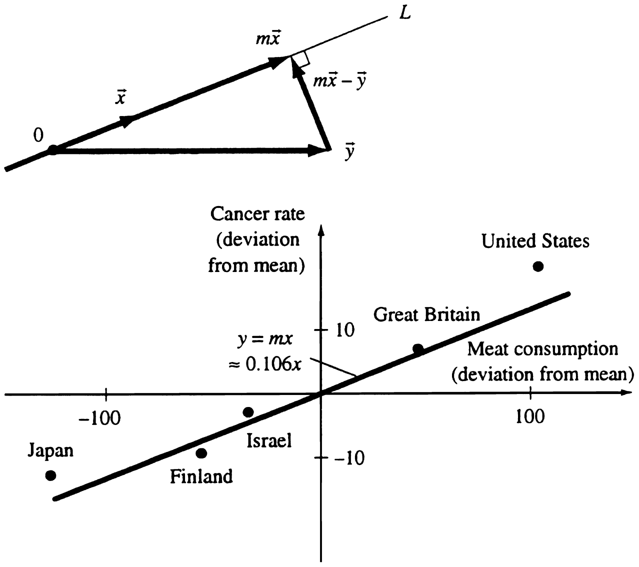



Figure 5.3: for Problem 5.1.20.

5.1.**21** Call the three given vectors $\vec{v}_1, \vec{v}_2$, and $\vec{v}_3$. Since $\vec{v}_2$ is required to be a unit vector, we must have $b = g = 0$. Now $\vec{v}_1 \cdot \vec{v}_2 = d$ must be zero, so that $d = 0$.

Likewise, $\vec{v}_2 \cdot \vec{v}_3 = e$ must be zero, so that $e = 0$.

Since $\vec{v}_3$ must be a unit vector, we have $\|\vec{v}_3\|^2 = c^2 + \frac{1}{4} = 1$, so that $c = \pm\frac{\sqrt{3}}{2}$.

Since we are asked to find just one solution, let us pick $c = \frac{\sqrt{3}}{2}$.

The condition $\vec{v}_1 \cdot \vec{v}_3 = 0$ now implies that $\frac{\sqrt{3}}{2}a + \frac{1}{2}f = 0$, or $f = -\sqrt{3}a$.

Finally, it is required that $\|\vec{v}_1\|^2 = a^2 + f^2 = a^2 + 3a^2 = 4a^2 = 1$, so that $a = \pm\frac{1}{2}$.

Let us pick $a = \frac{1}{2}$, so that $f = -\frac{\sqrt{3}}{2}$.

Summary:

$$\vec{v}_1 = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{\sqrt{3}}{2} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$$

There are other solutions; some components will have different signs.

5.1.**22** Let $W = \{\vec{x} \text{ in } \mathbb{R}^n : \vec{x} \cdot \vec{v}_i = 0 \text{ for all } i = 1, \ldots, m\}$. We are asked to show that $V^\perp = W$, that is, any $\vec{x}$ in $V^\perp$ is in $W$, and vice versa.

If $\vec{x}$ is in $V^\perp$, then $\vec{x} \cdot \vec{v} = 0$ for all $\vec{v}$ in $V$; in particular, $x \cdot \vec{v}_i = 0$ for all $i$ (since the $\vec{v}_i$ are in $V$), so that $\vec{x}$ is in $W$.

Conversely, consider a vector $\vec{x}$ in $W$. To show that $\vec{x}$ is in $V^\perp$, we have to verify that $\vec{x} \cdot \vec{v} = 0$ for all $\vec{v}$ in $V$. Pick a particular $\vec{v}$ in $V$. Since the $\vec{v}_i$ span $V$, we can write $\vec{v} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$, for some scalars $c_i$. Then $\vec{x} \cdot \vec{v} = c_1(\vec{x} \cdot \vec{v}_1) + \cdots + c_m(\vec{x} \cdot \vec{v}_m) = 0$, as claimed.

5.1.**23** We will follow the hint. Let $\vec{v}$ be a vector in $V$. Then $\vec{v} \cdot \vec{x} = 0$ for all $\vec{x}$ in $V^\perp$. Since $(V^\perp)^\perp$ contains all vectors $\vec{y}$ such that $\vec{y} \cdot \vec{x} = 0$, $\vec{v}$ is in $(V^\perp)^\perp$. So $V$ is a subspace of $(V^\perp)^\perp$.

Then, by Theorem 5.1.8c, $\dim(V) + \dim(V^\perp) = n$ and $\dim(V^\perp) + \dim((V^\perp)^\perp) = n$, so $\dim(V) + \dim(V^\perp) = \dim(V^\perp) + \dim((V^\perp)^\perp)$ and $\dim(V) = \dim((V^\perp)^\perp)$. Since $V$ is a subspace of $(V^\perp)^\perp$, it follows that $V = (V^\perp)^\perp$, by Exercise 3.3.61.

5.1.**24** Write $T(\vec{x}) = \text{proj}_V(\vec{x})$ for simplicity.

To prove the linearity of $T$ we will use the definition of a projection: $T(\vec{x})$ is in $V$, and $\vec{x} - T(\vec{x})$ is in $V^\perp$.

To show that $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$, note that $T(\vec{x}) + T(\vec{y})$ is in $V$ (since $V$ is a subspace), and $\vec{x} + \vec{y} - (T(\vec{x}) + T(\vec{y})) = (\vec{x} - T(\vec{x})) + (\vec{y} - T(\vec{y}))$ is in $V^\perp$ (since $V^\perp$ is a subspace, by Theorem 5.1.8a).

To show that $T(k\vec{x}) = kT(\vec{x})$, note that $kT(\vec{x})$ is in $V$ (since $V$ is a subspace), and $k\vec{x} - kT(\vec{x}) = k(\vec{x} - T(\vec{x}))$ is in $V^\perp$ (since $V^\perp$ is a subspace).

5.1.**25** a $\|k\vec{v}\|^2 = (k\vec{v}) \cdot (k\vec{v}) = k^2(\vec{v} \cdot \vec{v}) = k^2\|\vec{v}\|^2$

Now take square roots of both sides; note that $\sqrt{k^2} = |k|$, the absolute value of $k$ (think about the case when $k$ is negative). $\|k\vec{v}\| = |k|\|\vec{v}\|$, as claimed.

b $\|\vec{u}\| = \left\|\frac{1}{\|\vec{v}\|}\vec{v}\right\| = \frac{1}{\|\vec{v}\|}\|\vec{v}\| = 1$, as claimed.

$$\uparrow$$

by part a

5.1.**26** The two given vectors spanning the subspace are orthogonal, but they are not unit vectors: both have length 7. To obtain an orthonormal basis $\vec{u}_1, \vec{u}_2$ of the subspace, we divide by 7:

$$\vec{u}_1 = \frac{1}{7}\begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}, \vec{u}_2 = \frac{1}{7}\begin{bmatrix} 3 \\ -6 \\ 2 \end{bmatrix}.$$

Now we can use Theorem 5.1.5, with $\vec{x} = \begin{bmatrix} 49 \\ 49 \\ 49 \end{bmatrix}$:

$\text{proj}_V \vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + (\vec{u}_2 \cdot \vec{x})\vec{u}_2 = 11 \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 3 \\ -6 \\ 2 \end{bmatrix} = \begin{bmatrix} 19 \\ 39 \\ 64 \end{bmatrix}.$

5.1.**27** Since the two given vectors in the subspace are orthogonal, we have the orthonormal basis

$\vec{u}_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$

Now we can use Theorem 5.1.5, with $\vec{x} = 9\vec{e}_1 : \text{proj}_V \vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + (\vec{u}_2 \cdot \vec{x})\vec{u}_2$

$= 2 \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix} - 2 \begin{bmatrix} -2 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ 2 \\ -2 \end{bmatrix}.$

5.1.**28** Since the three given vectors in the subspace are orthogonal, we have the orthonormal basis

$\vec{u}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{u}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \vec{u}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$

Now we can use Theorem 5.1.5, with $\vec{x} = \vec{e}_1 : \text{proj}_V \vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + (\vec{u}_2 \cdot \vec{x})\vec{u}_2 + (\vec{u}_3 \cdot \vec{x})\vec{u}_3 = \frac{1}{4} \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$

5.1.**29** By the Pythagorean theorem (Theorem 5.1.9),

$$\begin{aligned} \|\vec{x}\|^2 &= \|7\vec{u}_1 - 3\vec{u}_2 + 2\vec{u}_3 + \vec{u}_4 - \vec{u}_5\|^2 \\ &= \|7\vec{u}_1\|^2 + \|3\vec{u}_2\|^2 + \|2\vec{u}_3\|^2 + \|\vec{u}_4\|^2 + \|\vec{u}_5\|^2 \\ &= 49 + 9 + 4 + 1 + 1 \\ &= 64, \text{ so that } \|\vec{x}\| = 8. \end{aligned}$$

5.1.**30** Since $\vec{y} = \text{proj}_V \vec{x}$, the vector $\vec{x} - \vec{y}$ is orthogonal to $\vec{y}$, by definition of an orthogonal projection (see Theorem 5.1.4): $(\vec{x} - \vec{y}) \cdot \vec{y} = 0$ or $\vec{x} \cdot \vec{y} - \|\vec{y}\|^2 = 0$ or $\vec{x} \cdot \vec{y} = \|\vec{y}\|^2$. See Figure 5.4.

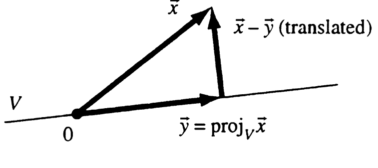



Figure 5.4: for Problem 5.1.30.

5.1.**31** If $V = \text{span}(\vec{u}_1, \ldots, \vec{u}_m)$, then $\text{proj}_V\vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$, by Theorem 5.1.5, and $\|\text{proj}_V\vec{x}\|^2 = (\vec{u}_1 \cdot \vec{x})^2 + \cdots + (\vec{u}_m \cdot \vec{x})^2 = p$, by the Pythagorean theorem (Theorem 5.1.9). Therefore $p \leq \|\vec{x}\|^2$, by Theorem 5.1.10. The two quantities are equal if (and only if) $\vec{x}$ is in $V$.

5.1.**32** By Theorem 2.4.9a, the matrix $G$ is invertible if (and only if) $(\vec{v}_1 \cdot \vec{v}_1)(\vec{v}_2 \cdot \vec{v}_2) - (\vec{v}_1 \cdot \vec{v}_2)^2 = \|\vec{v}_1\|^2\|\vec{v}_2\|^2 - (\vec{v}_1 \cdot \vec{v}_2)^2 \neq 0$. The Cauchy-Schwarz inequality (Theorem 5.1.11) tells us that $\|\vec{v}_1\|^2\|\vec{v}_2\|^2 - (\vec{v}_1 \cdot \vec{v}_2)^2 \geq 0$; equality holds if (and only if) $\vec{v}_1$ and $\vec{v}_2$ are parallel (that is, linearly dependent).

5.1.**33** Let $\vec{x} = \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix}$ be a vector in $\mathbb{R}^n$ whose components add up to 1, that is, $x_1 + \cdots + x_n = 1$. Let $\vec{y} = \begin{bmatrix} 1 \\ \cdots \\ 1 \end{bmatrix}$ (all $n$ components are 1). The Cauchy-Schwarz inequality (Theorem 5.1.11) tells us that $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$, or, $|x_1 + \cdots + x_n| \leq \|\vec{x}\|\sqrt{n}$, or $\|\vec{x}\| \geq \frac{1}{\sqrt{n}}$. By Theorem 5.1.11, the equation $\|\vec{x}\| = \frac{1}{\sqrt{n}}$ holds if (and only if) the vectors $\vec{x}$ and $\vec{y}$ are parallel, that is, $x_1 = x_2 = \cdots = x_n = \frac{1}{n}$. Thus the vector of minimal length is $\vec{x} = \begin{bmatrix} \frac{1}{n} \\ \cdots \\ \frac{1}{n} \end{bmatrix}$ $\left(\text{all components are } \frac{1}{n}\right)$.

Figure 5.5 illustrates the case $n = 2$.

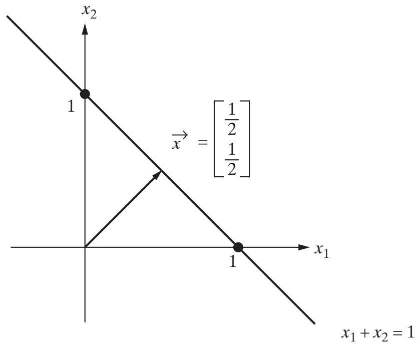



Figure 5.5: for Problem 5.1.33.

5.1.**34** Let $\vec{x}$ be a unit vector in $\mathbb{R}^n$, that is, $\|\vec{x}\| = 1$. Let $\vec{y} = \begin{bmatrix} 1 \\ \ldots \\ 1 \end{bmatrix}$ (all $n$ components are 1). The Cauchy-Schwarz inequality (Theorem 5.1.11) tells us that $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$, or, $|x_1 + \ldots + x_n| \leq \|\vec{x}\|\sqrt{n} = \sqrt{n}$. By Theorem 5.1.11, the equation $x_1 + \ldots + x_n = \sqrt{n}$ holds if $\vec{x} = k\vec{y}$ for positive $k$. Thus $\vec{x}$ must be a unit vector of the form $\vec{x} = \begin{bmatrix} k \\ \ldots \\ k \end{bmatrix}$ for some positive $k$. It is required that $nk^2 = 1$, or, $k = \frac{1}{\sqrt{n}}$. Thus $\vec{x} = \begin{bmatrix} \frac{1}{\sqrt{n}} \\ \cdots \\ \frac{1}{\sqrt{n}} \end{bmatrix}$ $\left(\text{all components are } \frac{1}{\sqrt{n}}\right)$.

Figure 5.6 illustrates the case $n = 2$.

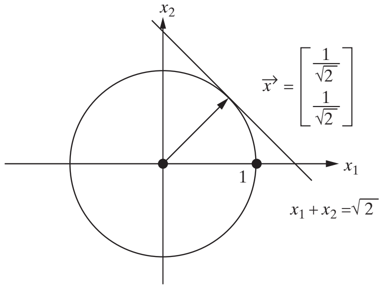



Figure 5.6: for Problem 5.1.34.

5.1.**35** Applying the Cauchy-Schwarz inequality to $\vec{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ gives $|\vec{u}\cdot\vec{v}| \le \|\vec{u}\|\|\vec{v}\|$, or $|x+2y+3z| \le \sqrt{14}$. The minimal value $x+2y+3z = -\sqrt{14}$ is attained when $\vec{u} = k\vec{v}$ for negative $k$. Thus $\vec{u}$ must be a unit vector of the form $\vec{u} = \begin{bmatrix} k \\ 2k \\ 3k \end{bmatrix}$, for negative $k$. It is required that $14k^2 = 1$, or, $k = -\frac{1}{\sqrt{14}}$. Thus $\vec{u} = \begin{bmatrix} -\frac{1}{\sqrt{14}} \\ -\frac{2}{\sqrt{14}} \\ -\frac{3}{\sqrt{14}} \end{bmatrix}$.

5.1.**36** Let $\vec{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}$. It is required that $\vec{x}\cdot\vec{y} = 0.2a + 0.3b + 0.5c = 76$. Our goal is to minimize quantity $\vec{x}\cdot\vec{x} = a^2+b^2+c^2$. The Cauchy-Schwarz inequality (squared) tells us that $(\vec{x}\cdot\vec{y})^2 \le \|\vec{x}\|^2\|\vec{y}\|^2$, or $76^2 \le (a^2+b^2+c^2)(0.2^2+0.3^2+0.5^2)$ or $a^2+b^2+c^2 \ge \frac{76^2}{0.38}$. The quantity $a^2+b^2+c^2$ is minimal when $a^2+b^2+c^2 = \frac{76^2}{0.38}$. This is the case when $\vec{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.2k \\ 0.3k \\ 0.5k \end{bmatrix}$ for some positive constant $k$. It is required that $0.2a+0.3b+0.5c = (0.2)^2k + (0.3)^2k + (0.5)^2k = 0.38k = 76$, so that $k = 200$. Thus $a = 40, b = 60, c = 100$: The student must study 40 hours for the first exam, 60 hours for the second, and 100 hours for the third.

5.1.**37** Using Definition 2.2.2 as a guide, we find that $\text{ref}_V\vec{x} = 2(\text{proj}_V\vec{x}) - \vec{x} = 2(\vec{u}_1\cdot\vec{x})\vec{u}_1 + 2(\vec{u}_2\cdot\vec{x})\vec{u}_2 - \vec{x}$.

5.1.**38** Since $\vec{v}_1$ and $\vec{v}_2$ are unit vectors, the condition $\vec{v}_1\cdot\vec{v}_2 = \|\vec{v}_1\|\|\vec{v}_2\|\cos(\alpha) = \cos(\alpha) = \frac{1}{2}$ implies that $\vec{v}_1$ and $\vec{v}_2$ enclose an angle of $60^\circ$ $\left(=\frac{\pi}{3}\right)$. The vectors $\vec{v}_1$ and $\vec{v}_3$ enclose an angle of $60^\circ$ as well.
In the case $n = 2$ there are two possible scenarios: either $\vec{v}_2 = \vec{v}_3$, or $\vec{v}_2$ and $\vec{v}_3$ enclose an angle of $120^\circ$. Therefore, either $\vec{v}_2\cdot\vec{v}_3 = 1$ or $\vec{v}_2\cdot\vec{v}_3 = \cos(120^\circ) = -\frac{1}{2}$.
In the case $n = 3$, the vectors $\vec{v}_2$ and $\vec{v}_3$ could enclose any angle between $0^\circ$ (if $\vec{v}_2 = \vec{v}_3$) and $120^\circ$, as illustrated in Figure 5.7. We have $-\frac{1}{2} \le \vec{v}_2\cdot\vec{v}_3 \le 1$.

For example, consider $\vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} \left(\frac{\sqrt{3}}{2}\right)\cos\theta \\ \left(\frac{\sqrt{3}}{2}\right)\sin\theta \\ \frac{1}{2} \end{bmatrix}$

Note that $\vec{v}_2\cdot\vec{v}_3 = \left(\frac{3}{4}\right)\sin\theta + \frac{1}{4}$ could be anything between $-\frac{1}{2}$ (when $\sin\theta = -1$) and 1 (when $\sin\theta = 1$), as claimed.

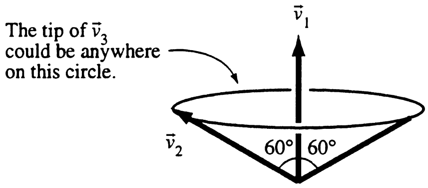



Figure 5.7: for Problem 5.1.38.

If $n$ exceeds three, we can consider the orthogonal projection $\vec{w}$ of $\vec{v}_3$ onto the plane $E$ spanned by $\vec{v}_1$ and $\vec{v}_2$.

Since $\text{proj}_{\vec{v}_1}\vec{w} = (\vec{v}_1 \cdot \vec{w})\vec{v}_1 = \frac{1}{2}\vec{v}_1$, and since $\|\vec{w}\| \leq \|\vec{v}_3\| = 1$, (by Theorem 5.1.10), the tip of $\vec{w}$ will be on the line segment in Figure 5.8. Note that the angle $\phi$ enclosed by the vectors $\vec{v}_2$ and $\vec{w}$ is between $0°$ and $120°$, so that $\cos\phi$ is between $-\frac{1}{2}$ and 1.

Therefore, $\vec{v}_2 \cdot \vec{v}_3 = \vec{v}_2 \cdot \vec{w} = \|\vec{w}\|\cos\phi$ is between $-\frac{1}{2}$ and 1.

This implies that $\angle(\vec{v}_2, \vec{v}_3)$ is between $0°$ and $120°$ as well. To see that all these values are attained, add $(n-3)$ zeros to the three vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ in $\mathbb{R}^3$ given above.

