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,

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

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 $y = .9 + .4x$.

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

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix},$$

and one can compute

$$X^T X = \begin{bmatrix} 4 & 12 \\ 12 & 46 \end{bmatrix}, X^T \mathbf{y} = \begin{bmatrix} 6 \\ 25 \end{bmatrix}, \hat{\beta} = (X^T X)^{-1} X^T \mathbf{y} = \begin{bmatrix} -.6 \\ .7 \end{bmatrix}$$

The least-squares line $y = \beta_0 + \beta_1 x$ is thus $y = -.6 + .7x$.

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

$$X = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 4 \end{bmatrix},$$

and one can compute

$$X^T X = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}, X^T \mathbf{y} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}, \hat{\beta} = (X^T X)^{-1} X^T \mathbf{y} = \begin{bmatrix} 1.1 \\ 1.3 \end{bmatrix}$$

The least-squares line $y = \beta_0 + \beta_1 x$ is thus $y = 1.1 + 1.3x$.

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

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix},$$

and one can compute

$$X^T X = \begin{bmatrix} 4 & 16 \\ 16 & 74 \end{bmatrix}, X^T \mathbf{y} = \begin{bmatrix} 6 \\ 17 \end{bmatrix}, \hat{\beta} = (X^T X)^{-1} X^T \mathbf{y} = \begin{bmatrix} 4.3 \\ -.7 \end{bmatrix}$$

The least-squares line $y = \beta_0 + \beta_1 x$ is thus $y = 4.3 - .7x$.

5. If two data points have different x-coordinates, then the two columns of the design matrix X cannot be multiples of each other and hence are linearly independent. By Theorem 14 in Section 6.5, the normal equations have a unique solution.

6. If the columns of X were linearly dependent, then the same dependence relation would hold for the vectors in $\mathbb{R}^3$ formed from the top three entries in each column. That is, the columns of the matrix

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}$$ would also be linearly dependent, and so this matrix (called a Vandermonde matrix)

would be noninvertible. Note that the determinant of this matrix is $(x_2 - x_1)(x_3 - x_1)(x_3 - x_2) \neq 0$ since x_1, x_2, and x_3 are distinct. Thus this matrix is invertible, which means that the columns of X are in fact linearly independent. By Theorem 14 in Section 6.5, the normal equations have a unique solution.

7. **a**. The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \\ 5 & 25 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 1.8 \\ 2.7 \\ 3.4 \\ 3.8 \\ 3.9 \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \text{ and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$$

b. **[M]** One computes that (to two decimal places) $\hat{\beta} = \begin{bmatrix} 1.76 \\ -.20 \end{bmatrix}$, so the desired least-squares equation is $y = 1.76x - .20x^2$.

8. **a**. The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} x_1 & x_1^2 & x_1^3 \\ \vdots & \vdots & \vdots \\ x_n & x_n^2 & x_n^3 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \text{ and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

b. **[M]** For the given data,

$$X = \begin{bmatrix} 4 & 16 & 64 \\ 6 & 36 & 216 \\ 8 & 64 & 512 \\ 10 & 100 & 1000 \\ 12 & 144 & 1728 \\ 14 & 196 & 2744 \\ 16 & 256 & 4096 \\ 18 & 324 & 5832 \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} 1.58 \\ 2.08 \\ 2.5 \\ 2.8 \\ 3.1 \\ 3.4 \\ 3.8 \\ 4.32 \end{bmatrix}$$

so $\hat{\beta} = (X^TX)^{-1}X^T\mathbf{y} = \begin{bmatrix} .5132 \\ -.03348 \\ .001016 \end{bmatrix}$, and the least-squares curve is $y = .5132x - .03348x^2 + .001016x^3$.

9. The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} \cos 1 & \sin 1 \\ \cos 2 & \sin 2 \\ \cos 3 & \sin 3 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 7.9 \\ 5.4 \\ -.9 \end{bmatrix}, \beta = \begin{bmatrix} A \\ B \end{bmatrix}, \text{ and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}$$

10. **a**. The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} e^{-.02(10)} & e^{-.07(10)} \\ e^{-.02(11)} & e^{-.07(11)} \\ e^{-.02(12)} & e^{-.07(12)} \\ e^{-.02(14)} & e^{-.07(14)} \\ e^{-.02(15)} & e^{-.07(15)} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 21.34 \\ 20.68 \\ 20.05 \\ 18.87 \\ 18.30 \end{bmatrix}, \beta = \begin{bmatrix} M_A \\ M_B \end{bmatrix}, \text{ and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix},$$

b. [M] One computes that (to two decimal places) $\hat{\beta} = \begin{bmatrix} 19.94 \\ 10.10 \end{bmatrix}$, so the desired least-squares equation is $y = 19.94e^{-.02t} + 10.10e^{-.07t}$.

11. **[M]** The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} 1 & 3\cos .88 \\ 1 & 2.3\cos 1.1 \\ 1 & 1.65\cos 1.42 \\ 1 & 1.25\cos 1.77 \\ 1 & 1.01\cos 2.14 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 3 \\ 2.3 \\ 1.65 \\ 1.25 \\ 1.01 \end{bmatrix}, \beta = \begin{bmatrix} \beta \\ e \end{bmatrix}, \text{and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$$

One computes that (to two decimal places) $\hat{\beta} = \begin{bmatrix} 1.45 \\ .811 \end{bmatrix}$. Since $e = .811 < 1$ the orbit is an ellipse. The equation $r = \beta / (1 - e \cos \vartheta)$ produces $r = 1.33$ when $\vartheta = 4.6$.

12. **[M]** The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} 1 & 3.78 \\ 1 & 4.11 \\ 1 & 4.41 \\ 1 & 4.73 \\ 1 & 4.88 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 91 \\ 98 \\ 103 \\ 110 \\ 112 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \text{and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$$

One computes that (to two decimal places) $\hat{\beta} = \begin{bmatrix} 18.56 \\ 19.24 \end{bmatrix}$, so the desired least-squares equation is $p = 18.56 + 19.24 \ln w$. When $w = 100$, $p \approx 107$ millimeters of mercury.

13. **[M]**

a. The model that produces the correct least-squares fit is $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, where

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 2^2 & 2^3 \\ 1 & 3 & 3^2 & 3^3 \\ 1 & 4 & 4^2 & 4^3 \\ 1 & 5 & 5^2 & 5^3 \\ 1 & 6 & 6^2 & 6^3 \\ 1 & 7 & 7^2 & 7^3 \\ 1 & 8 & 8^2 & 8^3 \\ 1 & 9 & 9^2 & 9^3 \\ 1 & 10 & 10^2 & 10^3 \\ 1 & 11 & 11^2 & 11^3 \\ 1 & 12 & 12^2 & 12^3 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0 \\ 8.8 \\ 29.9 \\ 62.0 \\ 104.7 \\ 159.1 \\ 222.0 \\ 294.5 \\ 380.4 \\ 471.1 \\ 571.7 \\ 686.8 \\ 809.2 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \text{and } \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \\ \epsilon_{12} \end{bmatrix}$$

One computes that (to four decimal places) $\hat{\beta} = \begin{bmatrix} -.8558 \\ 4.7025 \\ 5.5554 \\ -.0274 \end{bmatrix}$, so the desired least-squares polynomial is

$y(t) = -.8558 + 4.7025t + 5.5554t^2 - .0274t^3.$

b. The velocity $v(t)$ is the derivative of the position function $y(t)$, so $v(t) = 4.7025 + 11.1108t - .0822t^2$, and $v(4.5) = 53.0$ ft/sec.

14. Write the design matrix as $[\mathbf{1} \quad \mathbf{x}]$. Since the residual vector $\boldsymbol{\epsilon} = \mathbf{y} - X\hat{\beta}$ is orthogonal to Col X,

$$0 = \mathbf{1} \cdot \boldsymbol{\epsilon} = \mathbf{1} \cdot (\mathbf{y} - X\hat{\beta}) = \mathbf{1}^T \mathbf{y} - (\mathbf{1}^T X)\hat{\beta}$$

$$= (y_1 + \ldots + y_n) - \begin{bmatrix} n & \sum x \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \sum y - n\hat{\beta}_0 - \hat{\beta}_1 \sum x = n\overline{y} - n\hat{\beta}_0 - n\hat{\beta}_1 \overline{x}$$

This equation may be solved for $\overline{y}$ to find $\overline{y} = \hat{\beta}_0 + \hat{\beta}_1 \overline{x}$.

15. From equation (1) on page 420,

$$X^T X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x \\ \sum x & (\sum x)^2 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum xy \end{bmatrix}$$

The equations (7) in the text follow immediately from the normal equations $X^T X\beta = X^T \mathbf{y}$.

16. The determinant of the coefficient matrix of the equations in (7) is $n\sum x^2 - (\sum x)^2$. Using the 2×2 formula for the inverse of the coefficient matrix,

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \frac{1}{n\sum x^2 - (\sum x)^2} \begin{bmatrix} \sum x^2 & -\sum x \\ -\sum x & n \end{bmatrix} \begin{bmatrix} \sum y \\ \sum xy \end{bmatrix}$$

Hence

$$\hat{\beta}_0 = \frac{(\sum x^2)(\sum y) - (\sum x)(\sum xy)}{n\sum x^2 - (\sum x)^2}, \quad \hat{\beta}_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

Note: A simple algebraic calculation shows that $\sum y - (\sum x)\hat{\beta}_1 = n\hat{\beta}_0$, which provides a simple formula for $\hat{\beta}_0$ once $\hat{\beta}_1$ is known.

17. **a**. The mean of the data in Example 1 is $\overline{x} = 5.5$, so the data in mean-deviation form are (–3.5, 1), (–.5, 2), (1.5, 3), (2.5, 3), and the associated design matrix is $X = \begin{bmatrix} 1 & -3.5 \\ 1 & -.5 \\ 1 & 1.5 \\ 1 & 2.5 \end{bmatrix}$. The columns of X are orthogonal because the entries in the second column sum to 0.

b. The normal equations are $X^TX\beta = X^T\mathbf{y}$, or $\begin{bmatrix} 4 & 0 \\ 0 & 21 \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 7.5 \end{bmatrix}$. One computes that $\hat{\beta} = \begin{bmatrix} 9/4 \\ 5/14 \end{bmatrix}$, so the desired least-squares line is $y = (9/4) + (5/14)x^* = (9/4) + (5/14)(x - 5.5)$.

18. Since

$$X^TX = \begin{bmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \end{bmatrix}\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x \\ \sum x & (\sum x)^2 \end{bmatrix}$$

X^TX is a diagonal matrix when $\sum x = 0$.

19. The residual vector $\boldsymbol{\epsilon} = \mathbf{y} - X\hat{\beta}$ is orthogonal to Col X, while $\hat{\mathbf{y}} = X\hat{\beta}$ is in Col X. Since $\boldsymbol{\epsilon}$ and $\hat{\mathbf{y}}$ are thus orthogonal, apply the Pythagorean Theorem to these vectors to obtain

$$\text{SS(T)} = \|\mathbf{y}\|^2 = \|\hat{\mathbf{y}} + \boldsymbol{\epsilon}\|^2 = \|\hat{\mathbf{y}}\|^2 + \|\boldsymbol{\epsilon}\|^2 = \|X\hat{\beta}\|^2 + \|\mathbf{y} - X\hat{\beta}\|^2 = \text{SS(R)} + \text{SS(E)}$$

20. Since $\hat{\beta}$ satisfies the normal equations, $X^TX\hat{\beta} = X^T\mathbf{y}$, and

$$\|X\hat{\beta}\|^2 = (X\hat{\beta})^T(X\hat{\beta}) = \hat{\beta}^TX^TX\hat{\beta} = \hat{\beta}^TX^T\mathbf{y}$$

Since $\|X\hat{\beta}\|^2 = \text{SS(R)}$ and $\mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|^2 = \text{SS(T)}$, Exercise 19 shows that

$$\text{SS(E)} = \text{SS(T)} - \text{SS(R)} = \mathbf{y}^T\mathbf{y} - \hat{\beta}^TX^T\mathbf{y}$$

6.7 SOLUTIONS

Notes: The three types of inner products described here (in Examples 1, 2, and 7) are matched by examples in Section 6.8. It is possible to spend just one day on selected portions of both sections. Example 1 matches the weighted least squares in Section 6.8. Examples 2–6 are applied to trend analysis in Seciton 6.8. This material is aimed at students who have not had much calculus or who intend to take more than one course in statistics.

For students who have seen some calculus, Example 7 is needed to develop the Fourier series in Section 6.8. Example 8 is used to motivate the inner product on $C[a, b]$. The Cauchy-Schwarz and triangle inequalities are not used here, but they should be part of the training of every mathematics student.

1. The inner product is $\langle x, y\rangle = 4x_1y_1 + 5x_2y_2$. Let $\mathbf{x}$ = (1, 1), $\mathbf{y}$ = (5, –1).

a. Since $\|\mathbf{x}\|^2 = \langle x, x\rangle = 9$, $\|\mathbf{x}\| = 3$. Since $\|\mathbf{y}\|^2 = \langle y, y\rangle = 105$, $\|\mathbf{x}\| = \sqrt{105}$. Finally, $|\langle x, y\rangle|^2 = 15^2 = 225$.

b. A vector $\mathbf{z}$ is orthogonal to $\mathbf{y}$ if and only if $\langle x, y\rangle = 0$, that is, $20z_1 - 5z_2 = 0$, or $4z_1 = z_2$. Thus all multiples of $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ are orthogonal to $\mathbf{y}$.

2. The inner product is $\langle x, y\rangle = 4x_1y_1 + 5x_2y_2$. Let $\mathbf{x}$ = (3, –2), $\mathbf{y}$ = (–2, 1). Compute that $\|\mathbf{x}\|^2 = \langle x, x\rangle = 56$, $\|\mathbf{y}\|^2 = \langle y, y\rangle = 21$, $\|\mathbf{x}\|^2\|\mathbf{y}\|^2 = 56 \cdot 21 = 1176$, $\langle x, y\rangle = –34$, and $|\langle x, y\rangle|^2 = 1156$. Thus $|\langle x, y\rangle|^2 \le \|\mathbf{x}\|^2\|\mathbf{y}\|^2$, as the Cauchy-Schwarz inequality predicts.

3. The inner product is $\langle p, q\rangle = p(–1)q(–1) + p(0)q(0) + p(1)q(1)$, so $\langle 4 + t, 5 - 4t^2\rangle = 3(1) + 4(5) + 5(1) = 28$.

4. The inner product is $\langle p, q\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$, so $\langle 3t-t^2, 3+2t^2\rangle =$ $(-4)(5)+0(3)+2(5)=-10$.

5. The inner product is $\langle p, q\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$, so $\langle p,q\rangle = \langle 4+t, 4+t\rangle = 3^2+4^2+5^2 = 50$ and $\| p \| = \sqrt{\langle p,p\rangle} = \sqrt{50} = 5\sqrt{2}$. Likewise $\langle q,q\rangle = \langle 5-4t^2, 5-4t^2\rangle = 1^2+5^2+1^2 = 27$ and $\| q \| = \sqrt{\langle q,q\rangle} = \sqrt{27} = 3\sqrt{3}$.

6. The inner product is $\langle p, q\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$, so $\langle p,p\rangle = \langle 3t-t^2, 3t-t^2\rangle =$ $(-4)^2+0^2+2^2 = 20$ and $\| p \| = \sqrt{\langle p,p\rangle} = \sqrt{20} = 2\sqrt{5}$. Likewise $\langle q,q\rangle = \langle 3+2t^2, 3+2t^2\rangle =$ $5^2+3^2+5^2 = 59$ and $\| q \| = \sqrt{\langle q,q\rangle} = \sqrt{59}$.

7. The orthogonal projection $\hat{q}$ of q onto the subspace spanned by p is

$$\hat{q} = \frac{\langle q,p\rangle}{\langle p,p\rangle} p = \frac{28}{50}(4+t) = \frac{56}{25} + \frac{14}{25}t$$

8. The orthogonal projection $\hat{q}$ of q onto the subspace spanned by p is

$$\hat{q} = \frac{\langle q,p\rangle}{\langle p,p\rangle} p = -\frac{10}{20}(3t-t^2) = -\frac{3}{2}t + \frac{1}{2}t^2$$

9. The inner product is $\langle p, q\rangle = p(-3)q(-3) + p(-1)q(-1) + p(1)q(1) + p(3)q(3)$.

a. The orthogonal projection $\hat{p}_2$ of p_2 onto the subspace spanned by p_0 and p_1 is

$$\hat{p}_2 = \frac{\langle p_2,p_0\rangle}{\langle p_0,p_0\rangle} p_0 + \frac{\langle p_2,p_1\rangle}{\langle p_1,p_1\rangle} p_1 = \frac{20}{4}(1) + \frac{0}{20}t = 5$$

b. The vector $q = p_3 - \hat{p}_2 = t^2 - 5$ will be orthogonal to both p_0 and p_1 and $\{p_0, p_1, q\}$ will be an orthogonal basis for $\text{Span}\{p_0, p_1, p_2\}$. The vector of values for q at (–3, –1, 1, 3) is (4, –4, –4, 4), so scaling by 1/4 yields the new vector $q = (1/4)(t^2-5)$.

10. The best approximation to $p = t^3$ by vectors in $W = \text{Span}\{p_0, p_1, q\}$ will be

$$\hat{p} = \text{proj}_W p = \frac{\langle p,p_0\rangle}{\langle p_0,p_0\rangle} p_0 + \frac{\langle p,p_1\rangle}{\langle p_1,p_1\rangle} p_1 + \frac{\langle p,q\rangle}{\langle q,q\rangle} q = \frac{0}{4}(1) + \frac{164}{20}(t) + \frac{0}{4}\left(\frac{t^2-5}{4}\right) = \frac{41}{5}t$$

11. The orthogonal projection of $p = t^3$ onto $W = \text{Span}\{p_0, p_1, p_2\}$ will be

$$\hat{p} = \text{proj}_W p = \frac{\langle p,p_0\rangle}{\langle p_0,p_0\rangle} p_0 + \frac{\langle p,p_1\rangle}{\langle p_1,p_1\rangle} p_1 + \frac{\langle p,p_2\rangle}{\langle p_2,p_2\rangle} p_2 = \frac{0}{5}(1) + \frac{34}{10}(t) + \frac{0}{14}(t^2-2) = \frac{17}{5}t$$

12. Let $W = \text{Span}\{p_0, p_1, p_2\}$. The vector $p_3 = p - \text{proj}_W p = t^3 - (17/5)t$ will make $\{p_0, p_1, p_2, p_3\}$ an orthogonal basis for the subspace $\mathbb{P}_3$ of $\mathbb{P}_4$. The vector of values for p_3 at (–2, –1, 0, 1, 2) is (–6/5, 12/5, 0, –12/5, 6/5), so scaling by 5/6 yields the new vector $p_3 = (5/6)(t^3 - (17/5)t) =$ $(5/6)t^3 - (17/6)t$.

13. Suppose that A is invertible and that $\langle \mathbf{u}, \mathbf{v}\rangle = (A\mathbf{u}) \cdot (A\mathbf{v})$ for $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. Check each axiom in the definition on page 428, using the properties of the dot product.

i. $\langle \mathbf{u}, \mathbf{v}\rangle = (A\mathbf{u}) \cdot (A\mathbf{v}) = (A\mathbf{v}) \cdot (A\mathbf{u}) = \langle \mathbf{v}, \mathbf{u}\rangle$

ii. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w}\rangle = (A(\mathbf{u} + \mathbf{v})) \cdot (A\mathbf{w}) = (A\mathbf{u} + A\mathbf{v}) \cdot (A\mathbf{w}) = (A\mathbf{u}) \cdot (A\mathbf{w}) + (A\mathbf{v}) \cdot (A\mathbf{w}) = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$

iii. $\langle c\mathbf{u}, \mathbf{v}\rangle = (A(c\mathbf{u})) \cdot (A\mathbf{v}) = (c(A\mathbf{u})) \cdot (A\mathbf{v}) = c((A\mathbf{u}) \cdot (A\mathbf{v})) = c\langle \mathbf{u}, \mathbf{v}\rangle$

iv. $\langle c\mathbf{u},\mathbf{u}\rangle = (A\mathbf{u})\cdot(A\mathbf{u}) = \| A\mathbf{u}\|^2 \ge 0,$ and this quantity is zero if and only if the vector $A\mathbf{u}$ is $\mathbf{0}$. But $A\mathbf{u} = \mathbf{0}$ if and only $\mathbf{u} = \mathbf{0}$ because A is invertible.

