8. Since $\mathbf{u}_1 \cdot \mathbf{u}_2 = -1 + 3 - 2 = 0$, $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal set. By the Orthogonal Decomposition Theorem,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = 2\mathbf{u}_1 + \frac{1}{2}\mathbf{u}_2 = \begin{bmatrix} 3/2 \\ 7/2 \\ 1 \end{bmatrix}, \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -5/2 \\ 1/2 \\ 2 \end{bmatrix}$$

and $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

9. Since $\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_1 \cdot \mathbf{u}_3 = \mathbf{u}_2 \cdot \mathbf{u}_3 = 0$, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. By the Orthogonal Decomposition Theorem,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3 = 2\mathbf{u}_1 + \frac{2}{3}\mathbf{u}_2 - \frac{2}{3}\mathbf{u}_3 = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}, \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix}$$

and $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

10. Since $\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathbf{u}_1 \cdot \mathbf{u}_3 = \mathbf{u}_2 \cdot \mathbf{u}_3 = 0$, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. By the Orthogonal Decomposition Theorem,

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3 = \frac{1}{3}\mathbf{u}_1 + \frac{14}{3}\mathbf{u}_2 - \frac{5}{3}\mathbf{u}_3 = \begin{bmatrix} 5 \\ 2 \\ 3 \\ 6 \end{bmatrix}, \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -2 \\ 2 \\ 2 \\ 0 \end{bmatrix}$$

and $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$.

11. Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. The Best Approximation Theorem says that $\hat{\mathbf{y}}$, which is the orthogonal projection of $\mathbf{y}$ onto $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, is the closest point to $\mathbf{y}$ in W. This vector is

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{1}{2}\mathbf{v}_1 + \frac{3}{2}\mathbf{v}_2 = \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

12. Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. The Best Approximation Theorem says that $\hat{\mathbf{y}}$, which is the orthogonal projection of y onto $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, is the closest point to $\mathbf{y}$ in W. This vector is

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = 3\mathbf{v}_1 + 1\mathbf{v}_2 = \begin{bmatrix} -1 \\ -5 \\ -3 \\ 9 \end{bmatrix}$$

13. Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. By the Best Approximation Theorem, the closest point in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\mathbf{z}$ is

$$\hat{\mathbf{z}} = \frac{\mathbf{z} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{z} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{2}{3}\mathbf{v}_1 - \frac{7}{3}\mathbf{v}_2 = \begin{bmatrix} -1 \\ -3 \\ -2 \\ 3 \end{bmatrix}$$

mposition Theorem.

al Decomposition

Decomposition

ich is the
ector is

ch is the
tor is

nt in

14. Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. By the Best Approximation Theorem, the closest point in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\mathbf{z}$ is

$$\hat{\mathbf{z}} = \frac{\mathbf{z}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{z}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 = \frac{1}{2}\mathbf{v}_1 + 0\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1/2 \\ -3/2 \end{bmatrix}$$

15. The distance from the point $\mathbf{y}$ in $\mathbb{R}^3$ to a subspace W is defined as the distance from $\mathbf{y}$ to the closest in W. Since the closest point in W to $\mathbf{y}$ is $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$, the desired distance is $\|\mathbf{y} - \hat{\mathbf{y}}\|$. One comput

$$\hat{\mathbf{y}} = \begin{bmatrix} 3 \\ -9 \\ -1 \end{bmatrix}, \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix}, \text{ and } \|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{40} = 2\sqrt{10}.$$

16. The distance from the point $\mathbf{y}$ in $\mathbb{R}^4$ to a subspace W is defined as the distance from $\mathbf{y}$ to the closest in W. Since the closest point in W to $\mathbf{y}$ is $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$, the desired distance is $\|\mathbf{y} - \hat{\mathbf{y}}\|$. One compute

$$\hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ -3 \\ 9 \end{bmatrix}, \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix}, \text{ and } \|\mathbf{y} - \hat{\mathbf{y}}\| = 8.$$

17. **a.** $U^TU = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, UU^T = \begin{bmatrix} 8/9 & -2/9 & 2/9 \\ -2/9 & 5/9 & 4/9 \\ 2/9 & 4/9 & 5/9 \end{bmatrix}$

b. Since $U^TU = I_2$, the columns of U form an orthonormal basis for W, and by Theorem 10

$$\text{proj}_W \mathbf{y} = UU^T\mathbf{y} = \begin{bmatrix} 8/9 & -2/9 & 2/9 \\ -2/9 & 5/9 & 4/9 \\ 2/9 & 4/9 & 5/9 \end{bmatrix}\begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