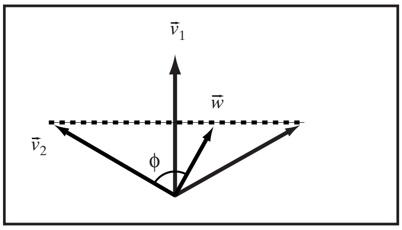



Figure 5.8: for Problem 5.1.38.

5.1.**39** No! By definition of a projection, the vector $\vec{x} - \text{proj}_L\vec{x}$ is perpendicular to $\text{proj}_L\vec{x}$, so that $(\vec{x} - \text{proj}_L\vec{x}) \cdot (\text{proj}_L\vec{x}) = \vec{x} \cdot \text{proj}_L\vec{x} - \|\text{proj}_L\vec{x}\|^2 = 0$ and $\vec{x} \cdot \text{proj}_L\vec{x} = \|\text{proj}_L\vec{x}\|^2 \geq 0$. (See Figure 5.9.)

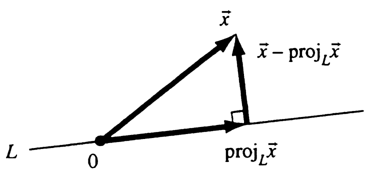



Figure 5.9: for Problem 5.1.39.

5.1.**40** $||\vec{v}_2|| = \sqrt{\vec{v}_2 \cdot \vec{v}_2} = \sqrt{a_{22}} = 3$.

5.1.**41** $\theta = \arccos(\frac{\vec{v}_2 \cdot \vec{v}_3}{||\vec{v}_2||||\vec{v}_3||}) = \arccos(\frac{a_{23}}{\sqrt{a_{22}}\sqrt{a_{33}}}) = \arccos(\frac{20}{21}) \approx 0.31$ radians.

5.1.**42** $||\vec{v}_1 + \vec{v}_2|| = \sqrt{(\vec{v}_1 + \vec{v}_2) \cdot (\vec{v}_1 + \vec{v}_2)} = \sqrt{a_{11} + 2a_{12} + a_{22}} = \sqrt{22}$.

5.1.**43** Let $\vec{u} = \frac{\vec{v}_2}{||\vec{v}_2||} = \frac{\vec{v}_2}{3}$. Then, $\vec{u}$ is an orthonormal basis for span$(\vec{v}_2)$. Using Theorem 5.1.5, $\text{proj}_{\vec{v}_2}(\vec{v}_1) = (\vec{u}\cdot\vec{v}_1)\vec{u} = (\frac{\vec{v}_2}{3}\cdot\vec{v}_1)\frac{\vec{v}_2}{3} = \frac{1}{3}(\vec{v}_2\cdot\vec{v}_1)\frac{\vec{v}_2}{3} = \frac{1}{3}(a_{12})\frac{\vec{v}_2}{3} = \frac{5}{9}\vec{v}_2$.

5.1.**44** One method to solve this is to take $\vec{v} = \vec{v}_2 - \text{proj}_{\vec{v}_3}\vec{v}_2 = \vec{v}_2 - \frac{20}{49}\vec{v}_3$.

5.1.**45** Write the projection as a linear combination of $\vec{v}_2$ and $\vec{v}_3$, $c_2\vec{v}_2 + c_3\vec{v}_3$. Now you want $\vec{v}_1 - c_2\vec{v}_2 - c_3\vec{v}_3$ to be perpendicular to $V$, that is, perpendicular to both $\vec{v}_2$ and $\vec{v}_3$. Using dot products, this boils down to two linear equation in two unknowns, $9c_2 + 20c_3 = 5$, and $20c_2 + 49c_3 = 11$, with the solution $c_2 = \frac{25}{41}$ and $c_3 = -\frac{1}{41}$. Thus the answer is $\frac{25}{41}\vec{v}_2 - \frac{1}{41}\vec{v}_3$.

5.1.**46** Write the projection as a linear combination of $\vec{v}_1$ and $\vec{v}_2 : c_1\vec{v}_1 + c_2\vec{v}_2$. Now we want $\vec{v}_3 - c_1\vec{v}_1 + c_2\vec{v}_2$ to be perpendicular to $V$, that is, perpendicular to both $\vec{v}_1$ and $\vec{v}_2$. Using dot products, this boils down to two linear equations in two unknowns, $11 = 3c_1 + 5c_2$ and $20 = 5c_1 + 9c_2$, with the solution $c_1 = -\frac{1}{2}, c_2 = \frac{5}{2}$. Thus, the answer is $-\frac{1}{2}\vec{v}_1 + \frac{5}{2}\vec{v}_2$.

## Section 5.2

In Exercises 1–14, we will refer to the given vectors as $\vec{v}_1, \ldots, \vec{v}_m$, where $m = 1, 2$, or 3.

5.2.**1** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$

5.2.**2** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{7}\begin{bmatrix} 6 \\ 3 \\ 2 \end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2 - (\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2 - (\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{7}\begin{bmatrix} 2 \\ -6 \\ 3 \end{bmatrix}$$

Note that $\vec{u}_1 \cdot \vec{v}_2 = 0$.

5.2.**3** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{5}\begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2 - (\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2 - (\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{5}\begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}$$

5.2.**4** $\vec{u}_1 = \frac{1}{5}\begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}$ and $\vec{u}_2 = \frac{1}{5}\begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}$ as in Exercise 3.

Since $\vec{v}_3$ is orthogonal to $\vec{u}_1$ and $\vec{u}_2, \vec{u}_3 = \frac{1}{\|\vec{v}_3\|}\vec{v}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$.

**5.2.5** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$

$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{\sqrt{18}}\begin{bmatrix} -1 \\ -1 \\ 4 \end{bmatrix} = \frac{1}{3\sqrt{2}}\begin{bmatrix} -1 \\ -1 \\ 4 \end{bmatrix}$

**5.2.6** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \vec{e}_1$

$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \vec{e}_2$

$\vec{u}_3 = \frac{\vec{v}_3^{\perp}}{\|\vec{v}_3^{\perp}\|} = \frac{\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2}{\|\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \vec{e}_3$

**5.2.7** Note that $\vec{v}_1$ and $\vec{v}_2$ are orthogonal, so that $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ and $\vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$. Then

$\vec{u}_3 = \frac{\vec{v}_3^{\perp}}{\|\vec{v}_3^{\perp}\|} = \frac{\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2}{\|\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2\|} = \frac{1}{\sqrt{36}}\begin{bmatrix} 2 \\ -4 \\ 4 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$.

**5.2.8** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{7}\begin{bmatrix} 5 \\ 4 \\ 2 \\ 2 \end{bmatrix}$

$\vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \frac{1}{7}\begin{bmatrix} -2 \\ 2 \\ 5 \\ -4 \end{bmatrix}$

**5.2.9** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{10}\begin{bmatrix} -1 \\ 7 \\ -7 \\ 1 \end{bmatrix}$

**5.2.10** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}$$

5.2.**11** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{5}\begin{bmatrix}4\\0\\0\\3\end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{\sqrt{225}}\begin{bmatrix}-3\\2\\14\\4\end{bmatrix} = \frac{1}{15}\begin{bmatrix}-3\\2\\14\\4\end{bmatrix}$$

5.2.**12** $\vec{u}_1 = \frac{1}{7}\begin{bmatrix}2\\3\\0\\6\end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{3}\begin{bmatrix}0\\-2\\2\\1\end{bmatrix}$$

5.2.**13** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$$

$$\vec{u}_3 = \frac{\vec{v}_3^{\perp}}{\|\vec{v}_3^{\perp}\|} = \frac{\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2}{\|\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2\|} = \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$$

5.2.**14** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{10}\begin{bmatrix}1\\7\\1\\7\end{bmatrix}$

$$\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\0\\1\\0\end{bmatrix}$$

$$\vec{u}_3 = \frac{\vec{v}_3^{\perp}}{\|\vec{v}_3^{\perp}\|} = \frac{\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2}{\|\vec{v}_3-(\vec{u}_1\cdot\vec{v}_3)\vec{u}_1-(\vec{u}_2\cdot\vec{v}_3)\vec{u}_2\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

In Exercises 15–28, we will use the results of Exercises 1–14 (note that Exercise $k$, where $k = 1, \ldots, 14$, gives the $QR$ factorization of the matrix in Exercise $(k+14)$). We can set $Q = [\vec{u}_1 \ldots \vec{u}_m]$; the entries of $R$ are

$$\begin{array}{lll} r_{11} & = \|\vec{v}_1\| & \\ r_{22} & = \|\vec{v}_2^{\perp}\| & = \|\vec{v}_2 - (\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\| \\ r_{33} & = \|\vec{v}_3^{\perp}\| & = \|\vec{v}_3 - (\vec{u}_1\cdot\vec{v}_3)\vec{u}_1 - (\vec{u}_2\cdot\vec{v}_3)\vec{u}_2\| \\ r_{ij} & & = \vec{u}_i\cdot\vec{v}_j, \text{ where } i<j. \end{array}$$

5.2.**15** $Q = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, R = [3]$

5.2.**16** $Q = \frac{1}{7}\begin{bmatrix} 6 & 2 \\ 5 & -6 \\ 2 & 3 \end{bmatrix}, R = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$

5.2.**17** $Q = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 0 & 0 \\ 5 & -4 \end{bmatrix}, R = \begin{bmatrix} 5 & 5 \\ 0 & 35 \end{bmatrix}$

5.2.**18** $Q = \frac{1}{5}\begin{bmatrix} 4 & 3 & 0 \\ 0 & 0 & -5 \\ 5 & -4 & 0 \end{bmatrix}, R = \begin{bmatrix} 5 & 5 & 0 \\ 0 & 35 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

5.2.**19** $Q = \frac{1}{3}\begin{bmatrix} 2 & -\frac{1}{\sqrt{2}} \\ 2 & -\frac{1}{\sqrt{2}} \\ 1 & \frac{4}{\sqrt{2}} \end{bmatrix}, R = 3\begin{bmatrix} 1 & 1 \\ 0 & \sqrt{2} \end{bmatrix}$

5.2.**20** $Q = I_3$, $R = [\,\vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3\,] = \begin{bmatrix} 2 & 3 & 5 \\ 0 & 4 & 6 \\ 0 & 0 & 7 \end{bmatrix}$

5.2.**21** $Q = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{bmatrix}, R = \begin{bmatrix} 3 & 0 & 12 \\ 0 & 3 & -12 \\ 0 & 0 & 6 \end{bmatrix}$

5.2.**22** $Q = \frac{1}{7}\begin{bmatrix} 5 & -2 \\ 4 & 2 \\ 2 & 5 \\ 2 & -4 \end{bmatrix}, R = \begin{bmatrix} 7 & 7 \\ 0 & 7 \end{bmatrix}$

**5.2.23** $Q = \begin{bmatrix} 0.5 & -0.1 \\ 0.5 & 0.7 \\ 0.5 & -0.7 \\ 0.5 & 0.1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 4 \\ 0 & 10 \end{bmatrix}$

**5.2.24** $Q = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 10 \\ 0 & 2 \end{bmatrix}$

**5.2.25** $Q = \frac{1}{15}\begin{bmatrix} 12 & -3 \\ 0 & 2 \\ 0 & 14 \\ 9 & 4 \end{bmatrix}$, $R = \begin{bmatrix} 5 & 10 \\ 0 & 15 \end{bmatrix}$

**5.2.26** $Q = \begin{bmatrix} \frac{2}{7} & 0 \\ \frac{3}{7} & -\frac{2}{3} \\ 0 & \frac{2}{3} \\ \frac{6}{7} & \frac{1}{3} \end{bmatrix}$, $R = \begin{bmatrix} 7 & 14 \\ 0 & 3 \end{bmatrix}$

**5.2.27** $Q = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$

**5.2.28** $Q = \begin{bmatrix} \frac{1}{10} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{7}{10} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{10} & \frac{1}{\sqrt{2}} & 0 \\ \frac{7}{10} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$, $R = \begin{bmatrix} 10 & 10 & 10 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$

**5.2.29** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{5}\begin{bmatrix} -3 \\ 4 \end{bmatrix}$ $\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \frac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix}$. (See Figure 5.10.)

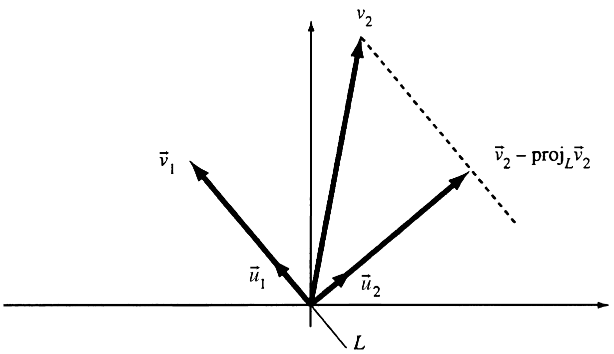



Figure 5.10: for Problem 5.2.29.

5.2.**30** See Figure 5.11.

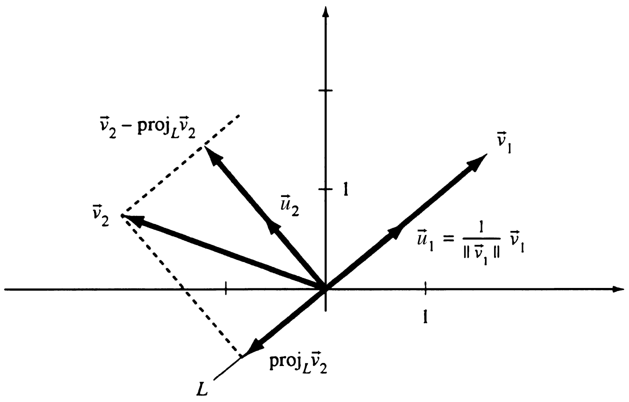



Figure 5.11: for Problem 5.2.30.

5.2.**31** $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \vec{e}_1$

$\vec{v}_2^{\perp} = \vec{v}_2 - \text{proj}_{V_1}\vec{v}_2 = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix} - \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ c \\ 0 \end{bmatrix}$, so that $\vec{u}_2 = \frac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \vec{e}_2$

Here $V_1 = \text{span}(\vec{e}_1) = x$ axis.

$\vec{v}_3^{\perp} = \vec{v}_3 - \text{proj}_{V_2}\vec{v}_3 = \begin{bmatrix} d \\ e \\ f \end{bmatrix} - \begin{bmatrix} d \\ e \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ f \end{bmatrix}$, so that $\vec{u}_3 = \frac{\vec{v}_3^{\perp}}{\|\vec{v}_3^{\perp}\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \vec{e}_3$.

Here $V_2 = \text{span}(\vec{e}_1, \vec{e}_2) = x$-$y$ plane. (See Figure 5.12.)

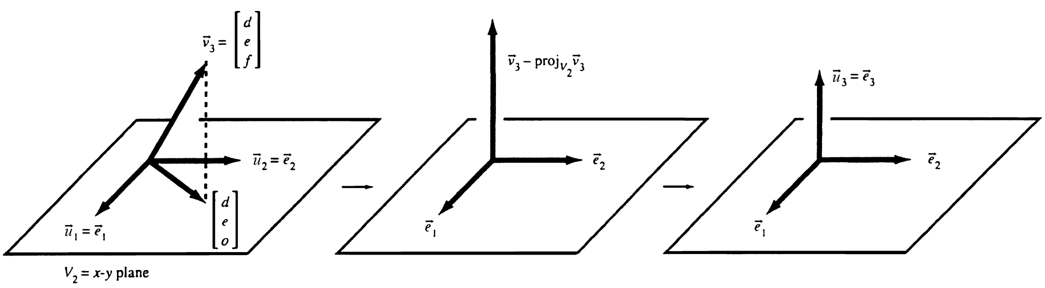



Figure 5.12: for Problem 5.2.31.

5.2.**32** A basis of the plane is $\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Now apply the Gram-Schmidt process.

$$\vec{u}_1 = \tfrac{1}{\|\vec{v}_1\|}\vec{v}_1 = \tfrac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix}$$

$$\vec{u}_2 = \tfrac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \tfrac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \tfrac{1}{\sqrt{6}}\begin{bmatrix}-1\\-1\\2\end{bmatrix}$$

Your solution may be different if you start with a different basis $\vec{v}_1, \vec{v}_2$ of the plane.

5.2.**33** $\text{rref}(A) = \begin{bmatrix}1&0&0&1\\0&1&1&0\end{bmatrix}$

A basis of $\ker(A)$ is $\vec{v}_1 = \begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}0\\-1\\1\\0\end{bmatrix}$.

Since $\vec{v}_1$ and $\vec{v}_2$ are orthogonal already, we obtain $\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$, $\vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\-1\\1\\0\end{bmatrix}$.

5.2.**34** $\text{rref}(A) = \begin{bmatrix}1&0&-1&-2\\0&1&2&3\end{bmatrix}$

A basis of $\ker(A)$ is $\vec{v}_1 = \begin{bmatrix}1\\-2\\1\\0\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}2\\-3\\0\\1\end{bmatrix}$.

We apply the Gram-Schmidt process and obtain

$$\vec{u}_1 = \tfrac{1}{\|\vec{v}_1\|}\vec{v}_1 = \tfrac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\\0\end{bmatrix}$$

$$\vec{u}_2 = \tfrac{\vec{v}_2^{\perp}}{\|\vec{v}_2^{\perp}\|} = \tfrac{\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1}{\|\vec{v}_2-(\vec{u}_1\cdot\vec{v}_2)\vec{u}_1\|} = \tfrac{1}{\sqrt{30}}\begin{bmatrix}2\\-1\\-4\\3\end{bmatrix}$$

5.2.**35** $\text{rref}(A) = \begin{bmatrix}1&0&\frac{1}{3}\\0&1&\frac{1}{3}\\0&0&0\end{bmatrix}$

The non-redundant columns of $A$ give us a basis of $\text{im}(A)$:

$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$

Since $\vec{v}_1$ and $\vec{v}_2$ are orthogonal already, we obtain $\vec{u}_1 = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, $\vec{u}_2 = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$.

5.2.**36** Write $M = \underbrace{\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}}_{\uparrow \atop Q_0} \underbrace{\begin{bmatrix} 2 & 3 & 5 \\ 0 & -4 & 6 \\ 0 & 0 & 7 \end{bmatrix}}_{\uparrow \atop R_0}$

This is almost the $QR$ factorization of $M$: the matrix $Q_0$ has orthonormal columns and $R_0$ is upper triangular; the only problem is the entry $-4$ on the diagonal of $R_0$. Keeping in mind how matrices are multiplied, we can change all the signs in the second column of $Q_0$ and in the second row of $R_0$ to fix this problem:

$$M = \underbrace{\frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}}_{\uparrow \atop Q} \underbrace{\begin{bmatrix} 2 & 3 & 5 \\ 0 & 4 & -6 \\ 0 & 0 & 7 \end{bmatrix}}_{\uparrow \atop R}$$

5.2.**37** Write $M = \underbrace{\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}}_{\uparrow \atop Q_0} \underbrace{\begin{bmatrix} 3 & 4 \\ 0 & 5 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\uparrow \atop R_0}$

Note that the last two columns of $Q_0$ and the last two rows of $R_0$ have no effect on the product $Q_0R_0$; if we drop them, we have the Q$R$ factorization of $M$:

$$M = \underbrace{\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\uparrow \atop Q} \underbrace{\begin{bmatrix} 3 & 4 \\ 0 & 5 \end{bmatrix}}_{\uparrow \atop R}$$

5.2.**38** Since $\vec{v}_1 = 2\vec{e}_3$, $\vec{v}_2 = -3\vec{e}_1$ and $\vec{v}_3 = 4\vec{e}_4$ are orthogonal, we have

$$Q = \begin{bmatrix} \frac{\vec{v}_1}{\|\vec{v}_1\|} & \frac{\vec{v}_2}{\|\vec{v}_2\|} & \frac{\vec{v}_3}{\|\vec{v}_3\|} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } R = \begin{bmatrix} \|\vec{v}_1\| & 0 & 0 \\ 0 & \|\vec{v}_2\| & 0 \\ 0 & 0 & \|\vec{v}_3\| \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

5.2.**39** $\vec{u}_1 = \frac{1}{\sqrt{14}}\begin{bmatrix}1\\2\\3\end{bmatrix}$, $\vec{u}_2 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\-1\end{bmatrix}$, $\vec{u}_3 = \vec{u}_1 \times \vec{u}_2 = \frac{1}{\sqrt{42}}\begin{bmatrix}-5\\4\\-1\end{bmatrix}$

5.2.**40** If $\vec{v}_1, \ldots, \vec{v}_n$ are the columns of $A$, then $Q = \begin{bmatrix} \frac{\vec{v}_1}{\|\vec{v}_1\|} & \cdots & \frac{\vec{v}_n}{\|\vec{v}_n\|}\end{bmatrix}$ and $R = \begin{bmatrix}\|\vec{v}_1\| & & 0\\ & \ddots & \\ 0 & & \|\vec{v}_n\|\end{bmatrix}$.