14. Suppose that T is a one-to-one linear transformation from a vector space V into $\mathbb{R}^n$ and that $\langle \mathbf{u}, \mathbf{v}\rangle = T(\mathbf{u}) \cdot T(\mathbf{v})$ for $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. Check each axiom in the definition on page 428, using the properties of the dot product and T. The linearity of T is used often in the following.

i. $\langle \mathbf{u}, \mathbf{v}\rangle = T(\mathbf{u}) \cdot T(\mathbf{v}) = T(\mathbf{v}) \cdot T(\mathbf{u}) = \langle \mathbf{v}, \mathbf{u}\rangle$

ii. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w}\rangle = T(\mathbf{u} + \mathbf{v}) \cdot T(\mathbf{w}) = (T(\mathbf{u}) + T(\mathbf{v})) \cdot T(\mathbf{w}) = T(\mathbf{u}) \cdot T(\mathbf{w}) + T(\mathbf{v}) \cdot T(\mathbf{w}) = \langle \mathbf{u}, w\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$

iii. $\langle c\mathbf{u}, \mathbf{v}\rangle = T(c\mathbf{u}) \cdot T(\mathbf{v}) = (cT(\mathbf{u})) \cdot T(\mathbf{v}) = c(T(\mathbf{u}) \cdot T(\mathbf{v})) = c\langle \mathbf{u}, \mathbf{v}\rangle$

iv. $\langle \mathbf{u},\mathbf{u}\rangle = T(\mathbf{u})\cdot T(\mathbf{u}) = \|T(\mathbf{u})\|^2 \ge 0,$ and this quantity is zero if and only if $\mathbf{u} = \mathbf{0}$ since T is a one-to-one transformation.

15. Using Axioms 1 and 3, $\langle \mathbf{u}, c\mathbf{v}\rangle = \langle c\mathbf{v}, \mathbf{u}\rangle = c\langle \mathbf{v}, \mathbf{u}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle$.

16. Using Axioms 1, 2 and 3,

$$\begin{aligned}\|\mathbf{u}-\mathbf{v}\|^2 &= \langle \mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v}\rangle = \langle \mathbf{u},\mathbf{u}-\mathbf{v}\rangle - \langle \mathbf{v},\mathbf{u}-\mathbf{v}\rangle \\ &= \langle \mathbf{u},\mathbf{u}\rangle - \langle \mathbf{u},\mathbf{v}\rangle - \langle \mathbf{v},\mathbf{u}\rangle + \langle \mathbf{v},\mathbf{v}\rangle = \langle \mathbf{u},\mathbf{u}\rangle - 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 - 2\langle \mathbf{u},\mathbf{v}\rangle + \|\mathbf{v}\|^2\end{aligned}$$

Since $\{\mathbf{u}, \mathbf{v}\}$ is orthonormal, $\|\mathbf{u}\|^2 = \|\mathbf{v}\|^2 = 1$ and $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. So $\|\mathbf{u}-\mathbf{v}\|^2 = 2$.

17. Following the method in Exercise 16,

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u},\mathbf{u}+\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{u}+\mathbf{v}\rangle \\ &= \langle \mathbf{u},\mathbf{u}\rangle + \langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{u}\rangle + \langle \mathbf{v},\mathbf{v}\rangle = \langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u},\mathbf{v}\rangle + \|\mathbf{v}\|^2\end{aligned}$$

Subtracting these results, one finds that $\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 = 4\langle \mathbf{u},\mathbf{v}\rangle,$ and dividing by 4 gives the desired identity.

18. In Exercises 16 and 17, it has been shown that $\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u},\mathbf{v}\rangle + \|\mathbf{v}\|^2$ and $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\langle \mathbf{u},\mathbf{v}\rangle + \|\mathbf{v}\|^2$. Adding these two results gives $\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$.

19. let $\mathbf{u} = \begin{bmatrix}\sqrt{a}\\ \sqrt{b}\end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix}\sqrt{b}\\ \sqrt{a}\end{bmatrix}$. Then $\|\mathbf{u}\|^2 = a+b,$ $\|\mathbf{v}\|^2 = a+b,$ and $\langle \mathbf{u},\mathbf{v}\rangle = 2\sqrt{ab}.$ Since a and b are nonnegative, $\|\mathbf{u}\| = \sqrt{a+b},$ $\|\mathbf{v}\| = \sqrt{a+b}.$ Plugging these values into the Cauchy-Schwarz inequality gives

$$2\sqrt{ab} = |\langle \mathbf{u},\mathbf{v}\rangle| \le \|\mathbf{u}\|\|\mathbf{v}\| = \sqrt{a+b}\sqrt{a+b} = a+b$$

Dividing both sides of this equation by 2 gives the desired inequality.

20. The Cauchy-Schwarz inequality may be altered by dividing both sides of the inequality by 2 and then squaring both sides of the inequality. The result is

$$\left(\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{2}\right)^2 \le \frac{\|\mathbf{u}\|^2\|\mathbf{v}\|^2}{4}$$

Now let $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then $\|\mathbf{u}\|^2 = a^2 + b^2$, $\|\mathbf{v}\|^2 = 2$, and $\langle \mathbf{u}, \mathbf{v}\rangle = a + b$. Plugging these values into the inequality above yields the desired inequality.

21. The inner product is $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$. Let $f(t) = 1 - 3t^2$, $g(t) = t - t^3$. Then

$$\langle f, g\rangle = \int_0^1 (1-3t^2)(t-t^3)\, dt = \int_0^1 3t^5 - 4t^3 + t\, dt = 0$$

22. The inner product is $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$. Let $f(t) = 5t - 3$, $g(t) = t^3 - t^2$. Then

$$\langle f, g\rangle = \int_0^1 (5t-3)(t^3-t^2)dt = \int_0^1 5t^4 - 8t^3 + 3t^2 dt = 0$$

23. The inner product is $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$, so $\langle f, f\rangle = \int_0^1 (1-3t^2)^2 dt = \int_0^1 9t^4 - 6t^2 + 1\, dt = 4/5$, and $\| f \| = \sqrt{\langle f, f\rangle} = 2/\sqrt{5}$.

24. The inner product is $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$, so $\langle g, g\rangle = \int_0^1 (t^3-t^2)^2 dt = \int_0^1 t^6 - 2t^5 + t^4 dt = 1/105$, and $\| g \| = \sqrt{\langle g, g\rangle} = 1/\sqrt{105}$.

25. The inner product is $\langle f, g\rangle = \int_{-1}^1 f(t)g(t)dt$. Then 1 and t are orthogonal because $\langle 1, t\rangle = \int_{-1}^1 t\, dt = 0$. So 1 and t can be in an orthogonal basis for $\text{Span}\{1, t, t^2\}$. By the Gram-Schmidt process, the third basis element in the orthogonal basis can be

$$t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\langle t^2, t\rangle}{\langle t, t\rangle}t$$

Since $\langle t^2, 1\rangle = \int_{-1}^1 t^2 dt = 2/3$, $\langle 1, 1\rangle = \int_{-1}^1 1\, dt = 2$, and $\langle t^2, t\rangle = \int_{-1}^1 t^3 dt = 0$, the third basis element can be written as $t^2 - (1/3)$. This element can be scaled by 3, which gives the orthogonal basis as $\{1, t, 3t^2 - 1\}$.

26. The inner product is $\langle f, g\rangle = \int_{-2}^2 f(t)g(t)dt$. Then 1 and t are orthogonal because $\langle 1, t\rangle = \int_{-2}^2 t\, dt = 0$. So 1 and t can be in an orthogonal basis for $\text{Span}\{1, t, t^2\}$. By the Gram-Schmidt process, the third basis element in the orthogonal basis can be

$$t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\langle t^2, t\rangle}{\langle t, t\rangle}t$$

Since $\langle t^2, 1\rangle = \int_{-2}^2 t^2 dt = 16/3$, $\langle 1, 1\rangle = \int_{-2}^2 1\, dt = 4$, and $\langle t^2, t\rangle = \int_{-2}^2 t^3 dt = 0$, the third basis element can be written as $t^2 - (4/3)$. This element can be scaled by 3, which gives the orthogonal basis as $\{1, t, 3t^2 - 4\}$.

27. **[M]** The new orthogonal polynomials are multiples of $-17t+5t^3$ and $72-155t^2+35t^4$. These polynomials may be scaled so that their values at –2, –1, 0, 1, and 2 are small integers.

28. **[M]** The orthogonal basis is $f_0(t)=1$, $f_1(t)=\cos t$, $f_2(t)=\cos^2 t-(1/2)=(1/2)\cos 2t$, and $f_3(t)=\cos^3 t-(3/4)\cos t=(1/4)\cos 3t$.

6.8 SOLUTIONS

Notes: The connections between this section and Section 6.7 are described in the notes for that section. For my junior-senior class, I spend three days on the following topics: Theorems 13 and 15 in Section 6.5, plus Examples 1, 3, and 5; Example 1 in Section 6.6; Examples 2 and 3 in Section 6.7, with the motivation for the definite integral; and Fourier series in Section 6.8.

1. The weighting matrix W, design matrix X, parameter vector β, and observation vector $\mathbf{y}$ are:

$$W=\begin{bmatrix}1&0&0&0&0\\0&2&0&0&0\\0&0&2&0&0\\0&0&0&2&0\\0&0&0&0&1\end{bmatrix},\ X=\begin{bmatrix}1&-2\\1&-1\\1&0\\1&1\\1&2\end{bmatrix},\ \beta=\begin{bmatrix}\beta_0\\\beta_1\end{bmatrix},\ \mathbf{y}=\begin{bmatrix}0\\0\\2\\4\\4\end{bmatrix}$$

The design matrix X and the observation vector $\mathbf{y}$ are scaled by W:

$$WX=\begin{bmatrix}1&-2\\2&-2\\2&0\\2&2\\1&2\end{bmatrix},\ W\mathbf{y}=\begin{bmatrix}0\\0\\4\\8\\4\end{bmatrix}$$

Further compute

$$(WX)^TWX=\begin{bmatrix}14&0\\0&16\end{bmatrix},\ (WX)^TW\mathbf{y}=\begin{bmatrix}28\\24\end{bmatrix}$$

and find that

$$\hat{\beta}=((WX)^TWX)^{-1}(WX)^TW\mathbf{y}=\begin{bmatrix}1/14&0\\0&1/16\end{bmatrix}\begin{bmatrix}28\\24\end{bmatrix}=\begin{bmatrix}2\\3/2\end{bmatrix}$$

Thus the weighted least-squares line is $y = 2 + (3/2)x$.

2. Let X be the original design matrix, and let $\mathbf{y}$ be the original observation vector. Let W be the weighting matrix for the first method. Then $2W$ is the weighting matrix for the second method. The weighted least-squares by the first method is equivalent to the ordinary least-squares for an equation whose normal equation is

$$(WX)^TWX\hat{\beta}=(WX)^TW\mathbf{y} \qquad (1)$$

while the second method is equivalent to the ordinary least-squares for an equation whose normal equation is

$$(2WX)^T(2W)X\hat{\beta}=(2WX)^T(2W)\mathbf{y} \qquad (2)$$

Since equation (2) can be written as $4(WX)^TWX\hat{\beta}=4(WX)^TW\mathbf{y}$, it has the same solutions as equation (1).

3. From Example 2 and the statement of the problem, $p_0(t)=1$, $p_1(t)=t$, $p_2(t)=t^2-2$, $p_3(t)=(5/6)t^3-(17/6)t$, and g = (3, 5, 5, 4, 3). The cubic trend function for g is the orthogonal projection $\hat{p}$ of g onto the subspace spanned by p_0, p_1, p_2, and p_3:

$$\hat{p}=\frac{\langle g,p_0\rangle}{\langle p_0,p_0\rangle}p_0+\frac{\langle g,p_1\rangle}{\langle p_1,p_1\rangle}p_1+\frac{\langle g,p_2\rangle}{\langle p_2,p_2\rangle}p_2+\frac{\langle g,p_3\rangle}{\langle p_3,p_3\rangle}p_3$$

$$=\frac{20}{5}(1)+\frac{-1}{10}t+\frac{-7}{14}\left(t^2-2\right)+\frac{2}{10}\left(\frac{5}{6}t^3-\frac{17}{6}t\right)$$

$$=4-\frac{1}{10}t-\frac{1}{2}\left(t^2-2\right)+\frac{1}{5}\left(\frac{5}{6}t^3-\frac{17}{6}t\right)=5-\frac{2}{3}t-\frac{1}{2}t^2+\frac{1}{6}t^3$$

This polynomial happens to fit the data exactly.

4. The inner product is $\langle p, q\rangle = p(-5)q(-5) + p(-3)q(-3) + p(-1)q(-1) + p(1)q(1) + p(3)q(3) + p(5)q(5)$.

a. Begin with the basis $\{1, t, t^2\}$ for $\mathbb{P}_2$. Since 1 and t are orthogonal, let $p_0(t)=1$ and $p_1(t)=t$. Then the Gram-Schmidt process gives

$$p_2(t)=t^2-\frac{\langle t^2,1\rangle}{\langle 1,1\rangle}1-\frac{\langle t^2,t\rangle}{\langle t,t\rangle}t=t^2-\frac{70}{6}=t^2-\frac{35}{3}$$

The vector of values for p_2 is (40/3, –8/3, –32/3, –32/3, –8/3, 40/3), so scaling by 3/8 yields the new function $p_2=(3/8)(t^2-(35/3))=(3/8)t^2-(35/8)$.

b. The data vector is g = (1, 1, 4, 4, 6, 8). The quadratic trend function for g is the orthogonal projection $\hat{p}$ of g onto the subspace spanned by p_0, p_1 and p_2:

$$\hat{p}=\frac{\langle g,p_0\rangle}{\langle p_0,p_0\rangle}p_0+\frac{\langle g,p_1\rangle}{\langle p_1,p_1\rangle}p_1+\frac{\langle g,p_2\rangle}{\langle p_2,p_2\rangle}p_2=\frac{24}{6}(1)+\frac{50}{70}t+\frac{6}{84}\left(\frac{3}{8}t^2-\frac{35}{8}\right)$$

$$=4+\frac{5}{7}t+\frac{1}{14}\left(\frac{3}{8}t^2-\frac{35}{8}\right)=\frac{59}{16}+\frac{5}{7}t+\frac{3}{112}t^2$$

5. The inner product is $\langle f, g\rangle=\int_0^{2\pi}f(t)g(t)dt$. Let $m \neq n$. Then

$$\langle \sin mt, \sin nt\rangle=\int_0^{2\pi}\sin mt\sin nt\,dt=\frac{1}{2}\int_0^{2\pi}\cos((m-n)t)-\cos((m+n)t)dt=0$$

Thus sin mt and sin nt are orthogonal.

6. The inner product is $\langle f, g\rangle=\int_0^{2\pi}f(t)g(t)dt$. Let m and n be positive integers. Then

$$\langle \sin mt,\cos nt\rangle=\int_0^{2\pi}\sin mt\cos nt\,dt=\frac{1}{2}\int_0^{2\pi}\sin((m+n)t)+\sin((m-n)t)dt=0$$

Thus sinmt and cosnt are orthogonal.

7. The inner product is $\langle f, g \rangle = \int_0^{2\pi} f(t)g(t)dt$. Let k be a positive integer. Then

$$\| \cos kt \|^2 = \langle \cos kt, \cos kt \rangle = \int_0^{2\pi} \cos^2 kt\, dt = \frac{1}{2}\int_0^{2\pi} 1 + \cos 2kt\, dt = \pi$$

and

$$\| \sin kt \|^2 = \langle \sin kt, \sin kt \rangle = \int_0^{2\pi} \sin^2 kt\, dt = \frac{1}{2}\int_0^{2\pi} 1 - \cos 2kt\, dt = \pi$$

8. Let $f(t) = t - 1$. The Fourier coefficients for f are:

$$\frac{a_0}{2} = \frac{1}{2}\frac{1}{\pi}\int_0^{2\pi} f(t)\, dt = \frac{1}{2\pi}\int_0^{2\pi} t - 1\, dt = -1 + \pi$$

and for $k > 0$,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\, dt = \frac{1}{\pi}\int_0^{2\pi} (t-1)\cos kt\, dt = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\, dt = \frac{1}{\pi}\int_0^{2\pi} (t-1)\sin kt\, dt = -\frac{2}{k}$$

The third-order Fourier approximation to f is thus

$$\frac{a_0}{2} + b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t = -1 + \pi - 2\sin t - \sin 2t - \frac{2}{3}\sin 3t$$

9. Let $f(t) = 2\pi - t$. The Fourier coefficients for f are:

$$\frac{a_0}{2} = \frac{1}{2}\frac{1}{\pi}\int_0^{2\pi} f(t)\, dt = \frac{1}{2\pi}\int_0^{2\pi} 2\pi - t\, dt = \pi$$

and for $k > 0$,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\, dt = \frac{1}{\pi}\int_0^{2\pi} (2\pi - t)\cos kt\, dt = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\, dt = \frac{1}{\pi}\int_0^{2\pi} (2\pi - t)\sin kt\, dt = \frac{2}{k}$$

The third-order Fourier approximation to f is thus

$$\frac{a_0}{2} + b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t = \pi + 2\sin t + \sin 2t + \frac{2}{3}\sin 3t$$

10. Let $f(t) = \begin{cases} 1 & \text{for } 0 \le t < \pi \\ -1 & \text{for } \pi \le t < 2\pi \end{cases}$. The Fourier coefficients for f are:

$$\frac{a_0}{2} = \frac{1}{2}\frac{1}{\pi}\int_0^{2\pi} f(t)\, dt = \frac{1}{2\pi}\int_0^{\pi} dt - \frac{1}{2\pi}\int_{\pi}^{2\pi} dt = 0$$

and for $k > 0$,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\, dt = \frac{1}{\pi}\int_0^{\pi} \cos kt\, dt - \frac{1}{\pi}\int_{\pi}^{2\pi} \cos kt\, dt = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\, dt = \frac{1}{\pi}\int_0^{\pi} \sin kt\, dt - \frac{1}{\pi}\int_{\pi}^{2\pi} \sin kt\, dt = \begin{cases} 4/(k\pi) & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even} \end{cases}$$

The third-order Fourier approximation to f is thus

$$b_1 \sin t + b_3 \sin 3t = \frac{4}{\pi}\sin t + \frac{4}{3\pi}\sin 3t$$

11. The trigonometric identity $\cos 2t = 1 - 2\sin^2 t$ shows that

$$\sin^2 t = \frac{1}{2} - \frac{1}{2}\cos 2t$$

The expression on the right is in the subspace spanned by the trigonometric polynomials of order 3 or less, so this expression is the third-order Fourier approximation to $\cos^3 t$.

12. The trigonometric identity $\cos 3t = 4\cos^3 t - 3\cos t$ shows that

$$\cos^3 t = \frac{3}{4}\cos t + \frac{1}{4}\cos 3t$$

The expression on the right is in the subspace spanned by the trigonometric polynomials of order 3 or less, so this expression is the third-order Fourier approximation to $\cos^3 t$.

13. Let f and g be in $C\,[0, 2\pi]$ and let m be a nonnegative integer. Then the linearity of the inner product shows that

$\langle (f+g), \cos mt\rangle = \langle f, \cos mt\rangle + \langle g, \cos mt\rangle$, $\langle (f+g), \sin mt\rangle = \langle f, \sin mt\rangle + \langle g, \sin mt\rangle$

Dividing these identities respectively by $\langle \cos mt, \cos mt\rangle$ and $\langle \sin mt, \sin mt\rangle$ shows that the Fourier coefficients a_m and b_m for $f+g$ are the sums of the corresponding Fourier coefficients of f and of g.

14. Note that g and h are both in the subspace H spanned by the trigonometric polynomials of order 2 or less. Since h is the second-order Fourier approximation to f, it is closer to f than any other function in the subspace H.

15. **[M]** The weighting matrix W is the 13 × 13 diagonal matrix with diagonal entries 1, 1, 1, .9, .9, .8, .7, .6, .5, .4, .3, .2, .1. The design matrix X, parameter vector β, and observation vector **y** are:

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 2^2 & 2^3 \\ 1 & 3 & 3^2 & 3^3 \\ 1 & 4 & 4^2 & 4^3 \\ 1 & 5 & 5^2 & 5^3 \\ 1 & 6 & 6^2 & 6^3 \\ 1 & 7 & 7^2 & 7^3 \\ 1 & 8 & 8^2 & 8^3 \\ 1 & 9 & 9^2 & 9^3 \\ 1 & 10 & 10^2 & 10^3 \\ 1 & 11 & 11^2 & 11^3 \\ 1 & 12 & 12^2 & 12^3 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0.0 \\ 8.8 \\ 29.9 \\ 62.0 \\ 104.7 \\ 159.1 \\ 222.0 \\ 294.5 \\ 380.4 \\ 471.1 \\ 571.7 \\ 686.8 \\ 809.2 \end{bmatrix}$$

The design matrix X and the observation vector $\mathbf{y}$ are scaled by W:

$$WX = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 1.0 & 1.0 \\ 1.0 & 2.0 & 4.0 & 8.0 \\ .9 & 2.7 & 8.1 & 24.3 \\ .9 & 3.6 & 14.4 & 57.6 \\ .8 & 4.0 & 20.0 & 100.0 \\ .7 & 4.2 & 25.2 & 151.2 \\ .6 & 4.2 & 29.4 & 205.8 \\ .5 & 4.0 & 32.0 & 256.0 \\ .4 & 3.6 & 32.4 & 291.6 \\ .3 & 3.0 & 30.0 & 300.0 \\ .2 & 2.2 & 24.2 & 266.2 \\ .1 & 1.2 & 14.4 & 172.8 \end{bmatrix}, W\mathbf{y} = \begin{bmatrix} 0.00 \\ 8.80 \\ 29.90 \\ 55.80 \\ 94.23 \\ 127.28 \\ 155.40 \\ 176.70 \\ 190.20 \\ 188.44 \\ 171.51 \\ 137.36 \\ 80.92 \end{bmatrix}$$

Further compute

$$(WX)^T WX = \begin{bmatrix} 6.66 & 22.23 & 120.77 & 797.19 \\ 22.23 & 120.77 & 797.19 & 5956.13 \\ 120.77 & 797.19 & 5956.13 & 48490.23 \\ 797.19 & 5956.13 & 48490.23 & 420477.17 \end{bmatrix}, (WX)^T W\mathbf{y} = \begin{bmatrix} 747.844 \\ 4815.438 \\ 35420.468 \\ 285262.440 \end{bmatrix}$$

and find that

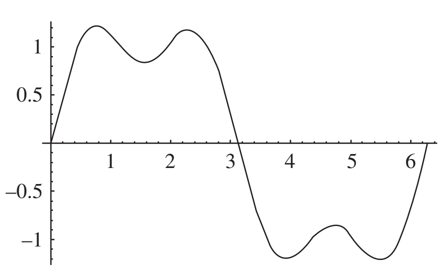
$$\hat{\beta} = ((WX)^T WX)^{-1}(WX)^T W\mathbf{y} = \begin{bmatrix} -0.2685 \\ 3.6095 \\ 5.8576 \\ -0.0477 \end{bmatrix}$$

Thus the weighted least-squares cubic is $y = g(t) = -.2685 + 3.6095t + 5.8576t^2 - .0477t^3$. The velocity at $t = 4.5$ seconds is $g'(4.5) = 53.4$ ft./sec. This is about 0.7% faster than the estimate obtained in Exercise 13 of Section 6.6.