18. **a.** $U^TU = [1] = 1, UU^T = \begin{bmatrix} 1/10 & -3/10 \\ -3/10 & 9/10 \end{bmatrix}$

b. Since $U^TU = 1$, $\{\mathbf{u}_1\}$ forms an orthonormal basis for W, and by Theorem 10

$$\text{proj}_W \mathbf{y} = UU^T\mathbf{y} = \begin{bmatrix} 1/10 & -3/10 \\ -3/10 & 9/10 \end{bmatrix}\begin{bmatrix} 7 \\ 9 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix}.$$

19. By the Orthogonal Decomposition Theorem, $\mathbf{u}_3$ is the sum of a vector in $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ and a vector $\mathbf{v}$ orthogonal to W. This exercise asks for the vector $\mathbf{v}$:

$$\mathbf{v} = \mathbf{u}_3 - \text{proj}_W \mathbf{u}_3 = \mathbf{u}_3 - \left(-\frac{1}{3}\mathbf{u}_1 + \frac{1}{15}\mathbf{u}_2\right) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -2/5 \\ 4/5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/5 \\ 1/5 \end{bmatrix}$$

Any multiple of the vector $\mathbf{v}$ will also be in $W^\perp$.

20. By the Orthogonal Decomposition Theorem, $\mathbf{u}_4$ is the sum of a vector in $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ and a vector $\mathbf{v}$ orthogonal to W. This exercise asks for the vector $\mathbf{v}$:

$$\mathbf{v} = \mathbf{u}_4 - \text{proj}_W \mathbf{u}_4 = \mathbf{u}_4 - \left(\frac{1}{6}\mathbf{u}_1 - \frac{1}{30}\mathbf{u}_2\right) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1/5 \\ -2/5 \end{bmatrix} = \begin{bmatrix} 0 \\ 4/5 \\ 2/5 \end{bmatrix}$$

Any multiple of the vector $\mathbf{v}$ will also be in $W^\perp$.

21. a. True. See the calculations for $\mathbf{z}_2$ in Example 1 or the box after Example 6 in Section 6.1.

b. True. See the Orthogonal Decomposition Theorem.

c. False. See the last paragraph in the proof of Theorem 8, or see the second paragraph after the statement of Theorem 9.

d. True. See the box before the Best Approximation Theorem.

e. True. Theorem 10 applies to the column space W of U because the columns of U are linearly independent and hence form a basis for W.

22. a. True. See the proof of the Orthogonal Decomposition Theorem.

b. True. See the subsection "A Geometric Interpretation of the Orthogonal Projection."

c. True. The orthgonal decomposition in Theorem 8 is unique.

d. False. The Best Approximation Theorem says that the best approximation to $\mathbf{y}$ is $\text{proj}_W \mathbf{y}$.

e. False. This statement is only true if $\mathbf{x}$ is in the column space of U. If $n > p$, then the column space of U will not be all of $\mathbb{R}^n$, so the statement cannot be true for all $\mathbf{x}$ in $\mathbb{R}^n$.

23. By the Orthogonal Decomposition Theorem, each $\mathbf{x}$ in $\mathbb{R}^n$ can be written uniquely as $\mathbf{x} = \mathbf{p} + \mathbf{u}$, with $\mathbf{p}$ in Row A and $\mathbf{u}$ in $(\text{Row } A)^\perp$. By Theorem 3 in Section 6.1, $(\text{Row } A)^\perp = \text{Nul } A$, so $\mathbf{u}$ is in Nul$A$.