(See Exercise 38 as an example.)

5.2.**41** If all diagonal entries of $A$ are positive, then we have $Q = I_n$ and $R = A$. A small modification is necessary if $A$ has negative entries on the diagonal: if $a_{ii} < 0$ we let $r_{ij} = -a_{ij}$ for all $j$, and we let $q_{ii} = -1$; if $a_{ii} > 0$ we let $r_{ij} = a_{ij}$ and $q_{ii} = 1$. Furthermore, $q_{ij} = 0$ if $i \neq j$ (that is, $Q$ is diagonal).

For example, $\underbrace{\begin{bmatrix}-1 & 2 & 3\\0 & 4 & 5\\0 & 0 & -6\end{bmatrix}}_{A} = \underbrace{\begin{bmatrix}-1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{bmatrix}}_{Q}\underbrace{\begin{bmatrix}1 & -2 & -3\\0 & 4 & 5\\0 & 0 & 6\end{bmatrix}}_{R}$

5.2.**42** We have $r_{11} = \|\vec{v}_1\|$ and $r_{22} = \|\vec{v}_2^{\perp}\| = \|\vec{v}_2 - \text{proj}_L \vec{v}_2\|$, so that $r_{11}r_{22}$ is the area of the parallelogram defined by $\vec{v}_1$ and $\vec{v}_2$. See Figure 5.13.

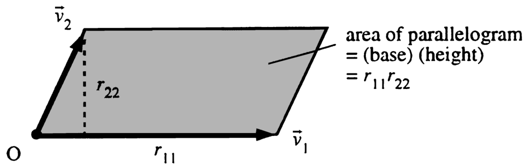



Figure 5.13: for Problem 5.2.42.

5.2.**43** Partition the matrices $Q$ and $R$ in the $QR$ factorization of $A$ as follows:

$$[\,A_1 \quad A_2\,] = A = QR = [\,Q_1 \quad Q_2\,]\begin{bmatrix}R_1 & R_2\\0 & R_3\end{bmatrix} = [\,Q_1R_1 \quad Q_1R_2 + Q_2R_3\,],$$

where $Q_1$ is $n \times m_1, Q_2$ is $n \times m_2, R_1$ is $m_1 \times m_1$, and $R_3$ is $m_2 \times m_2$.

Then, $A_1 = Q_1R_1$ is the $QR$ factorization of $A_1$: note that the columns of $A_1$ are orthonormal, and $R_1$ is upper triangular with positive diagonal entries.

5.2.**44** No! If $m$ exceeds $n$, then there is no $n \times m$ matrix $Q$ with orthonormal columns (if the columns of a matrix are orthonormal, then they are linearly independent).

5.2.**45** Yes. Let $A = [\,\vec{v}_1 \quad \cdots \quad \vec{v}_m\,]$. The idea is to perform the Gram-Schmidt process in reversed order, starting with $\vec{u}_m = \frac{1}{\|\vec{v}_m\|}\vec{v}_m$.

Then we can express $\vec{v}_j$ as a linear combination of $\vec{u}_j, \ldots, \vec{u}_m$, so that $[\,\vec{v}_1 \quad \cdots \quad \vec{v}_j \quad \cdots \quad \vec{v}_m\,] = [\,\vec{u}_1 \quad \cdots \quad \vec{u}_j \quad \cdots \quad \vec{u}_m\,]\,L$ for some *lower* triangular matrix $L$, with

$$\vec{v}_j = [\,\vec{u}_1 \quad \cdots \quad \vec{u}_j \quad \cdots \quad \vec{u}_m\,] \begin{bmatrix} l_{1j} \\ \cdots \\ l_{jj} \\ \cdots \\ l_{mj} \end{bmatrix} = l_{jj}\vec{u}_j + \cdots + l_{mj}\vec{u}_m.$$

## Section 5.3

**5.3.1** Not orthogonal, the column vectors fail to be perpendicular to each other.

**5.3.2** This matrix is orthogonal. Check that the column vectors are unit vectors, and that they are perpendicular to each other.

**5.3.3** This matrix is orthogonal. Check that the column vectors are unit vectors, and that they are perpendicular to each other.

**5.3.4** Not orthogonal, the first and third column vectors fail to be perpendicular to each other.

**5.3.5** $3A$ will not be orthogonal, because the length of the column vectors will be 3 instead of 1, and they will fail to be unit vectors.

**5.3.6** $-B$ will certainly be orthogonal, since the columns will be perpendicular unit vectors.

**5.3.7** $AB$ is orthogonal by Theorem 5.3.4a.

**5.3.8** $A + B$ will not necessarily be orthogonal, because the columns may not be unit vectors. For example, if $A = B = I_n$, then $A + B = 2I_n$, which is not orthogonal.

**5.3.9** $B^{-1}$ is orthogonal by Theorem 5.3.4b.

**5.3.10** This matrix will be orthogonal, by Theorem 5.3.4.

**5.3.11** $A^T$ is orthogonal. $A^T = A^{-1}$, by Theorem 5.3.7, and $A^{-1}$ is orthogonal by Theorem 5.3.4b.

**5.3.13** $3A$ is symmetric, since $(3A)^T = 3A^T = 3A$.

**5.3.14** $-B$ is symmetric, since $(-B)^T = -B^T = -B$.

**5.3.15** $AB$ is not necessarily symmetric, since $(AB)^T = B^T A^T = BA$, which is not necessarily the same as $AB$. (Here we used Theorem 5.3.9a.)

**5.3.16** $A + B$ is symmetric, since $(A + B)^T = A^T + B^T = A + B$.

**5.3.17** $B^{-1}$ is symmetric, because $(B^{-1})^T = (B^T)^{-1} = B^{-1}$. In the first step we have used 5.3.9b.

**5.3.18** $A^{10}$ is symmetric, since $(A^{10})^T = (A^T)^{10} = A^{10}$.

**5.3.19** This matrix is symmetric. First note that $(A^2)^T = (A^T)^2 = A^2$ for a symmetric matrix $A$. Now we can use the linearity of the transpose, $(2I_n + 3A - 4A^2)^T = 2I_n^T + 3A^T - (4A^2)^T = 2I_n + 3A - 4(A^T)^2 = 2I_n + 3A - 4A^2$.

**5.3.20** $AB^2A$ is symmetric, since $(AB^2A)^T = (ABBA)^T = (BA)^T(AB)^T = A^TB^TB^TA^T = AB^2A$.

**5.3.21** Symmetric. $(A^TA)^T = A^T(A^T)^T = A^TA$.

**5.3.22** $BB^T$ is symmetric: $(BB^T)^T = (B^T)^TB^T = BB^T$.

**5.3.23** Not necessarily symmetric. $(A - A^T)^T = A^T - A = -(A - A^T)$.

**5.3.24** Not necessarily symmetric. $(A^TBA)^T = A^T(A^TB)^T = A^TB^TA$.

**5.3.25** Symmetric, because $(A^TB^TBA)^T = A^TB^T(B^T)^T(A^T)^T = A^TB^TBA$.

**5.3.26** Symmetric, since $(B(A + A^T)B^T)^T = ((A + A^T)B^T)^TB^T = B(A + A^T)^TB^T$

$= B(A^T + A)^TB^T = B((A^T)^T + A^T)B^T = B(A + A^T)B^T$.

**5.3.27** Using Theorems 5.3.6 and 5.3.9a, we find that $(A\vec{v}) \cdot \vec{w} = (A\vec{v})^T\vec{w} = \vec{v}^TA^T\vec{w} = \vec{v} \cdot (A^T\vec{w})$, as claimed.

**5.3.28** We will follow the hint.

$(iv) \Rightarrow (vi)$ : If $A^TA = I_n$, then $(A\vec{x}) \cdot (A\vec{y}) = (A\vec{x})^T(A\vec{y}) = \vec{x}^TA^TA\vec{y} = \vec{x}^TI_n\vec{y} = \vec{x} \cdot \vec{y}$ for all $\vec{x}$ and $\vec{y}$.

$(vi) \Rightarrow (ii)$ : If $(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y}$ for all $\vec{x}$ and $\vec{y}$, then $\|A\vec{x}\| = \sqrt{(A\vec{x}) \cdot (A\vec{x})} = \sqrt{\vec{x} \cdot \vec{x}} = \|\vec{x}\|$ for all $\vec{x}$.

Recall that the equivalence of statements $(i)$ through $(v)$ is proven in the text.

**5.3.29** We will use the fact that $L$ preserves length (by Definition 5.3.1) and the dot product, by Summary 5.3.8 (vi).

$\angle(L(\vec{v}), L(\vec{w})) = \arccos \frac{L(\vec{v}) \cdot L(\vec{w})}{\|L(\vec{v})\|\|L(\vec{w})\|} = \arccos \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\|\|\vec{w}\|} = \angle(\vec{v}, \vec{w})$.

**5.3.30** If $L(\vec{x}) = \vec{0}$, then $\|L(\vec{x})\| = \|\vec{x}\| = 0$, so that $\vec{x} = \vec{0}$. Therefore, $\ker(L) = \{\vec{0}\}$.

By Theorem 3.3.7, $\dim(\text{im}(L)) = m - \dim(\ker(L)) = m$.

Since $\mathbb{R}^n$ has an $m$-dimensional subspace (namely, $\text{im}(L)$), the inequality $m \leq n$ holds.

The transformation $L$ preserves right angles (the proof of Theorem 5.3.2 applies), so that the columns of $A$ are orthonormal (since they are $L(\vec{e}_1), \ldots, L(\vec{e}_m)$).

Therefore, we have $A^TA = I_m$ (the proof of Theorem 5.3.7 applies).

Since the vectors $\vec{v}_1, \ldots, \vec{v}_m$ form an orthonormal basis of $\text{im}(A)$, the matrix $AA^T$ represents the orthogonal projection onto $\text{im}(A)$, by Theorem 5.3.10.

A simple example of such a transformation is $L(\vec{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \vec{x}$, that is, $L \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$.

5.3.**31** Yes! If $A$ is orthogonal, then so is $A^T$, by Exercise 11. Since the columns of $A^T$ are orthogonal, so are the rows of $A$.

5.3.**32** a No! As a counterexample, consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ (see Exercise 30).

b Yes! More generally, if $A$ and $B$ are $n \times n$ matrices such that $BA = I_n$, then $AB = I_n$, by Theorem 2.4.8c.

5.3.**33** Write $A = [\,\vec{v}_1 \quad \vec{v}_2\,]$. The unit vector $\vec{v}_1$ can be expressed as $\vec{v} = \begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix}$, for some $\phi$. Then $\vec{v}_2$ will be one of the two unit vectors orthogonal to $\vec{v}_1 : \vec{v}_2 = \begin{bmatrix} -\sin(\phi) \\ \cos(\phi) \end{bmatrix}$ or $\vec{v}_2 = \begin{bmatrix} \sin(\phi) \\ -\cos(\phi) \end{bmatrix}$. (See Figure 5.7.)

Therefore, an orthogonal $2 \times 2$ matrix is either of the form $A = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$ or $A = \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ \sin(\phi) & -\cos(\phi) \end{bmatrix}$, representing a rotation or a reflection. Compare with Exercise 2.2.24. See Figure 5.14.

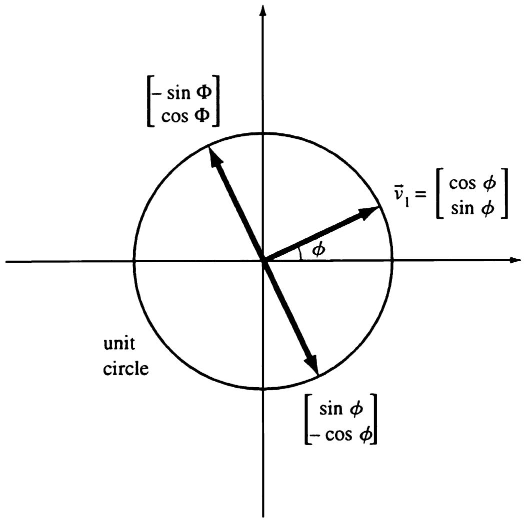



Figure 5.14: for Problem 5.3.33.

5.3.**34** Since the first two columns are orthogonal to the third, we have $c = d = 0$. Then $\begin{bmatrix} a & b \\ e & f \end{bmatrix}$ is an orthogonal $2 \times 2$ matrix; By Exercise 33, the $3 \times 3$ matrix $A$ is either of the form $A = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ 0 & 0 & 1 \\ \sin(\phi) & \cos(\phi) & 0 \end{bmatrix}$ or $A = \begin{bmatrix} \cos(\phi) & \sin(\phi) & 0 \\ 0 & 0 & 1 \\ \sin(\phi) & -\cos(\phi) & 0 \end{bmatrix}$.

5.3.**35** Let us first think about the inverse $L = T^{-1}$ of $T$.

Write $L(\vec{x}) = A\vec{x} = [\,\vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3\,]\,\vec{x}$. It is required that $L(\vec{e}_3) = \vec{v}_3 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$.

Furthermore, the vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$ must form an orthonormal basis of $\mathbb{R}^3$. By inspection, we find $\vec{v}_1 = \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$.

Then $\vec{v}_2 = \vec{v}_1 \times \vec{v}_3 = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}$ does the job. In summary, we have $L(\vec{x}) = \frac{1}{3}\begin{bmatrix} -2 & -1 & 2 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{bmatrix}\vec{x}$.

Since the matrix of $L$ is orthogonal, the matrix of $T = L^{-1}$ is the transpose of the matrix of $L$:

$T(\vec{x}) = \frac{1}{3}\begin{bmatrix} -2 & 1 & 2 \\ -1 & 2 & -2 \\ 2 & 2 & 1 \end{bmatrix}\vec{x}$.

There are many other answers (since there are many choices for the vector $\vec{v}_1$ above).

5.3.**36** Let the third column be the cross product of the first two: $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} \\ \frac{2}{3} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} \\ \frac{1}{3} & 0 & -\frac{4}{\sqrt{18}} \end{bmatrix}$.

There is another solution, with the signs in the last column reversed.

5.3.**37** No, since the vectors $\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 2 \\ 0 \end{bmatrix}$ are orthogonal, whereas $\begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$ are not (see Theorem 5.3.2).

5.3.**38** a The general form of a skew-symmetric $3 \times 3$ matrix is $A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$, with

$A^2 = \begin{bmatrix} -a^2 - b^2 & -bc & ac \\ -bc & -a^2 - c^2 & -ab \\ ac & -ab & -b^2 - c^2 \end{bmatrix}$, a symmetric matrix.

b By Theorem 5.3.9.a, $(A^2)^T = (A^T)^2 = (-A)^2 = A^2$, so that $A^2$ is symmetric.

5.3.**39** By Theorem 5.3.10, the matrix of the projection is $\vec{u}\vec{u}^T$; the $ij$th entry of this matrix is $u_i u_j$.

5.3.**40** An orthonormal basis of $W$ is $\vec{u}_1 = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} -0.1 \\ 0.7 \\ -0.7 \\ 0.1 \end{bmatrix}$ (see Exercise 5.2.9).

By Theorem 5.3.10, the matrix of the projection onto $W$ is $QQ^T$, where $Q = [\,\vec{u}_1 \quad \vec{u}_2\,]$.

$$QQ^T = \frac{1}{100}\begin{bmatrix} 26 & 18 & 32 & 24 \\ 18 & 74 & -24 & 32 \\ 32 & -24 & 74 & 18 \\ 24 & 32 & 18 & 26 \end{bmatrix}$$

5.3.**41** A unit vector on the line is $\vec{u} = \frac{1}{\sqrt{n}}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.

The matrix of the orthogonal projection is $\vec{u}\vec{u}^T$, the $n \times n$ matrix whose entries are all $\frac{1}{n}$ (compare with Exercise 39).

5.3.**42** a Suppose we are projecting onto a subspace $W$ of $\mathbb{R}^n$. Since $A\vec{x}$ is in $W$ already, the orthogonal projection of $A\vec{x}$ onto $W$ is just $A\vec{x}$ itself: $A(A\vec{x}) = A\vec{x}$, or $A^2\vec{x} = A\vec{x}$.

Since this equation holds for all $\vec{x}$, we have $A^2 = A$.

b $A = QQ^T$, for some matrix $Q$ with orthonormal columns $\vec{u}_1, \dots, \vec{u}_m$. Note that $Q^TQ = I_m$, since the $ij$th entry of $Q^TQ$ is $\vec{u}_i \cdot \vec{u}_j$. Then $A^2 = QQ^TQQ^T = Q(Q^TQ)Q^T = QI_mQ^T = QQ^T = A$.

5.3.**43** Examine how $A$ acts on $\vec{u}$, and on a vector $\vec{v}$ orthogonal to $\vec{u}$:

$A\vec{u} = (2\vec{u}\vec{u}^T - I_3)\vec{u} = 2\vec{u}\vec{u}^T\vec{u} - \vec{u} = \vec{u},$ since $\vec{u}^T\vec{u} = \vec{u}\cdot\vec{u} = \|\vec{u}\|^2 = 1$.

$A\vec{v} = (2\vec{u}\vec{u}^T - I_3)\vec{v} = 2\vec{u}\vec{u}^T\vec{v} - \vec{v} = -\vec{v},$ since $\vec{u}^T\vec{v} = \vec{u}\cdot\vec{v} = 0$.

Since $A$ leaves the vectors in $L = \text{span}(\vec{u})$ unchanged and reverses the vectors in $V = L^\perp$, it represents the *reflection about* $L$.

Note that $B = -A$, so that $B$ reverses the vectors in $L$ and leaves the vectors in $V$ unchanged; that is, $B$ represents the reflection about $V$.

5.3.**44** Note that $A^T$ is an $m \times n$ matrix. By Theorems 3.3.7 and 5.3.9c we have

$\dim(\ker(A^T)) = n - \text{rank}(A^T) = n - \text{rank}(A)$.

By Theorem 3.3.6, $\dim(\text{im}(A)) = \text{rank}(A)$, so that $\dim(\text{im}(A)) + \dim(\ker(A^T)) = n$.

5.3.**45** Note that $A^T$ is an $m \times n$ matrix. By Theorems 3.3.7 and 5.3.9c, we have

$\dim(\ker(A)) = m - \text{rank}(A)$ and $\dim(\ker(A^T)) = n - \text{rank}(A^T) = n - \text{rank}(A)$,

so that $\dim(\ker(A)) = \dim(\ker(A^T))$ if (and only if) $A$ is a square matrix.

5.3.**46** By Theorem 5.2.2, the columns $\vec{u}_1, \dots, \vec{u}_m$ of $Q$ are orthonormal. Therefore, $Q^TQ = I_m$, since the $ij$th entry of $Q^TQ$ is $\vec{u}_i \cdot \vec{u}_j$.

If we multiply the equation $M = QR$ by $Q^T$ from the left then $Q^TM = Q^TQR = R$, as claimed.

5.3.**47** By Theorem 5.2.2, the columns $\vec{u}_1, \dots, \vec{u}_m$ of $Q$ are orthonormal. Therefore, $Q^TQ = I_m$, since the $ij$th entry of $Q^TQ$ is $\vec{u}_i \cdot \vec{u}_j$.