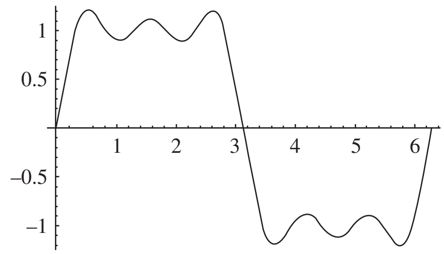
16. **[M]** Let $f(t) = \begin{cases} 1 & \text{for } 0 \le t < \pi \\ -1 & \text{for } \pi \le t < 2\pi \end{cases}$. The Fourier coefficients for f have already been found to be $a_k = 0$ for all $k \ge 0$ and $b_k = \begin{cases} 4/(k\pi) & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even} \end{cases}$. Thus

$$f_4(t) = \frac{4}{\pi}\sin t + \frac{4}{3\pi}\sin 3t \text{ and } f_5(t) = \frac{4}{\pi}\sin t + \frac{4}{3\pi}\sin 3t + \frac{4}{5\pi}\sin 5t$$

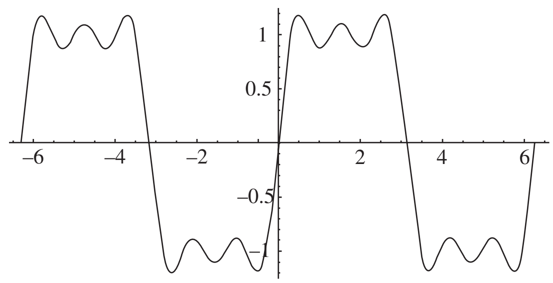
A graph of f_4 over the interval $[0, 2\pi]$ is



A graph of f_5 over the interval $[0, 2\pi]$ is



A graph of f_5 over the interval $[-2\pi, 2\pi]$ is



Chapter 6 SUPPLEMENTARY EXERCISES

1. **a**. False. The length of the zero vector is zero.

b. True. By the displayed equation before Example 2 in Section 6.1, with $c = -1$, $\|-\mathbf{x}\| = \|(-1)\mathbf{x}\| = |-1|\|\mathbf{x}\| = \|\mathbf{x}\|$.

c. True. This is the definition of distance.

d. False. This equation would be true if $r\|\mathbf{v}\|$ were replaced by $|r|\|\mathbf{v}\|$.

e. False. Orthogonal *nonzero* vectors are linearly independent.

f. True. If $\mathbf{x} \cdot \mathbf{u} = 0$ and $\mathbf{x} \cdot \mathbf{v} = 0$, then $\mathbf{x} \cdot (\mathbf{u} - \mathbf{v}) = \mathbf{x} \cdot \mathbf{u} - \mathbf{x} \cdot \mathbf{v} = 0$.

g. True. This is the "only if" part of the Pythagorean Theorem in Section 6.1.

h. True. This is the "only if" part of the Pythagorean Theorem in Section 6.1 where $\mathbf{v}$ is replaced by $-\mathbf{v}$, because $\|-\mathbf{v}\|^2$ is the same as $\|\mathbf{v}\|^2$.

i. False. The orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ is a scalar multiple of $\mathbf{u}$, not $\mathbf{y}$ (except when $\mathbf{y}$ itself is already a multiple of $\mathbf{u}$).

j. True. The orthogonal projection of any vector $\mathbf{y}$ onto W is always a vector in W.

k. True. This is a special case of the statement in the box following Example 6 in Section 6.1 (and proved in Exercise 30 of Section 6.1).

l. False. The zero vector is in both W and $W^\perp$.

m. True. See Exercise 32 in Section 6.2. If $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, then $(c_i\mathbf{v}_i)\cdot(c_j\mathbf{v}_j) = c_ic_j(\mathbf{v}_i \cdot \mathbf{v}_j) = c_ic_j0 = 0$.

n. False. This statement is true only for a *square* matrix. See Theorem 10 in Section 6.3.

o. False. An orthogonal matrix is square and has *orthonormal* columns.

p. True. See Exercises 27 and 28 in Section 6.2. If U has orthonormal columns, then $U^TU = I$. If U is also square, then the Invertible Matrix Theorem shows that U is invertible and $U^{-1} = U^T$. In this case, $U^TU = I$, which shows that the columns of U^T are orthonormal; that is, the rows of U are orthonormal.

q. True. By the Orthogonal Decomposition Theorem, the vectors $\text{proj}_W \mathbf{v}$ and $\mathbf{v} - \text{proj}_W \mathbf{v}$ are orthogonal, so the stated equality follows from the Pythagorean Theorem.

r. False. A least-squares solution is a vector $\hat{\mathbf{x}}$ (not $A\hat{\mathbf{x}}$) such that $A\hat{\mathbf{x}}$ is the closest point to **b** in Col A.

s. False. The equation $\hat{\mathbf{x}} = (\mathrm{A}^\mathrm{T}\mathrm{A})^{-1}\mathrm{A}^\mathrm{T}\mathbf{b}$ describes the *solution* of the normal equations, not the matrix form of the normal equations. Furthermore, this equation makes sense only when A^TA is invertible.

2. If $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal set and $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then the vectors $c_1\mathbf{v}_1$ and $c_2\mathbf{v}_2$ are orthogonal (Exercise 32 in Section 6.2). By the Pythagorean Theorem and properties of the norm

$$\|\mathbf{x}\|^2 = \|c_1\mathbf{v}_1 + c_2\mathbf{v}_2\|^2 = \|c_1\mathbf{v}_1\|^2 + \|c_2\mathbf{v}_2\|^2 = (c_1\|\mathbf{v}_1\|)^2 + (c_2\|\mathbf{v}_2\|)^2 = |c_1|^2 + |c_2|^2$$

So the stated equality holds for $p = 2$. Now suppose the equality holds for $p = k$, with $k \geq 2$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\}$ be an orthonormal set, and consider $\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} = \mathbf{u}_k + c_{k+1}\mathbf{v}_{k+1}$, where $\mathbf{u}_k = c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k$. Observe that $\mathbf{u}_k$ and $c_{k+1}\mathbf{v}_{k+1}$ are orthogonal because $\mathbf{v}_j \cdot \mathbf{v}_{k+1} = 0$ for $j = 1,\ldots,k$. By the Pythagorean Theorem and the assumption that the stated equality holds for k, and because $\|c_{k+1}\mathbf{v}_{k+1}\|^2 = |c_{k+1}|^2\|\mathbf{v}_{k+1}\|^2 = |c_{k+1}|^2$,

$$\|\mathbf{x}\|^2 = \|\mathbf{u}_k + c_{k+1}\mathbf{v}_{k+1}\|^2 = \|\mathbf{u}_k\|^2 + \|c_{k+1}\mathbf{v}_{k+1}\|^2 = |c_1|^2 + \ldots + |c_{k+1}|^2$$

Thus the truth of the equality for $p = k$ implies its truth for $p = k + 1$. By the principle of induction, the equality is true for all integers $p \geq 2$.

3. Given **x** and an orthonormal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$, let $\hat{\mathbf{x}}$ be the orthogonal projection of **x** onto the subspace spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_p$. By Theorem 10 in Section 6.3, $\hat{\mathbf{x}} = (\mathbf{x}\cdot\mathbf{v}_1)\mathbf{v}_1 + \ldots + (\mathbf{x}\cdot\mathbf{v}_p)\mathbf{v}_p$. By Exercise 2, $\|\hat{\mathbf{x}}\|^2 = |\mathbf{x}\cdot\mathbf{v}_1|^2 + \ldots + |\mathbf{x}\cdot\mathbf{v}_p|^2$. Bessel's inequality follows from the fact that $\|\hat{\mathbf{x}}\|^2 \leq \|\mathbf{x}\|^2$, which is noted before the proof of the Cauchy-Schwarz inequality in Section 6.7.

4. By parts (a) and (c) of Theorem 7 in Section 6.2, $\{U\mathbf{v}_1, \ldots, U\mathbf{v}_k\}$ is an orthonormal set in $\mathbb{R}^n$. Since there are n vectors in this linearly independent set, the set is a basis for $\mathbb{R}^n$.

5. Suppose that $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ for all **x**, **y** in $\mathbb{R}^n$, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis for $\mathbb{R}^n$. For $j = 1, \ldots, n$, $U\mathbf{e}_j$ is the jth column of U. Since $\|U\mathbf{e}_j\|^2 = (U\mathbf{e}_j)\cdot(U\mathbf{e}_j) = \mathbf{e}_j\cdot\mathbf{e}_j = 1$, the columns of U are unit vectors; since $(U\mathbf{e}_j)\cdot(U\mathbf{e}_k) = \mathbf{e}_j\cdot\mathbf{e}_k = 0$ for $j \neq k$, the columns are pairwise orthogonal.

6. If $U\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \neq \mathbf{0}$, then by Theorem 7(a) in Section 6.2 and by a property of the norm, $\|\mathbf{x}\| = \|U\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$, which shows that $|\lambda| = 1$, because $\mathbf{x} \neq \mathbf{0}$.

7. Let **u** be a unit vector, and let $Q = I - 2\mathbf{u}\mathbf{u}^T$. Since $(\mathbf{u}\mathbf{u}^T)^T = \mathbf{u}^{TT}\mathbf{u}^T = \mathbf{u}\mathbf{u}^T$,

$$Q^T = (I - 2\mathbf{u}\mathbf{u}^T)^T = I - 2(\mathbf{u}\mathbf{u}^T)^T = I - 2\mathbf{u}\mathbf{u}^T = Q$$

Then

$$QQ^T = Q^2 = (I - 2\mathbf{u}\mathbf{u}^T)^2 = I - 2\mathbf{u}\mathbf{u}^T - 2\mathbf{u}\mathbf{u}^T + 4(\mathbf{u}\mathbf{u}^T)(\mathbf{u}\mathbf{u}^T)$$

Since $\mathbf{u}$ is a unit vector, $\mathbf{u}^T\mathbf{u} = \mathbf{u}\cdot\mathbf{u} = 1$, so $(\mathbf{u}\mathbf{u}^T)(\mathbf{u}\mathbf{u}^T) = \mathbf{u}(\mathbf{u}^T)(\mathbf{u})\mathbf{u}^T = \mathbf{u}\mathbf{u}^T$, and

$$QQ^T = I - 2\mathbf{u}\mathbf{u}^T - 2\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T = I$$

Thus Q is an orthogonal matrix.

8. **a**. Suppose that $\mathbf{x}\cdot\mathbf{y} = 0$. By the Pythagorean Theorem, $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x}+\mathbf{y}\|^2$. Since T preserves lengths and is linear,

$$\|T(\mathbf{x})\|^2 + \|T(\mathbf{y})\|^2 = \|T(\mathbf{x}+\mathbf{y})\|^2 = \|T(\mathbf{x})+T(\mathbf{y})\|^2$$

This equation shows that $T(\mathbf{x})$ and $T(\mathbf{y})$ are orthogonal, because of the Pythagorean Theorem. Thus T preserves orthogonality.

b. The standard matrix of T is $\begin{bmatrix} T(\mathbf{e}_1) & \dots & T(\mathbf{e}_n)\end{bmatrix}$, where $\mathbf{e}_1,\dots,\mathbf{e}_n$ are the columns of the identity matrix. Then $\{T(\mathbf{e}_1),\dots,T(\mathbf{e}_n)\}$ is an orthonormal set because T preserves both orthogonality and lengths (and because the columns of the identity matrix form an orthonormal set). Finally, a square matrix with orthonormal columns is an orthogonal matrix, as was observed in Section 6.2.

9. Let $W = \text{Span}\{\mathbf{u}, \mathbf{v}\}$. Given $\mathbf{z}$ in $\mathbb{R}^n$, let $\hat{\mathbf{z}} = \text{proj}_W \mathbf{z}$. Then $\hat{\mathbf{z}}$ is in Col A, where $A = \begin{bmatrix}\mathbf{u} & \mathbf{v}\end{bmatrix}$. Thus there is a vector, say, $\hat{\mathbf{x}}$ in $\mathbb{R}^2$, with $A\hat{\mathbf{x}} = \hat{\mathbf{z}}$. So, $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{z}$. The normal equations may be solved to find $\hat{\mathbf{x}}$, and then $\hat{\mathbf{z}}$ may be found by computing $A\hat{\mathbf{x}}$.

10. Use Theorem 14 in Section 6.5. If $c \neq 0$, the least-squares solution of $A\mathbf{x} = c\mathbf{b}$ is given by $(A^TA)^{-1}A^T(c\mathbf{b})$, which equals $c(A^TA)^{-1}A^T\mathbf{b}$, by linearity of matrix multiplication. This solution is c times the least-squares solution of $A\mathbf{x} = \mathbf{b}$.

11. Let $\mathbf{x} = \begin{bmatrix} x \\ y \\ z\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} a \\ b \\ c\end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 5\end{bmatrix}$, and $A = \begin{bmatrix}\mathbf{v}^T \\ \mathbf{v}^T \\ \mathbf{v}^T\end{bmatrix} = \begin{bmatrix} 1 & -2 & 5 \\ 1 & -2 & 5 \\ 1 & -2 & 5\end{bmatrix}$. Then the given set of equations is $A\mathbf{x} = \mathbf{b}$, and the set of all least-squares solutions coincides with the set of solutions of the normal equations $A^TA\mathbf{x} = A^T\mathbf{b}$. The column-row expansions of A^TA and $A^T\mathbf{b}$ give

$$A^TA = \mathbf{v}\mathbf{v}^T + \mathbf{v}\mathbf{v}^T + \mathbf{v}\mathbf{v}^T = 3\mathbf{v}\mathbf{v}^T, A^T\mathbf{b} = a\mathbf{v} + b\mathbf{v} + c\mathbf{v} = (a+b+c)\mathbf{v}$$

Thus $A^TA\mathbf{x} = 3(\mathbf{v}\mathbf{v}^T)\mathbf{x} = 3\mathbf{v}(\mathbf{v}^T\mathbf{x}) = 3(\mathbf{v}^T\mathbf{x})\mathbf{v}$ since $\mathbf{v}^T\mathbf{x}$ is a scalar, and the normal equations have become $3(\mathbf{v}^T\mathbf{x})\mathbf{v} = (a+b+c)\mathbf{v}$, so $3(\mathbf{v}^T\mathbf{x}) = a+b+c$, or $\mathbf{v}^T\mathbf{x} = (a+b+c)/3$. Computing $\mathbf{v}^T\mathbf{x}$ gives the equation $x - 2y + 5z = (a + b + c)/3$ which must be satisfied by all least-squares solutions to $A\mathbf{x} = \mathbf{b}$.

12. The equation (1) in the exercise has been written as $V\lambda = \mathbf{b}$, where V is a single nonzero column vector $\mathbf{v}$, and $\mathbf{b} = A\mathbf{v}$. The least-squares solution $\hat{\lambda}$ of $V\lambda = \mathbf{b}$ is the exact solution of the normal equations $V^TV\lambda = V^T\mathbf{b}$. In the original notation, this equation is $\mathbf{v}^T\mathbf{v}\lambda = \mathbf{v}^TA\mathbf{v}$. Since $\mathbf{v}^T\mathbf{v}$ is nonzero, the least squares solution $\hat{\lambda}$ is $\mathbf{v}^TA\mathbf{v}/(\mathbf{v}^T\mathbf{v})$. This expression is the Rayleigh quotient discussed in the Exercises for Section 5.8.

13. **a**. The row-column calculation of $A\mathbf{u}$ shows that each row of A is orthogonal to every $\mathbf{u}$ in Nul A. So each row of A is in $(\text{Nul } A)^\perp$. Since $(\text{Nul } A)^\perp$ is a subspace, it must contain all linear combinations of the rows of A; hence $(\text{Nul } A)^\perp$ contains Row A.

b. If rank $A = r$, then dimNul $A = n - r$ by the Rank Theorem. By Exercsie 24(c) in Section 6.3, $\text{dimNul } A + \dim(\text{Nul } A)^\perp = n$, so $\dim(\text{Nul } A)^\perp$ must be r. But Row A is an r-dimensional subspace of $(\text{Nul } A)^\perp$ by the Rank Theorem and part (a). Therefore, $\text{Row } A = (\text{Nul } A)^\perp$.

c. Replace A by A^T in part (b) and conclude that $\text{Row } A^T = (\text{Nul } A^T)^\perp$. Since $\text{Row } A^T = \text{Col } A$, $\text{Col } A = (\text{Nul } A^T)^\perp$.

14. The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is in Col A. By Exercise 13(c), $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is orthogonal to $\text{Nul } A^T$. This happens if and only if $\mathbf{b}$ is orthogonal to all solutions of $A^T\mathbf{x} = \mathbf{0}$.

15. If $A = URU^T$ with U orthogonal, then A is similar to R (because U is invertible and $U^T = U^{-1}$), so A has the same eigenvalues as R by Theorem 4 in Section 5.2. Since the eigenvalues of R are its n real diagonal entries, A has n real eigenvalues.

16. **a**. If $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}$, then $AU = \begin{bmatrix} \lambda_1\mathbf{u}_1 & A\mathbf{u}_2 & \cdots & A\mathbf{u}_n \end{bmatrix}$. Since $\mathbf{u}_1$ is a unit vector and $\mathbf{u}_2, \ldots, \mathbf{u}_n$ are orthogonal to $\mathbf{u}_1$, the first column of U^TAU is $U^T(\lambda_1\mathbf{u}_1) = \lambda_1 U^T\mathbf{u}_1 = \lambda_1\mathbf{e}_1$.

b. From (a),

$$U^TAU = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

View U^TAU as a 2 × 2 block upper triangular matrix, with A_1 as the (2, 2)-block. Then from Supplementary Exercise 12 in Chapter 5,

$$\det(U^TAU - \lambda I_n) = \det((\lambda_1 - \lambda)I_1) \cdot \det(A_1 - \lambda I_{n-1}) = (\lambda_1 - \lambda) \cdot \det(A_1 - \lambda I_{n-1})$$

This shows that the eigenvalues of U^TAU, namely, $\lambda_1, \ldots, \lambda_n$, consist of λ_1 and the eigenvalues of A_1. So the eigenvalues of A_1 are $\lambda_2, \ldots, \lambda_n$.

17. **[M]** Compute that $\|\Delta\mathbf{x}\|/\|\mathbf{x}\| = .4618$ and $\text{cond}(A) \times (\|\Delta\mathbf{b}\|/\|\mathbf{b}\|) = 3363 \times (1.548 \times 10^{-4}) = .5206$. In this case, $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$ is almost the same as $\text{cond}(A) \times \|\Delta b\|/\|\mathbf{b}\|$.

18. **[M]** Compute that $\|\Delta\mathbf{x}\|/\|\mathbf{x}\| = .00212$ and $\text{cond}(A) \times (\|\Delta\mathbf{b}\|/\|\mathbf{b}\|) = 3363 \times (.00212) \approx 7.130$. In this case, $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$ is almost the same as $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$, even though the large condition number suggests that $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$ could be much larger.

19. **[M]** Compute that $\|\Delta\mathbf{x}\|/\|\mathbf{x}\| = 7.178 \times 10^{-8}$ and $\text{cond}(A) \times (\|\Delta\mathbf{b}\|/\|\mathbf{b}\|) = 23683 \times (2.832 \times 10^{-4}) =$ 6.707. Observe that the realtive change in $\mathbf{x}$ is *much* smaller than the relative change in $\mathbf{b}$. In fact the theoretical bound on the realtive change in $\mathbf{x}$ is 6.707 (to four significant figures). This exercise shows that even when a condition number is large, the relative error in the solution need not be as large as you suspect.