Next, suppose $A\mathbf{x} = \mathbf{b}$ is consistent. Let $\mathbf{x}$ be a solution and write $\mathbf{x} = \mathbf{p} + \mathbf{u}$ as above. Then $A\mathbf{p} = A(\mathbf{x} - \mathbf{u}) = A\mathbf{x} - A\mathbf{u} = \mathbf{b} - \mathbf{0} = \mathbf{b}$, so the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution $\mathbf{p}$ in Row A. Finally, suppose that $\mathbf{p}$ and $\mathbf{p}_1$ are both in Row$A$ and both satisfy $A\mathbf{x} = \mathbf{b}$. Then $\mathbf{p} - \mathbf{p}_1$ is in $\text{Nul } A = (\text{Row } A)^\perp$, since $A(\mathbf{p} - \mathbf{p}_1) = A\mathbf{p} - A\mathbf{p}_1 = \mathbf{b} - \mathbf{b} = \mathbf{0}$. The equations $\mathbf{p} = \mathbf{p}_1 + (\mathbf{p} - \mathbf{p}_1)$ and $\mathbf{p} = \mathbf{p} + \mathbf{0}$ both then decompose $\mathbf{p}$ as the sum of a vector in Row$A$ and a vector in $(\text{Row } A)^\perp$. By the uniqueness of the orthogonal decomposition (Theorem 8), $\mathbf{p} = \mathbf{p}_1$, and $\mathbf{p}$ is unique.

24. a. By hypothesis, the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_p$ are pairwise orthogonal, and the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_q$ are pairwise orthogonal. Since $\mathbf{w}_i$ is in W for any i and $\mathbf{v}_j$ is in $W^\perp$ for any j, $\mathbf{w}_i \cdot \mathbf{v}_j = 0$ for any i and j. Thus $\{\mathbf{w}_1, \ldots, \mathbf{w}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$ forms an orthogonal set.

b. For any $\mathbf{y}$ in $\mathbb{R}^n$, write $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ as in the Orthogonal Decomposition Theorem, with $\hat{\mathbf{y}}$ in W and $\mathbf{z}$ in $W^\perp$. Then there exist scalars $c_1, \ldots, c_p$ and $d_1, \ldots, d_q$ such that $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = c_1\mathbf{w}_1 + \ldots + c_p\mathbf{w}_p + d_1\mathbf{v}_1 + \ldots + d_q\mathbf{v}_q$. Thus the set $\{\mathbf{w}_1, \ldots, \mathbf{w}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$ spans $\mathbb{R}^n$.

c. The set $\{\mathbf{w}_1, \ldots, \mathbf{w}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$ is linearly independent by (a) and spans $\mathbb{R}^n$ by (b), and is thus a basis for $\mathbb{R}^n$. Hence $\dim W + \dim W^\perp = p + q = \dim \mathbb{R}^n$.

2. Set $\mathbf{v}_1 = \mathbf{x}_1$ and compute that $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - \frac{1}{2}\mathbf{v}_1 = \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix}$. Thus an orthogonal basis for W is

$$\left\{ \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ -8 \end{bmatrix} \right\}.$$

3. Set $\mathbf{v}_1 = \mathbf{x}_1$ and compute that $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - \frac{1}{2}\mathbf{v}_1 = \begin{bmatrix} 3 \\ 3/2 \\ 3/2 \end{bmatrix}$. Thus an orthogonal basis for W is

$$\left\{ \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3/2 \\ 3/2 \end{bmatrix} \right\}.$$

4. Set $\mathbf{v}_1 = \mathbf{x}_1$ and compute that $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - (-2)\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix}$. Thus an orthogonal basis for W is

$$\left\{ \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix} \right\}.$$

5. Set $\mathbf{v}_1 = \mathbf{x}_1$ and compute that $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - 2\mathbf{v}_1 = \begin{bmatrix} 5 \\ 1 \\ -4 \\ -1 \end{bmatrix}$. Thus an orthogonal basis for W is

$$\left\{ \begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \\ -1 \end{bmatrix} \right\}.$$

6. Set $\mathbf{v}_1 = \mathbf{x}_1$ and compute that $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - (-3)\mathbf{v}_1 = \begin{bmatrix} 4 \\ 6 \\ -3 \\ 0 \end{bmatrix}$. Thus an orthogonal basis for W is

$$\left\{ \begin{bmatrix} 3 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ -3 \\ 0 \end{bmatrix} \right\}.$$

7. Since $\|\mathbf{v}_1\| = \sqrt{30}$ and $\|\mathbf{v}_2\| = \sqrt{27/2} = 3\sqrt{6}/2$, an orthonormal basis for W is

$$\left\{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}\right\} = \left\{\begin{bmatrix} 2/\sqrt{30} \\ -5/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}\right\}.$$

8. Since $\|\mathbf{v}_1\| = \sqrt{50}$ and $\|\mathbf{v}_2\| = \sqrt{54} = 3\sqrt{6}$, an orthonormal basis for W is

$$\left\{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}\right\} = \left\{\begin{bmatrix} 3/\sqrt{50} \\ -4/\sqrt{50} \\ 5/\sqrt{50} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}\right\}.$$

9. Call the columns of the matrix $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ and perform the Gram-Schmidt process on these vector

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - (-2)\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \mathbf{x}_3 - \frac{3}{2}\mathbf{v}_1 - \left(-\frac{1}{2}\right)\mathbf{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}$$

Thus an orthogonal basis for W is $\left\{\begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}\right\}$.