By Theorem 5.3.9a, we now have $A^TA = (QR)^TQR = R^TQ^TQR = R^TR$.

5.3.**48** As suggested, we consider the $QR$ factorization

$A^T = PR$

of $A^T$, where $P$ is orthogonal and $R$ is upper triangular with positive diagonal entries.

By Theorem 5.3.9a, $A = (PR)^T = R^T P^T$.

Note that $L = R^T$ is lower triangular and $Q = P^T$ is orthogonal.

5.3.**49** Yes! By Exercise 5.2.45, we can write $A^T = PL$, where $P$ is orthogonal and $L$ is lower triangular.

By Theorem 5.3.9a, $A = (PL)^T = L^T P^T$.

Note that $R = L^T$ is upper triangular, and $Q = P^T$ is orthogonal (by Exercise 11).

5.3.**50** a If an $n \times n$ matrix $A$ is orthogonal and upper triangular, then $A^{-1}$ is both lower triangular (since $A^{-1} = A^T$) and upper triangular (being the inverse of an upper triangular matrix; compare with Exercise 2.4.35c).

Therefore, $A^{-1} = A^T$ is a diagonal matrix, and so is $A$ itself. Since $A$ is orthogonal with positive diagonal entries, all the diagonal entries must be 1, so that $A = I_n$.

b Using the terminology suggested in the hint, we observe that $Q_2^{-1}Q_1$ is orthogonal (by Theorem 5.3.4) and $R_2R_1^{-1}$ is upper triangular with positive diagonal entries. By part a, the matrix $Q_2^{-1}Q_1 = R_2R_1^{-1}$ is $I_n$, so that $Q_1 = Q_2$ and $R_1 = R_2$, as claimed.

5.3.**51** a Using the terminology suggested in the hint, we observe that $I_m = Q_1^T Q_1 = (Q_2S)^T Q_2 S = S^T Q_2^T Q_2 S = S^T S$, so that $S$ is orthogonal, by Theorem 5.3.7.

b Using the terminology suggested in the hint, we observe that $R_2R_1^{-1}$ is both orthogonal (let $S = R_2R_1^{-1}$ in part a) and upper triangular, with positive diagonal entries. By Exercise 50a, we have $R_2R_1^{-1} = I_m$, so that $R_1 = R_2$. Then $Q_1 = Q_2R_2R_1^{-1} = Q_2$, as claimed.

5.3.**52** Applying the strategy outlined in Summary 4.1.6 to the general element $\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$ of $V$, we find the basis

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, so that $\dim(V) = 6$.

5.3.**53** Applying the strategy outlined in Summary 4.1.6 to the general element $\begin{bmatrix} 0 & b & c \\ -b & 0 & e \\ -c & -e & 0 \end{bmatrix}$ of $V$, we find the basis $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$, so that $\dim(V) = 3$.

5.3.**54** To write the general form of a skew-symmetric $n \times n$ matrix $A$, we can place arbitrary constants above the diagonal, the opposite entries below the diagonal ($a_{ij} = -a_{ji}$), and zeros on the diagonal (since $a_{ii} = -a_{ii}$). See Exercise 53 for the case $n = 3$. Thus the dimension of the space equals the number of entries above the diagonal of an $n \times n$ matrix. In Exercise 55 we will see that there are $(n^2 - n)/2$ such entries. Thus $\dim(V) = (n^2 - n)/2$.

5.3.**55** To write the general form of a symmetric $n \times n$ matrix $A$, we can place arbitrary constants on and above the diagonal, and then write the corresponding entries below the diagonal ($a_{ij} = a_{ji}$). See Exercise 52 for the case $n = 3$. Thus the dimension of the space equals the number of entries on and above the diagonal of an $n \times n$ matrix. Now there are $n^2$ entries in the matrix, $n^2 - n$ off the diagonal, and half of them, $(n^2 - n)/2$, above the diagonal. Since there are $n$ entries on the diagonal, we have $\dim(V) = (n^2 - n)/2 + n = (n^2 + n)/2$.

5.3.**56** Yes and yes (see Exercise 57).

5.3.**57** Yes, $L$ is linear, since $L(A+B) = (A+B)^T = A^T + B^T = L(A) + L(B)$ and $L(kA) = (kA)^T = kA^T = kL(A)$. Yes, $L$ is an isomorphism; the inverse is the transformation $R(A) = A^T$ from $\mathbb{R}^{n\times m}$ to $\mathbb{R}^{m\times n}$.

5.3.**58** Adapting the solution of Exercise 59, we see that the kernel consists of all skew-symmetric matrices, and the image consists of all symmetric matrices.

5.3.**59** The kernel consists of all matrixes $A$ such that $L(A) = \frac{1}{2}(A - A^T) = 0$, that is, $A^T = A$; those are the symmetric matrices.

Following the hint, let's apply $L$ to a skew-symmetric matrix $A$, with $A^T = -A$. Then $L(A) = (1/2)(A - A^T) = (1/2)2A = A$, so that $A$ is in the image of $L$. Conversely, if $A$ is any $2 \times 2$ matrix, then $L(A)$ will be skew-symmetric, since $(L(A))^T = (1/2)(A - A^T)^T = (1/2)(A^T - A) = -L(A)$. In conclusion: The kernel of $L$ consists of all symmetric matrices, and the image consists of all skew-symmetric matrices.

5.3.**60** Using Theorem 4.3.2, we find the matrix $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

5.3.**61** Note that the first three matrices of the given basis $\mathcal{B}$ are symmetric, so that $L(A) = A - A^T = 0$, and the coordinate vector $[L(A)]_{\mathcal{B}}$ is $\vec{0}$ for all three of them. The last matrix of the basis is skew-symmetric, so that $L(A) = 2A$, and $[L(A)]_{\mathcal{B}} = 2\vec{e}_4$. Using Theorem 4.3.2, we find that the $\mathcal{B}$-matrix of $L$ is $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

5.3.**62** By Theorem 5.3.9a, $A^T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

5.3.**63** By Exercise 2.4.94b, the given LDU factorization of $A$ is unique.

By Theorem 5.3.9a, $A = A^T = (LDU)^T = U^T D^T L^T = U^T D L^T$ is another way to write the LDU factorization of $A$ (since $U^T$ is lower triangular and $L^T$ is upper triangular). By the uniqueness of the LDU factorization, we have $U = L^T$ (and $L = U^T$), as claimed.

5.3.**64** a $\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix} + \begin{bmatrix} C & -D^T \\ D & C^T \end{bmatrix} = \begin{bmatrix} A+C & -B^T - D^T \\ B+D & A^T + C^T \end{bmatrix}$
$= \begin{bmatrix} A+C & -(B+D)^T \\ B+D & (A+C)^T \end{bmatrix}$ is of the required form.

b $k\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix} = \begin{bmatrix} kA & -kB^T \\ kB & kA^T \end{bmatrix} = \begin{bmatrix} kA & -(kB)^T \\ kB & (kA)^T \end{bmatrix}$ is of the required form.

c The general element of $H$ is $M = \begin{bmatrix} p & -q & -r & -s \\ q & p & s & -r \\ r & -s & p & q \\ s & r & -q & p \end{bmatrix}$, with four arbitrary constants, $r$,$s$,$p$, and $q$. Thus $\dim(H) = 4$; use the strategy outlined in Summary 4.1.6 to construct a basis.

d $\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix}\begin{bmatrix} C & -D^T \\ D & C^T \end{bmatrix} = \begin{bmatrix} AC - B^TD & -AD^T - B^TC^T \\ BC + A^TD & -BD^T + A^TC^T \end{bmatrix} = \begin{bmatrix} AC - B^TD & -(BC + A^TD)^T \\ BC + A^TD & (AC - B^TD)^T \end{bmatrix}$ is of the required form.

Note that $A$, $B$, $C$, $D$, and their transposes are rotation-dilation matrices, so that they all commute.

e $\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix}^T = \begin{bmatrix} A^T & B^T \\ -B & (A^T)^T \end{bmatrix}$ is of the required form.

f Note that the columns $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ or $M$ are orthogonal, and they all have length $\sqrt{p^2 + q^2 + r^2 + s^2}$. Now $M^TM$ is the $4 \times 4$ matrix whose $ij$th entry is $\vec{v}_i \cdot \vec{v}_j$, so that $M^TM = (p^2 + q^2 + r^2 + s^2)I_4$.

g If $M \neq 0$, then $k = p^2 + q^2 + r^2 + s^2 > 0$, and $\left(\frac{1}{k}M^T\right)M = I_4$, so that $M$ is invertible, with $M^{-1} = \frac{1}{p^2+q^2+r^2+s^2}M^T$.

By parts b and e, $M^{-1}$ is in $H$ as well.

h No! $A = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ do not commute ($AB = -BA$).

5.3.**65** Write $10A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$; it is required that $a, b, c$ and $d$ be integers. Now $A = \begin{bmatrix} \frac{a}{10} & \frac{b}{10} \\ \frac{c}{10} & \frac{d}{10} \end{bmatrix}$ must be an orthogonal matrix, implying that $(\frac{a}{10})^2 + (\frac{c}{10})^2 = 1$, or $a^2 + c^2 = 100$. Checking the squares of all integers from 1 to 9, we see that there are only two ways to write 100 as a sum of two positive perfect squares: $100 = 36 + 64 = 64 + 36$. Since $a$ and $c$ are required to be positive, we have either $a = 6$ and $c = 8$ or $a = 8$ and $c = 6$. In each case we have two options for the second column of $A$, namely, the two unit vectors perpendicular to the first column vector. Thus we end up with four solutions:

$A = \begin{bmatrix} .6 & -.8 \\ .8 & .6 \end{bmatrix}, \begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}, \begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}$ or $\begin{bmatrix} .8 & .6 \\ .6 & -.8 \end{bmatrix}$.

5.3.**66** One approach is to take one of the solutions from Exercise 65, say, the rotation matrix $B = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$, and then let $A = B^2 = \begin{bmatrix} 0.28 & -0.96 \\ 0.96 & 0.28 \end{bmatrix}$. Matrix $A$ is orthogonal by Theorems 5.3.4a.

5.3.**67** a We need to show that $A^TA\vec{c} = A^T\vec{x}$, or, equivalently, that $A^T(\vec{x} - A\vec{c}) = \vec{0}$. But $A^T(\vec{x} - A\vec{c}) = A^T(\vec{x} - c_1\vec{v}_1 - \cdots - c_m\vec{v}_m)$ is the vector whose $i^{th}$ component is

$(\vec{v}_i)^T(\vec{x} - c_1\vec{v}_1 - \cdots - c_m\vec{v}_m) = \vec{v}_i \cdot (\vec{x} - c_1\vec{v}_1 - \cdots - c_m\vec{v}_m)$, which we know to be zero.

b The system $A^TA\vec{c} = A^T\vec{x}$ has a unique solution $\vec{c}$ for a given $\vec{x}$, since $\vec{c}$ is the coordinate vector of $\text{proj}_V\vec{x}$ with respect to the basis $\vec{v}_1, \ldots, \vec{v}_m$. Thus the coefficient matrix $A^TA$ must be invertible, so that we can solve for $\vec{c}$ and write $\vec{c} = (A^TA)^{-1}A^T\vec{x}$. Then $\text{proj}_V\vec{x} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = A\vec{c} = A(A^TA)^{-1}A^T\vec{x}$.

5.3.**68** If $A = QR$, then $A(A^TA)^{-1}A^T = QR(R^TQ^TQR)^{-1}R^TQ^T = QR(R^TR)^{-1}R^TQ^T = QRR^{-1}(R^T)^{-1}R^TQ^T = QQ^T$, as in Theorem 5.3.10. The equation $Q^TQ = I_m$ holds since the columns of $Q$ are orthonormal.

5.3.**69** We will use the terminology introduced in Theorem 5.3.10. Since the two given vectors $\vec{v}_1$ and $\vec{v}_2$ are orthogonal (verify that $\vec{v}_1 \cdot \vec{v}_2 = 0$), with $\|\vec{v}_1\| = \|\vec{v}_2\| = \sqrt{3}$, we have the orthonormal basis $\vec{u}_1 = \frac{1}{\sqrt{3}}\vec{v}_1$, $\vec{u}_2 = \frac{1}{\sqrt{3}}\vec{v}_2$ , so

$$Q = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} \text{ and } P_W = QQ^T = \frac{1}{3}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & -1 & 0 \\ 1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 1 \end{bmatrix}.$$

5.3.**70** a. We will use the terminology introduced in Theorem 5.3.10. Consider the basis $\mathcal{B} = (\vec{v}_1, \vec{v}_2)$ of $V$ presented in Exercise 4.3.73c. Since the two vectors $\vec{v}_1$ and $\vec{v}_2$ are orthogonal (verify that $\vec{v}_1 \cdot \vec{v}_2 = 0$), with $\|\vec{v}_1\| = \|\vec{v}_2\| = \sqrt{6}$, we have the orthonormal basis $\vec{u}_1 = \frac{1}{\sqrt{6}}\vec{v}_1$, $\vec{u}_2 = \frac{1}{\sqrt{6}}\vec{v}_2$ , so

$$Q = \frac{1}{\sqrt{6}}\begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 1 & 1 \\ 2 & 0 \end{bmatrix} \text{ and } P_V = QQ^T = \frac{1}{6}\begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & -1 & 1 & 0 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 4 & -2 & 2 & 0 \\ -2 & 2 & 0 & 2 \\ 2 & 0 & 2 & 2 \\ 0 & 2 & 2 & 4 \end{bmatrix}.$$

b. If $\vec{x}$ is any vector in $V$, then $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 + x_2 - x_3 = 0$ and $\begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 + x_3 - x_4 = 0$, by definition of $V$. Since the given basis vectors of $W$ are orthogonal to $V$, the space $W$ is the orthogonal compliment of $V$, meaning that $W = V^\perp$ and $V = W^\perp$. By definition of a projection, we have $\vec{x} = \text{proj}_W\vec{x} + \text{proj}_V\vec{x}$ for all $\vec{x}$ in $\mathbb{R}^4$, so that $P_W + P_V = I_4$.

Also, $\text{proj}_W(\text{proj}_V\vec{x}) = \text{proj}_V(\text{proj}_W\vec{x}) = \vec{0}$ for all $\vec{x}$ in $\mathbb{R}^4$, meaning that $P_VP_W = P_WP_V = 0$, the zero matrix of size $4 \times 4$.

5.3.**71** If $A$ and $B$ are Hankel matrices of size $n \times n$, and $C = A + B$, then $c_{ij} = a_{ij} + b_{ij} = a_{i+1,j-1} + b_{i+1,j-1} = c_{i+1,j-1}$for all $i = 1, ..., n-1$ and for all $j = 2, ..., n$, showing that $C$ is a Hankel matrix as well. An analogous argument shows that the Hankel matrices are closed under scalar multiplication.

Now, what is the dimension of the space $H_n$ of the Hankel matrices of size $n \times n$? In the case $n = 4$, the dimension is 7, since there are 7 free variables in the matrix

$$A = \begin{bmatrix} a & b & c & d \\ b & c & d & e \\ c & d & e & f \\ d & e & f & g \end{bmatrix}.$$

For an arbitrary $n$, we can choose the entries in the first column and those in the last row freely; the other entries are then determined by those choices. Because $a_{n1}$ belongs both to the first column and to the last row, we have $\dim H_n = 2n - 1$.

5.3.**72** We will use the terminology introduced in Theorem 5.3.10. Let

$$Q = \vec{u} = \frac{1}{\|\vec{v}\|}\vec{v} = \frac{1}{\|\vec{v}\|}\begin{bmatrix} 1 \\ a \\ a^2 \\ ... \\ a^{n-1} \end{bmatrix} \text{ and } P = QQ^T = \frac{1}{\|\vec{v}\|^2}\begin{bmatrix} 1 & a & a^2 & ... & a^{n-1} \\ a & a^2 & a^3 & ... & a^n \\ a^2 & a^3 & a^4 & ... & a^{n+1} \\ ... & ... & ... & ... & ... \\ a^{n-1} & a^n & a^{n+1} & ... & a^{2n-2} \end{bmatrix},$$

where $\|\vec{v}\|^2 = 1 + a^2 + a^4 + ... + a^{2n-2}$. Note that the $ij^{th}$ entry of $P$ is $p_{ij} = \frac{1}{\|\vec{v}\|^2}a^{i+j-2}$ so that $P$ is a Hankel matrix since $p_{ij} = \frac{1}{\|\vec{v}\|^2}a^{i+j-2} = \frac{1}{\|\vec{v}\|^2}a^{(i+1)+(j-1)-2} = p_{i+1,j-1}$.

In the case of $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ we have $\|\vec{v}\|^2 = 1 + 4 + 16 = 21$, so $P = \frac{1}{21}\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 4 & 8 & 16 \end{bmatrix}$. This is indeed a Hankel matrix.

5.3.**73** a. Note that $ab = \frac{(1+\sqrt{5})(1-\sqrt{5})}{4} = -1$. Now

$$\vec{a} \cdot \vec{b} = 1 + ab + (ab)^2 + (ab)^3 + ... + (ab)^{n-2} + (ab)^{n-1} = 1 - 1 + 1 - 1 + ... + 1 - 1 = 0,$$

so that $\vec{a}$ is orthogonal to $\vec{b}$, as claimed.

b. If we let $\vec{u}_1 = \frac{1}{\|\vec{a}\|}\vec{a}$ and $\vec{u}_2 = \frac{1}{\|\vec{b}\|}\vec{b}$ , then $P = \vec{u}_1\vec{u}_1^T + \vec{u}_2\vec{u}_2^T$ ; see the paragraph preceding Theorem 5.3.10. In Exercise 72 we saw that both $\vec{u}_1\vec{u}_1^T$ and $\vec{u}_2\vec{u}_2^T$ are Hankel matrices, and in Exercise 71 we saw that the set of Hankel matrices is closed under addition. Thus $P$ is a Hankel matrix as well.

5.3.**74** a.

| $m$ | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_m$ | 5 | -3 | 2 | -1 | 1 | 0 | 1 | 1 | 2 | 3 | 5 |

b. We conjecture that $f_m = f_{-m}$ for odd $m$ and $f_m = -f_{-m}$ for even $m$.

We will prove this conjecture by induction, with the base case, $f_1 = f_{-1}$ , being established in the table above. Assuming the result for positive integers $< m$, we will prove it for $m$.