20. **[M]** Compute that $\|\Delta\mathbf{x}\|/\|\mathbf{x}\| = .2597$ and $\text{cond}(A) \times (\|\Delta\mathbf{b}\|/\|\mathbf{b}\|) = 23683 \times (1.097 \times 10^{-5}) = .2598$. This calculation shows that the relative change in $\mathbf{x}$, for this particular $\mathbf{b}$ and $\Delta\mathbf{b}$, should not exceed .2598. In this case, the theoretical maximum change is almost acheived.

7 Symmetric Matrices and Quadratic Forms

7.1 SOLUTIONS

Notes: Students can profit by reviewing Section 5.3 (focusing on the Diagonalization Theorem) before working on this section. Theorems 1 and 2 and the calculations in Examples 2 and 3 are important for the sections that follow. Note that *symmetric matrix* means *real symmetric matrix*, because all matrices in the text have real entries, as mentioned at the beginning of this chapter. The exercises in this section have been constructed so that mastery of the Gram-Schmidt process is not needed.

Theorem 2 is easily proved for the 2×2 case:

$$\text{If } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ then } \lambda = \frac{1}{2}\left(a + d \pm \sqrt{(a-d)^2 + 4b^2}\right).$$

If $b = 0$ there is nothing to prove. Otherwise, there are two distinct eigenvalues, so A must be diagonalizable. In each case, an eigenvector for λ is $\begin{bmatrix} d-\lambda \\ -b \end{bmatrix}$.

1. Since $A = \begin{bmatrix} 3 & 5 \\ 5 & -7 \end{bmatrix} = A^T$, the matrix is symmetric.

2. Since $A = \begin{bmatrix} -3 & 5 \\ -5 & 3 \end{bmatrix} \neq A^T$, the matrix is not symmetric.

3. Since $A = \begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix} \neq A^T$, the matrix is not symmetric.

4. Since $A = \begin{bmatrix} 0 & 8 & 3 \\ 8 & 0 & -2 \\ 3 & -2 & 0 \end{bmatrix} = A^T$, the matrix is symmetric.

5. Since $A = \begin{bmatrix} -6 & 2 & 0 \\ 0 & -6 & 2 \\ 0 & 0 & -6 \end{bmatrix} \neq A^T$, the matrix is not symmetric.

6. Since A is not a square matrix $A \neq A^T$ and the matrix is not symmetric.

7. Let $P=\begin{bmatrix}.6 & .8\\ .8 & -.6\end{bmatrix}$, and compute that

$$P^TP=\begin{bmatrix}.6 & .8\\ .8 & -.6\end{bmatrix}\begin{bmatrix}.6 & .8\\ .8 & -.6\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I_2$$

Since P is a square matrix, P is orthogonal and $P^{-1}=P^T=\begin{bmatrix}.6 & .8\\ .8 & -.6\end{bmatrix}$.

8. Let $P=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$, and compute that

$$P^TP=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I_2$$

Since P is a square matrix, P is orthogonal and $P^{-1}=P^T=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$.

9. Let $P=\begin{bmatrix}-5 & 2\\ 2 & 5\end{bmatrix}$, and compute that

$$P^TP=\begin{bmatrix}-5 & 2\\ 2 & 5\end{bmatrix}\begin{bmatrix}-5 & 2\\ 2 & 5\end{bmatrix}=\begin{bmatrix}29 & 0\\ 0 & 29\end{bmatrix}\neq I_2$$

Thus P is not orthogonal.

10. Let $P=\begin{bmatrix}-1 & 2 & 2\\ 2 & -1 & 2\\ 2 & 2 & -1\end{bmatrix}$, and compute that

$$P^TP=\begin{bmatrix}-1 & 2 & 2\\ 2 & -1 & 2\\ 2 & 2 & -1\end{bmatrix}\begin{bmatrix}-1 & 2 & 2\\ 2 & -1 & 2\\ 2 & 2 & -1\end{bmatrix}=\begin{bmatrix}9 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 9\end{bmatrix}\neq I_3$$

Thus P is not orthogonal.

11. Let $P=\begin{bmatrix}2/3 & 2/3 & 1/3\\ 0 & 1/\sqrt{5} & -2/\sqrt{5}\\ \sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45}\end{bmatrix}$, and compute that

$$P^TP=\begin{bmatrix}2/3 & 0 & \sqrt{5}/3\\ 2/3 & 1/\sqrt{5} & -4/\sqrt{45}\\ 1/3 & -2/\sqrt{5} & -2/\sqrt{45}\end{bmatrix}\begin{bmatrix}2/3 & 2/3 & 1/3\\ 0 & 1/\sqrt{5} & -2/\sqrt{5}\\ \sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45}\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=I_3$$

Since P is a square matrix, P is orthogonal and $P^{-1} = P^T = \begin{bmatrix} 2/3 & 0 & \sqrt{5}/3 \\ 2/3 & 1/\sqrt{5} & -4/\sqrt{45} \\ 1/3 & -2/\sqrt{5} & -2/\sqrt{45} \end{bmatrix}.$

12. Let $P = \begin{bmatrix} .5 & .5 & -.5 & -.5 \\ -.5 & .5 & -.5 & .5 \\ .5 & .5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix}$, and compute that

$$P^T P = \begin{bmatrix} .5 & -.5 & .5 & -.5 \\ .5 & .5 & .5 & .5 \\ -.5 & -.5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix} \begin{bmatrix} .5 & .5 & -.5 & -.5 \\ -.5 & .5 & -.5 & .5 \\ .5 & .5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_4$$

Since P is a square matrix, P is orthogonal and $P^{-1} = P^T = \begin{bmatrix} .5 & -.5 & .5 & -.5 \\ .5 & .5 & .5 & .5 \\ -.5 & -.5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix}.$

13. Let $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Then the characteristic polynomial of A is $(3-\lambda)^2 - 1 = \lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2),$ so the eigenvalues of A are 4 and 2. For $\lambda = 4$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. For $\lambda = 2$, one computes that a basis for the eigenspace is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. Let

$$P = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \text{and } D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

14. Let $A = \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$. Then the characteristic polynomial of A is $(1-\lambda)^2 - 25 = \lambda^2 - 2\lambda - 24 = (\lambda - 6)(\lambda + 4),$ so the eigenvalues of A are 6 and –4. For $\lambda = 6$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. For $\lambda = -4$, one computes that a basis for the eigenspace is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \text{ and } D = \begin{bmatrix} 6 & 0 \\ 0 & -4 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

15. Let $A = \begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$. Then the characteristic polynomial of A is $(16-\lambda)(1-\lambda)-16 = \lambda^2 - 17\lambda = (\lambda-17)\lambda$, so the eigenvalues of A are 17 and 0. For $\lambda = 17$, one computes that a basis for the eigenspace is $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} -4/\sqrt{17} \\ 1/\sqrt{17} \end{bmatrix}$. For $\lambda = 0$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{17} \\ 4/\sqrt{17} \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} -4/\sqrt{17} & 1/\sqrt{17} \\ 1/\sqrt{17} & 4/\sqrt{17} \end{bmatrix} \text{ and } D = \begin{bmatrix} 17 & 0 \\ 0 & 0 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

16. Let $A = \begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$. Then the characteristic polynomial of A is $(-7-\lambda)(7-\lambda)-576 = \lambda^2 - 625 =$ $(\lambda-25)(\lambda+25)$, so the eigenvalues of A are 25 and –25. For $\lambda = 25$, one computes that a basis for the eigenspace is $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$. For $\lambda = -25$, one computes that a basis for the eigenspace is $\begin{bmatrix} -4 \\ 3 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix} \text{ and } D = \begin{bmatrix} 25 & 0 \\ 0 & -25 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

17. Let $A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$. The eigenvalues of A are 5, 2, and –2. For $\lambda = 5$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$. For $\lambda = 2$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$. For λ = –2, one computes that a basis for the eigenspace is $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & 1\sqrt{2} \end{bmatrix} \text{ and } D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

18. Let $A = \begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. The eigenvalues of A are 25, 3, and –50. For λ = 25, one computes that a basis for the eigenspace is $\begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} -4/5 \\ 3/5 \\ 0 \end{bmatrix}$. For λ = 3, one computes that a basis for the eigenspace is $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, which is of length 1, so $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. For λ = –50, one computes that a basis for the eigenspace is $\begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} 3/5 \\ 4/5 \\ 0 \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} -4/5 & 0 & 3/5 \\ 3/5 & 0 & 4/5 \\ 0 & 1 & 0 \end{bmatrix} \text{ and } D = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -50 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

19. Let $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$. The eigenvalues of A are 7 and –2. For λ = 7, one computes that a basis for the eigenspace is $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$. This basis may be converted via orthogonal projection to an orthogonal basis for the eigenspace: $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix} \right\}$. These vectors can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 4/\sqrt{45} \\ 2/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$. For $\lambda = -2$, one computes that a basis for the eigenspace is $\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} -1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 2/\sqrt{5} & 2/\sqrt{45} & -1/3 \\ 0 & 5/\sqrt{45} & 2/3 \end{bmatrix} \text{ and } D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

20. Let $A = \begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$. The eigenvalues of A are 13, 7, and 1. For $\lambda = 13$, one computes that a basis for the eigenspace is $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$. For $\lambda = 7$, one computes that a basis for the eigenspace is $\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$. For $\lambda = 1$, one computes that a basis for the eigenspace is $\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} 2/3 & -1/3 & 2/3 \\ -1/3 & 2/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix} \text{ and } D = \begin{bmatrix} 13 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

21. Let $A=\begin{bmatrix}4&1&3&1\\1&4&1&3\\3&1&4&1\\1&3&1&4\end{bmatrix}$. The eigenvalues of A are 9, 5, and 1. For λ = 9, one computes that a basis for the eigenspace is $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$, which can be normalized to get $\mathbf{u}_1=\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}$. For λ = 5, one computes that a basis for the eigenspace is $\begin{bmatrix}-1\\1\\-1\\1\end{bmatrix}$, which can be normalized to get $\mathbf{u}_2=\begin{bmatrix}-1/2\\1/2\\-1/2\\1/2\end{bmatrix}$. For λ = 1, one computes that a basis for the eigenspace is $\left\{\begin{bmatrix}-1\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}\right\}$. This basis is an orthogonal basis for the eigenspace, and these vectors can be normalized to get $\mathbf{u}_3=\begin{bmatrix}-1/\sqrt{2}\\0\\1/\sqrt{2}\\0\end{bmatrix}$, $\mathbf{u}_4=\begin{bmatrix}0\\-1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix}$. Let

$$P=[\mathbf{u}_1\quad\mathbf{u}_2\quad\mathbf{u}_3\quad\mathbf{u}_4]=\begin{bmatrix}1/2&-1/2&-1/\sqrt{2}&0\\1/2&1/2&0&-1/\sqrt{2}\\1/2&-1/2&1/\sqrt{2}&0\\1/2&1/2&0&1/\sqrt{2}\end{bmatrix}\text{and } D=\begin{bmatrix}9&0&0&0\\0&5&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A=PDP^{-1}$.

22. Let $A=\begin{bmatrix}2&0&0&0\\0&1&0&1\\0&0&2&0\\0&1&0&1\end{bmatrix}$. The eigenvalues of A are 2 and 0. For λ = 2, one computes that a basis for the eigenspace is $\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\1\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right\}$. This basis is an orthogonal basis for the eigenspace, and these vectors can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$, and $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$. For λ = 0, one computes that a basis for the eigenspace is $\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_4 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \mathbf{u}_4] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 0 & 1 & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \text{and } D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

23. Let $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$. Since each row of A sums to 5,

$$A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

and 5 is an eigenvalue of A. The eigenvector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ may be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$. One may also compute that

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

so $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ is an eigenvector of A with associated eigenvalue λ = 2. For λ = 2, one computes that a basis for the eigenspace is $\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}\right\}$. This basis is an orthogonal basis for the eigenspace, and these vectors can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$ and $\mathbf{u}_3 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$.

Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix} \text{and}\, D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

24. Let $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$. One may compute that

$$A\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -20 \\ 20 \\ 10 \end{bmatrix} = 10\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

so $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ is an eigenvector of A with associated eigenvalue $\lambda_1 = 10$. Likewise one may compute that

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

so $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is an eigenvector of A with associated eigenvalue $\lambda_2 = 1$. For $\lambda_2 = 1$, one computes that a basis for the eigenspace is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}$. This basis may be converted via orthogonal projection to an orthogonal basis for the eigenspace: $\{\mathbf{v}_2, \mathbf{v}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \right\}$. The eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ may be normalized to get the vectors $\mathbf{u}_1 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$, and $\mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{18} \\ 1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix}$. Let

$$P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} -2/3 & 1/\sqrt{2} & 1/\sqrt{18} \\ 2/3 & 1/\sqrt{2} & -1/\sqrt{18} \\ 1/3 & 0 & 4/\sqrt{18} \end{bmatrix} \text{and}\, D = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

25. **a**. True. See Theorem 2 and the paragraph preceding the theorem.

b. True. This is a particular case of the statement in Theorem 1, where **u** and **v** are nonzero.

c. False. There are n real eigenvalues (Theorem 3), but they need not be distinct (Example 3).

d. False. See the paragraph following formula (2), in which each **u** is a unit vector.

26. **a**. True. See Theorem 2.

b. True. See the displayed equation in the paragraph before Theorem 2.

c. False. An orthogonal matrix can be symmetric (and hence orthogonally diagonalizable), but not every orthogonal matrix is symmetric. See the matrix P in Example 2.

d. True. See Theorem 3(b).

27. Since A is symmetric, $(B^TAB)^T = B^TA^TB^{TT} = B^TAB$, and B^TAB is symmetric. Applying this result with $A = I$ gives B^TB is symmetric. Finally, $(BB^T)^T = B^{TT}B^T = BB^T$, so BB^T is symmetric.

28. Let A be an $n \times n$ symmetric matrix. Then

$$(A\mathbf{x})\cdot\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^TA^T\mathbf{y} = \mathbf{x}^TA\mathbf{y} = \mathbf{x}\cdot(A\mathbf{y})$$

since $A^T = A$.

29. Since A is orthogonally diagonalizable, $A = PDP^{-1}$, where P is orthogonal and D is diagonal. Since A is invertible, $A^{-1} = (PDP^{-1})^{-1} = PD^{-1}P^{-1}$. Notice that D^{-1} is a diagonal matrix, so A^{-1} is orthogonally diagonalizable.

30. If A and B are orthogonally diagonalizable, then A and B are symmetric by Theorem 2. If $AB = BA$, then $(AB)^T = (BA)^T = A^TB^T = AB$. So AB is symmetric and hence is orthogonally diagonalizable by Theorem 2.

31. The Diagonalization Theorem of Section 5.3 says that the columns of P are linearly independent eigenvectors corresponding to the eigenvalues of A listed on the diagonal of D. So P has exactly k columns of eigenvectors corresponding to λ. These k columns form a basis for the eigenspace.

32. If $A = PRP^{-1}$, then $P^{-1}AP = R$. Since P is orthogonal, $R = P^TAP$. Hence $R^T = (P^TAP)^T = P^TA^TP^{TT} =$ $P^TAP = R$, which shows that R is symmetric. Since R is also upper triangular, its entries above the diagonal must be zeros to match the zeros below the diagonal. Thus R is a diagonal matrix.

33. It is previously been found that A is orthogonally diagonalized by P, where

$$P = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3\end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix} \text{and } D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Thus the spectral decomposition of A is

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \lambda_3\mathbf{u}_3\mathbf{u}_3^T = 8\mathbf{u}_1\mathbf{u}_1^T + 6\mathbf{u}_2\mathbf{u}_2^T + 3\mathbf{u}_3\mathbf{u}_3^T$$

$$= 8\begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 6\begin{bmatrix} 1/6 & 1/6 & -2/6 \\ 1/6 & 1/6 & -2/6 \\ -2/6 & -2/6 & 4/6 \end{bmatrix} + 3\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

34. It is previously been found that A is orthogonally diagonalized by P, where

$$P=[\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3]=\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix} \text{ and } D=\begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Thus the spectral decomposition of A is

$$A=\lambda_1\mathbf{u}_1\mathbf{u}_1^T+\lambda_2\mathbf{u}_2\mathbf{u}_2^T+\lambda_3\mathbf{u}_3\mathbf{u}_3^T=7\mathbf{u}_1\mathbf{u}_1^T+7\mathbf{u}_2\mathbf{u}_2^T-2\mathbf{u}_3\mathbf{u}_3^T$$

$$=7\begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}+7\begin{bmatrix} 1/18 & -4/18 & -1/18 \\ -4/18 & 16/18 & 4/18 \\ -1/18 & 4/18 & 1/18 \end{bmatrix}-2\begin{bmatrix} 4/9 & 2/9 & -4/9 \\ 2/9 & 1/9 & -2/9 \\ -4/9 & -2/9 & 4/9 \end{bmatrix}$$

35. **a**. Given $\mathbf{x}$ in $\mathbb{R}^n$, $b\mathbf{x}=(\mathbf{u}\mathbf{u}^T)\mathbf{x}=\mathbf{u}(\mathbf{u}^T\mathbf{x})=(\mathbf{u}^T\mathbf{x})\mathbf{u}$, because $\mathbf{u}^T\mathbf{x}$ is a scalar. So $B\mathbf{x}=(\mathbf{x}\cdot\mathbf{u})\mathbf{u}$. Since $\mathbf{u}$ is a unit vector, $B\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}$.

b. Since $B^T=(\mathbf{u}\mathbf{u}^T)^T=\mathbf{u}^{TT}\mathbf{u}^T=\mathbf{u}\mathbf{u}^T=B$, B is a symmetric matrix. Also, $B^2=(\mathbf{u}\mathbf{u}^T)(\mathbf{u}\mathbf{u}^T)=\mathbf{u}(\mathbf{u}^T\mathbf{u})\mathbf{u}^T=\mathbf{u}\mathbf{u}^T=B$ because $\mathbf{u}^T\mathbf{u}=1$.

c. Since $\mathbf{u}^T\mathbf{u}=1$, $B\mathbf{u}=(\mathbf{u}\mathbf{u}^T)\mathbf{u}=\mathbf{u}(\mathbf{u}^T\mathbf{u})=\mathbf{u}(1)=\mathbf{u}$, so $\mathbf{u}$ is an eigenvector of B with corresponding eigenvalue 1.

36. Given any $\mathbf{y}$ in $\mathbb{R}^n$, let $\hat{\mathbf{y}}=B\mathbf{y}$ and $\mathbf{z}=\mathbf{y}-\hat{\mathbf{y}}$. Suppose that $B^T=B$ and $B^2=B$. Then $B^TB=BB=B$.

a. Since $\mathbf{z}\cdot\hat{\mathbf{y}}=(\mathbf{y}-\hat{\mathbf{y}})\cdot(B\mathbf{y})=\mathbf{y}\cdot(B\mathbf{y})-\hat{\mathbf{y}}\cdot(B\mathbf{y})=\mathbf{y}^TB\mathbf{y}-(B\mathbf{y})^TB\mathbf{y}=\mathbf{y}^TB\mathbf{y}-\mathbf{y}^TB^TB\mathbf{y}=0$, $\mathbf{z}$ is orthogonal to $\hat{\mathbf{y}}$.

b. Any vector in $W=\text{Col } B$ has the form $B\mathbf{u}$ for some $\mathbf{u}$. Noting that B is symmetric, Exercise 28 gives

$$(\mathbf{y}-\hat{\mathbf{y}})\cdot(B\mathbf{u})=[B(\mathbf{y}-\hat{\mathbf{y}})]\cdot\mathbf{u}=[B\mathbf{y}-BB\mathbf{y}]\cdot\mathbf{u}=0$$

since $B^2=B$. So $\mathbf{y}-\hat{\mathbf{y}}$ is in $W^{\perp}$, and the decomposition $\mathbf{y}=\hat{\mathbf{y}}+(\mathbf{y}-\hat{\mathbf{y}})$ expresses $\mathbf{y}$ as the sum of a vector in W and a vector in $W^{\perp}$. By the Orthogonal Decomposition Theorem in Section 6.3, this decomposition is unique, and so $\hat{\mathbf{y}}$ must be $\text{proj}_W\mathbf{y}$.

37. **[M]** Let $A=\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$. The eigenvalues of A are 18, 10, 4, and –12. For $\lambda=18$, one computes that a basis for the eigenspace is $\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1=\begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$. For $\lambda=10$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2=\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$.