10. Call the columns of the matrix $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ and perform the Gram-Schmidt process on these vectors:

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - (-3)\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \mathbf{x}_3 - \frac{1}{2}\mathbf{v}_1 - \frac{5}{2}\mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix}$$

Thus an orthogonal basis for W is $\left\{\begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix}\right\}$.

a. The normal equations are $(A^TA)\mathbf{x} = A^T\mathbf{b}$: $\begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$

b. Compute

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} = \begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}^{-1}\begin{bmatrix} 6 \\ -6 \end{bmatrix} = \frac{1}{216}\begin{bmatrix} 42 & -6 \\ -6 & 6 \end{bmatrix}\begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

$$= \frac{1}{216}\begin{bmatrix} 288 \\ -72 \end{bmatrix} = \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix}$$

4. To find the normal equations and to find $\hat{\mathbf{x}}$, compute

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

a. The normal equations are $(A^TA)\mathbf{x} = A^T\mathbf{b}$: $\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$

b. Compute

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}^{-1}\begin{bmatrix} 6 \\ 14 \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 11 & -3 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

$$= \frac{1}{24}\begin{bmatrix} 24 \\ 24 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

5. To find the least squares solutions to $A\mathbf{x} = \mathbf{b}$, compute and row reduce the augmented matrix for the system $A^TA\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} A^TA & A^T\mathbf{b} \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 & 14 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so all vectors of the form $\hat{\mathbf{x}} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ are the least-squares solutions of $A\mathbf{x} = \mathbf{b}$.

6. To find the least squares solutions to $A\mathbf{x} = \mathbf{b}$, compute and row reduce the augmented matrix for the system $A^TA\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} A^TA & A^T\mathbf{b} \end{bmatrix} = \begin{bmatrix} 6 & 3 & 3 & 27 \\ 3 & 3 & 0 & 12 \\ 3 & 0 & 3 & 15 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

so all vectors of the form $\hat{\mathbf{x}} = \begin{bmatrix} 5 \\ -1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ are the least-squares solutions of $A\mathbf{x} = \mathbf{b}$.

7. From Exercise 3, $A=\begin{bmatrix}1&-2\\-1&2\\0&3\\2&5\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}3\\1\\-4\\2\end{bmatrix}$, and $\hat{\mathbf{x}}=\begin{bmatrix}4/3\\-1/3\end{bmatrix}$. Since

$$A\hat{\mathbf{x}}-\mathbf{b}=\begin{bmatrix}1&-2\\-1&2\\0&3\\2&5\end{bmatrix}\begin{bmatrix}4/3\\-1/3\end{bmatrix}-\begin{bmatrix}3\\1\\-4\\2\end{bmatrix}=\begin{bmatrix}2\\-2\\-1\\1\end{bmatrix}-\begin{bmatrix}3\\1\\-4\\2\end{bmatrix}=\begin{bmatrix}-1\\-3\\3\\-1\end{bmatrix}$$

the least squares error is $\|A\hat{\mathbf{x}}-\mathbf{b}\|=\sqrt{20}=2\sqrt{5}$.

8. From Exercise 4, $A=\begin{bmatrix}1&3\\1&-1\\1&1\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}5\\1\\0\end{bmatrix}$, and $\hat{\mathbf{x}}=\begin{bmatrix}1\\1\end{bmatrix}$. Since

$$A\hat{\mathbf{x}}-\mathbf{b}=\begin{bmatrix}1&3\\1&-1\\1&1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}-\begin{bmatrix}5\\1\\0\end{bmatrix}=\begin{bmatrix}4\\0\\2\end{bmatrix}-\begin{bmatrix}5\\1\\0\end{bmatrix}=\begin{bmatrix}-1\\-1\\2\end{bmatrix}$$

the least squares error is $\|A\hat{\mathbf{x}}-\mathbf{b}\|=\sqrt{6}$.