For odd $m$, we have $f_m = f_{m-1} + f_{m-2} = -f_{1-m} + f_{2-m} = -f_{1-m} + (f_{1-m} + f_{-m)} = f_{-m}$

For even $m$, we have $f_m = f_{m-1} + f_{m-2} = f_{1-m} - f_{2-m} = f_{1-m} - (f_{1-m} + f_{-m)} = -f_{-m}$.

c. $\vec{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}$, $\quad \vec{w} = \begin{bmatrix} 5 \\ -3 \\ 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}$, with $\|\vec{v}\|^2 = \|\vec{w}\|^2 = 40 = 5 \cdot 8 = f_5 f_6$ and $\vec{v} \cdot \vec{w} = 0$.

d. If $k$ is an odd integer between 0 and $n$, then the $k$th summand of $\vec{v} \cdot \vec{w}$ is $f_{k-1}f_{-n+k} = f_{k-1}f_{n-k}$, while the $(n + 1 - k)$th summand is $f_{n-k}f_{1-k} = -f_{n-k}f_{k-1}$; these two summands add up to 0. As we sum over all odd integers from 1 to $n - 1$, we see that $\vec{v} \cdot \vec{w} = 0$.

e. We have $P = \frac{1}{\|\vec{v}\|^2}\vec{v}\vec{v}^T + \frac{1}{\|\vec{w}\|^2}\vec{w}\vec{w}^T = \frac{1}{f_{n-1}f_n}(\vec{v}\vec{v}^T + \vec{w}\vec{w}^T)$; see the paragraph preceding Theorem 5.3.10. Considering the first and last components of $\vec{v}$ and $\vec{w}$, we can write the first and the last columns of $P$:

$$P = \frac{1}{f_{n-1}f_n}\left(\begin{bmatrix} \vec{0} & ... & f_{n-1}\vec{v} \end{bmatrix} + \begin{bmatrix} f_{n-1}\vec{w} & ... & \vec{0} \end{bmatrix}\right) = \frac{1}{f_{n-1}f_n}\begin{bmatrix} f_{n-1}\vec{w} & .... & f_{n-1}\vec{v} \end{bmatrix} = \frac{1}{f_n}\begin{bmatrix} \vec{w} & .... & \vec{v} \end{bmatrix}.$$

Since $P$ is a Hankel matrix, it is determined by its first and last columns, so that $P$ has the form given in the exercise.

f. Using the result $\|\vec{v}\|^2 = \|\vec{w}\|^2 = 40 = 5 \cdot 8 = f_5 f_6$ from part c, we find

$$P = \frac{1}{40}\begin{bmatrix} 0 & 5 \\ 1 & -3 \\ 1 & 2 \\ 2 & -1 \\ 3 & 1 \\ 5 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 1 & 2 & 3 & 5 \\ 5 & -3 & 2 & -1 & 0 & 0 \end{bmatrix} = \frac{1}{40}\begin{bmatrix} 25 & -15 & 10 & -5 & 5 & 0 \\ -15 & 10 & -5 & 5 & 0 & 5 \\ 10 & -5 & 5 & 0 & 5 & 5 \\ -5 & 5 & 0 & 5 & 5 & 10 \\ 5 & 0 & 5 & 5 & 10 & 15 \\ 0 & 5 & 5 & 10 & 15 & 25 \end{bmatrix}$$

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

## Section 5.4

5.4.**1** A basis of $\ker(A^T)$ is $\begin{bmatrix} -3 \\ 2 \end{bmatrix}$. (See Figure 5.15.)

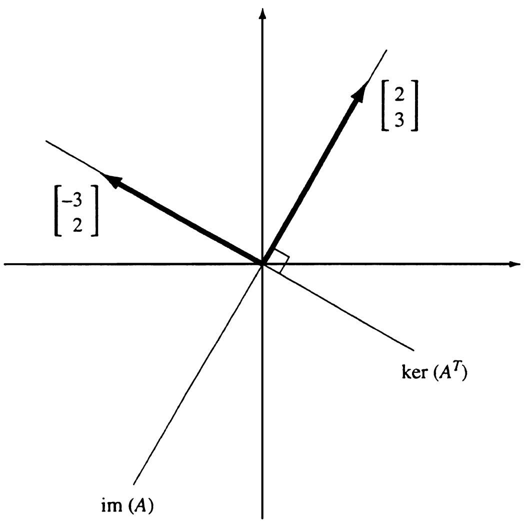



Figure 5.15: for Problem 5.4.1.

5.4.**2** A basis of $\ker(A^T)$ is $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$. im$(A)$ is the plane perpendicular to this line.

5.4.**3** We will first show that the vectors $\vec{v}_1, \ldots, \vec{v}_p, \vec{w}_1, \ldots, \vec{w}_q$ span $\mathbb{R}^n$. Any vector $\vec{v}$ in $\mathbb{R}^n$ can be written as $\vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp}$, where $\vec{v}_{\parallel}$ is in $V$ and $\vec{v}_{\perp}$ is in $V^{\perp}$ (by definition of orthogonal projection, Theorem 5.1.4).

Now $\vec{v}_{\parallel}$ is a linear combination of $\vec{v}_1, \ldots, \vec{v}_p$, and $\vec{v}_{\perp}$ is a linear combination of $\vec{w}_1, \ldots, \vec{w}_q$, showing that the vectors $\vec{v}_1, \ldots, \vec{v}_p, \vec{w}_1, \ldots, \vec{w}_q$ span $\mathbb{R}^n$.

Note that $p + q = n$, by Theorem 5.1.8c; therefore, the vectors $\vec{v}_1, \ldots, \vec{v}_p, \vec{w}_1, \ldots, \vec{w}_q$ form a basis of $\mathbb{R}^n$, by Theorem 3.3.4d.

5.4.**4** By Theorem 5.4.1, the equation $(\text{im } B)^{\perp} = \ker(B^T)$ holds for any matrix $B$. Now let $B = A^T$. Then $(\text{im}(A^T))^{\perp} = \ker(A)$. Taking transposes of both sides and using Theorem 5.1.8d we obtain $\text{im}(A^T) = (\ker A)^{\perp}$, as claimed.

5.4.**5** $V = \ker(A)$, where $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 5 & 4 \end{bmatrix}$.

Then $V^{\perp} = (\ker A)^{\perp} = \text{im}(A^T)$, by Exercise 4.

The two columns of $A^T$ form a basis of $V^{\perp}$:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 5 \\ 4 \end{bmatrix}$$

5.4.**6** Yes! For any matrix $A$,

$$\text{im}(A) \underset{\substack{\uparrow \\ \text{Theorem 5.4.1}}}{=} (\ker(A^T))^{\perp} \underset{\substack{\uparrow \\ \text{Theorem 5.4.2a}}}{=} (\ker(AA^T))^{\perp} = (\ker(AA^T)^T)^{\perp} \underset{\substack{\uparrow \\ \text{Theorems 5.4.1 and 5.1.8d.}}}{=} \text{im}(AA^T).$$

5.4.**7** $\text{im}(A)$ and $\ker(A)$ are orthogonal complements by Theorem 5.4.1:

$(\text{im}A)^{\perp} = \ker(A^T) = \ker(A)$

5.4.**8** a By Theorem 5.4.6, $L^+(\vec{y}) = (A^TA)^{-1}A^T\vec{y}$.

The transformation $L^+$ is linear since it is "given by a matrix," by Definition 2.1.1.

b If $L$ (and therefore $A$) is invertible, then $L^+(\vec{y}) = A^{-1}(A^T)^{-1}A^T\vec{y} = A^{-1}\vec{y} = L^{-1}\vec{y}$, so that $L^+ = L^{-1}$.

c $L^+(L(\vec{x})) = \big($ the unique least-squares solution $\vec{u}$ of $L(\vec{u}) = L(\vec{x})\big) = \vec{x}$.

d $L(L^+(\vec{y})) = A(A^TA)^{-1}A^T\vec{y} = \text{proj}_V\vec{y}$, where $V = \text{im}(A)$, by Theorem 5.4.7.

e Here $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $L^+(\vec{y}) = (A^TA)^{-1}A^T\vec{y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\vec{y}$.

5.4.**9** $\vec{x}_0$ is the shortest of all the vectors in $S$. (See Figure 5.16.)

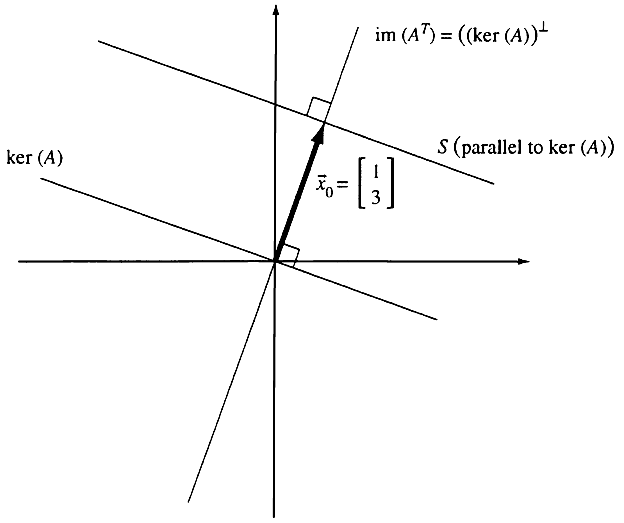



Figure 5.16: for Problem 5.4.9.

5.4.**10** a If $\vec{x}$ is an arbitrary solution of the system $A\vec{x} = \vec{b}$, let $\vec{x}_h = \text{proj}_V \vec{x}$, where $V = \ker(A)$, and $\vec{x}_0 = \vec{x} - \text{proj}_V \vec{x}$. Note that $\vec{b} = A\vec{x} = A(\vec{x}_h + \vec{x}_0) = A\vec{x}_h + A\vec{x}_0 = A\vec{x}_0$, since $\vec{x}_h$ is in $\ker(A)$.

b If $\vec{x}_0$ and $\vec{x}_1$ are two solutions of the system $A\vec{x} = \vec{b}$, both from $(\ker A)^\perp$, then $\vec{x}_1 - \vec{x}_0$ is in the subspace $(\ker A)^\perp$ as well. Also, $A(\vec{x}_1 - \vec{x}_0) = A\vec{x}_1 - A\vec{x}_0 = \vec{b} - \vec{b} = \vec{0}$, so that $\vec{x}_1 - \vec{x}_0$ is in $\ker(A)$. By Theorem 5.1.8b, it follows that $\vec{x}_1 - \vec{x}_0 = \vec{0}$, or $\vec{x}_1 = \vec{x}_0$, as claimed.

c Write $\vec{x}_1 = \vec{x}_h + \vec{x}_0$ as in part a; note that $\vec{x}_h$ is orthogonal to $\vec{x}_0$. The claim now follows from the Pythagorean Theorem (Theorem 5.1.9).

5.4.**11** a Note that $L^+(\vec{y}) = A^T(AA^T)^{-1}\vec{y}$; indeed, this vector is in $\text{im}(A^T) = (\ker A)^\perp$, and it is a solution of $L(\vec{x}) = A\vec{x} = \vec{y}$.

b $L(L^+(\vec{y})) = \vec{y}$, by definition of $L^+$.

c $L^+(L(\vec{x})) = A^T(AA^T)^{-1}A\vec{x} = \text{proj}_V \vec{x}$, where $V = \text{im}(A^T) = (\ker A)^\perp$, by Theorem 5.4.7.

d $\text{im}(L^+) = \text{im}(A^T)$, by part c, and $\ker(L^+) = \{\vec{0}\}$ (if $\vec{y}$ is in $\ker(L^+)$, then $\vec{y} = L(L^+(\vec{y})) = L(\vec{0}) = \vec{0}$, by part b).

e Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$; then the matrix of $L^+$ is $A^T(AA^T)^{-1} = A^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

5.4.**12** By Theorem 5.4.5, the least-squares solutions of the linear system $A\vec{x} = \vec{b}$ are the exact solutions of the (consistent) system $A^T A\vec{x} = A^T\vec{b}$. The minimal solution of this normal equation (in the sense of Exercise 10) is called the minimal least-squares solution of the system $A\vec{x} = \vec{b}$.

Equivalently, the minimal least-squares solution of $A\vec{x} = \vec{b}$ can be defined as the minimal solution of the consistent system $A\vec{x} = \text{proj}_V \vec{b}$, where $V = \text{im}(A)$.

5.4.**13** a Suppose that $L^+(\vec{y}_1) = \vec{x}_1$ and $L^+(\vec{y}_2) = \vec{x}_2$; this means that $\vec{x}_1$ and $\vec{x}_2$ are both in $(\ker A)^\perp = \text{im}(A^T)$, $A^TA\vec{x}_1 = A^T\vec{y}_1$, and $A^TA\vec{x}_2 = A^T\vec{y}_2$. Then $\vec{x}_1 + \vec{x}_2$ is in $\text{im}(A^T)$ as well, and $A^TA(\vec{x}_1 + \vec{x}_2) = A^T(\vec{y}_1 + \vec{y}_2)$, so that $L^+(\vec{y}_1 + \vec{y}_2) = \vec{x}_1 + \vec{x}_2$.

The verification of the property $L^+(k\vec{y}) = kL^+(\vec{y})$ is analogous.

b $L^+(L(\vec{x}))$ is the orthogonal projection of $\vec{x}$ onto $(\ker A)^\perp = \text{im}(A^T)$.

c $L(L^+(\vec{y}))$ is the orthogonal projection of $\vec{y}$ onto $\text{im}(A) = (\ker A^T))^\perp$.

d $\text{im}(L^+) = \text{im}(A^T)$ and $\ker(L^+) = \ker(A^T)$, by parts b and c.

e $L^+\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{y_1}{2} \\ 0 \\ 0 \end{bmatrix}$, so that the matrix of $L^+$ is $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

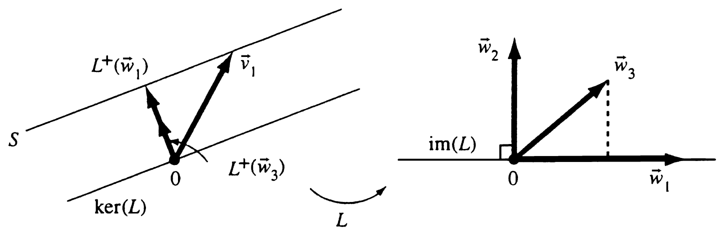



Figure 5.17: for Problem 5.4.14.

5.4.**14** $L^+(\vec{w}_1)$ is the minimal solution of the system $L(\vec{x}) = \vec{w}_1$. The line $S$ in Figure 5.17 shows all solutions of the system

$L(\vec{x}) = \vec{w}_1$ (compare with Exercise 9). The minimal solution, $L^+(\vec{w}_1)$, is perpendicular to $\ker(L)$.

$L^+(\vec{w}_2) = L^+(\text{proj}_{\text{im}(L)}\vec{w}_2) = L^+(\vec{0}) = \vec{0}$

$L^+(\vec{w}_3) = L^+(\text{proj}_{\text{im}(L)}\vec{w}_3) \approx L^+(0.55\vec{w}_1) = 0.55L^+(\vec{w}_1)$

5.4.**15** Note that $(A^TA)^{-1}A^TA = I_n$; let $B = (A^TA)^{-1}A^T$.

5.4.**16** If $A$ is an $m \times n$ matrix, then

$$\dim(\text{im}A)^\perp \underset{\substack{\uparrow \\ \text{Theorem 5.1.8c}}}{=} m - \dim(\text{im}A) \underset{\substack{\uparrow \\ \text{Theorem 3.3.6}}}{=} m - \text{rank}(A)$$

and $\dim(\ker(A^T)) \underset{\substack{\uparrow \\ \text{Theorem 3.3.7}}}{=} m - \text{rank}(A^T)$.

It follows that $\text{rank}(A) = \text{rank}(A^T)$, as claimed.

5.4.**17** Yes! By Theorem 5.4.2, $\ker(A) = \ker(A^TA)$. Taking dimensions of both sides and using Theorem 3.3.7, we find that $n - \text{rank}(A) = n - \text{rank}(A^TA)$; the claim follows.

5.4.**18** Yes! By Exercise 17, $\text{rank}(A) = \text{rank}(A^TA)$. Substituting $A^T$ for $A$ in Exercise 17 and using Theorem 5.3.9c, we find that $\text{rank}(A) = \text{rank}(A^T) = \text{rank}(AA^T)$. The claim follows.

5.4.**19** $\vec{x}^* = (A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, by Theorem 5.4.6.

5.4.**20** Using Theorem 5.4.6, we find $\vec{x}^* = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $\vec{b} - A\vec{x}^* = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$.

Note that $\vec{b} - A\vec{x}^*$ is perpendicular to the two columns of $A$.

5.4.**21** Using Theorem 5.4.6, we find $\vec{x}^* = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\vec{b} - A\vec{x}^* = \begin{bmatrix} -12 \\ 36 \\ -18 \end{bmatrix}$, so that $\|\vec{b} - A\vec{x}^*\| = 42$.

5.4.**22** Using Theorem 5.4.6, we find $\vec{x}^* = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ and $\vec{b} - A\vec{x}^* = \vec{0}$. This system is in fact consistent and $\vec{x}^*$ is the exact solution; the error $\|\vec{b} - A\vec{x}^*\|$ is 0.

5.4.**23** Using Theorem 5.4.6, we find $\vec{x}^* = \vec{0}$; here $\vec{b}$ is perpendicular to $\text{im}(A)$.

5.4.**24** Using Theorem 5.4.6, we find $\vec{x}^* = [2]$.

5.4.**25** In this case, the normal equation $A^TA\vec{x} = A^T\vec{b}$ is $\begin{bmatrix} 5 & 15 \\ 15 & 45 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \end{bmatrix}$, which simplifies to $x_1 + 3x_2 = 1$, or $x_1 = 1 - 3x_2$. The solutions are of the form $\vec{x}^* = \begin{bmatrix} 1-3t \\ t \end{bmatrix}$, where $t$ is an arbitrary constant.

5.4.**26** Here, the normal equation $A^TA\vec{x} = A^T\vec{b}$ is $\begin{bmatrix} 66 & 78 & 90 \\ 78 & 93 & 108 \\ 90 & 108 & 126 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, with solutions $\vec{x}^* = \begin{bmatrix} t - \frac{7}{6} \\ 1-2t \\ t \end{bmatrix}$, where $t$ is an arbitrary constant.

5.4.**27** The least-squares solutions of the system $SA\vec{x} = S\vec{b}$ are the exact solutions of the normal equation $(SA)^TSA\vec{x} = (SA)^TS\vec{b}$.

Note that $S^TS = I_n$, since $S$ is orthogonal; therefore, the normal equation simplifies as follows: $(SA)^TSA\vec{x} = A^TS^TSA\vec{x} = A^TA\vec{x}$ and $(SA)^TS\vec{b} = A^TS^TS\vec{b} = A^T\vec{b}$, so that the normal equation is $A^TA\vec{x} = A^T\vec{b}$, the same as the normal equation of the system $A\vec{x} = \vec{b}$. Therefore, the systems $A\vec{x} = \vec{b}$ and $SA\vec{x} = S\vec{b}$ have the same least-squares solution, $\vec{x}^* = \begin{bmatrix} 7 \\ 11 \end{bmatrix}$.

5.4.**28** The least-squares solutions of the system $A\vec{x} = \vec{u}_n$ are the exact solutions of $A\vec{x} = \text{proj}_{\text{im}(A)}\vec{u}_n$. Note that $\vec{u}_n$ is orthogonal to $\text{im}(A)$, so that $\text{proj}_{\text{im}(A)}\vec{u}_n = \vec{0}$, and the unique least-squares solution is $\vec{x}^* = \vec{0}$.

5.4.**29** By Theorem 5.4.6, $\vec{x}^* = (A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} 1+\varepsilon & 1 \\ 1 & 1+\varepsilon \end{bmatrix}^{-1} \begin{bmatrix} 1+\varepsilon \\ 1+\varepsilon \end{bmatrix} = \frac{1}{2+\varepsilon} \begin{bmatrix} 1+\varepsilon \\ 1+\varepsilon \end{bmatrix} \approx \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$, where $\varepsilon = 10^{-20}$.

If we use a hand-held calculator, due to roundoff errors we find the normal equation $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, with infinitely many solutions.

5.4.**30** We attempt to solve the system

$$\begin{matrix} c_0 + 0c_1 = 0 \\ c_0 + 0c_1 = 1 \\ c_0 + 1c_1 = 1 \end{matrix}, \text{ or } \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

This system cannot be solved exactly; the least-squares solution is $\begin{bmatrix} c_0^* \\ c_1^* \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$. The line that fits the data points best is $f^*(t) = \frac{1}{2} + \frac{1}{2}t$.

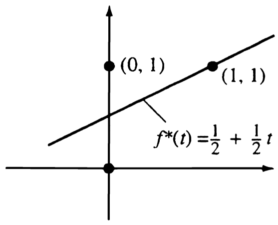



Figure 5.18: for Problem 5.4.30.

The line goes through the point (1, 1) and "splits the difference" between (0, 0) and (0, 1). See Figure 5.18.

5.4.**31** We want $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ such that

$$\begin{matrix} 3 = c_0 + 0c_1 \\ 3 = c_0 + 1c_1 \\ 6 = c_0 + 1c_1 \end{matrix} \text{ or } \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 6 \end{bmatrix}.$$

Since $\ker \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \{\vec{0}\}$, $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}^* = \left( \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 6 \end{bmatrix}$

$= \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 12 \\ 9 \end{bmatrix} = \begin{bmatrix} 3 \\ \frac{3}{2} \end{bmatrix}$ so $f^*(t) = 3 + \frac{3}{2}t$. (See Figure 5.19.)