For $\lambda = 4$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}$. For $\lambda = -12$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_4 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$. Let $P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \mathbf{u}_4] = \begin{bmatrix} -1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & 1/2 \end{bmatrix}$ and $D = \begin{bmatrix} 18 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & -12 \end{bmatrix}$. Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

38. **[M]** Let $A = \begin{bmatrix} .38 & -.18 & -.06 & -.04 \\ -.18 & .59 & -.04 & .12 \\ -.06 & -.04 & .47 & -.12 \\ -.04 & .12 & -.12 & .41 \end{bmatrix}$. The eigenvalues of A are .25, .30, .55, and .75. For $\lambda = .25$, one computes that a basis for the eigenspace is $\begin{bmatrix} 4 \\ 2 \\ 2 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} .8 \\ .4 \\ .4 \\ .2 \end{bmatrix}$. For $\lambda = .30$, one computes that a basis for the eigenspace is $\begin{bmatrix} -1 \\ -2 \\ 2 \\ 4 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_2 = \begin{bmatrix} -.2 \\ -.4 \\ .4 \\ .8 \end{bmatrix}$. For $\lambda = .55$, one computes that a basis for the eigenspace is $\begin{bmatrix} 2 \\ -1 \\ -4 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} .4 \\ -.2 \\ -.8 \\ .4 \end{bmatrix}$. For $\lambda = .75$, one computes that a basis for the eigenspace is $\begin{bmatrix} -2 \\ 4 \\ -1 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_4 = \begin{bmatrix} -.4 \\ .8 \\ -.2 \\ .4 \end{bmatrix}$. Let $P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \mathbf{u}_4] = \begin{bmatrix} .8 & -.2 & .4 & -.4 \\ .4 & -.4 & -.2 & .8 \\ .4 & .4 & -.8 & -.2 \\ .2 & .8 & .4 & .4 \end{bmatrix}$ and

$D = \begin{bmatrix} .25 & 0 & 0 & 0 \\ 0 & .30 & 0 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & 0 & .75 \end{bmatrix}$. Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

39. **[M]** Let $A = \begin{bmatrix} .31 & .58 & .08 & .44 \\ .58 & -.56 & .44 & -.58 \\ .08 & .44 & .19 & -.08 \\ .44 & -.58 & -.08 & .31 \end{bmatrix}$. The eigenvalues of A are .75, 0, and –1.25. For λ = .75, one computes that a basis for the eigenspace is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 2 \\ 0 \end{bmatrix} \right\}$. This basis may be converted via orthogonal projection to the orthogonal basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 4 \\ -3 \end{bmatrix} \right\}$. These vectors can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$,

$\mathbf{u}_2 = \begin{bmatrix} 3/\sqrt{50} \\ 4/\sqrt{50} \\ 4/\sqrt{50} \\ -3/\sqrt{50} \end{bmatrix}$. For λ = 0, one computes that a basis for the eigenspace is $\begin{bmatrix} -2 \\ -1 \\ 4 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} -.4 \\ -.2 \\ .8 \\ .4 \end{bmatrix}$. For $\lambda = -1.25$, one computes that a basis for the eigenspace is $\begin{bmatrix} -2 \\ 4 \\ -1 \\ 2 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_4 = \begin{bmatrix} -.4 \\ .8 \\ -.2 \\ .4 \end{bmatrix}$.

Let $P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \mathbf{u}_4] = \begin{bmatrix} 1/\sqrt{2} & 3/\sqrt{50} & -.4 & -.4 \\ 0 & 4/\sqrt{50} & -.2 & .8 \\ 0 & 4/\sqrt{50} & .8 & -.2 \\ 1/\sqrt{2} & -3/\sqrt{50} & .4 & .4 \end{bmatrix}$ and $D = \begin{bmatrix} .75 & 0 & 0 & 0 \\ 0 & .75 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1.25 \end{bmatrix}$. Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

40. **[M]** Let $A=\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$. The eigenvalues of A are 8, 32, –28, and 17. For $\lambda = 8$, one computes that a basis for the eigenspace is $\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right\}$. This basis may be converted via orthogonal projection to the orthogonal basis $\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \\ 0 \end{bmatrix}\right\}$. These vectors can be normalized to get $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \\ 0 \\ 0 \end{bmatrix}$. For $\lambda = 32$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \\ 0 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ -3/\sqrt{12} \\ 0 \end{bmatrix}$. For $\lambda = -28$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -4 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_4 = \begin{bmatrix} 1/\sqrt{20} \\ 1/\sqrt{20} \\ 1/\sqrt{20} \\ 1/\sqrt{20} \\ -4/\sqrt{20} \end{bmatrix}$. For $\lambda = 17$, one computes that a basis for the eigenspace is $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, which can be normalized to get $\mathbf{u}_5 = \begin{bmatrix} 1/\sqrt{5} \\ 1/\sqrt{5} \\ 1/\sqrt{5} \\ 1/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$.

Let $P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \quad \mathbf{u}_4 \quad \mathbf{u}_5] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ 0 & -2/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ 0 & 0 & -3/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ 0 & 0 & 0 & -4/\sqrt{20} & 1/\sqrt{5} \end{bmatrix}$ and

$D = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 32 & 0 & 0 \\ 0 & 0 & 0 & -28 & 0 \\ 0 & 0 & 0 & 0 & 17 \end{bmatrix}$. Then P orthogonally diagonalizes A, and $A = PDP^{-1}$.

7.2 SOLUTIONS

Notes: This section can provide a good conclusion to the course, because the mathematics here is widely used in applications. For instance, Exercises 23 and 24 can be used to develop the second derivative test for functions of two variables. However, if time permits, some interesting applications still lie ahead. Theorem 4 is used to prove Theorem 6 in Section 7.3, which in turn is used to develop the singular value decomposition.

1. **a**. $\mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 5 & 1/3 \\ 1/3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5x_1^2 + (2/3)x_1x_2 + x_2^2$

b. When $\mathbf{x} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$, $\mathbf{x}^T A\mathbf{x} = 5(6)^2 + (2/3)(6)(1) + (1)^2 = 185$.

c. When $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{x}^T A\mathbf{x} = 5(1)^2 + (2/3)(1)(3) + (3)^2 = 16$.

2. **a**. $\mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 4 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4x_1^2 + 2x_2^2 + x_3^2 + 6x_1x_2 + 2x_2x_3$

b. When $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{x}^T A\mathbf{x} = 4(2)^2 + 2(-1)^2 + (5)^2 + 6(2)(-1) + 2(-1)(5) = 21$.

c. When $\mathbf{x} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$, $\mathbf{x}^T A\mathbf{x} = 4(1/\sqrt{3})^2 + 2(1/\sqrt{3})^2 + (1/\sqrt{3})^2 + 6(1/\sqrt{3})(1/\sqrt{3}) + 2(1/\sqrt{3})(1/\sqrt{3}) = 5$.

3. **a**. The matrix of the quadratic form is $\begin{bmatrix} 10 & -3 \\ -3 & -3 \end{bmatrix}$.

b. The matrix of the quadratic form is $\begin{bmatrix} 5 & 3/2 \\ 3/2 & 0 \end{bmatrix}$.

4. **a**. The matrix of the quadratic form is $\begin{bmatrix} 20 & 15/2 \\ 15/2 & -10 \end{bmatrix}$.

b. The matrix of the quadratic form is $\begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}$.

5. **a**. The matrix of the quadratic form is $\begin{bmatrix} 8 & -3 & 2 \\ -3 & 7 & -1 \\ 2 & -1 & -3 \end{bmatrix}$.

b. The matrix of the quadratic form is $\begin{bmatrix} 0 & 2 & 3 \\ 2 & 0 & -4 \\ 3 & -4 & 0 \end{bmatrix}$.

6. **a**. The matrix of the quadratic form is $\begin{bmatrix} 5 & 5/2 & -3/2 \\ 5/2 & -1 & 0 \\ -3/2 & 0 & 7 \end{bmatrix}$.

b. The matrix of the quadratic form is $\begin{bmatrix} 0 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & 1 \end{bmatrix}$.

7. The matrix of the quadratic form is $A = \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$. The eigenvalues of A are 6 and –4. An eigenvector for $\lambda = 6$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. An eigenvector for $\lambda = -4$ is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. Then $A = PDP^{-1}$, where $P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and $D = \begin{bmatrix} 6 & 0 \\ 0 & -4 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 6y_1^2 - 4y_2^2$$

8. The matrix of the quadratic form is $A = \begin{bmatrix} 9 & -4 & 4 \\ -4 & 7 & 0 \\ 4 & 0 & 11 \end{bmatrix}$. The eigenvalues of A are 3, 9, and 15. An eigenvector for $\lambda = 3$ is $\begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} -2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$. An eigenvector for $\lambda = 9$ is $\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$. An eigenvector for $\lambda = 15$ is $\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$, which may be normalized to $\mathbf{u}_3 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$. Then $A = PDP^{-1}$, where $P = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] = \begin{bmatrix} -2/3 & -1/3 & 2/3 \\ -2/3 & 2/3 & -1/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix}$ and

$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 15 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 3y_1^2 + 9y_2^2 + 15y_3^2$$

9. The matrix of the quadratic form is $A = \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix}$. The eigenvalues of A are 7 and 2, so the quadratic form is positive definite. An eigenvector for $\lambda = 7$ is $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$.

An eigenvector for $\lambda = 2$ is $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Then $A = PDP^{-1}$, where

$P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$ and $D = \begin{bmatrix} 7 & 0 \\ 0 & 2 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 7y_1^2 + 2y_2^2$$

10. The matrix of the quadratic form is $A = \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix}$. The eigenvalues of A are 11 and 1, so the quadratic form is positive definite. An eigenvector for $\lambda = 11$ is $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$.

An eigenvector for $\lambda = 1$ is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$. Then $A = PDP^{-1}$, where

$P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$ and $D = \begin{bmatrix} 11 & 0 \\ 0 & 1 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 11y_1^2 + y_2^2$$

11. The matrix of the quadratic form is $A = \begin{bmatrix} 2 & 5 \\ 5 & 2 \end{bmatrix}$. The eigenvalues of A are 7 and –3, so the quadratic form is indefinite. An eigenvector for $\lambda = 7$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. An eigenvector for $\lambda = -3$ is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. Then $A = PDP^{-1}$, where $P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and $D = \begin{bmatrix} 7 & 0 \\ 0 & -3 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 7y_1^2 - 3y_2^2$$

12. The matrix of the quadratic form is $A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$. The eigenvalues of A are –1 and –6, so the quadratic form is negative definite. An eigenvector for $\lambda = -1$ is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$. An eigenvector for $\lambda = -6$ is $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Then $A = PDP^{-1}$, where $P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$ and $D = \begin{bmatrix} -1 & 0 \\ 0 & -6 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = -y_1^2 - 6y_2^2$$

13. The matrix of the quadratic form is $A = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}$. The eigenvalues of A are 10 and 0, so the quadratic form is positive semidefinite. An eigenvector for $\lambda = 10$ is $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$. An eigenvector for $\lambda = 0$ is $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$. Then $A = PDP^{-1}$, where $P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ -3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix}$ and $D = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 10y_1^2$$

14. The matrix of the quadratic form is $A = \begin{bmatrix} 8 & 3 \\ 3 & 0 \end{bmatrix}$. The eigenvalues of A are 9 and –1, so the quadratic form is indefinite. An eigenvector for $\lambda = 9$ is $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, which may be normalized to $\mathbf{u}_1 = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$. An eigenvector for $\lambda = -1$ is $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$, which may be normalized to $\mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix}$. Then $A = PDP^{-1}$, where $P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix}$ and $D = \begin{bmatrix} 9 & 0 \\ 0 & -1 \end{bmatrix}$. The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 9y_1^2 - y_2^2$$

15. **[M]** The matrix of the quadratic form is $A=\begin{bmatrix} -2 & 2 & 2 & 2 \\ 2 & -6 & 0 & 0 \\ 2 & 0 & -9 & 3 \\ 2 & 0 & 3 & -9 \end{bmatrix}$. The eigenvalues of A are 0, –6, –8, and –12, so the quadratic form is negative semidefinite. The corresponding eigenvectors may be computed:

$$\lambda=0:\begin{bmatrix} 3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \lambda=-6:\begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}, \lambda=-8:\begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \lambda=-12:\begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

These eigenvectors may be normalized to form the columns of P, and $A=PDP^{-1}$, where

$$P=\begin{bmatrix} 3/\sqrt{12} & 0 & -1/2 & 0 \\ 1/\sqrt{12} & -2/\sqrt{6} & 1/2 & 0 \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & -1/\sqrt{2} \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & 1/\sqrt{2} \end{bmatrix} \text{ and } D=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & -12 \end{bmatrix}$$

The desired change of variable is $\mathbf{x}=P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x}=(P\mathbf{y})^T A(P\mathbf{y})=\mathbf{y}^T P^T AP\mathbf{y}=\mathbf{y}^T D\mathbf{y}=-6y_2^2-8y_3^2-12y_4^2$$

16. **[M]** The matrix of the quadratic form is $A=\begin{bmatrix} 4 & 3/2 & 0 & -2 \\ 3/2 & 4 & 2 & 0 \\ 0 & 2 & 4 & 3/2 \\ -2 & 0 & 3/2 & 4 \end{bmatrix}$. The eigenvalues of A are 13/2 and 3/2, so the quadratic form is positive definite. The corresponding eigenvectors may be computed:

$$\lambda=13/2:\left\{\begin{bmatrix} -4 \\ 0 \\ 3 \\ 5 \end{bmatrix},\begin{bmatrix} 3 \\ 5 \\ 4 \\ 0 \end{bmatrix}\right\}, \lambda=3/2:\left\{\begin{bmatrix} 4 \\ 0 \\ -3 \\ 5 \end{bmatrix},\begin{bmatrix} 3 \\ -5 \\ 4 \\ 0 \end{bmatrix}\right\}$$

Each set of eigenvectors above is already an orthogonal set, so they may be normalized to form the columns of P, and $A=PDP^{-1}$, where

$$P=\begin{bmatrix} 3/\sqrt{50} & -4/\sqrt{50} & 3/\sqrt{50} & 4/\sqrt{50} \\ 5/\sqrt{50} & 0 & -5/\sqrt{50} & 0 \\ 4/\sqrt{50} & 3/\sqrt{50} & 4/\sqrt{50} & -3/\sqrt{50} \\ 0 & 5/\sqrt{50} & 0 & 5/\sqrt{50} \end{bmatrix} \text{ and } D=\begin{bmatrix} 13/2 & 0 & 0 & 0 \\ 0 & 13/2 & 0 & 0 \\ 0 & 0 & 3/2 & 0 \\ 0 & 0 & 0 & 3/2 \end{bmatrix}$$

The desired change of variable is $\mathbf{x}=P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x}=(P\mathbf{y})^T A(P\mathbf{y})=\mathbf{y}^T P^T AP\mathbf{y}=\mathbf{y}^T D\mathbf{y}=\frac{13}{2}y_1^2+\frac{13}{2}y_2^2+\frac{3}{2}y_3^2+\frac{3}{2}y_4^2$$

17. **[M]** The matrix of the quadratic form is $A=\begin{bmatrix} 1 & 9/2 & 0 & -6 \\ 9/2 & 1 & 6 & 0 \\ 0 & 6 & 1 & 9/2 \\ -6 & 0 & 9/2 & 1 \end{bmatrix}$. The eigenvalues of A are 17/2 and –13/2, so the quadratic form is indefinite. The corresponding eigenvectors may be computed:

$$\lambda = 17/2: \left\{ \begin{bmatrix} -4 \\ 0 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 4 \\ 0 \end{bmatrix} \right\}, \lambda = -13/2: \left\{ \begin{bmatrix} 4 \\ 0 \\ -3 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 4 \\ 0 \end{bmatrix} \right\}$$

Each set of eigenvectors above is already an orthogonal set, so they may be normalized to form the columns of P, and $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 3/\sqrt{50} & -4/\sqrt{50} & 3/\sqrt{50} & 4/\sqrt{50} \\ 5/\sqrt{50} & 0 & -5/\sqrt{50} & 0 \\ 4/\sqrt{50} & 3/\sqrt{50} & 4/\sqrt{50} & -3/\sqrt{50} \\ 0 & 5/\sqrt{50} & 0 & 5/\sqrt{50} \end{bmatrix} \text{ and } D = \begin{bmatrix} 17/2 & 0 & 0 & 0 \\ 0 & 17/2 & 0 & 0 \\ 0 & 0 & -13/2 & 0 \\ 0 & 0 & 0 & -13/2 \end{bmatrix}$$

The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = \frac{17}{2}y_1^2 + \frac{17}{2}y_2^2 - \frac{13}{2}y_3^2 - \frac{13}{2}y_4^2$$

18. **[M]** The matrix of the quadratic form is $A=\begin{bmatrix} 11 & -6 & -6 & -6 \\ -6 & -1 & 0 & 0 \\ -6 & 0 & 0 & -1 \\ -6 & 0 & -1 & 0 \end{bmatrix}$. The eigenvalues of A are 17, 1, –1, and –7, so the quadratic form is indefinite. The corresponding eigenvectors may be computed:

$$\lambda = 17: \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \lambda = 1: \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \lambda = -1: \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}, \lambda = -7: \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

These eigenvectors may be normalized to form the columns of P, and $A = PDP^{-1}$, where

$$P = \begin{bmatrix} -3/\sqrt{12} & 0 & 0 & 1/2 \\ 1/\sqrt{12} & 0 & 2/\sqrt{6} & 1/2 \\ 1/\sqrt{12} & -1/\sqrt{2} & 1/\sqrt{6} & 1/2 \\ 1/\sqrt{12} & 1/\sqrt{2} & 1/\sqrt{6} & 1/2 \end{bmatrix} \text{ and } D = \begin{bmatrix} 17 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -7 \end{bmatrix}$$

The desired change of variable is $\mathbf{x} = P\mathbf{y}$, and the new quadratic form is

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} = 17y_1^2 + y_2^2 - y_3^2 - 7y_4^2$$

19. Since 8 is larger than 5, the x_2^2 term should be as large as possible. Since $x_1^2 + x_2^2 = 1$, the largest value that x_2 can take is 1, and $x_1 = 0$ when $x_2 = 1$. Thus the largest value the quadratic form can take when $\mathbf{x}^T\mathbf{x} = 1$ is 5(0) + 8(1) = 8.

20. Since 5 is larger in absolute value than –3, the x_1^2 term should be as large as possible. Since $x_1^2 + x_2^2 = 1$, the largest value that x_1 can take is 1, and $x_2 = 0$ when $x_1 = 1$. Thus the largest value the quadratic form can take when $\mathbf{x}^T\mathbf{x} = 1$ is 5(1) – 3(0) = 5.

21. **a**. True. See the definition before Example 1, even though a nonsymmetric matrix could be used to compute values of a quadratic form.

b. True. See the paragraph following Example 3.

c. True. The columns of P in Theorem 4 are eigenvectors of A. See the Diagonalization Theorem in Section 5.3.

d. False. $Q(\mathbf{x}) = 0$ when $\mathbf{x} = \mathbf{0}$.

e. True. See Theorem 5(a).

f. True. See the Numerical Note after Example 6.

22. **a**. True. See the paragraph before Example 1.

b. False. The matrix P must be orthogonal and make P^TAP diagonal. See the paragraph before Example 4.

c. False. There are also “degenerate” cases: a single point, two intersecting lines, or no points at all. See the subsection “A Geometric View of Principal Axes.”

d. False. See the definition before Theorem 5.

e. True. See Theorem 5(b). If $\mathbf{x}^TA\mathbf{x}$ has only negative values for $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{x}^TA\mathbf{x}$ is negative definite.

23. The characteristic polynomial of A may be written in two ways:

$$\det(A - \lambda I) = \det\begin{bmatrix} a-\lambda & b \\ b & d-\lambda \end{bmatrix} = \lambda^2 - (a+d)\lambda + ad - b^2$$

and

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$$

The coefficients in these polynomials may be equated to obtain $\lambda_1 + \lambda_2 = a + d$ and $\lambda_1\lambda_2 = ad - b^2 = \det A$.

24. If det A > 0, then by Exercise 23, $\lambda_1\lambda_2 > 0$, so that λ_1 and λ_2 have the same sign; also, $ad = \det A + b^2 > 0$.

a. If det A > 0 and a > 0, then d > 0 also, since ad > 0. By Exercise 23, $\lambda_1 + \lambda_2 = a + d > 0$. Since λ_1 and λ_2 have the same sign, they are both positive. So Q is positive definite by Theorem 5.

b. If det A > 0 and a < 0, then d < 0 also, since ad > 0. By Exercise 23, $\lambda_1 + \lambda_2 = a + d < 0$. Since λ_1 and λ_2 have the same sign, they are both negative. So Q is negative definite by Theorem 5.

c. If det A < 0, then by Exercise 23, $\lambda_1\lambda_2 < 0$. Thus λ_1 and λ_2 have opposite signs. So Q is indefinite by Theorem 5.

25. Exercise 27 in Section 7.1 showed that B^TB is symmetric. Also $\mathbf{x}^TB^TB\mathbf{x} = (B\mathbf{x})^TB\mathbf{x} = \|B\mathbf{x}\| \geq 0$, so the quadratic form is positive semidefinite, and the matrix B^TB is positive semidefinite. Suppose that B is square and invertible. Then if $\mathbf{x}^TB^TB\mathbf{x} = 0$, $\|B\mathbf{x}\| = 0$ and $B\mathbf{x} = 0$. Since B is invertible, $\mathbf{x} = \mathbf{0}$. Thus if $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}^TB^TB\mathbf{x} > 0$ and B^TB is positive definite.

26. Let $A = PDP^T$, where $P^T = P^{-1}$. The eigenvalues of A are all positive: denote them $\lambda_1, \dots, \lambda_n$. Let C be the diagonal matrix with $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ on its diagonal. Then $D = C^2 = C^T C$. If $B = PCP^T$, then B is positive definite because its eigenvalues are the positive numbers on the diagonal of C. Also

$$B^T B = (PCP^T)^T (PCP^T) = (P^{TT} C^T P^T)(PCP^T) = PC^T CP^T = PDP^T = A$$

since $P^T P = I$.