9. (a) Because the columns $\mathbf{a}_1$ and $\mathbf{a}_2$ of A are orthogonal, the method of Example 4 may be used to find $\hat{\mathbf{b}}$, the orthogonal projection of $\mathbf{b}$ onto Col A:

$$\hat{\mathbf{b}}=\frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1+\frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2=\frac{2}{7}\mathbf{a}_1+\frac{1}{7}\mathbf{a}_2=\frac{2}{7}\begin{bmatrix}1\\3\\-2\end{bmatrix}+\frac{1}{7}\begin{bmatrix}5\\1\\4\end{bmatrix}=\begin{bmatrix}1\\1\\0\end{bmatrix}$$

(b) The vector $\hat{\mathbf{x}}$ contains the weights which must be placed on $\mathbf{a}_1$ and $\mathbf{a}_2$ to produce $\hat{\mathbf{b}}$. These weights are easily read from the above equation, so $\hat{\mathbf{x}}=\begin{bmatrix}2/7\\1/7\end{bmatrix}$.

10. (a) Because the columns $\mathbf{a}_1$ and $\mathbf{a}_2$ of A are orthogonal, the method of Example 4 may be used to find $\hat{\mathbf{b}}$, the orthogonal projection of $\mathbf{b}$ onto Col A:

$$\hat{\mathbf{b}}=\frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1+\frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2=3\mathbf{a}_1+\frac{1}{2}\mathbf{a}_2=3\begin{bmatrix}1\\-1\\1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}2\\4\\2\end{bmatrix}=\begin{bmatrix}4\\-1\\4\end{bmatrix}$$

(b) The vector $\hat{\mathbf{x}}$ contains the weights which must be placed on $\mathbf{a}_1$ and $\mathbf{a}_2$ to produce $\hat{\mathbf{b}}$. These weights are easily read from the above equation, so $\hat{\mathbf{x}}=\begin{bmatrix}3\\1/2\end{bmatrix}$.

14. One computes that

$$A\mathbf{u}=\begin{bmatrix}3\\8\\2\end{bmatrix},\ \mathbf{b}-A\mathbf{u}=\begin{bmatrix}2\\-4\\2\end{bmatrix},\ \|\mathbf{b}-A\mathbf{u}\|=\sqrt{24}$$

$$A\mathbf{v}=\begin{bmatrix}7\\2\\8\end{bmatrix},\ \mathbf{b}-A\mathbf{v}=\begin{bmatrix}-2\\2\\-4\end{bmatrix},\ \|\mathbf{b}-A\mathbf{v}\|=\sqrt{24}$$

Since $A\mathbf{u}$ and $A\mathbf{u}$ are equally close to $\mathbf{b}$, and the orthogonal projection is the *unique* closest point in Col A to $\mathbf{b}$, neither $A\mathbf{u}$ nor $A\mathbf{v}$ can be the closest point in Col A to $\mathbf{b}$. Thus neither $\mathbf{u}$ nor $\mathbf{v}$ can be a least-squares solution of $A\mathbf{x}=\mathbf{b}$.

15. The least squares solution satisfies $R\hat{\mathbf{x}}=Q^T\mathbf{b}$. Since $R=\begin{bmatrix}3&5\\0&1\end{bmatrix}$ and $Q^T\mathbf{b}=\begin{bmatrix}7\\-1\end{bmatrix}$, the augmented matrix for the system may be row reduced to find

$$\begin{bmatrix}R & Q^T\mathbf{b}\end{bmatrix}=\begin{bmatrix}3&5&7\\0&1&-1\end{bmatrix}\sim\begin{bmatrix}1&0&4\\0&1&-1\end{bmatrix}$$

and so $\hat{\mathbf{x}}=\begin{bmatrix}4\\-1\end{bmatrix}$ is the least squares solution of $A\mathbf{x}=\mathbf{b}$.

16. The least squares solution satisfies $R\hat{\mathbf{x}}=Q^T\mathbf{b}$. Since $R=\begin{bmatrix}2&3\\0&5\end{bmatrix}$ and $Q^T\mathbf{b}=\begin{bmatrix}17/2\\9/2\end{bmatrix}$, the augmented matrix for the system may be row reduced to find

$$\begin{bmatrix}R & Q^T\mathbf{b}\end{bmatrix}=\begin{bmatrix}2&3&17/2\\0&5&9/2\end{bmatrix}\sim\begin{bmatrix}1&0&2.9\\0&1&.9\end{bmatrix}$$

and so $\hat{\mathbf{x}}=\begin{bmatrix}2.9\\.9\end{bmatrix}$ is the least squares solution of $A\mathbf{x}=\mathbf{b}$.