5.4.**32** We want $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$ of $f(t) = c_0 + c_1t + c_2t^2$ such that

$$\begin{matrix} 27 = c_0 + 0c_1 + 0c_2 \\ 0 = c_0 + 1c_1 + 1c_2 \\ 0 = c_0 + 2c_1 + 4c_2 \\ 0 = c_0 + 3c_1 + 9c_2 \end{matrix} \text{ or } \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 27 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

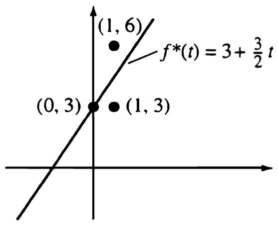



Figure 5.19: for Problem 5.4.31.

If we call the coefficient matrix $A$, we notice that $\ker(A) = \{\vec{0}\}$ so

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}^* = (A^TA)^{-1}A^T \begin{bmatrix} 27 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 25.65 \\ -28.35 \\ 6.75 \end{bmatrix} \text{ so } f^*(t) = 25.65 - 28.35t + 6.75t^2.$$

5.4.**33** We want $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$ such that

$$\begin{matrix} 0 = c_0 + \sin(0)c_1 + \cos(0)c_2 \\ 1 = c_0 + \sin(1)c_1 + \cos(1)c_2 \\ 2 = c_0 + \sin(2)c_1 + \cos(2)c_2 \\ 3 = c_0 + \sin(3)c_1 + \cos(3)c_2 \end{matrix} \text{ or } \begin{bmatrix} 1 & 0 & 1 \\ 1 & \sin(1) & \cos(1) \\ 1 & \sin(2) & \cos(2) \\ 1 & \sin(3) & \cos(3) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}.$$

Since the coefficient matrix has kernel $\{\vec{0}\}$, we compute $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}^*$ using Theorem 5.4.6, obtaining

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}^* \approx \begin{bmatrix} 1.5 \\ 0.1 \\ -1.41 \end{bmatrix} \text{ so } f^*(t) \approx 1.5 + 0.1\sin t - 1.41\cos t.$$

5.4.**34** We want $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$ such that

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & \sin(0.5) & \cos(0.5) & \sin(1) & \cos(1) \\ 1 & \sin(1) & \cos(1) & \sin(2) & \cos(2) \\ 1 & \sin(1.5) & \cos(1.5) & \sin(3) & \cos(3) \\ 1 & \sin(2) & \cos(2) & \sin(4) & \cos(4) \\ 1 & \sin(2.5) & \cos(2.5) & \sin(5) & \cos(5) \\ 1 & \sin(3) & \cos(3) & \sin(6) & \cos(6) \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.5 \\ 1 \\ 1.5 \\ 2 \\ 2.5 \\ 3 \end{bmatrix}$$

Since the columns of the coefficient matrix are linearly independent, its kernel is $\{\vec{0}\}$. We can use Theorem 5.4.6

to compute $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} \approx \begin{bmatrix} 1.5 \\ 0.109 \\ -1.537 \\ 0.303 \\ 0.043 \end{bmatrix}$ so $f^*(t) \approx 1.5 + 0.109\sin(t) - 1.537\cos(t) + 0.303\sin(2t) + 0.043\cos(2t)$.

5.4.**35** a The $ij$th entry of $A_n^T A_n$ is the dot product of the $i$th row of $A_n^T$ and the $j$th column of $A_n$, i.e.

$$A_n^T A_n = \begin{bmatrix} n & \sum_{i=1}^n \sin a_i & \sum_{i=1}^n \cos a_i \\ \sum_{i=1}^n \sin a_i & \sum_{i=1}^n \sin^2 a_i & \sum_{i=1}^n \sin a_i \cos a_i \\ \sum_{i=1}^n \cos a_i & \sum_{i=1}^n \sin a_i \cos a_i & \sum_{i=1}^n \cos^2 a_i \end{bmatrix} \text{ and } A_n^T \vec{b} = \begin{bmatrix} \sum_{i=1}^n g(a_i) \\ \sum_{i=1}^n g(a_i)\sin a_i \\ \sum_{i=1}^n g(a_i)\cos a_i \end{bmatrix}.$$

b $$\lim_{n\to\infty} \tfrac{2\pi}{n} A_n^T A_n = \begin{bmatrix} 2\pi & \int_0^{2\pi} \sin t\,dt & \int_0^{2\pi} \cos t\,dt \\ \int_0^{2\pi} \sin t\,dt & \int_0^{2\pi} \sin^2 t\,dt & \int_0^{2\pi} \sin t\cos t\,dt \\ \int_0^{2\pi} \cos t\,dt & \int_0^{2\pi} \sin t\cos t\,dt & \int_0^{2\pi} \cos^2 t\,dt \end{bmatrix} = \begin{bmatrix} 2\pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{bmatrix}$$

and $$\lim_{n\to\infty} \tfrac{2\pi}{n} A_n^T \vec{b} = \begin{bmatrix} \int_0^{2\pi} g(t)\,dt \\ \int_0^{2\pi} g(t)\sin t\,dt \\ \int_0^{2\pi} g(t)\cos t\,dt \end{bmatrix}$$

(Here $\frac{2\pi}{n} = \Delta t$ so $\lim_{n\to\infty} \frac{2\pi}{n} \sum_{i=1}^n \cos(t_i) = \lim_{n\to\infty} \sum_{i=1}^n \cos(t_i)\Delta t = \int_0^{2\pi} \cos t\,dt$ for instance. All other limits are obtained similarly. )

c $$\begin{bmatrix} c \\ p \\ q \end{bmatrix} = \lim_{n\to\infty} \begin{bmatrix} c_n \\ p_n \\ q_n \end{bmatrix} = \begin{bmatrix} 2\pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{bmatrix}^{-1} \begin{bmatrix} \int_0^{2\pi} g(t)\,dt \\ \int_0^{2\pi} g(t)\sin t\,dt \\ \int_0^{2\pi} g(t)\cos t\,dt \end{bmatrix} = \begin{bmatrix} \frac{1}{2\pi}\int_0^{2\pi} g(t)\,dt \\ \frac{1}{\pi}\int_0^{2\pi} g(t)\sin t\,dt \\ \frac{1}{\pi}\int_0^{2\pi} g(t)\cos t\,dt \end{bmatrix}$$ and $f(t) = c + p\sin t + q\cos t$,

where $c$, $p$, $q$ are given above.

5.4.**36** We want $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ such that

$$\begin{aligned}
a + b\sin\left(\frac{2\pi}{366}32\right) + c\cos\left(\frac{2\pi}{366}32\right) &= 10 \\
a + b\sin\left(\frac{2\pi}{366}77\right) + c\cos\left(\frac{2\pi}{366}77\right) &= 12 \\
a + b\sin\left(\frac{2\pi}{366}121\right) + c\cos\left(\frac{2\pi}{366}121\right) &= 14 \\
a + b\sin\left(\frac{2\pi}{366}152\right) + c\cos\left(\frac{2\pi}{366}152\right) &= 15
\end{aligned}$$

Using $A = \begin{bmatrix} 1 & \sin\left(\frac{2\pi}{366}32\right) & \cos\left(\frac{2\pi}{366}32\right) \\ 1 & \sin\left(\frac{2\pi}{366}77\right) & \cos\left(\frac{2\pi}{366}77\right) \\ 1 & \sin\left(\frac{2\pi}{366}121\right) & \cos\left(\frac{2\pi}{366}121\right) \\ 1 & \sin\left(\frac{2\pi}{366}152\right) & \cos\left(\frac{2\pi}{366}152\right) \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 10 \\ 12 \\ 14 \\ 15 \end{bmatrix}$, we compute $\begin{bmatrix} a \\ b \\ c \end{bmatrix}^*$

$= (A^TA)^{-1}A^T\vec{b} \approx \begin{bmatrix} 12.26 \\ 0.431 \\ -2.899 \end{bmatrix}$ and $f^*(t) \approx 12.26 + 0.431\sin\left(\frac{2\pi}{366}t\right) - 2.899\cos\left(\frac{2\pi}{366}t\right)$.

5.4.**37** a We want $c_0, c_1$ such that

$$\begin{aligned}
c_0 + c_1(35) &= \log 35 \\
c_0 + c_1(46) &= \log 46 \\
c_0 + c_1(59) &= \log 77 \\
c_0 + c_1(69) &= \log 133
\end{aligned}
\quad \text{or} \quad
\underset{\substack{\uparrow \\ A}}{\begin{bmatrix} 1 & 35 \\ 1 & 46 \\ 1 & 59 \\ 1 & 69 \end{bmatrix}} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \underset{\substack{\uparrow \\ \vec{b}}}{\begin{bmatrix} \log 35 \\ \log 46 \\ \log 77 \\ \log 133 \end{bmatrix}}$$

so $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}^* = (A^TA)^{-1}A^T\vec{b} \approx \begin{bmatrix} 0.915 \\ 0.017 \end{bmatrix}$ so $\log(d) \approx 0.915 + 0.017t$.

b $d \approx 10^{0.915} \cdot 10^{0.017t} \approx 8.22 \cdot 10^{0.017t}$

c If $t = 88$ then $d \approx 258$. Since the Airbus has only 93 displays, new technologies must have rendered the old trends obsolete.

5.4.**38** We want $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$ such that

$$\begin{aligned}
110 &= c_0 + 2c_1 + c_2 \\
180 &= c_0 + 12c_1 + 0c_2 \\
120 &= c_0 + 5c_1 + c_2 \\
160 &= c_0 + 11c_1 + c_2 \\
160 &= c_0 + 6c_1 + 0c_2
\end{aligned}
\quad \text{or} \quad
\begin{bmatrix} 1 & 2 & 1 \\ 1 & 12 & 0 \\ 1 & 5 & 1 \\ 1 & 11 & 1 \\ 1 & 6 & 0 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 110 \\ 180 \\ 120 \\ 160 \\ 160 \end{bmatrix}.$$

The least-squares solution is $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}^* = \begin{bmatrix} 125 \\ 5 \\ -25 \end{bmatrix}$, so that $w^* = 125 + 5h - 25g$.

For a general population, we expect $c_0$ and $c_1$ to be positive, since $c_0$ gives the weight of a $5'$ male, and increased height should contribute positively to the weight. We expect $c_2$ to be negative, since females tend to be lighter than males of equal height.

5.4.**39** a We want $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ such that

$\log(250) = c_0 + c_1 \log(600,000)$

$\log(60) = c_0 + c_1 \log(200,000)$

$\log(25) = c_0 + c_1 \log(60,000)$

$\log(12) = c_0 + c_1 \log(10,000)$

$\log(5) = c_0 + c_1 \log(2500)$

The least-squares solution to the above system is $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}^* \approx \begin{bmatrix} -1.616 \\ 0.664 \end{bmatrix}$ so $\log z \approx -1.616 + 0.664 \log g$.

b Exponentiating both sides of the answer to a, we get $z \approx 10^{-1.616} \cdot g^{0.664} \approx 0.0242 \cdot g^{0.664}$.

c This model is close since $\sqrt{g} = g^{0.5}$.

5.4.**40** First we look for $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ such that $\log D = c_0 + c_1 \log a$.

Proceeding as in Exercise 39, we get $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}^* \approx \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}$, i.e. $\log D \approx 1.5 \log a$, hence $D \approx 10^{1.5 \log a} = a^{1.5}$.

Note that the formula $D = a^{1.5}$ is Kepler's third law of planetary motion.

5.4.**41** a We want $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$ such that $\log D = c_0 + c_1 t$ ($t$ in years since 1975), i.e.

$$\begin{aligned} \log 533 &= c_0 + c_1(0) \\ \log 1,823 &= c_0 + c_1(10) \\ \log 4,974 &= c_0 + c_1(20) \\ \log 7,933 &= c_0 + c_1(30) \end{aligned}$$

The least-squares solution to the system is $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}^* \approx \begin{bmatrix} 2.8 \\ 0.040 \end{bmatrix}$, i.e. $\log D \approx 2.8 + 0.04$ or $D \approx 10^{2.8} \cdot 10^{0.040t}$.

b In the year 2015, we have $t = 40$ and $D \approx 24,200$. The formula predicts a debt of about US$24 trillion.

5.4.**42** Clearly, $L$ is a linear transformation. We will use Theorem 4.2.4a and show that $\ker(L) = \{\vec{0}\}$ and $\text{im}(L) = \text{im}(A)$. Now $\ker(L) = \ker(A) \cap \text{im}(A^T) = \{\vec{0}\}$, by Theorems 5.4.1 and 5.1.8b. Also, $\text{im}(L) = \{A\vec{v} : \vec{v} \text{ in } \text{im}(A^T)\} = \text{im}(AA^T) = \text{im}(A)$, by Exercise 6.

# Section 5.5

5.5.**1** Since $f$ is nonzero, there is a $c$ on $[a,\ b]$ such that $f(c)=d\neq 0$. By continuity, there is an $\varepsilon>0$ such that $|f(x)|>\frac{|d|}{2}$ for all $x$ on the open interval $(c-\varepsilon, c+\varepsilon)$ where $f$ is defined (see any good Calculus text).

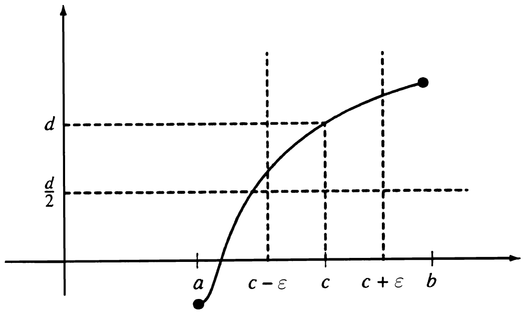



Figure 5.20: for Problem 5.5.1.

We assume that $c-\varepsilon\geq a$ and $c+\varepsilon\leq b$ and leave the other cases to the reader. Then $\langle f,f\rangle=\displaystyle\int_a^b (f(t))^2\,dt \geq \int_{c-\varepsilon}^{c+\varepsilon}\left(\frac{d}{2}\right)^2 dt=\frac{d^2\varepsilon}{2}>0$, as claimed. (See Figure 5.20.)

5.5.**2** We perform the following operations:

$$\langle f,g+h\rangle \underset{\substack{\uparrow\\ \text{Definition}\\ 5.5.1\text{a}}}{=} \langle g+h,f\rangle \underset{\substack{\uparrow\\ \text{Definition}\\ 5.5.1\text{b}}}{=} \langle g,f\rangle+\langle h,f\rangle \underset{\substack{\uparrow\\ \text{Definition}\\ 5.5.1\text{a}}}{=} \langle f,g\rangle+\langle f,h\rangle$$

5.5.**3** a Note that $\langle\vec{x},\vec{y}\rangle=(S\vec{x})^TS\vec{y}=S\vec{x}\cdot S\vec{y}$. We will check the four parts of Definition 5.5.1

$\alpha$. $\langle\vec{x},\vec{y}\rangle=S\vec{x}\cdot S\vec{y}=S\vec{y}\cdot S\vec{x}=\langle\vec{y},\vec{x}\rangle$

$\beta$. $\langle\vec{x}+\vec{y},\vec{z}\rangle=S(\vec{x}+\vec{y})\cdot S\vec{z}=(S\vec{x}+S\vec{y})\cdot S\vec{z}=(S\vec{x}\cdot S\vec{z})+(S\vec{y}\cdot S\vec{z})=\langle\vec{x},\vec{z}\rangle+\langle\vec{y},\vec{z}\rangle$

$\gamma$. $\langle c\vec{x},\vec{y}\rangle=S(c\vec{x})\cdot S\vec{y}=c(S\vec{x})\cdot S\vec{y}=c\langle\vec{x},\vec{y}\rangle$

$\delta$. If $\vec{x}\neq 0$, then $\langle\vec{x},\vec{x}\rangle=S\vec{x}\cdot S\vec{x}=\|S\vec{x}\|^2$ is positive if $S\vec{x}\neq\vec{0}$, that is, if $\vec{x}$ is not in the kernel of $S$. It is required that $S\vec{x}\neq\vec{0}$ whenever $\vec{x}\neq\vec{0}$, that is, $\ker(S)=\{\vec{0}\}$.

*Answer:* $S$ must be invertible.

b It is required that $\langle\vec{x},\vec{y}\rangle=(S\vec{x})^TS\vec{y}=\vec{x}^TS^TS\vec{y}$ equal $\vec{x}\cdot\vec{y}=\vec{x}^T\vec{y}$ for all $\vec{x}$ and $\vec{y}$. This is the case if and only if $S^TS=I_n$, that is, $S$ *is orthogonal.*

5.5.**4** a For column vectors $\vec{v},\vec{w}$, we have $\langle\vec{v},\vec{w}\rangle=\text{trace}(\vec{v}^T\vec{w})=\text{trace}(\vec{v}\cdot\vec{w})=\vec{v}\cdot\vec{w}$, the dot product.

b For row vectors $\vec{v}, \vec{w}$, the $ij$th entry of $\vec{v}^T\vec{w}$ is $v_i w_j$, so that $\langle \vec{v}, \vec{w} \rangle = \text{trace}(\vec{v}^T\vec{w}) = \sum_{i=1}^{m} v_i w_i = \vec{v} \cdot \vec{w}$, again the dot product.

5.5.**5** a $\langle\langle A, B \rangle\rangle = \text{tr}(AB^T) = \text{tr}((AB^T)^T) = \text{tr}(BA^T) = \langle\langle B, A \rangle\rangle$ In the second step we have used the fact that $\text{tr}(M) = \text{tr}(M^T)$, for any square matrix $M$.

b $\langle\langle A + B, C \rangle\rangle = \text{tr}((A + B)C^T) = \text{tr}(AC^T + BC^T) = \text{tr}(AC^T) + \text{tr}(BC^T)$

$= \langle\langle A, C \rangle\rangle + \langle\langle B, C \rangle\rangle$

c $\langle\langle cA, B \rangle\rangle = \text{tr}(cAB^T) = c\text{tr}(AB^T) = c\langle\langle A, B \rangle\rangle$

d $\langle\langle A, A \rangle\rangle = \text{tr}(AA^T) = \sum_{i=1}^{n} \|\vec{v}_i\|^2 > 0$ if $A \neq 0$, where $\vec{v}_i$ is the $i$th row of $A$.

We have shown that $\langle\langle \cdot, \cdot \rangle\rangle$ does indeed define an inner product.

5.5.**6** a The $ii$th entry of $PQ$ is $\sum_{k=1}^{m} p_{ik}q_{ki}$, so that $\text{tr}(PQ) = \sum_{i=1}^{n}\sum_{k=1}^{m} p_{ik}q_{ki}$.

Likewise, $\text{tr}(QP) = \sum_{i=1}^{m}\sum_{k=1}^{n} q_{ik}p_{ki}$.

Reversing the roles of $i$ and $k$ and the order of summation we see that $\text{tr}(PQ) = \text{tr}(QP)$, as claimed.

b Using part a and the fact that $\text{tr}(M) = \text{tr}(M^T)$ for any square matrix $M$, we find that $\langle A, B \rangle = \text{tr}(A^TB) = \text{tr}(BA^T) = \text{tr}((BA^T)^T) = \text{tr}(AB^T) = \langle\langle A, B \rangle\rangle$

5.5.**7** Axioms *a, b*, and *c* hold for any choice of $k$ (check this!). Also, it is required that $\langle\langle v, v \rangle\rangle = k\langle v, v \rangle$ be positive for nonzero $v$. Since $\langle v, v \rangle$ is positive, this is the case if (and only if) $k$ is *positive.*

5.5.**8** By parts b and c of Definition 5.5.1, we have $T(u + v) = \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle = T(u) + T(v)$ and $T(cv) = \langle cv, w \rangle = c\langle v, w \rangle = cT(v)$, so that $T$ is linear. If $w = 0$, then $\text{im}(T) = \{0\}$ and $\ker(T) = V$. If $w \neq 0$, then $\text{im}(T) = \mathbb{R}$ and $\ker(T)$ consists of all $v$ perpendicular to $w$.