27. Since the eigenvalues of A and B are all positive, the quadratic forms $\mathbf{x}^T A\mathbf{x}$ and $\mathbf{x}^T B\mathbf{x}$ are positive definite by Theorem 5. Let $\mathbf{x} \neq \mathbf{0}$. Then $\mathbf{x}^T A\mathbf{x} > 0$ and $\mathbf{x}^T B\mathbf{x} > 0$, so $\mathbf{x}^T (A+B)\mathbf{x} = \mathbf{x}^T A\mathbf{x} + \mathbf{x}^T B\mathbf{x} > 0$, and the quadratic form $\mathbf{x}^T (A+B)\mathbf{x}$ is positive definite. Note that $A + B$ is also a symmetric matrix. Thus by Theorem 5 all the eigenvalues of $A + B$ must be positive.

28. The eigenvalues of A are all positive by Theorem 5. Since the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A (see Exercise 25 in Section 5.1), the eigenvalues of A^{-1} are all positive. Note that A^{-1} is also a symmetric matrix. By Theorem 5, the quadratic form $\mathbf{x}^T A^{-1}\mathbf{x}$ is positive definite.

7.3 SOLUTIONS

Notes: Theorem 6 is the main result needed in the next two sections. Theorem 7 is mentioned in Example 2 of Section 7.4. Theorem 8 is needed at the very end of Section 7.5. The economic principles in Example 6 may be familiar to students who have had a course in macroeconomics.

1. The matrix of the quadratic form on the left is $A = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 6 & -2 \\ 0 & -2 & 7 \end{bmatrix}$. The equality of the quadratic forms implies that the eigenvalues of A are 9, 6, and 3. An eigenvector may be calculated for each eigenvalue and normalized:

$$\lambda = 9: \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}, \lambda = 6: \begin{bmatrix} 2/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \lambda = 3: \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$

The desired change of variable is $\mathbf{x} = P\mathbf{y}$, where $P = \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$.

2. The matrix of the quadratic form on the left is $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}$. The equality of the quadratic forms implies that the eigenvalues of A are 5, 2, and 0. An eigenvector may be calculated for each eigenvalue and normalized:

$$\lambda = 5: \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \lambda = 2: \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \lambda = 0: \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

The desired change of variable is $\mathbf{x} = P\mathbf{y}$, where $P = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix}$.

3. (a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ is the greatest eigenvalue λ_1 of A. By Exercise 1, $\lambda_1 = 9$.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. By Exercise 1, $\mathbf{u} = \pm\begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$ is the second greatest eigenvalue λ_2 of A. By Exercise 1, $\lambda_2 = 6$.

4. (a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ is the greatest eigenvalue λ_1 of A. By Exercise 2, $\lambda_1 = 5$.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. By Exercise 2, $\mathbf{u} = \pm\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$ is the second greatest eigenvalue λ_2 of A. By Exercise 2, $\lambda_2 = 2$.

5. The matrix of the quadratic form is $A = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$. The eigenvalues of A are $\lambda_1 = 7$ and $\lambda_2 = 3$.

(a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ is the greatest eigenvalue λ_1 of A, which is 7.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 7$, so $\mathbf{u} = \pm\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$ is the second greatest eigenvalue λ_2 of A, which is 3.

6. The matrix of the quadratic form is $A = \begin{bmatrix} 7 & 3/2 \\ 3/2 & 3 \end{bmatrix}$. The eigenvalues of A are $\lambda_1 = 15/2$ and $\lambda_2 = 5/2$.

(a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ is the greatest eigenvalue λ_1 of A, which is 15/2.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 7$, so $\mathbf{u} = \pm\begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x}=1$ and $\mathbf{x}^T\mathbf{u}=0$ is the second greatest eigenvalue λ_2 of A, which is 5/2.

7. The eigenvalues of the matrix of the quadratic form are $\lambda_1 = 2$, $\lambda_2 = -1$, and $\lambda_3 = -4$. By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} 1/2 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 2$, so $\mathbf{u} = \pm\begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$.

8. The eigenvalues of the matrix of the quadratic form are $\lambda_1 = 9$, and $\lambda_2 = -3$. By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ are linearly independent eigenvectors corresponding to $\lambda_1 = 2$, so $\mathbf{u}$ can be any unit vector which is a linear combination of $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$. Alternatively, $\mathbf{u}$ can be any unit vector which is orthogonal to the eigenspace corresponding to the eigenvalue $\lambda_2 = -3$. Since multiples of $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ are eigenvectors corresponding to $\lambda_2 = -3$, $\mathbf{u}$ can be any unit vector orthogonal to $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

9. This is equivalent to finding the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$. By Theorem 6, this value is the greatest eigenvalue λ_1 of the matrix of the quadratic form. The matrix of the quadratic form is $A = \begin{bmatrix} 7 & -1 \\ -1 & 3 \end{bmatrix}$, and the eigenvalues of A are $\lambda_1 = 5+\sqrt{5}$, $\lambda_2 = 5-\sqrt{5}$. Thus the desired constrained maximum value is $\lambda_1 = 5+\sqrt{5}$.

10. This is equivalent to finding the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$. By Theorem 6, this value is the greatest eigenvalue λ_1 of the matrix of the quadratic form. The matrix of the quadratic form is $A=\begin{bmatrix}-3 & -1\\ -1 & 5\end{bmatrix}$, and the eigenvalues of A are $\lambda_1=1+\sqrt{17},\ \lambda_2=1-\sqrt{17}$. Thus the desired constrained maximum value is $\lambda_1=1+\sqrt{17}$.

11. Since $\mathbf{x}$ is an eigenvector of A corresponding to the eigenvalue 3, $A\mathbf{x}=3\mathbf{x}$, and $\mathbf{x}^T A\mathbf{x}=\mathbf{x}^T(3\mathbf{x})=3(\mathbf{x}^T\mathbf{x})=3\|\mathbf{x}\|^2=3$ since $\mathbf{x}$ is a unit vector.

12. Let $\boldsymbol{x}$ be a unit eigenvector for the eigenvalue λ. Then $\mathbf{x}^T A\mathbf{x}=\mathbf{x}^T(\lambda\mathbf{x})=\lambda(\mathbf{x}^T\mathbf{x})=\lambda$ since $\mathbf{x}^T\mathbf{x}=1$. So λ must satisfy $m\le\lambda\le M$.

13. If $m=M$, then let $t=(1-0)m+0M=m$ and $\mathbf{x}=\mathbf{u}_n$. Theorem 6 shows that $\mathbf{u}_n^T A\mathbf{u}_n=m$. Now suppose that $m<M$, and let t be between m and M. Then $0\le t-m\le M-m$ and $0\le(t-m)/(M-m)\le 1$. Let $\alpha=(t-m)/(M-m)$, and let $\mathbf{x}=\sqrt{1-\alpha}\,\mathbf{u}_n+\sqrt{\alpha}\,\mathbf{u}_1$. The vectors $\sqrt{1-\alpha}\,\mathbf{u}_n$ and $\sqrt{\alpha}\,\mathbf{u}_1$ are orthogonal because they are eigenvectors for different eigenvectors (or one of them is $\mathbf{0}$). By the Pythagorean Theorem

$$\mathbf{x}^T\mathbf{x}=\|\mathbf{x}\|^2=\|\sqrt{1-\alpha}\,\mathbf{u}_n\|^2+\|\sqrt{\alpha}\,\mathbf{u}_1\|^2=|1-\alpha|\|\mathbf{u}_n\|^2+|\alpha|\|\mathbf{u}_1\|^2=(1-\alpha)+\alpha=1$$

since $\mathbf{u}_n$ and $\mathbf{u}_1$ are unit vectors and $0\le\alpha\le 1$. Also, since $\mathbf{u}_n$ and $\mathbf{u}_1$ are orthogonal,

$$\begin{aligned}\mathbf{x}^T A\mathbf{x}&=(\sqrt{1-\alpha}\,\mathbf{u}_n+\sqrt{\alpha}\,\mathbf{u}_1)^T A(\sqrt{1-\alpha}\,\mathbf{u}_n+\sqrt{\alpha}\,\mathbf{u}_1)\\&=(\sqrt{1-\alpha}\,\mathbf{u}_n+\sqrt{\alpha}\,\mathbf{u}_1)^T(m\sqrt{1-\alpha}\,\mathbf{u}_n+M\sqrt{\alpha}\,\mathbf{u}_1)\\&=|1-\alpha|\,m\mathbf{u}_n^T\mathbf{u}_n+|\alpha|\,M\mathbf{u}_1^T\mathbf{u}_1=(1-\alpha)m+\alpha M=t\end{aligned}$$

Thus the quadratic form $\mathbf{x}^T A\mathbf{x}$ assumes every value between m and M for a suitable unit vector $\boldsymbol{x}$.

14. **[M]** The matrix of the quadratic form is $A=\begin{bmatrix}0 & 1/2 & 3/2 & 15\\ 1/2 & 0 & 15 & 3/2\\ 3/2 & 15 & 0 & 1/2\\ 15 & 3/2 & 1/2 & 0\end{bmatrix}$. The eigenvalues of A are $\lambda_1=17,\ \lambda_2=13,\ \lambda_3=-14,$ and $\lambda_4=-16$.

(a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ is the greatest eigenvalue λ_1 of A, which is 17.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ is an eigenvector corresponding to $\lambda_1=17$, so $\mathbf{u}=\pm\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x}=1$ and $\mathbf{x}^T\mathbf{u}=0$ is the second greatest eigenvalue λ_2 of A, which is 13.

15. **[M]** The matrix of the quadratic form is $A = \begin{bmatrix} 0 & 3/2 & 5/2 & 7/2 \\ 3/2 & 0 & 7/2 & 5/2 \\ 5/2 & 7/2 & 0 & 3/2 \\ 7/2 & 5/2 & 3/2 & 0 \end{bmatrix}$. The eigenvalues of A are $\lambda_1 = 15/2,\ \lambda_2 = -1/2,\ \lambda_3 = -5/2,$ and $\lambda_4 = -9/2.$

(a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ is the greatest eigenvalue λ_1 of A, which is 15/2.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ occurs at a unit eigenvector **u** corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 15/2$, so $\mathbf{u} = \pm\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$ is the second greatest eigenvalue λ_2 of A, which is –1/2.

16. **[M]** The matrix of the quadratic form is $A = \begin{bmatrix} 4 & -3 & -5 & -5 \\ -3 & 0 & -3 & -3 \\ -5 & -3 & 0 & -1 \\ -5 & -3 & -1 & 0 \end{bmatrix}$. The eigenvalues of A are $\lambda_1 = 9$, $\lambda_2 = 3,\ \lambda_3 = 1,$ and $\lambda_4 = -9.$

(a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ is the greatest eigenvalue λ_1 of A, which is 9.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$ occurs at a unit eigenvector **u** corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 9$, so $\mathbf{u} = \pm\begin{bmatrix} -2/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$ is the second greatest eigenvalue λ_2 of A, which is 3.

17. **[M]** The matrix of the quadratic form is $A=\begin{bmatrix} -6 & -2 & -2 & -2 \\ -2 & -10 & 0 & 0 \\ -2 & 0 & -13 & 3 \\ -2 & 0 & 3 & -13 \end{bmatrix}$. The eigenvalues of A are $\lambda_1=-4$, $\lambda_2=-10$, $\lambda_3=-12$, and $\lambda_4=-16$.

(a) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ is the greatest eigenvalue λ_1 of A, which is –4.

(b) By Theorem 6, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x}=1$ occurs at a unit eigenvector $\mathbf{u}$ corresponding to the greatest eigenvalue λ_1 of A. One may compute that $\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1=-4$, so $\mathbf{u}=\pm\begin{bmatrix} -3/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \end{bmatrix}$.

(c) By Theorem 7, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints $\mathbf{x}^T\mathbf{x}=1$ and $\mathbf{x}^T\mathbf{u}=0$ is the second greatest eigenvalue λ_2 of A, which is –10.

7.4 SOLUTIONS

Notes: The section presents a modern topic of great importance in applications, particularly in computer calculations. An understanding of the singular value decomposition is essential for advanced work in science and engineering that requires matrix computations. Moreover, the singular value decomposition explains much about the structure of matrix transformations. The SVD does for an arbitrary matrix almost what an orthogonal decomposition does for a symmetric matrix.

1. Let $A=\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$. Then $A^T A=\begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}$, and the eigenvalues of $A^T A$ are seen to be (in decreasing order) $\lambda_1=9$ and $\lambda_2=1$. Thus the singular values of A are $\sigma_1=\sqrt{9}=3$ and $\sigma_2=\sqrt{1}=1$.

2. Let $A=\begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$. Then $A^T A=\begin{bmatrix} 25 & 0 \\ 0 & 0 \end{bmatrix}$, and the eigenvalues of $A^T A$ are seen to be (in decreasing order) $\lambda_1=25$ and $\lambda_2=0$. Thus the singular values of A are $\sigma_1=\sqrt{25}=5$ and $\sigma_2=\sqrt{0}=0$.

3. Let $A=\begin{bmatrix} \sqrt{6} & 1 \\ 0 & \sqrt{6} \end{bmatrix}$. Then $A^T A=\begin{bmatrix} 6 & \sqrt{6} \\ \sqrt{6} & 7 \end{bmatrix}$, and the characteristic polynomial of $A^T A$ is $\lambda^2-13\lambda+36=(\lambda-9)(\lambda-4)$, and the eigenvalues of $A^T A$ are (in decreasing order) $\lambda_1=9$ and $\lambda_2=4$. Thus the singular values of A are $\sigma_1=\sqrt{9}=3$ and $\sigma_2=\sqrt{4}=2$.

4. Let $A=\begin{bmatrix}\sqrt{3} & 2\\ 0 & \sqrt{3}\end{bmatrix}$. Then $A^TA=\begin{bmatrix}3 & 2\sqrt{3}\\ 2\sqrt{3} & 7\end{bmatrix}$, and the characteristic polynomial of A^TA is $\lambda^2-10\lambda+9=(\lambda-9)(\lambda-1)$, and the eigenvalues of A^TA are (in decreasing order) $\lambda_1=9$ and $\lambda_2=1$. Thus the singular values of *A* are $\sigma_1=\sqrt{9}=3$ and $\sigma_2=\sqrt{1}=1$.

5. Let $A=\begin{bmatrix}-3 & 0\\ 0 & 0\end{bmatrix}$. Then $A^TA=\begin{bmatrix}9 & 0\\ 0 & 0\end{bmatrix}$, and the eigenvalues of A^TA are seen to be (in decreasing order) $\lambda_1=9$ and $\lambda_2=0$. Associated unit eigenvectors may be computed:

$$\lambda=9:\begin{bmatrix}1\\0\end{bmatrix},\lambda=0:\begin{bmatrix}0\\1\end{bmatrix}$$

Thus one choice for *V* is $V=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$. The singular values of *A* are $\sigma_1=\sqrt{9}=3$ and $\sigma_2=\sqrt{0}=0$. Thus the matrix Σ is $\Sigma=\begin{bmatrix}3 & 0\\ 0 & 0\end{bmatrix}$. Next compute

$$\mathbf{u}_1=\frac{1}{\sigma_1}A\mathbf{v}_1=\begin{bmatrix}-1\\0\end{bmatrix}$$

Because $A\mathbf{v}_2=\mathbf{0}$, the only column found for *U* so far is $\mathbf{u}_1$. Find the other column of *U* is found by extending $\{\mathbf{u}_1\}$ to an orthonormal basis for $\mathbb{R}^2$. An easy choice is $\mathbf{u}_2=\begin{bmatrix}0\\1\end{bmatrix}$.

Let $U=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$. Thus

$$A=U\Sigma V^T=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

6. Let $A=\begin{bmatrix}-2 & 0\\ 0 & -1\end{bmatrix}$. Then $A^TA=\begin{bmatrix}4 & 0\\ 0 & 1\end{bmatrix}$, and the eigenvalues of A^TA are seen to be (in decreasing order) $\lambda_1=4$ and $\lambda_2=1$. Associated unit eigenvectors may be computed:

$$\lambda=4:\begin{bmatrix}1\\0\end{bmatrix},\lambda=1:\begin{bmatrix}0\\1\end{bmatrix}$$

Thus one choice for *V* is $V=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$. The singular values of *A* are $\sigma_1=\sqrt{4}=2$ and $\sigma_2=\sqrt{1}=1$. Thus the matrix Σ is $\Sigma=\begin{bmatrix}2 & 0\\ 0 & 1\end{bmatrix}$. Next compute

$$\mathbf{u}_1=\frac{1}{\sigma_1}A\mathbf{v}_1=\begin{bmatrix}-1\\0\end{bmatrix},\mathbf{u}_2=\frac{1}{\sigma_2}A\mathbf{v}_2=\begin{bmatrix}0\\-1\end{bmatrix}$$

Since $\{\mathbf{u}_1,\mathbf{u}_2\}$ is a basis for $\mathbb{R}^2$, let $U=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$. Thus

$$A=U\Sigma V^T=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}2 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

7. Let $A=\begin{bmatrix}2 & -1\\ 2 & 2\end{bmatrix}$. Then $A^TA=\begin{bmatrix}8 & 2\\ 2 & 5\end{bmatrix}$, and the characteristic polynomial of A^TA is $\lambda^2-13\lambda+36=(\lambda-9)(\lambda-4)$, and the eigenvalues of A^TA are (in decreasing order) $\lambda_1=9$ and $\lambda_2=4$. Associated unit eigenvectors may be computed:

$$\lambda=9:\begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix},\lambda=4:\begin{bmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}$$

Thus one choice for V is $V=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$. The singular values of A are $\sigma_1=\sqrt{9}=3$ and $\sigma_2=\sqrt{4}=2$. Thus the matrix Σ is $\Sigma=\begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix}$. Next compute

$$\mathbf{u}_1=\frac{1}{\sigma_1}A\mathbf{v}_1=\begin{bmatrix}1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix},\mathbf{u}_2=\frac{1}{\sigma_2}A\mathbf{v}_2=\begin{bmatrix}-2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}$$

Since $\{\mathbf{u}_1,\mathbf{u}_2\}$ is a basis for $\mathbb{R}^2$, let $U=\begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}$. Thus

$$A=U\,\Sigma\,V^T=\begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$$

8. Let $A=\begin{bmatrix}2 & 3\\ 0 & 2\end{bmatrix}$. Then $A^TA=\begin{bmatrix}4 & 6\\ 6 & 13\end{bmatrix}$, and the characteristic polynomial of A^TA is $\lambda^2-17\lambda+16=(\lambda-16)(\lambda-1)$, and the eigenvalues of A^TA are (in decreasing order) $\lambda_1=16$ and $\lambda_2=1$. Associated unit eigenvectors may be computed:

$$\lambda=16:\begin{bmatrix}1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix},\lambda=1:\begin{bmatrix}-2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}$$

Thus one choice for V is $V=\begin{bmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}$. The singular values of A are $\sigma_1=\sqrt{16}=4$ and $\sigma_2=\sqrt{1}=1$. Thus the matrix Σ is $\Sigma=\begin{bmatrix}4 & 0\\ 0 & 1\end{bmatrix}$. Next compute

$$\mathbf{u}_1=\frac{1}{\sigma_1}A\mathbf{v}_1=\begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix},\mathbf{u}_2=\frac{1}{\sigma_2}A\mathbf{v}_2=\begin{bmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}$$

Since $\{\mathbf{u}_1,\mathbf{u}_2\}$ is a basis for $\mathbb{R}^2$, let $U=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$. Thus

$$A=U\,\Sigma\,V^T=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1/\sqrt{5} & 2/\sqrt{5}\\ -2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}$$

9. Let $A=\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$. Then $A^T A=\begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$, and the characteristic polynomial of $A^T A$ is $\lambda^2 - 100\lambda + 900 = (\lambda - 90)(\lambda - 10)$, and the eigenvalues of $A^T A$ are (in decreasing order) $\lambda_1 = 90$ and $\lambda_2 = 10$. Associated unit eigenvectors may be computed:

$$\lambda = 90: \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}, \lambda = 10: \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Thus one choice for V is $V=\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$. The singular values of A are $\sigma_1 = \sqrt{90} = 3\sqrt{10}$ and $\sigma_2 = \sqrt{10}$. Thus the matrix Σ is $\Sigma=\begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$$

Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is not a basis for $\mathbb{R}^3$, we need a unit vector $\mathbf{u}_3$ that is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$. The vector $\mathbf{u}_3$ must satisfy the set of equations $\mathbf{u}_1^T\mathbf{x} = 0$ and $\mathbf{u}_2^T\mathbf{x} = 0$. These are equivalent to the linear equations

$$\begin{aligned} x_1 + 0x_2 + x_3 &= 0 \\ -x_1 + 0x_2 + x_3 &= 0 \end{aligned}, \text{ so } \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Therefore let $U=\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$. Thus

$$A = U\Sigma V^T = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

10. Let $A=\begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$. Then $A^T A=\begin{bmatrix} 20 & -10 \\ -10 & 5 \end{bmatrix}$, and the characteristic polynomial of $A^T A$ is $\lambda^2 - 25\lambda = \lambda(\lambda - 25)$, and the eigenvalues of $A^T A$ are (in decreasing order) $\lambda_1 = 25$ and $\lambda_2 = 0$. Associated unit eigenvectors may be computed:

$$\lambda = 25: \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}, \lambda = 0: \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Thus one choice for V is $V=\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$. The singular values of A are $\sigma_1=\sqrt{25}=5$ and

