Thus an orthogonal basis for $W$ is $\left\{ \begin{bmatrix} -10 \\ 2 \\ -6 \\ 16 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ -3 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ 6 \\ 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 0 \\ 0 \\ -5 \end{bmatrix} \right\}$.

**25.** **[M]** The columns of $Q$ will be normalized versions of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ found in Exercise 24. Thus

$$Q = \begin{bmatrix} -1/2 & 1/2 & 1/\sqrt{3} & 0 \\ 1/10 & 1/2 & 0 & 1/\sqrt{2} \\ -3/10 & -1/2 & 1/\sqrt{3} & 0 \\ 4/5 & 0 & 1/\sqrt{3} & 0 \\ 1/10 & 1/2 & 0 & -1/\sqrt{2} \end{bmatrix}, R = Q^T A = \begin{bmatrix} 20 & -20 & -10 & 10 \\ 0 & 6 & -8 & -6 \\ 0 & 0 & 6\sqrt{3} & -3\sqrt{3} \\ 0 & 0 & 0 & 5\sqrt{2} \end{bmatrix}$$

**26.** **[M]** In MATLAB, when $A$ has $n$ columns, suitable commands are

```
Q = A(:,1)/norm(A(:,1))
    %  The first column of Q
    for j=2: n
       v=A(:,j) – Q*(Q'*A(:,j))
       Q(:,j)=v/norm(v)
       % Add a new column to Q
    end
```

## 6.5 SOLUTIONS

**Notes:** This is a core section – the basic geometric principles in this section provide the foundation for all the applications in Sections 6.6–6.8. Yet this section need not take a full day. Each example provides a stopping place. Theorem 13 and Example 1 are all that is needed for Section 6.6. Theorem 15, however, gives an illustration of why the QR factorization is important. Example 4 is related to Exercise 17 in Section 6.6.

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

$$A^T A = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 6 & -11 \\ -11 & 22 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 11 \end{bmatrix}$$

**a.** The normal equations are $(A^T A)\mathbf{x} = A^T \mathbf{b}$: $\begin{bmatrix} 6 & -11 \\ -11 & 22 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 11 \end{bmatrix}$.

**b.** Compute

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} 6 & -11 \\ -11 & 22 \end{bmatrix}^{-1} \begin{bmatrix} -4 \\ 11 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 22 & 11 \\ 11 & 6 \end{bmatrix} \begin{bmatrix} -4 \\ 11 \end{bmatrix}$$
$$= \frac{1}{11} \begin{bmatrix} 33 \\ 22 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

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

$$A^T A = \begin{bmatrix} 2 & -2 & 2 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 12 & 8 \\ 8 & 10 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 2 & -2 & 2 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix} = \begin{bmatrix} -24 \\ -2 \end{bmatrix}$$

**a.** The normal equations are $(A^T A)\mathbf{x} = A^T \mathbf{b}$: $\begin{bmatrix} 12 & 8 \\ 8 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -24 \\ -2 \end{bmatrix}$.

**b.** Compute

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} 12 & 8 \\ 8 & 10 \end{bmatrix}^{-1} \begin{bmatrix} -24 \\ -2 \end{bmatrix} = \frac{1}{56} \begin{bmatrix} 10 & -8 \\ -8 & 12 \end{bmatrix} \begin{bmatrix} -24 \\ -2 \end{bmatrix}$$
$$= \frac{1}{56} \begin{bmatrix} -224 \\ 168 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

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

$$A^T A = \begin{bmatrix} 1 & -1 & 0 & 2 \\ -2 & 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & -1 & 0 & 2 \\ -2 & 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

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

**b.** Compute

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

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

**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

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

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

**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 $\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$.

**11.** (a) Because the columns $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ 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 $A$:

$$\hat{\mathbf{b}} = \frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2 + \frac{\mathbf{b}\cdot\mathbf{a}_3}{\mathbf{a}_3\cdot\mathbf{a}_3}\mathbf{a}_3 = \frac{2}{3}\mathbf{a}_1 + 0\mathbf{a}_2 + \frac{1}{3}\mathbf{a}_3$$

$$= \frac{2}{3}\begin{bmatrix}4\\1\\6\\1\end{bmatrix} + 0\begin{bmatrix}0\\-5\\1\\-1\end{bmatrix} + \frac{1}{3}\begin{bmatrix}1\\1\\0\\-5\end{bmatrix} = \begin{bmatrix}3\\1\\4\\-1\end{bmatrix}$$

(b) The vector $\hat{\mathbf{x}}$ contains the weights which must be placed on $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ to produce $\hat{\mathbf{b}}$. These weights are easily read from the above equation, so $\hat{\mathbf{x}} = \begin{bmatrix}2/3\\0\\1/3\end{bmatrix}$.

**12.** (a) Because the columns $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ 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 $A$:

$$\hat{\mathbf{b}} = \frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2 + \frac{\mathbf{b}\cdot\mathbf{a}_3}{\mathbf{a}_3\cdot\mathbf{a}_3}\mathbf{a}_3 = \frac{1}{3}\mathbf{a}_1 + \frac{14}{3}\mathbf{a}_2 + \left(-\frac{5}{3}\right)\mathbf{a}_3$$

$$= \frac{1}{3}\begin{bmatrix}1\\1\\0\\-1\end{bmatrix} + \frac{14}{3}\begin{bmatrix}1\\0\\1\\1\end{bmatrix} - \frac{5}{3}\begin{bmatrix}0\\-1\\1\\-1\end{bmatrix} = \begin{bmatrix}5\\2\\3\\6\end{bmatrix}$$

(b) The vector $\hat{\mathbf{x}}$ contains the weights which must be placed on $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ to produce $\hat{\mathbf{b}}$. These weights are easily read from the above equation, so $\hat{\mathbf{x}} = \begin{bmatrix}1/3\\14/3\\-5/3\end{bmatrix}$.

**13.** One computes that

$$A\mathbf{u} = \begin{bmatrix}11\\-11\\11\end{bmatrix}, \mathbf{b} - A\mathbf{u} = \begin{bmatrix}0\\2\\-6\end{bmatrix}, \|\mathbf{b} - A\mathbf{u}\| = \sqrt{40}$$

$$A\mathbf{v} = \begin{bmatrix}7\\-12\\7\end{bmatrix}, \mathbf{b} - A\mathbf{v} = \begin{bmatrix}4\\3\\-2\end{bmatrix}, \|\mathbf{b} - A\mathbf{v}\| = \sqrt{29}$$

Since $A\mathbf{v}$ is closer to $\mathbf{b}$ than $A\mathbf{u}$ is, $A\mathbf{u}$ is not the closest point in Col $A$ to $\mathbf{b}$. Thus $\mathbf{u}$ cannot be a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

**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 19, 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 $\sqrt{2}/2$, $a_0 = a_2 \approx .35355339$, $a_1 = .5$.