5.5.**9** If $f$ is even and $g$ is odd, then $fg$ is odd, so that $\langle f, g \rangle = \int_{-1}^{1} fg = 0$.

5.5.**10** A function $g(t) = a + bt + ct^2$ is orthogonal to $f(t) = t$ if

$\langle f, g \rangle = \int_{-1}^{1} (at + bt^2 + ct^3)\, dt = [\, \frac{a}{2}t^2 + \frac{b}{3}t^3 + \frac{c}{4}t^4\, ]_{-1}^{1} = \dfrac{2}{3}b = 0$, that is, if $b = 0$.

Thus, the functions 1 and $t^2$ form a basis of the space of all functions in $P_2$ orthogonal to $f(t) = t$. To find an *orthonormal* basis $g_1(t), g_2(t)$, we apply Gram-Schmidt. Now $\|1\| = \frac{1}{2}\int_{-1}^{1} 1\, dt = 1$, so that we can let $g_1(t) = 1$.

Then $g_2(t) = \frac{t^2 - \langle 1, t^2 \rangle 1}{\|t^2 - \langle 1, t^2 \rangle 1\|} = \frac{t^2 - \frac{1}{3}}{\|t^2 - \frac{1}{3}\|} = \frac{\sqrt{5}}{2}(3t^2 - 1)$

*Answer*: $g_1(t) = 1, g_2(t) = \frac{\sqrt{5}}{2}(3t^2 - 1)$

5.5.**11** $\langle f, g\rangle = \langle \cos(t), \cos(t+\delta)\rangle = \langle \cos(t), \cos(t)\cos(\delta) - \sin(t)\sin(\delta)\rangle = \cos(\delta)\langle \cos(t), \cos(t)\rangle - \sin(\delta)$ $\langle \cos(t), \sin(t)\rangle = \cos(\delta)$, by Theorem 5.5.4.

Also, $\langle g, g\rangle = 1$ (left to reader).

Thus, $\angle(f, g) = \arccos\left(\frac{\langle f,g\rangle}{\|f\|\|g\|}\right) = \arccos(\cos\delta) = \delta$.

5.5.**12** By Theorem 5.5.5

$a_0 = \left\langle |t|, \frac{1}{\sqrt{2}}\right\rangle = \frac{1}{\sqrt{2}\pi}\int_{-\pi}^{\pi} |t|\,dt = \frac{\pi}{\sqrt{2}}, b_k = \langle |t|, \sin(kt)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} |t|\sin(kt)\,dt = 0$, since the integrand is an odd function.

$$\begin{aligned} c_k &= \langle |t|, \cos(kt)\rangle = \tfrac{1}{\pi}\int_{-\pi}^{\pi} |t|\cos(kt)\,dt = \frac{2}{\pi}\int_0^{\pi} t\cos(kt)\,dt \\ &= \tfrac{2}{\pi}\left\{\left[\tfrac{1}{k}t\sin(kt)\right]_0^{\pi} - \textstyle\int_0^{\pi}\tfrac{1}{k}\sin(kt)\,dt\right\} \\ &= \tfrac{2}{\pi}\tfrac{1}{k^2}[\cos(kt)]_0^{\pi} = \tfrac{2}{\pi k^2}(\cos(k\pi)-1) = \begin{cases} 0 & \text{if } k \text{ is even} \\ -\frac{4}{k^2\pi} & \text{if } k \text{ is odd}\end{cases} \end{aligned}$$

In summary:

$$\begin{aligned} a_0 &= 0 \\ b_k &= 0 \\ c_k &= \begin{cases} 0 & \text{if } k \text{ is even} \\ -\frac{4}{k^2\pi} & \text{if } k \text{ is odd}\end{cases} \end{aligned}$$

5.5.**13** The sequence $(a_0, b_1, c_1, b_2, c_2, \ldots)$ is "square-summable" by Theorem 5.5.6, so that it is in $\ell_2$. Also, $\|(a_0, b_1, c_1, b_2, c_2, \ldots)$ $a_0^2 + b_1^2 + c_1^2 + b_2^2 + c_2^2 + \cdots = \|f\|^2$, by Theorem 5.5.6, so that the two norms are equal.

5.5.**14** a This is not an inner product since there are nonzero polynomials $f(t)$ in $P_2$ with $f(1) = f(2) = 0$, so that $\langle f, f\rangle = (f(1))^2 + (f(2))^2 = 0$. (For example, let $f(t) = (t-1)(t-2)$.)

b This is an inner product. We leave it to the reader to check axioms *a* to *c*. As for *d*: A nonzero polynomial $f$ in $P_2$ has at most two zeros, so that $f(1) \neq 0$ or $f(2) \neq 0$ or $f(3) \neq 0$, and $\langle f, f\rangle = (f(1))^2 + (f(2))^2 + (f(3))^2 > 0$.

5.5.**15** First note that $b = \left\langle \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\rangle = \left\langle \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}\right\rangle = c$, by part a of Definition 5.5.1, so that $b = c$. Check that if $b = c$ then $\langle \vec{v}, \vec{w}\rangle = \langle \vec{w}, \vec{v}\rangle$ for all $\vec{v}, \vec{w}$ in $\mathbb{R}^2$. Check that parts (b) and (c) of Definition 5.5.1 are satisfied for all values of constants *c,d*. Next, we need to worry about part (d) of that definition. It is required that $\left\langle \begin{bmatrix}x_1\\x_2\end{bmatrix}, \begin{bmatrix}x_1\\x_2\end{bmatrix}\right\rangle = x_1^2 + 2bx_1x_2 + dx_2^2$ be positive for all nonzero $\begin{bmatrix}x_1\\x_2\end{bmatrix}$. We can complete the square and write $x_1^2 + 2bx_1x_2 + dx_2^2 = (x_1 + bx_2)^2 + (d - b^2)x_2^2$; this quantity is positive for all nonzero $\begin{bmatrix}x_1\\x_2\end{bmatrix}$ if (and only if) $d - b^2 > 0$, or, $d > b^2$. In summary, the function is an inner product if (and only if) $b = c$ and $d > b^2$.

5.5.**16** a We start with the standard basis 1, $t$ and use the Gram-Schmidt process to construct an *orthonormal* basis $g_1(t), g_2(t)$.

$\|1\| = \sqrt{\int_0^1 dt} = 1$, so that we can let $g_1(t) = 1$. Then $g_2(t) = \frac{t - \langle 1,t\rangle 1}{\|t - \langle 1,t\rangle 1\|} = \frac{t - \frac{1}{2}}{\|t - \frac{1}{2}\|} = \sqrt{3}(2t-1)$.

*Summary*: $g_1(t) = 1$ and $g_2(t) = \sqrt{3}(2t-1)$ is an orthonormal basis.

b We are looking for $\text{proj}_{P_1}(t^2) = \langle g_1(t), t^2 \rangle g_1(t) + \langle g_2(t), t^2 \rangle g_2(t)$, by Theorem 5.5.3.

We find that $\langle g_1(t), t^2 \rangle = \int_0^1 t^2 \, dt = \dfrac{1}{3}$ and $\langle g_2(t), t^2 \rangle = \sqrt{3} \int_0^1 (2t^3 - t^2) \, dt = \dfrac{\sqrt{3}}{6}$, so that $\text{proj}_{P_1} t^2 = \frac{1}{3} + \frac{1}{2}(2t - 1) = t - \frac{1}{6}$. See Figure 5.21.

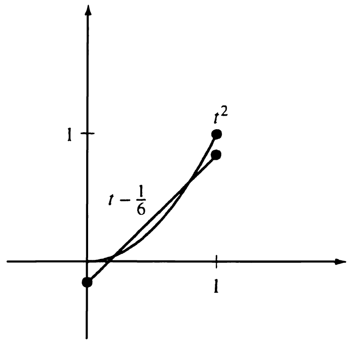



Figure 5.21: for Problem 5.5.16b.

5.5.**17** We leave it to the reader to check that the first three axioms are satisfied for any such $T$. As for axiom d: It is required that $\langle v, v \rangle = T(v) \cdot T(v) = \|T(v)\|^2$ be positive for any nonzero $v$, that is, $T(v) \neq \vec{0}$. This means that the *kernel of* $T$ *must be* $\{0\}$.

5.5.**18** Let the orthonormal basis be $f_1, \ldots, f_n$ and $f = c_1 f_1 + \cdots + c_n f_n$; then $[f]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$. The Pythagorean Theorem tells us that $\|f\|^2 = c_1^2 + \cdots + c_n^2 = \|[f]_{\mathcal{B}}\|^2$, so that $\|f\| = \|[f]_{\mathcal{B}}\|$.

5.5.**19** If we write $\vec{v} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \vec{w} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, and $A = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$, then $\left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle$

$= [x_1 x_2] \begin{bmatrix} p & q \\ r & s \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = px_1y_1 + qx_1y_2 + rx_2y_1 + sx_2y_2$. Note that in Exercise 15 we considered the special case $p = 1$. First it is required that $p = \left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle$ be positive.

Now we can write $\left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\rangle = p \left[ x_1y_1 + \frac{q}{p}x_1y_2 + \frac{r}{p}x_2y_1 + \frac{s}{p}x_2y_2 \right]$ and use our work in Exercise 15 $\Big($with $b = \frac{q}{p}, c = \frac{r}{p},$ to see that the conditions $q = r$ and $q^2 < ps$ must hold. In summary, the function is an inner product if (and only if) the entries of matrix $A = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ satisfy the conditions $p > 0$, $q = r$ and $\det(A) = ps - q^2 > 0$.

5.5.**20** a $\left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle = [x_1 x_2] \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = x_1 + 2x_2 = 0$ when $x_1 = -2x_2$. This is the line spanned by vector $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

b Since vectors $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}-2\\1\end{bmatrix}$ are orthogonal, we merely have to multiply each of them with the reciprocal of its norm. Now $\left\|\begin{bmatrix}1\\0\end{bmatrix}\right\|^2 = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\2 & 8\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = 1$, so that $\begin{bmatrix}1\\0\end{bmatrix}$ is a unit vector, and $\left\|\begin{bmatrix}-2\\1\end{bmatrix}\right\|^2 = \begin{bmatrix}-2 & 1\end{bmatrix}\begin{bmatrix}1 & 2\\2 & 8\end{bmatrix}\begin{bmatrix}-2\\1\end{bmatrix} = 4$, so that $\left\|\begin{bmatrix}-2\\1\end{bmatrix}\right\| = 2$. Thus $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}-1\\\frac{1}{2}\end{bmatrix}$ is an orthonormal basis.

5.5.**21** Note that $\langle\vec{v},\vec{w}\rangle = (\vec{v}+\vec{w})\cdot(\vec{v}+\vec{w}) - \vec{v}\cdot\vec{v} - \vec{w}\cdot\vec{w} = 2(\vec{v}\cdot\vec{w})$, double the dot product. Thus it's an inner product, by Exercise 7.

5.5.**22** Apply the Cauchy-Schwarz inequality to $f(t)$ and $g(t) = 1$; note that $\|g\| = 1$:

$$|\langle f,g\rangle| \le \|f\|\|g\| = \|f\| \text{ or } \langle f,g\rangle^2 \le \|f\|^2 \text{ or } \left(\int_0^1 f(t)\,dt\right)^2 \le \int_0^1 (f(t))^2\,dt.$$

5.5.**23** We start with the standard basis $1, t$ of $P_1$ and use the Gram-Schmidt process to construct and *orthonormal basis* $g_1(t), g_2(t)$.

$\|1\| = \sqrt{\frac{1}{2}(1\cdot 1) + 1\cdot 1} = 1$, so that we can let $g_1(t) = 1$. Then $g_2(t) = \frac{t-\langle 1,t\rangle 1}{\|t-\langle 1,t\rangle 1\|} = \frac{t-\frac{1}{2}}{\|t-\frac{1}{2}\|} = 2t-1$.

Summary: $g_1(t) = 1$ and $g_2(t) = 2t-1$ is an orthonormal basis.

5.5.**24** a $\langle f, g+h\rangle = \langle f,g\rangle + \langle f,h\rangle = 0 + 8 = 8$

b $\|g+h\| = \sqrt{\langle g+h, g+h\rangle} = \sqrt{\langle g,g\rangle + 2\langle g,h\rangle + \langle h,h\rangle} = \sqrt{1+6+50} = \sqrt{57}$

c Since $\langle f,g\rangle = 0, \|g\| = 1$, and $\|f\| = 2$, we know that $\frac{f}{2}, g$ is an orthonormal basis of span $(f,\ g)$.

Now $\text{proj}_E h = \left\langle \frac{f}{2}, h\right\rangle \frac{f}{2} + \langle g,h\rangle g = \frac{1}{4}\langle f,h\rangle f + \langle g,h\rangle g = 2f + 3g$.

d From part c we know that $\frac{1}{2}$ $f$, $g$ are orthonormal, so we apply Theorem 5.2.1 to obtain the third polynomial in an orthonormal basis of span$(f,\ g,\ h)$:

$\frac{h-\text{proj}_E h}{\|h-\text{proj}_E h\|} = \frac{h-2f-3g}{\|h-2f-3g\|} = \frac{h-2f-3g}{5} = -\frac{2}{5}f - \frac{3}{5}g + \frac{1}{5}h$

*Orthonormal basis*: $\frac{1}{2}f, g, -\frac{2}{5}f - \frac{3}{5}g + \frac{1}{5}h$

5.5.**25** Using the inner product defined in Example 2, we find that

$\|\vec{x}\| = \sqrt{\langle\vec{x},\vec{x}\rangle} = \sqrt{1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2} + \cdots} = \sqrt{\frac{\pi^2}{6}} = \frac{\pi}{\sqrt{6}}$ (see the text right after Theorem 5.5.6).

5.5.**26** $a_0 = \frac{1}{\sqrt{2}\pi}\displaystyle\int_{-\pi}^{\pi} f(t)\,dt = 0$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(kt)\,dt = \frac{1}{\pi}\left\{-\int_{-\pi}^{0}\sin(kt)\,dt + \int_0^{\pi}\sin(kt)\,dt\right\} = \frac{2}{\pi}\int_0^{\pi}\sin(kt)\,dt$$

$$= -\frac{2}{k\pi}[\cos(kt)]_0^{\pi} = \begin{cases} 0 & \text{if } k \text{ is even} \\ \frac{4}{\pi k} & \text{if } k \text{ is odd}\end{cases}$$

$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(kt)\,dt = 0,$ since the integrand is odd.

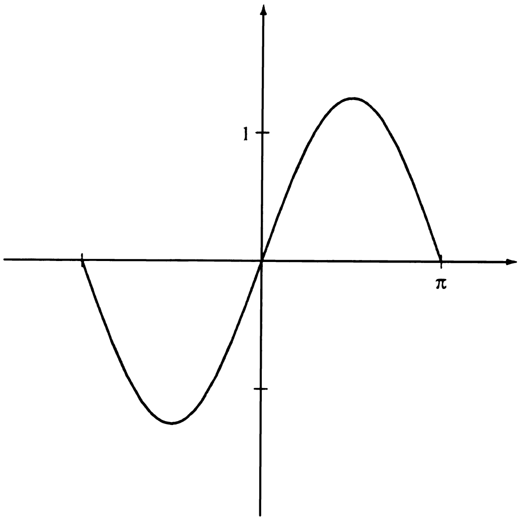



Figure 5.22: for Problem 5.5.26.

$f_1(t) = f_2(t) = \frac{4}{\pi}\sin(t)$. See Figure 5.22.

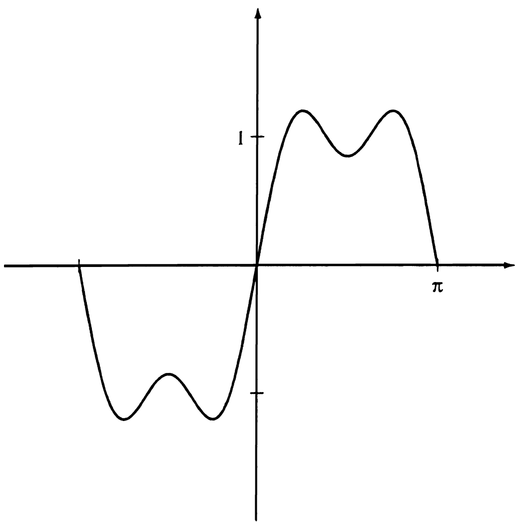



Figure 5.23: for Problem 5.5.26.

$f_3(t) = f_4(t) = \frac{4}{\pi}\sin(t) + \frac{4}{3\pi}\sin(3t)$. See Figure 5.23.

$f_5(t) = f_6(t) = \frac{4}{\pi}\sin(t) + \frac{4}{3\pi}\sin(3t) + \frac{4}{5\pi}\sin(5t)$. See Figure 5.24.

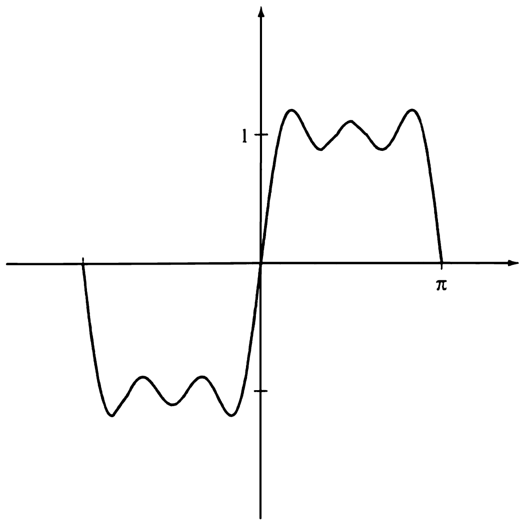



Figure 5.24: for Problem 5.5.26.

5.5.**27** $a_0 = \frac{1}{\sqrt{2}\pi}\int_{-\pi}^{\pi} f(t)\,dt = \dfrac{1}{\sqrt{2}}$.

$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(kt)\,dt = \dfrac{1}{\pi}\int_0^{\pi}\sin(kt)\,dt = -\dfrac{1}{k\pi}[\cos(kt)]_0^{\pi} = \begin{cases} 0 & \text{if } k \text{ is even} \\ \frac{2}{k\pi} & \text{if } k \text{ is odd} \end{cases}$

$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(kt)\,dt = \dfrac{1}{\pi}\int_0^{\pi}\cos(kt)\,dt = \dfrac{1}{k\pi}[\sin(kt)]_0^{\pi} = 0$

$f_1(t) = f_2(t) = \frac{1}{2} + \frac{2}{\pi}\sin(t)$, $f_3(t) = f_4(t) = \frac{1}{2} + \frac{2}{\pi}\sin(t) + \frac{2}{3\pi}\sin(3t)$

$\vdots$

5.5.**28** $\|f\|^2 = \langle f, f\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi}(f(t))^2\,dt = \dfrac{1}{\pi}\int_{-\pi}^{\pi} 1\,dt = 2$

Now Theorem 5.5.6 tells us that $\frac{16}{\pi^2} + \frac{16}{9\pi^2} + \frac{16}{25\pi^2} + \cdots = \frac{16}{\pi^2}\left(\sum_{k\,\text{odd}} \dfrac{1}{k^2}\right) = 2$, or $\sum_{k\,\text{odd}} \dfrac{1}{k^2} = 1 + \dfrac{1}{9} + \dfrac{1}{25} + \dfrac{1}{49} + \cdots = \dfrac{\pi^2}{8}$.

5.5.**29** $\|f\|^2 = \langle f, f\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi}(f(t))^2\,dt = \dfrac{1}{\pi}\int_0^{\pi} 1\,dt = 1$

Now Theorem 5.5.6 tells us that $\frac{1}{2} + \frac{4}{\pi^2} + \frac{4}{9\pi^2} + \frac{4}{25\pi^2} + \cdots = 1$, or $\frac{1}{2} + \frac{4}{\pi^2}\left(\sum_{k\,\text{odd}} \dfrac{1}{k^2}\right) = 1$ or

$\sum_{k\,\text{odd}} \dfrac{1}{k^2} = 1 + \dfrac{1}{9} + \dfrac{1}{25} + \dfrac{1}{49} + \cdots = \dfrac{\frac{1}{2}}{\frac{4}{\pi^2}} = \dfrac{\pi^2}{8}$.