$\sigma_2=\sqrt{0}=0$. Thus the matrix Σ is $\Sigma=\begin{bmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1=\frac{1}{\sigma_1}A\mathbf{v}_1=\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}$$

Because $A\mathbf{v}_2=\mathbf{0}$, the only column found for U so far is $\mathbf{u}_1$. Find the other columns of U found by extending $\{\mathbf{u}_1\}$ to an orthonormal basis for $\mathbb{R}^3$. In this case, we need two orthogonal unit vectors $\mathbf{u}_2$ and $\mathbf{u}_3$ that are orthogonal to $\mathbf{u}_1$. Each vector must satisfy the equation $\mathbf{u}_1^T\mathbf{x}=0$, which is equivalent to the equation $2x_1+x_2=0$. An orthonormal basis for the solution set of this equation is

$$\mathbf{u}_2=\begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \\ 0 \end{bmatrix}, \mathbf{u}_3=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, let $U=\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} & 0 \\ 1/\sqrt{5} & -2/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Thus

$$A=U\,\Sigma\,V^T=\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} & 0 \\ 1/\sqrt{5} & -2/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

11. Let $A=\begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$. Then $A^TA=\begin{bmatrix} 81 & -27 \\ -27 & 9 \end{bmatrix}$, and the characteristic polynomial of A^TA is

$\lambda^2-90\lambda=\lambda(\lambda-90)$, and the eigenvalues of A^TA are (in decreasing order) $\lambda_1=90$ and $\lambda_2=0$. Associated unit eigenvectors may be computed:

$$\lambda=90:\begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{bmatrix}, \lambda=0:\begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix}.$$

Thus one choice for V is $V=\begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix}$. The singular values of A are $\sigma_1=\sqrt{90}=3\sqrt{10}$ and

$\sigma_2=\sqrt{0}=0$. Thus the matrix Σ is $\Sigma=\begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1=\frac{1}{\sigma_1}A\mathbf{v}_1=\begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Because $A\mathbf{v}_2 = \mathbf{0}$, the only column found for U so far is $\mathbf{u}_1$. The other columns of U can be found by extending $\{\mathbf{u}_1\}$ to an orthonormal basis for $\mathbb{R}^3$. In this case, we need two orthogonal unit vectors $\mathbf{u}_2$ and $\mathbf{u}_3$ that are orthogonal to $\mathbf{u}_1$. Each vector must satisfy the equation $\mathbf{u}_1^T\mathbf{x} = 0$, which is equivalent to the equation $-x_1 + 2x_2 + 2x_3 = 0.$ An orthonormal basis for the solution set of this equation is

$$\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix}.$$

Therefore, let $U = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$. Thus

$$A = U\Sigma V^T = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix}$$

12. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, and the eigenvalues of A^TA are seen to be (in decreasing order) $\lambda_1 = 3$ and $\lambda_2 = 2.$ Associated unit eigenvectors may be computed:

$$\lambda = 3: \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \lambda = 2: \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Thus one choice for V is $V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The singular values of A are $\sigma_1 = \sqrt{3}$ and $\sigma_2 = \sqrt{2}$. Thus the matrix Σ is $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sigma_2}A\mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is not a basis for $\mathbb{R}^3$, we need a unit vector $\mathbf{u}_3$ that is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$. The vector $\mathbf{u}_3$ must satisfy the set of equations $\mathbf{u}_1^T\mathbf{x} = 0$ and $\mathbf{u}_2^T\mathbf{x} = 0$. These are equivalent to the linear equations

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + 0x_2 - x_3 &= 0 \end{aligned}, \text{so } \mathbf{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \text{and } \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

Therefore let $U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$. Thus

$$A = U\,\Sigma V^T = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

13. Let $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. Then $A^T = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^{TT}A^T = AA^T = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$, and the eigenvalues of $A^{TT}A^T$ are seen to be (in decreasing order) $\lambda_1 = 25$ and $\lambda_2 = 9$. Associated unit eigenvectors may be computed:

$$\lambda = 25: \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \lambda = 9: \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Thus one choice for V is $V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. The singular values of A^T are $\sigma_1 = \sqrt{25} = 5$ and $\sigma_2 = \sqrt{9} = 3$. Thus the matrix Σ is $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A^T \mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sigma_2} A^T \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 1/\sqrt{18} \\ -4/\sqrt{18} \end{bmatrix}$$

Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is not a basis for $\mathbb{R}^3$, we need a unit vector $\mathbf{u}_3$ that is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$. The vector $\mathbf{u}_3$ must satisfy the set of equations $\mathbf{u}_1^T\mathbf{x} = 0$ and $\mathbf{u}_2^T\mathbf{x} = 0$. These are equivalent to the linear equations

$$\begin{aligned} x_1 + x_2 + 0x_3 &= 0 \\ -x_1 + x_2 - 4x_3 &= 0 \end{aligned}, \text{ so } \mathbf{x} = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \text{ and } \mathbf{u}_3 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$

Therefore let $U = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \\ 0 & -4/\sqrt{18} & 1/3 \end{bmatrix}$. Thus

$$A^T = U\,\Sigma V^T = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \\ 0 & -4/\sqrt{18} & 1/3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

An SVD for A is computed by taking transposes:

$$A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{18} & 1/\sqrt{18} & -4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$$

14. From Exercise 7, $A = U\Sigma V^T$ with $V = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$. Since the first column of V is unit eigenvector associated with the greatest eigenvalue λ_1 of $A^T A$, so the first column of V is a unit vector at which $\| A\mathbf{x} \|$ is maximized.

15. **a**. Since A has 2 nonzero singular values, rank $A = 2$.

b. By Example 6, $\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{ \begin{bmatrix} .40 \\ .37 \\ -.84 \end{bmatrix}, \begin{bmatrix} -.78 \\ -.33 \\ -.52 \end{bmatrix} \right\}$ is a basis for Col A and $\{\mathbf{v}_3\} = \left\{ \begin{bmatrix} .58 \\ -.58 \\ .58 \end{bmatrix} \right\}$ is a basis for Nul A.

16. **a**. Since A has 2 nonzero singular values, rank $A = 2$.

b. By Example 6, $\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{ \begin{bmatrix} -.86 \\ .31 \\ .41 \end{bmatrix}, \begin{bmatrix} -.11 \\ .68 \\ -.73 \end{bmatrix} \right\}$ is a basis for Col A and $\{\mathbf{v}_3, \mathbf{v}_4\} = \left\{ \begin{bmatrix} .65 \\ .08 \\ -.16 \\ -.73 \end{bmatrix}, \begin{bmatrix} -.34 \\ .42 \\ -.84 \\ -.08 \end{bmatrix} \right\}$ is a basis for Nul A.

17. Let $A = U\Sigma V^T = U\Sigma V^{-1}$. Since A is square and invertible, rank $A = n$, and all of the entries on the diagonal of Σ must be nonzero. So $A^{-1} = (U\Sigma V^{-1})^{-1} = V\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$.

18. First note that the determinant of an orthogonal matrix is ± 1, because $1 = \det I = \det U^T U = (\det U^T)(\det U) = (\det U)^2$. Suppose that A is square and $A = U\Sigma V^T$. Then Σ is square, and $\det A = (\det U)(\det \Sigma)(\det V^T) = \pm\det\Sigma = \pm\sigma_1 \dots \sigma_n$.

19. Since U and V are orthogonal matrices,

$$A^T A = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T\Sigma)V^T = V(\Sigma^T\Sigma)V^{-1}$$

If $\sigma_1, \dots, \sigma_r$ are the diagonal entries in Σ, then $\Sigma^T\Sigma$ is a diagonal matrix with diagonal entries $\sigma_1^2, \dots, \sigma_r^2$ and possibly some zeros. Thus V diagonalizes $A^T A$ and the columns of V are eigenvectors of $A^T A$ by the Diagonalization Theorem in Section 5.3. Likewise

$$AA^T = U\Sigma V^T(U\Sigma V^T)^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T = U(\Sigma\Sigma^T)U^{-1}$$

so U diagonalizes AA^T and the columns of U must be eigenvectors of AA^T. Moreover, the Diagonalization Theorem states that $\sigma_1^2, \dots, \sigma_r^2$ are the nonzero eigenvalues of $A^T A$. Hence $\sigma_1, \dots, \sigma_r$ are the nonzero singular values of A.

20. If A is positive definite, then $A = PDP^T$, where P is an orthogonal matrix and D is a diagonal matrix. The diagonal entries of D are positive because they are the eigenvalues of a positive definite matrix. Since P is an orthogonal matrix, $PP^T = I$ and the square matrix P^T is invertible. Moreover,

$(P^T)^{-1} = (P^{-1})^{-1} = P = (P^T)^T$, so P^T is an orthogonal matrix. Thus the factorization $A = PDP^T$ has the properties that make it a singular value decomposition.

21. Let $A = U\Sigma V^T$. The matrix PU is orthogonal, because P and U are both orthogonal. (See Exercise 29 in Section 6.2). So the equation $PA = (PU)\Sigma V^T$ has the form required for a singular value decomposition. By Exercise 19, the diagonal entries in Σ are the singular values of PA.

22. The right singular vector $\mathbf{v}_1$ is an eigenvector for the largest eigenvector λ_1 of $A^T A$. By Theorem 7 in Section 7.3, the second largest eigenvalue λ_2 is the maximum of $\mathbf{x}^T(A^T A)\mathbf{x}$ over all unit vectors orthogonal to $\mathbf{v}_1$. Since $\mathbf{x}^T(A^T A)\mathbf{x} = \|A\mathbf{x}\|^2$, the square root of λ_2, which is the second largest singular value of A, is the maximum of $\|A\mathbf{x}\|$ over all unit vectors orthogonal to $\mathbf{v}_1$.

23. From the proof of Theorem 10, $U\Sigma = \begin{bmatrix}\sigma_1\mathbf{u}_1 & \dots & \sigma_r\mathbf{u}_r & \mathbf{0} & \dots & \mathbf{0}\end{bmatrix}$. The column-row expansion of the product $(U\Sigma)V^T$ shows that

$$A = (U\Sigma)V^T = (U\Sigma)\begin{bmatrix}\mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T\end{bmatrix} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \dots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T$$

where r is the rank of A.

24. From Exercise 23, $A^T = \sigma_1\mathbf{v}_1\mathbf{u}_1^T + \dots + \sigma_r\mathbf{v}_r\mathbf{u}_r^T$. Then since $\mathbf{u}_i^T\mathbf{u}_j = \begin{cases}0 & \text{for } i \neq j \\ 1 & \text{for } i = j\end{cases}$,

$$A^T\mathbf{u}_j = (\sigma_1\mathbf{v}_1\mathbf{u}_1^T + \dots + \sigma_r\mathbf{v}_r\mathbf{u}_r^T)\mathbf{u}_j = (\sigma_j\mathbf{v}_j\mathbf{u}_j^T)\mathbf{u}_j = \sigma_j\mathbf{v}_j(\mathbf{u}_j^T\mathbf{u}_j) = \sigma_j\mathbf{v}_j$$

25. Consider the SVD for the standard matrix A of T, say $A = U\Sigma V^T$. Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $C = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ be bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ constructed respectively from the columns of V and U. Since the columns of V are orthogonal, $V^T\mathbf{v}_j = \mathbf{e}_j$, where $\mathbf{e}_j$ is the jth column of the $n \times n$ identity matrix. To find the matrix of T relative to B and C, compute

$$T(\mathbf{v}_j) = A\mathbf{v}_j = U\Sigma V^T\mathbf{v}_j = U\Sigma\mathbf{e}_j = U\sigma_j\mathbf{e}_j = \sigma_j U\mathbf{e}_j = \sigma_j\mathbf{u}_j$$

so $[T(\mathbf{v}_j)]_C = \sigma_j\mathbf{e}_j$. Formula (4) in the discussion at the beginning of Section 5.4 shows that the "diagonal" matrix Σ is the matrix of T relative to B and C.

26. **[M]** Let $A = \begin{bmatrix}-18 & 13 & -4 & 4 \\ 2 & 19 & -4 & 12 \\ -14 & 11 & -12 & 8 \\ -2 & 21 & 4 & 8\end{bmatrix}$. Then $A^T A = \begin{bmatrix}528 & -392 & 224 & -176 \\ -392 & 1092 & -176 & 536 \\ 224 & -176 & 192 & -128 \\ -176 & 536 & -128 & 288\end{bmatrix}$, and the eigenvalues of $A^T A$ are found to be (in decreasing order) $\lambda_1 = 1600$, $\lambda_2 = 400$, $\lambda_3 = 100$, and $\lambda_4 = 0$. Associated unit eigenvectors may be computed:

$$\lambda_1 : \begin{bmatrix}-.4 \\ .8 \\ -.2 \\ .4\end{bmatrix}, \lambda_2 : \begin{bmatrix}.8 \\ .4 \\ .4 \\ .2\end{bmatrix}, \lambda_3 : \begin{bmatrix}.4 \\ -.2 \\ -.8 \\ .4\end{bmatrix}, \lambda_4 : \begin{bmatrix}-.2 \\ -.4 \\ .4 \\ .8\end{bmatrix}$$

Thus one choice for V is $V = \begin{bmatrix} -.4 & .8 & .4 & -.2 \\ .8 & .4 & -.2 & -.4 \\ -.2 & .4 & -.8 & .4 \\ .4 & .2 & .4 & .8 \end{bmatrix}$. The singular values of A are $\sigma_1 = 40$, $\sigma_1 = 20$, $\sigma_3 = 10$, and $\sigma_4 = 0$. Thus the matrix Σ is $\Sigma = \begin{bmatrix} 40 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \begin{bmatrix} .5 \\ .5 \\ .5 \\ .5 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \begin{bmatrix} -.5 \\ .5 \\ -.5 \\ .5 \end{bmatrix},$$

$$\mathbf{u}_3 = \frac{1}{\sigma_3} A\mathbf{v}_3 = \begin{bmatrix} -.5 \\ .5 \\ .5 \\ -.5 \end{bmatrix}$$

Because $A\mathbf{v}_4 = \mathbf{0}$, only three columns of U have been found so far. The last column of U can be found by extending $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ to an orthonormal basis for $\mathbb{R}^4$. The vector $\mathbf{u}_4$ must satisfy the set of equations $\mathbf{u}_1^T\mathbf{x} = 0$, $\mathbf{u}_2^T\mathbf{x} = 0$, and $\mathbf{u}_3^T\mathbf{x} = 0$. These are equivalent to the linear equations

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ -x_1 + x_2 - x_3 + x_4 &= 0 \\ -x_1 + x_2 + x_3 - x_4 &= 0 \end{aligned}, \text{ so } \mathbf{x} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \text{ and } \mathbf{u}_4 = \begin{bmatrix} -.5 \\ -.5 \\ .5 \\ .5 \end{bmatrix}.$$

Therefore, let $U = \begin{bmatrix} .5 & -.5 & -.5 & -.5 \\ .5 & .5 & .5 & -.5 \\ .5 & -.5 & .5 & .5 \\ .5 & .5 & -.5 & .5 \end{bmatrix}$. Thus

$$A = U\Sigma V^T = \begin{bmatrix} .5 & -.5 & -.5 & -.5 \\ .5 & .5 & .5 & -.5 \\ .5 & -.5 & .5 & .5 \\ .5 & .5 & -.5 & .5 \end{bmatrix} \begin{bmatrix} 40 & 0 & 0 & 0 \\ 0 & 20 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -.4 & .8 & -.2 & .4 \\ .8 & .4 & .4 & .2 \\ .4 & -.2 & -.8 & .4 \\ -.2 & -.4 & .4 & .8 \end{bmatrix}$$

27. **[M]** Let $A = \begin{bmatrix} 6 & -8 & -4 & 5 & -4 \\ 2 & 7 & -5 & -6 & 4 \\ 0 & -1 & -8 & 2 & 2 \\ -1 & -2 & 4 & 4 & -8 \end{bmatrix}$. Then $A^TA = \begin{bmatrix} 41 & -32 & -38 & 14 & -8 \\ -32 & 118 & -3 & -92 & 74 \\ -38 & -3 & 121 & 10 & -52 \\ 14 & -92 & 10 & 81 & -72 \\ -8 & 74 & -52 & -72 & 100 \end{bmatrix}$, and the eigenvalues of A^TA are found to be (in decreasing order) $\lambda_1 = 270.87$, $\lambda_2 = 147.85$, $\lambda_3 = 23.73$, $\lambda_4 = 18.55$, and $\lambda_5 = 0$. Associated unit eigenvectors may be computed:

$$\lambda_1 : \begin{bmatrix} -.10 \\ .61 \\ -.21 \\ -.52 \\ .55 \end{bmatrix}, \lambda_2 : \begin{bmatrix} -.39 \\ .29 \\ .84 \\ -.14 \\ -.19 \end{bmatrix}, \lambda_3 : \begin{bmatrix} -.74 \\ -.27 \\ -.07 \\ .38 \\ .49 \end{bmatrix}, \lambda_4 : \begin{bmatrix} .41 \\ -.50 \\ .45 \\ -.23 \\ .58 \end{bmatrix}, \lambda_5 : \begin{bmatrix} -.36 \\ -.48 \\ -.19 \\ -.72 \\ -.29 \end{bmatrix}$$

Thus one choice for V is $V = \begin{bmatrix} -.10 & -.39 & -.74 & .41 & -.36 \\ .61 & .29 & -.27 & -.50 & -.48 \\ -.21 & .84 & -.07 & .45 & -.19 \\ -.52 & -.14 & .38 & -.23 & -.72 \\ .55 & -.19 & .49 & .58 & -.29 \end{bmatrix}$. The nonzero singular values of A are

$\sigma_1 = 16.46,\ \sigma_1 = 12.16,\ \sigma_3 = 4.87,$ and $\sigma_4 = 4.31.$ Thus the matrix Σ is

$\Sigma = \begin{bmatrix} 16.46 & 0 & 0 & 0 & 0 \\ 0 & 12.16 & 0 & 0 & 0 \\ 0 & 0 & 4.87 & 0 & 0 \\ 0 & 0 & 0 & 4.31 & 0 \end{bmatrix}$. Next compute

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \begin{bmatrix} -.57 \\ .63 \\ .07 \\ -.51 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \begin{bmatrix} -.65 \\ -.24 \\ -.63 \\ .34 \end{bmatrix},$$

$$\mathbf{u}_3 = \frac{1}{\sigma_3} A\mathbf{v}_3 = \begin{bmatrix} -.42 \\ -.68 \\ .53 \\ -.29 \end{bmatrix}, \mathbf{u}_4 = \frac{1}{\sigma_4} A\mathbf{v}_4 = \begin{bmatrix} .27 \\ -.29 \\ -.56 \\ -.73 \end{bmatrix}$$

Since $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ is a basis for $\mathbb{R}^4$, let $U = \begin{bmatrix} -.57 & -.65 & -.42 & .27 \\ .63 & -.24 & -.68 & -.29 \\ .07 & -.63 & .53 & -.56 \\ -.51 & .34 & -.29 & -.73 \end{bmatrix}$. Thus

$$A = U\Sigma V^T$$

$$= \begin{bmatrix} -.57 & -.65 & -.42 & .27 \\ .63 & -.24 & -.68 & -.29 \\ .07 & -.63 & .53 & -.56 \\ -.51 & .34 & -.29 & -.73 \end{bmatrix} \begin{bmatrix} 16.46 & 0 & 0 & 0 & 0 \\ 0 & 12.16 & 0 & 0 & 0 \\ 0 & 0 & 4.87 & 0 & 0 \\ 0 & 0 & 0 & 4.31 & 0 \end{bmatrix} \begin{bmatrix} -.10 & .61 & -.21 & -.52 & .55 \\ -.39 & .29 & .84 & -.14 & -.19 \\ -.74 & -.27 & -.07 & .38 & .49 \\ .41 & -.50 & .45 & -.23 & .58 \\ -.36 & -.48 & -.19 & -.72 & -.29 \end{bmatrix}$$

28. **[M]** Let $A=\begin{bmatrix}4 & 0 & -7 & -7\\ -6 & 1 & 11 & 9\\ 7 & -5 & 10 & 19\\ -1 & 2 & 3 & -1\end{bmatrix}$. Then $A^TA=\begin{bmatrix}102 & -43 & -27 & 52\\ -43 & 30 & -33 & -88\\ -27 & -33 & 279 & 335\\ 52 & -88 & 335 & 492\end{bmatrix}$, and the eigenvalues of A^TA are found to be (in decreasing order) $\lambda_1=749.9785$, $\lambda_2=146.2009$, $\lambda_3=6.8206$, and $\lambda_4=1.3371\times10^{-6}$. The singular values of A are thus $\sigma_1=27.3857$, $\sigma_2=12.0914$, $\sigma_3=2.61163$, and $\sigma_4=.00115635$. The condition number $\sigma_1/\sigma_4=23{,}683$.