17. a. True. See the beginning of the section. The distance from $A\mathbf{x}$ to $\mathbf{b}$ is $\|A\mathbf{x}-\mathbf{b}\|$.
    b. True. See the comments about equation (1).
    c. False. The inequality points in the wrong direction. See the definition of a least-squares solution.
    d. True. See Theorem 13.
    e. True. See Theorem 14.

18. a. True. See the paragraph following the definition of a least-squares solution.
    b. False. If $\hat{\mathbf{x}}$ is the least-squares solution, then $A\hat{\mathbf{x}}$ is the point in the column space of A closest to $\mathbf{b}$. See Figure 1 and the paragraph preceding it.
    c. True. See the discussion following equation (1).
    d. False. The formula applies only when the columns of A are linearly independent. See Theorem 14.
    e. False. See the comments after Example 4.
    f. False. See the Numerical Note.

19. **a.** If $A\mathbf{x} = \mathbf{0}$, then $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$. This shows that Nul A is contained in Nul A^TA.

b. If $A^TA\mathbf{x} = \mathbf{0}$, then $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$. So $(A\mathbf{x})^T(A\mathbf{x}) = 0$, which means that $\|A\mathbf{x}\|^2 = 0$, and hence $A\mathbf{x} = \mathbf{0}$. This shows that Nul A^TA is contained in Nul A.

20. Suppose that $A\mathbf{x} = \mathbf{0}$. Then $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$. Since A^TA is invertible, $\mathbf{x}$ must be $\mathbf{0}$. Hence the columns of A are linearly independent.

21. **a.** If A has linearly independent columns, then the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. By Exercise 17, the equation $A^TA\mathbf{x} = \mathbf{0}$ also has only the trivial solution. Since A^TA is a square matrix, it must be invertible by the Invertible Matrix Theorem.

b. Since the n linearly independent columns of A belong to $\mathbb{R}^m$, m could not be less than n.

c. The n linearly independent columns of A form a basis for Col A, so the rank of A is n.

22. Note that A^TA has n columns because A does. Then by the Rank Theorem and Exercise 19,

$$\text{rank } A^TA = n - \dim \text{Nul } A^TA = n - \dim \text{Nul } A = \text{rank } A$$

23. By Theorem 14, $\hat{\mathbf{b}} = A\hat{\mathbf{x}} = A(A^TA)^{-1}A^T\mathbf{b}$. The matrix $A(A^TA)^{-1}A^T$ is sometimes called the *hat-matrix* in statistics.

24. Since in this case $A^TA = I$, the normal equations give $\hat{\mathbf{x}} = A^T\mathbf{b}$.

25. The normal equations are $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$, whose solution is the set of all (x, y) such that $x + y = 3$. The solutions correspond to the points on the line midway between the lines $x + y = 2$ and $x + y = 4$.

26. **[M]** Using .7 as an approximation for $\sqrt{2}/2$, $a_0 = a_2 \approx .353535$ and $a_1 = .5$. Using .707 as an approximation for $\sqrt{2}/2$, $a_0 = a_2 \approx .35355339$, $a_1 = .5$.

6.6 SOLUTIONS

Notes: This section is a valuable reference for any person who works with data that requires statistical analysis. Many graduate fields require such work. Science students in particular will benefit from Example 1. The general linear model and the subsequent examples are aimed at students who may take a multivariate statistics course. That may include more students than one might expect.

1. The design matrix X and the observation vector $\mathbf{y}$ are

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix},$$

and one can compute

$$X^TX = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}, X^T\mathbf{y} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}, \hat{\beta} = (X^TX)^{-1}X^T\mathbf{y} = \begin{bmatrix} .9 \\ .4 \end{bmatrix}$$

The least-squares line $y = \beta_0 + \beta_1 x$ is thus $y = .9 + .4x$.