5.5.**30** There is an invertible linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ that transforms $E$ into the unit circle. (If $E$ is parametrized by $\cos(t)\vec{w}_1 + \sin(t)\vec{w}_2$, let $A = [\,\vec{w}_1\ \vec{w}_2\,]^{-1}$. Compare with Exercises 2.2.54 and 2.2.55.)

Now $\vec{x}$ is on $E$ if $\|T(\vec{x})\| = 1$, or $T(\vec{x}) \cdot T(\vec{x}) = 1$. This means that the inner product $\langle \vec{x}, \vec{y} \rangle = T(\vec{x}) \cdot T(\vec{y})$ does the job (see Exercise 17). (There are other, very different approaches to this problem.)

5.5.**31** An orthonormal basis of $P_2$ of the desired form is $f_0(t) = \frac{1}{\sqrt{2}}, f_1(t) = \sqrt{\frac{3}{2}}t, f_2(t) = \frac{1}{2}\sqrt{\frac{5}{2}}(3t^2 - 1)$ (compare with Exercise 10), and the zeros of $f_2(t)$ are $a_{1,2} = \pm\frac{1}{\sqrt{3}}$.

Next we find the weights $w_1, w_2$ such that $\displaystyle\int_{-1}^{1} f(t)\,dt = \sum_{i=1}^{2} w_i f(a_i)$ for all $f$ in $P_1$. We need to make sure that the equation holds for 1 and $t$.

$\begin{bmatrix} 2 = w_1 + w_2 \\ 0 = \frac{1}{\sqrt{3}}w_1 - \frac{1}{\sqrt{3}}w_2 \end{bmatrix}$, with solution $w_1 = w_2 = 1$.

It follows that the equation $\displaystyle\int_{-1}^{1} f(t)\,dt = f(a_1) + f(a_2) = f\left(\frac{1}{\sqrt{3}}\right) + f\left(-\frac{1}{\sqrt{3}}\right)$ holds for all polynomials $f$ in $P_1$. We can check that it holds for $t^2$ and $t^3$ as well, that is, it holds in fact for all cubic polynomials.

$\displaystyle\int_{-1}^{1} t^2\,dt = \frac{2}{3}$ equals $\left(\frac{1}{\sqrt{3}}\right)^2 + \left(-\frac{1}{\sqrt{3}}\right)^2 = \frac{2}{3}$, and $\displaystyle\int_{-1}^{1} t^3\,dt = 0$ equals $\left(\frac{1}{\sqrt{3}}\right)^3 + \left(-\frac{1}{\sqrt{3}}\right)^3 = 0$.

5.5.**32** a. $\langle t^n, t^m \rangle = \frac{1}{2}\int_{-1}^{1} t^{n+m}dt = \frac{1}{2}\left[\frac{t^{n+m+1}}{n+m+1}\right]_{-1}^{1} = \begin{cases} \frac{1}{n+m+1} & \text{if } n+m \text{ is even} \\ 0 & \text{if } n+m \text{ is odd} \end{cases}$

b. $\|t^n\| = \sqrt{\langle t^n, t^n \rangle} = \sqrt{\frac{1}{2n+1}}$ by part a.

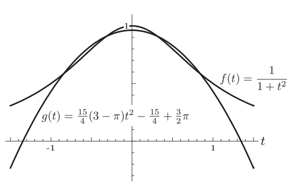
c. It helps to observe that 1 and $t^2$ are orthogonal to $t$ and $t^3$. Accordingly, many terms in the Gram-Schmidt formulas vanish.

$$g_0(t) = \frac{1}{\|1\|} = 1, g_1(t) = \frac{t}{\|t\|} = \sqrt{3}\,t, g_2(t) = \frac{t^2 - \langle t^2, 1\rangle 1}{\|t^2 - \langle t^2, 1\rangle 1\|} = \frac{\sqrt{5}}{2}\left(3t^2 - 1\right)$$

$$g_3(t) = \frac{t^3 - \langle t^3, \sqrt{3}\,t\rangle \sqrt{3}\,t}{\left\|t^3 - \langle t^3, \sqrt{3}\,t\rangle \sqrt{3}\,t\right\|} = \frac{t^3 - 3t/5}{\|t^3 - 3t/5\|} = \frac{\sqrt{7}}{2}\left(5t^3 - 3t\right)$$

d. The first few Legendre Polynomials are $\frac{g_0(t)}{g_1(1)} = 1$, $\frac{g_1(t)}{g_1(1)} = t$, $\frac{g_2(t)}{g_2(1)} = \left(3t^2 - 1\right)/2$, $\frac{g_3(t)}{g_3(1)} = \left(5t^3 - 3t\right)/2$.

e. Since $f(t) = \frac{1}{1+t^2}$ is an even function, we need to consider $g_0(t)$ and $g_2(t)$ only. The solution is $\langle f(t), g_0(t)\rangle g_0(t) + \langle f(t), g_2(t)\rangle g_2(t) = \langle f(t), 1\rangle 1 + \frac{5}{4}\left\langle f(t), 3t^2 - 1\right\rangle \left(3t^2 - 1\right) = \frac{\pi}{4} + \frac{5}{8}\left(6 - 2\pi\right)\left(3t^2 - 1\right) = \frac{15}{4}\left(3 - \pi\right)t^2 - \frac{15}{4} + \frac{3\pi}{2}$.



5.5.**33** a The first property of an inner product, $\langle f, g\rangle = \langle g, f\rangle$, follows from the fact that $f(t)g(t) = g(t)f(t)$. The property $\langle f + h, g\rangle = \langle f, g\rangle + \langle h, g\rangle$ follows from the sum rule for integrals, and the property $\langle cf, g\rangle = c\,\langle f, g\rangle$ follows from the constant multiple rule for integrals. Proceed as in Exercise 1 to show that $\langle f, f\rangle > 0$ for all nonzero $f$.

b If $f(t) = 1$ then $\|f\| = \sqrt{\langle f, f\rangle} = \sqrt{\int_a^b w(t)\cdot 1^2 dt} = \sqrt{\int_a^b w(t)dt} = 1.$

5.5.**34** a. Note that the graph of $\sqrt{1-t^2}$ is the upper half of the unit circle centered at the origin. Thus $\int_{-1}^{1} w(t)\,dt = \frac{2}{\pi}\int_{-1}^{1}\sqrt{1-t^2}dt = \frac{2}{\pi}$ (half the area of the unit circle) $= 1$

b. If $f(t) = 1$ then $\|f\| = \sqrt{\langle f, f\rangle} = 1$ by part a and Exercise 33b.

c. If $n + m$ is odd then $\langle t^n, t^m\rangle = \frac{2}{\pi}\int_{-1}^{1}\sqrt{1-t^2}\; t^{n+m}dt = 0$ since the integrand is an odd function.

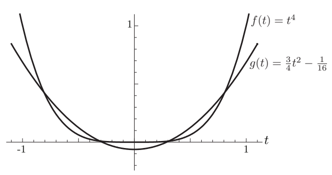
d. It follows from the definition of this inner product that $\langle t, t\rangle = \langle 1, t^2\rangle = 1/4$ and $\langle t^2, t^2\rangle = \langle t, t^3\rangle = 1/8$ so that $\|t\| = \sqrt{1/4} = 1/2$ and $\|t^2\| = \sqrt{1/8} = \sqrt{2}/4$.

e. It helps to observe that 1 and $t^2$ are orthogonal to $t$ and $t^3$. Accordingly, many terms in the Gram-Schmidt formulas vanish. The first few Chebyshev Polynomials of the Second Kind are

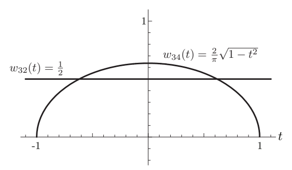
$$g_0(t) = \frac{1}{\|1\|} = 1, g_1(t) = \frac{t}{\|t\|} = 2\,t, g_2(t) = \frac{t^2 - \langle t^2, 1\rangle\,1}{\|t^2 - \langle t^2, 1\rangle\,1\|} = \frac{t^2 - 1/4}{\sqrt{1/16}} = 4t^2 - 1$$

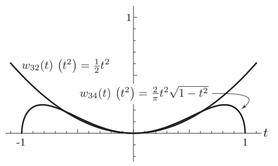
$$g_3(t) = \frac{t^3 - \langle t^3, 2\,t\rangle\,2\,t}{\|t^3 - \langle t^3, 2\,t\rangle\,2\,t\|} = \frac{t^3 - t/2}{\sqrt{1/64}} = 8t^3 - 4t.$$

f. Since $f(t) = t^4$ is an even function, we need to consider $g_0(t)$ and $g_2(t)$ only. The solution is $\langle f(t), g_0(t)\rangle\, g_0(t) + \langle f(t), g_2(t)\rangle\, g_2(t) = \langle t^4, 1\rangle\,1 + \langle t^4, 4t^2-1\rangle\,(4t^2-1) = \frac{1}{8} + \left(\frac{5}{16} - \frac{1}{8}\right)(4t^2-1) = \frac{3}{4}t^2 - \frac{1}{16}$.

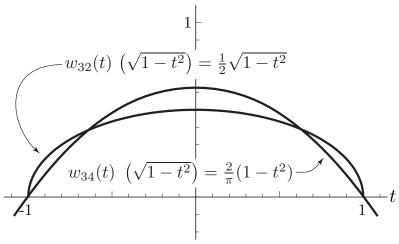



5.5.**35** a. $\|t\|_{32} = \sqrt{\frac{1}{2}\int_{-1}^{1} t^2 dt} = \sqrt{1/3}$ and $\|t\|_{34} = \sqrt{\langle t, t\rangle_{34}} = \sqrt{\langle 1, t^2\rangle_{34}} = 1/2$, so that $\|t\|_{32} > \|t\|_{34}$. Note that $\langle 1, t^2\rangle_{34} = 1/4$ is given in Exercise 34.



b. For $f(t) = \sqrt{1-t^2}$ we have $\|f\|_{32} = \sqrt{\frac{1}{2}\int_{-1}^{1}(1-t^2)\,dt} = \sqrt{2/3}$ and $\|f\|_{34} = \sqrt{\left\langle\sqrt{1-t^2}, \sqrt{1-t^2}\right\rangle_{34}} = \sqrt{\langle 1, 1-t^2\rangle_{34}} = \sqrt{1-1/4} = \sqrt{3/4}$.




## True or False

Ch 5.TF.**1** F. Consider $T(\vec{x}) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\vec{x}$.

Ch 5.TF.**2** T, by Theorem 5.3.9.b

Ch 5.TF.**3** T, by Theorem 5.3.4a

Ch 5.TF.**4** F. We have $(AB)^T = B^T A^T$, by Theorem 5.3.9a.

Ch 5.TF.**5** T, since $(A+B)^T = A^T + B^T = A + B$

Ch 5.TF.**6** T, by Theorem 5.3.4

Ch 5.TF.**7** F. Consider $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Ch 5.TF.**8** T. First note that $A^T = A^{-1}$, by Theorem 2.4.8. Thus $A$ is orthogonal, by Theorem 5.3.7.

Ch 5.TF.**9** F. The correct formula is $\text{proj}_L(\vec{x}) = (\vec{x}\cdot\vec{u})\vec{u}$, by Definition 2.2.1.

Ch 5.TF.**10** T, since $(7A)^T = 7A^T = 7A$.

Ch 5.TF.**11** F. The Pythagorean Theorem holds for orthogonal vectors $\vec{x}, \vec{y}$ only (Theorem 5.1.9)

Ch 5.TF.**12** T. $\det\begin{bmatrix} a & c \\ b & d \end{bmatrix} = ad - bc = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Ch 5.TF.**13** T. If $A$ is orthogonal, then $A^T = A^{-1}$, and $A^{-1}$ is orthogonal by Theorem 5.3.4b.

Ch 5.TF.**14** F. Consider $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Ch 5.TF.**15** F. Consider $A = B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $AB^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ isn't equal to $B^T A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

Ch 5.TF.**16** F. It is required that the columns of $A$ be orthonormal (Theorem 5.3.10). As a counterexample, consider $A = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, with $AA^T = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$.

Ch 5.TF.**17** T, since $(ABBA)^T = A^T B^T B^T A^T = ABBA$, by Theorem 5.3.9a

Ch 5.TF.**18** T, since $A^T B^T = (BA)^T = (AB)^T = B^T A^T$, by Theorem 5.3.9a

Ch 5.TF.**19** F. $\dim(V) + \dim(V^\perp) = 5$, by Theorem 5.1.8c. Thus one of the dimensions is even and the other odd.

Ch 5.TF.**20** T. Consider the $QR$ factorization (Theorem 5.2.2)

Ch 5.TF.**21** F. $\det \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = -1 - 0 = -1$, yet $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ is orthogonal.

Ch 5.TF.**22** T. $\left[\frac{1}{2}(A - A^T)\right]^T = \frac{1}{2}(A - A^T)^T = \frac{1}{2}(A^T - A) = -\left[\frac{1}{2}(A - A^T)\right]$.

Ch 5.TF.**23** T, since the columns are unit vectors.

Ch 5.TF.**24** T. Use the Gram-Schmidt process to construct such a basis (Theorem 5.2.1)

Ch 5.TF.**25** F. The columns fail to be unit vectors (use Theorem 5.3.3b)

Ch 5.TF.**26** T, by definition of an orthogonal projection (Theorem 5.1.4).

Ch 5.TF.**27** F. As a counterexample, consider $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Ch 5.TF.**28** T, by Theorem 5.4.1.

Ch 5.TF.**29** T, by Theorem 5.4.2a.

Ch 5.TF.**30** F. Consider $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, or any other symmetric matrix that fails to be orthogonal.

Ch 5.TF.**31** T. Try $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, so that $A + B = \begin{bmatrix} 1 + \cos\theta & -\sin\theta \\ \sin\theta & 1 + \cos\theta \end{bmatrix}$. It is required that $\begin{bmatrix} 1 + \cos\theta \\ \sin\theta \end{bmatrix}$ and $\begin{bmatrix} -\sin\theta \\ 1 + \cos\theta \end{bmatrix}$ be unit vectors, meaning that $1 + 2\cos\theta + \cos^2\theta + \sin^2\theta = 2 + 2\cos\theta = 1$, or $\cos\theta = -\frac{1}{2}$, and $\sin\theta = \pm\frac{\sqrt{3}}{2}$. Thus $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$ is a solution.

Ch 5.TF.**32** F. Consider $A = \begin{bmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$, for example, representing a rotation combined with a scaling.

Ch 5.TF.**33** F. Consider $A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$.

Ch 5.TF.**34** T. By Definition 5.1.12, quantity $\cos(\theta) = \frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\|\vec{w}\|}$ is positive, so that $\theta$ is an acute angle.

Ch 5.TF.**35** T. In Theorem 5.4.1, let $A = B^T$ to see that $(\text{im}(B^T))^\perp = \ker(B)$. Now take the orthogonal complements of both sides and use Theorem 5.1.8d.

Ch 5.TF.**36** T, since $(A^TA)^T = A^T(A^T)^T = A^TA$, by Theorem 5.3.9a.

Ch 5.TF.**37** F. Verify that matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$ are similar.

Ch 5.TF.**38** F. Consider $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. The correct formula $\text{im}(B) = \text{im}(BB^T)$ follows from Theorems 5.4.1 and 5.4.2.

Ch 5.TF.**39** T. We know that $A^T = A$ and $S^{-1} = S^T$. Now $(S^{-1}AS)^T = S^TA^T(S^{-1})^T = S^{-1}AS$, by Theorem 5.3.9a.

Ch 5.TF.**40** T. By Theorem 5.4.2, we have $\ker(A) = \ker(A^TA)$. Replacing $A$ by $A^T$ in this formula, we find that $\ker(A^T) = \ker(AA^T)$. Now $\ker(A) = \ker(A^TA) = \ker(AA^T) = \ker(A^T)$.

Ch 5.TF.**41** T. We attempt to write $A = S + Q$, where $S$ is symmetric and $Q$ is skew-symmetric. Then $A^T = S^T + Q^T = S - Q$. Adding the equations $A = S + Q$ and $A^T = S - Q$ together gives $2S = A + A^T$ and $S = \frac{1}{2}(A + A^T)$. Similarly we find $Q = \frac{1}{2}(A - A^T)$. Check that the decomposition $A = S + Q = (\frac{1}{2}(A + A^T)) + (\frac{1}{2}(A - A^T))$ does the job.

Ch 5.TF.**42** T. Apply the Cauchy-Schwarz inequality (squared), $(\vec{x}\cdot\vec{y})^2 \le \|\vec{x}\|^2\|\vec{y}\|^2$, to $\vec{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix}$ (all $n$ entries are 1).

Ch 5.TF.**43** T. Let $A = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$. We know that $AA^T = A^2$, or $\begin{bmatrix} x^2+y^2 & xz+yt \\ xz+yt & z^2+t^2 \end{bmatrix} = \begin{bmatrix} x^2+yz & xy+yt \\ zx+tz & yz+t^2 \end{bmatrix}$. We need to show that $y = z$. If $y \neq 0$, this follows from the equation $x^2 + y^2 = x^2 + yz$; if $z \neq 0$, it follows from $z^2 + t^2 = yz + t^2$; if both $y$ and $z$ are zero, we are all set.

Ch 5.TF.**44** T, since $\vec{x}\cdot(\text{proj}_V\vec{x}) = (\text{proj}_V\vec{x} + (\vec{x} - \text{proj}_V\vec{x}))\cdot\text{proj}_V\vec{x} = \|\text{proj}_V\vec{x}\|^2 \ge 0$. Note that $\vec{x} - \text{proj}_V\vec{x}$ is orthogonal to $\text{proj}_V\vec{x}$, by the definition of a projection.

Ch 5.TF.**45** T. Note that $1 = \left\|A\left(\frac{1}{\|\vec{x}\|}\vec{x}\right)\right\| = \left\|\frac{1}{\|\vec{x}\|}A\vec{x}\right\| = \frac{1}{\|\vec{x}\|}\|A\vec{x}\|$ for all nonzero $\vec{x}$, so that $\|A\vec{x}\| = \|\vec{x}\|$. See Definition 5.3.1.

Ch 5.TF.**46** T. If $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is a symmetric matrix, then $A - xI_2 = \begin{bmatrix} a-x & b \\ b & c-x \end{bmatrix}$. This matrix fails to be invertible if (and only if) $\det(A - xI_2) = (a-x)(c-x) - b^2 = 0$. We use the quadratic formula to find the (real) solutions $x = \frac{a+c\pm\sqrt{(a+c)^2-4ac+4b^2}}{2} = \frac{a+c\pm\sqrt{(a-c)^2+4b^2}}{2}$. Note that the discriminant $(a-c)^2 + 4b^2$ is positive or zero.

Ch 5.TF.**47** T; one basis is: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Ch 5.TF.**48** F; A direct computation or a geometrical argument shows that $Q = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$, representing a reflection, not a rotation.

Ch 5.TF.**49** F; $\dim(\mathbb{R}^{3\times3}) = 9$, $\dim(\mathbb{R}^{2\times2}) = 4$, so $\dim(\ker(L)) \geq 5$, but the space of all $3\times3$ skew-symmetric matrices has dimension of 3.

$$\left(\text{A basis is } \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.\right)$$

Ch 5.TF.**50** T; Consider an orthonormal basis $\vec{v}_1, \vec{v}_2$ of $V$, and a unit vector $\vec{v}_3$ perpendicular to $V$, and form the orthogonal matrix $S = [\vec{v}_1\ \vec{v}_2\ \vec{v}_3]$. Now $AS = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{0} \end{bmatrix} = S\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Since $S$ is orthogonal, we have $S^TAS = S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, a diagonal matrix.