Since $\det A = 15s^2 + 45 = 15(s^2+3) \neq 0$ for all values of s, the system will have a unique solution for all values of s. For such a system, the solution will be

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{15s+10}{15(s^2+3)} = \frac{3s+2}{3(s^2+3)},\ x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{6s-27}{15(s^2+3)} = \frac{2s-9}{5(s^2+3)}.$$

9. The system is equivalent to $A\mathbf{x}=\mathbf{b}$, where $A = \begin{bmatrix} s & -2s \\ 3 & 6s \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$. Compute

$$A_1(\mathbf{b}) = \begin{bmatrix} -1 & -2s \\ 4 & 6s \end{bmatrix},\ A_2(\mathbf{b}) = \begin{bmatrix} s & -1 \\ 3 & 4 \end{bmatrix},\ \det A_1(\mathbf{b}) = 2s,\ \det A_2(\mathbf{b}) = 4s+3.$$

Since $\det A = 6s^2 + 6s = 6s(s+1) = 0$ for $s = 0, -1$, the system will have a unique solution when $s \neq 0, -1$. For such a system, the solution will be

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{2s}{6s(s+1)} = \frac{1}{3(s+1)},\ x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{4s+3}{6s(s+1)}.$$

10. The system is equivalent to $A\mathbf{x}=\mathbf{b}$, where $A = \begin{bmatrix} 2s & 1 \\ 3s & 6s \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Compute

$$A_1(\mathbf{b}) = \begin{bmatrix} 1 & 1 \\ 2 & 6s \end{bmatrix},\ A_2(\mathbf{b}) = \begin{bmatrix} 2s & 1 \\ 3s & 2 \end{bmatrix},\ \det A_1(\mathbf{b}) = 6s-2,\ \det A_2(\mathbf{b}) = s.$$

Since $\det A = 12s^2 - 3s = 3s(4s-1) = 0$ for $s = 0, 1/4$, the system will have a unique solution when $s \neq 0, 1/4$. For such a system, the solution will be

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{6s-2}{3s(4s-1)},\ x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{s}{3s(4s-1)} = \frac{1}{3(4s-1)}.$$

11. Since $\det A = 3$ and the cofactors of the given matrix are

$$C_{11} = \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} = 0,\quad C_{12} = -\begin{vmatrix} 3 & 0 \\ -1 & 1 \end{vmatrix} = -3,\quad C_{13} = \begin{vmatrix} 3 & 0 \\ -1 & 1 \end{vmatrix} = 3,$$

$$C_{21} = -\begin{vmatrix} -2 & -1 \\ 1 & 1 \end{vmatrix} = 1,\quad C_{22} = \begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix} = -1,\quad C_{23} = -\begin{vmatrix} 0 & -2 \\ -1 & 1 \end{vmatrix} = 2,$$

$$C_{31} = \begin{vmatrix} -2 & -1 \\ 0 & 0 \end{vmatrix} = 0,\quad C_{32} = -\begin{vmatrix} 0 & -1 \\ 3 & 0 \end{vmatrix} = -3,\quad C_{33} = \begin{vmatrix} 0 & -2 \\ 3 & 0 \end{vmatrix} = 6,$$

$$\operatorname{adj} A = \begin{bmatrix} 0 & 1 & 0 \\ -3 & -1 & -3 \\ 3 & 2 & 6 \end{bmatrix} \text{ and } A^{-1} = \frac{1}{\det A}\operatorname{adj} A = \begin{bmatrix} 0 & 1/3 & 0 \\ -1 & -1/3 & -1 \\ 1 & 2/3 & 2 \end{bmatrix}.$$

12. Since $\det A = 5$ and the cofactors of the given matrix are

$$C_{11} = \begin{vmatrix} -2 & 1 \\ 1 & 0 \end{vmatrix} = -1,\quad C_{12} = -\begin{vmatrix} 2 & 1 \\ 0 & 0 \end{vmatrix} = 0,\quad C_{13} = \begin{vmatrix} 2 & -2 \\ 0 & 1 \end{vmatrix} = 2,$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 1 & 0 \end{vmatrix} = 3,\quad C_{22} = \begin{vmatrix} 1 & 3 \\ 0 & 0 \end{vmatrix} = 0,\quad C_{23} = -\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = -1,$$

$$C_{31} = \begin{vmatrix} 1 & 3 \end{vmatrix} \ldots,\quad C_{32} = -\begin{vmatrix} 1 & 3 \end{vmatrix} = 5,\quad C_{33} = \begin{vmatrix} 1 & 1 \end{vmatrix} = -4,$$