29. **[M]** Let $A=\begin{bmatrix}5 & 3 & 1 & 7 & 9\\ 6 & 4 & 2 & 8 & -8\\ 7 & 5 & 3 & 10 & 9\\ 9 & 6 & 4 & -9 & -5\\ 8 & 5 & 2 & 11 & 4\end{bmatrix}$. Then $A^TA=\begin{bmatrix}255 & 168 & 90 & 160 & 47\\ 168 & 111 & 60 & 104 & 30\\ 90 & 60 & 34 & 39 & 8\\ 160 & 104 & 39 & 415 & 178\\ 47 & 30 & 8 & 178 & 267\end{bmatrix}$, and the eigenvalues of A^TA are found to be (in decreasing order) $\lambda_1=672.589$, $\lambda_2=280.745$, $\lambda_3=127.503$, $\lambda_4=1.163$, and $\lambda_5=1.428\times10^{-7}$. The singular values of A are thus $\sigma_1=25.9343$, $\sigma_2=16.7554$, $\sigma_3=11.2917$, $\sigma_4=1.07853$, and $\sigma_5=.000377928$. The condition number $\sigma_1/\sigma_5=68{,}622$.

7.5 SOLUTIONS

Notes: The application presented here has turned out to be of interest to a wide variety of students, including engineers. I cover this in Course Syllabus 3 described above, but I only have time to mention the idea briefly to my other classes.

1. The matrix of observations is $X=\begin{bmatrix}19 & 22 & 6 & 3 & 2 & 20\\ 12 & 6 & 9 & 15 & 13 & 5\end{bmatrix}$ and the sample mean is $M=\frac{1}{6}\begin{bmatrix}72\\ 60\end{bmatrix}=\begin{bmatrix}12\\ 10\end{bmatrix}$. The mean-deviation form B is obtained by subtracting M from each column of X, so $B=\begin{bmatrix}7 & 10 & -6 & -9 & -10 & 8\\ 2 & -4 & -1 & 5 & 3 & -5\end{bmatrix}$. The sample covariance matrix is

$$S=\frac{1}{6-1}BB^T=\frac{1}{5}\begin{bmatrix}430 & -135\\ -135 & 80\end{bmatrix}=\begin{bmatrix}86 & -27\\ -27 & 16\end{bmatrix}$$

2. The matrix of observations is $X=\begin{bmatrix}1 & 5 & 2 & 6 & 7 & 3\\ 3 & 11 & 6 & 8 & 15 & 11\end{bmatrix}$ and the sample mean is $M=\frac{1}{6}\begin{bmatrix}24\\ 54\end{bmatrix}=\begin{bmatrix}4\\ 9\end{bmatrix}$. The mean-deviation form B is obtained by subtracting M from each column of X, so $B=\begin{bmatrix}-3 & 1 & -2 & 2 & 3 & -1\\ -6 & 2 & -3 & -1 & 6 & 2\end{bmatrix}$. The sample covariance matrix is

$$S=\frac{1}{6-1}BB^T=\frac{1}{5}\begin{bmatrix}28 & 40\\ 40 & 90\end{bmatrix}=\begin{bmatrix}5.6 & 8\\ 8 & 18\end{bmatrix}$$

3. The principal components of the data are the unit eigenvectors of the sample covariance matrix S. One computes that (in descending order) the eigenvalues of $S = \begin{bmatrix} 86 & -27 \\ -27 & 16 \end{bmatrix}$ are $\lambda_1 = 95.2041$ and $\lambda_2 = 6.79593$. One further computes that corresponding eigenvectors are $\mathbf{v}_1 = \begin{bmatrix} -2.93348 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} .340892 \\ 1 \end{bmatrix}$. These vectors may be normalized to find the principal components, which are $\mathbf{u}_1 = \begin{bmatrix} .946515 \\ -.322659 \end{bmatrix}$ for $\lambda_1 = 95.2041$ and $\mathbf{u}_2 = \begin{bmatrix} .322659 \\ .946515 \end{bmatrix}$ for $\lambda_2 = 6.79593$.

4. The principal components of the data are the unit eigenvectors of the sample covariance matrix S. One computes that (in descending order) the eigenvalues of $S = \begin{bmatrix} 5.6 & 8 \\ 8 & 18 \end{bmatrix}$ are $\lambda_1 = 21.9213$ and $\lambda_2 = 1.67874$. One further computes that corresponding eigenvectors are $\mathbf{v}_1 = \begin{bmatrix} .490158 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2.04016 \\ 1 \end{bmatrix}$. These vectors may be normalized to find the principal components, which are $\mathbf{u}_1 = \begin{bmatrix} .44013 \\ .897934 \end{bmatrix}$ for $\lambda_1 = 21.9213$ and $\mathbf{u}_2 = \begin{bmatrix} -.897934 \\ .44013 \end{bmatrix}$ for $\lambda_2 = 1.67874$.

5. **[M]** The largest eigenvalue of $S = \begin{bmatrix} 164.12 & 32.73 & 81.04 \\ 32.73 & 539.44 & 249.13 \\ 81.04 & 249.13 & 189.11 \end{bmatrix}$ is $\lambda_1 = 677.497$, and the first principal component of the data is the unit eigenvector corresponding to λ_1, which is $\mathbf{u}_1 = \begin{bmatrix} .129554 \\ .874423 \\ .467547 \end{bmatrix}$. The fraction of the total variance that is contained in this component is $\lambda_1 / \mathrm{tr}(S) = 677.497/(164.12 + 539.44 + 189.11) = .758956$, so 75.8956% of the variance of the data is contained in the first principal component.

6. **[M]** The largest eigenvalue of $S = \begin{bmatrix} 29.64 & 18.38 & 5.00 \\ 18.38 & 20.82 & 14.06 \\ 5.00 & 14.06 & 29.21 \end{bmatrix}$ is $\lambda_1 = 51.6957$, and the first principal component of the data is the unit eigenvector corresponding to λ_1, which is $\mathbf{u}_1 = \begin{bmatrix} .615525 \\ .599424 \\ .511683 \end{bmatrix}$. Thus one choice for the new variable is $y_1 = .615525x_1 + .599424x_2 + .511683x_3$. The fraction of the total variance that is contained in this component is $\lambda_1 / \mathrm{tr}(S) = 51.6957/(29.64 + 20.82 + 29.21) = .648872$, so 64.8872% of the variance of the data is explained by y_1.

7. Since the unit eigenvector corresponding to $\lambda_1 = 95.2041$ is $\mathbf{u}_1 = \begin{bmatrix} .946515 \\ -.322659 \end{bmatrix}$, one choice for the new variable is $y_1 = .946515x_1 - .322659x_2$. The fraction of the total variance that is contained in this component is $\lambda_1 / \text{tr}(S) = 95.2041/(86+16) = .933374$, so 93.3374% of the variance of the data is explained by y_1.

8. Since the unit eigenvector corresponding to $\lambda_1 = 21.9213$ is $\mathbf{u}_1 = \begin{bmatrix} .44013 \\ .897934 \end{bmatrix}$, one choice for the new variable is $y_1 = .44013x_1 + .897934x_2$. The fraction of the total variance that is contained in this component is $\lambda_1 / \text{tr}(S) = 21.9213/(5.6+18) = .928869$, so 92.8869% of the variance of the data is explained by y_1.

9. The largest eigenvalue of $S = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$ is $\lambda_1 = 9$, and the first principal component of the data is the unit eigenvector corresponding to λ_1, which is $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$. Thus one choice for y is $y = (1/3)x_1 + (2/3)x_2 + (2/3)x_3$, and the variance of y is $\lambda_1 = 9$.

10. **[M]** The largest eigenvalue of $S = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 11 & 4 \\ 2 & 4 & 5 \end{bmatrix}$ is $\lambda_1 = 15$, and the first principal component of the data is the unit eigenvector corresponding to λ_1, which is $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$. Thus one choice for y is $y = (1/\sqrt{6})x_1 + (2/\sqrt{6})x_2 + (1/\sqrt{6})x_3$, and the variance of y is $\lambda_1 = 15$.

11. **a**. If $\mathbf{w}$ is the vector in $\mathbb{R}^N$ with a 1 in each position, then $[\mathbf{X}_1 \ \ \dots \ \ \mathbf{X}_N]\mathbf{w} = \mathbf{X}_1 + \dots + \mathbf{X}_N = \mathbf{0}$ since the $\mathbf{X}_k$ are in mean-deviation form. Then

$$[\mathbf{Y}_1 \ \ \dots \ \ \mathbf{Y}_N]\mathbf{w} = \left[P^T\mathbf{X}_1 \ \ \dots \ \ P^T\mathbf{X}_N\right]\mathbf{w} = P^T[\mathbf{X}_1 \ \ \dots \ \ \mathbf{X}_N]\mathbf{w} = P^T\mathbf{0} = \mathbf{0}$$

Thus $\mathbf{Y}_1 + \dots + \mathbf{Y}_N = \mathbf{0}$, and the $\mathbf{Y}_k$ are in mean-deviation form.

b. By part a., the covariance matrix $S_{\mathbf{Y}}$ of $\mathbf{Y}_1, \dots, \mathbf{Y}_N$ is

$$\begin{aligned} S_{\mathbf{Y}} &= \frac{1}{N-1}[\mathbf{Y}_1 \ \ \dots \ \ \mathbf{Y}_N][\mathbf{Y}_1 \ \ \dots \ \ \mathbf{Y}_N]^T \\ &= \frac{1}{N-1}P^T[\mathbf{X}_1 \ \ \dots \ \ \mathbf{X}_N](P^T[\mathbf{X}_1 \ \ \dots \ \ \mathbf{X}_N])^T \\ &= P^T\left(\frac{1}{N-1}[\mathbf{X}_1 \ \ \dots \ \ \mathbf{X}_N][\mathbf{X}_1 \ \ \dots \ \ \mathbf{X}_N]^T\right)P = P^TSP \end{aligned}$$

since the $\mathbf{X}_k$ are in mean-deviation form.

12. By Exercise 11, the change of variables $\mathbf{X} = P\mathbf{Y}$ changes the covariance matrix S of X into the covariance matrix $P^T SP$ of $\mathbf{Y}$. The total variance of the data as described by $\mathbf{Y}$ is $\text{tr}(P^T SP)$. However, since $P^T SP$ is similar to S, they have the same trace (by Exercise 25 in Section 5.4). Thus the total variance of the data is unchanged by the change of variables $\mathbf{X} = P\mathbf{Y}$.

13. Let $\mathbf{M}$ be the sample mean for the data, and let $\hat{\mathbf{X}}_k = \mathbf{X}_k - \mathbf{M}$. Let $B = \begin{bmatrix} \hat{\mathbf{X}}_1 & \dots & \hat{\mathbf{X}}_N \end{bmatrix}$ be the matrix of observations in mean-deviation form. By the row-column expansion of BB^T, the sample covariance matrix is

$$\begin{aligned} S &= \frac{1}{N-1} BB^T \\ &= \frac{1}{N-1} \begin{bmatrix} \hat{\mathbf{X}}_1 & \dots & \hat{\mathbf{X}}_N \end{bmatrix} \begin{bmatrix} \hat{\mathbf{X}}_1^T \\ \vdots \\ \hat{\mathbf{X}}_N^T \end{bmatrix} \\ &= \frac{1}{N-1} \sum_{k=1}^{N} \hat{\mathbf{X}}_k \hat{\mathbf{X}}_k^T = \frac{1}{N-1} \sum_{k=1}^{N} (\mathbf{X}_k - \mathbf{M})(\mathbf{X}_k - \mathbf{M})^T \end{aligned}$$

Chapter 7 SUPPLEMENTARY EXERCISES

1. **a**. True. This is just part of Theorem 2 in Section 7.1. The proof appears just before the statement of the theorem.

b. False. A counterexample is $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

c. True. This is proved in the first part of the proof of Theorem 6 in Section 7.3. It is also a consequence of Theorem 7 in Section 6.2.

d. False. The principal axes of $\mathbf{x}^T A\mathbf{x}$ are the columns of any *orthogonal* matrix P that diagonalizes A. *Note*: When A has an eigenvalue whose eigenspace has dimension greater than 1, the principal axes are not uniquely determined.

e. False. A counterexample is $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. The columns here are orthogonal but not orthonormal.

f. False. See Example 6 in Section 7.2.

g. False. A counterexample is $A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then $\mathbf{x}^T A\mathbf{x} = 2 > 0$, but $\mathbf{x}^T A\mathbf{x}$ is an indefinite quadratic form.

h. True. This is basically the Principal Axes Theorem from Section 7.2. Any quadratic form can be written as $\mathbf{x}^T A\mathbf{x}$ for some symmetric matrix A.

i. False. See Example 3 in Section 7.3.

j. False. The maximum value must be computed over the set of *unit* vectors. Without a restriction on the norm of $\mathbf{x}$, the values of $\mathbf{x}^T A\mathbf{x}$ can be made as large as desired.

k. False. Any orthogonal change of variable $\mathbf{x} = P\mathbf{y}$ changes a positive definite quadratic form into another positive definite quadratic form. Proof: By Theorem 5 of Section 7.2., the classification of a quadratic form is determined by the eigenvalues of the matrix of the form. Given a form $\mathbf{x}^T A\mathbf{x}$, the matrix of the new quadratic form is $P^{-1}AP$, which is similar to A and thus has the same eigenvalues as A.

l. False. The term "definite eigenvalue" is undefined and therefore meaningless.

m. True. If $\mathbf{x} = P\mathbf{y}$, then $\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T P^{-1} AP\mathbf{y}$.

n. False. A counterexample is $U = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$. The columns of U must be *orthonormal* to make $UU^T\mathbf{x}$ the orthogonal projection of $\mathbf{x}$ onto Col U.

o. True. This follows from the discussion in Example 2 of Section 7.4., which refers to a proof given in Example 1.

p. True. Theorem 10 in Section 7.4 writes the decomposition in the form $U\Sigma V^T$, where U and V are orthogonal matrices. In this case, V^T is also an orthogonal matrix. Proof: Since V is orthogonal, V is invertible and $V^{-1} = V^T$. Then $(V^T)^{-1} = (V^{-1})^T = (V^T)^T$, and since V is square and invertible, V^T is an orthogonal matrix.

q. False. A counterexample is $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. The singular values of A are 2 and 1, but the singular values of $A^T A$ are 4 and 1.

2. **a**. Each term in the expansion of A is symmetric by Exercise 35 in Section 7.1. The fact that $(B + C)^T = B^T + C^T$ implies that any sum of symmetric matrices is symmetric, so A is symmetric.

b. Since $\mathbf{u}_1^T\mathbf{u}_1 = 1$ and $\mathbf{u}_j^T\mathbf{u}_1 = 0$ for $j \neq 1$,

$$A\mathbf{u}_1 = (\lambda_1\mathbf{u}_1\mathbf{u}_1^T)\mathbf{u}_1 + \ldots + (\lambda_n\mathbf{u}_n\mathbf{u}_n^T)\mathbf{u}_1 = \lambda_1\mathbf{u}_1(\mathbf{u}_1^T\mathbf{u}_1) + \ldots + \lambda_n\mathbf{u}_n(\mathbf{u}_n^T\mathbf{u}_1) = \lambda_1\mathbf{u}_1$$

Since $\mathbf{u}_1 \neq \mathbf{0}$, λ_1 is an eigenvalue of A. A similar argument shows that λ_j is an eigenvalue of A for $j = 2, \ldots, n$.

3. If rank $A = r$, then $\dim \text{Nul}\, A = n - r$ by the Rank Theorem. So 0 is an eigenvalue of A with multiplicity $n - r$, and of the n terms in the spectral decomposition of A exactly $n - r$ are zero. The remaining r terms (which correspond to nonzero eigenvalues) are all rank 1 matrices, as mentioned in the discussion of the spectral decomposition.

4. **a**. By Theorem 3 in Section 6.1, $(\text{Col}\, A)^\perp = \text{Nul}\, A^T = \text{Nul}\, A$ since $A^T = A$.

b. Let $\mathbf{y}$ be in $\mathbb{R}^n$. By the Orthogonal Decomposition Theorem in Section 6.3, $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in Col A and $\mathbf{z}$ is in $(\text{Col}\, A)^\perp$. By part a., $\mathbf{z}$ is in Nul A.

5. If $A\mathbf{v} = \lambda\mathbf{v}$ for some nonzero λ, then $\mathbf{v} = \lambda^{-1}A\mathbf{v} = A(\lambda^{-1}\mathbf{v})$, which shows that $\mathbf{v}$ is a linear combination of the columns of A.

6. Because A is symmetric, there is an orthonormal eigenvector basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for $\mathbb{R}^n$. Let r = rank A. If $r = 0$, then $A = O$ and the decomposition of Exercise 4(b) is $\mathbf{y} = \mathbf{0} + \mathbf{y}$ for each $\mathbf{y}$ in $\mathbb{R}^n$; if $r = n$ then the decomposition is $\mathbf{y} = \mathbf{y} + \mathbf{0}$ for each $\mathbf{y}$ in $\mathbb{R}^n$.

Assume that $0 < r < n$. Then $\dim \text{Nul}\, A = n - r$ by the Rank Theorem, and so 0 is an eigenvalue of A with multiplicity $n - r$. Hence there are r nonzero eigenvalues, counted according to their multiplicities.

Renumber the eigenvector basis if necessary so that $\mathbf{u}_1,\ldots,\mathbf{u}_r$ are the eigenvectors corresponding to the nonzero eigenvalues. By Exercise 5, $\mathbf{u}_1,\ldots,\mathbf{u}_r$ are in Col A. Also, $\mathbf{u}_{r+1},\ldots,\mathbf{u}_n$ are in Nul A because these vectors are eigenvectors corresponding to the eigenvalue 0. For $\mathbf{y}$ in $\mathbb{R}^n$, there are scalars $c_1,\ldots,c_n$ such that

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + \ldots + c_r\mathbf{u}_r}_{\hat{\mathbf{y}}} + \underbrace{c_{r+1}\mathbf{u}_{r+1} + \ldots + c_n\mathbf{u}_n}_{\mathbf{z}}$$

This provides the decomposition in Exercise 4(b).

7. If $A = R^TR$ and R is invertible, then A is positive definite by Exercise 25 in Section 7.2.

Conversely, suppose that A is positive definite. Then by Exercise 26 in Section 7.2, $A = B^TB$ for some positive definite matrix B. Since the eigenvalues of B are positive, 0 is not an eigenvalue of B and B is invertible. Thus the columns of B are linearly independent. By Theorem 12 in Section 6.4, $B = QR$ for some $n \times n$ matrix Q with orthonormal columns and some upper triangular matrix R with positive entries on its diagonal. Since Q is a square matrix, $Q^TQ = I$, and

$$A = B^TB = (QR)^T(QR) = R^TQ^TQR = R^TR$$

and R has the required properties.

8. Suppose that A is positive definite, and consider a Cholesky factorization of $A = R^TR$ with R upper triangular and having positive entries on its diagonal. Let D be the diagonal matrix whose diagonal entries are the entries on the diagonal of R. Since right-multiplication by a diagonal matrix scales the columns of the matrix on its left, the matrix $L = R^TD^{-1}$ is lower triangular with 1's on its diagonal. If $U = DR$, then $A = R^TD^{-1}DR = LU$.

9. If A is an $m \times n$ matrix and $\mathbf{x}$ is in $\mathbb{R}^n$, then $\mathbf{x}^TA^TA\mathbf{x} = (A\mathbf{x})^T(A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$. Thus A^TA is positive semidefinite. By Exercise 22 in Section 6.5, rank $A^TA =$ rank A.

10. If rank $G = r$, then dim Nul $G = n - r$ by the Rank Theorem. Hence 0 is an eigenvalue of G with multiplicity $n - r$, and the spectral decomposition of G is

$$G = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \ldots + \lambda_r\mathbf{u}_r\mathbf{u}_r^T$$

Also $\lambda_1,\ldots,\lambda_r$ are positive because G is positive semidefinite. Thus

$$G = \left(\sqrt{\lambda_1}\mathbf{u}_1\right)\left(\sqrt{\lambda_1}\mathbf{u}_1^T\right) + \ldots + \left(\sqrt{\lambda_r}\mathbf{u}_r\right)\left(\sqrt{\lambda_r}\mathbf{u}_r^T\right)$$

By the column-row expansion of a matrix product, $G = BB^T$ where B is the $n \times r$ matrix $B = \left[\sqrt{\lambda_1}\mathbf{u}_1 \quad \ldots \quad \sqrt{\lambda_r}\mathbf{u}_r\right]$. Finally, $G = A^TA$ for $A = B^T$.

11. Let $A = U\Sigma V^T$ be a singular value decomposition of A. Since U is orthogonal, $U^TU = I$ and $A = U\Sigma U^TUV^T = PQ$ where $P = U\Sigma U^T = U\Sigma U^{-1}$ and $Q = UV^T$. Since Σ is symmetric, P is symmetric, and P has nonnegative eigenvalues because it is similar to Σ, which is diagonal with nonnegative diagonal entries. Thus P is positive semidefinite. The matrix Q is orthogonal since it is the product of orthogonal matrices.

12. **a**. Because the columns of V_r are orthonormal,

$$AA^+\mathbf{y} = (U_rDV_r^T)(V_rD^{-1}U_r^T)\mathbf{y} = (U_rDD^{-1}U_r^T)\mathbf{y} = U_rU_r^T\mathbf{y}$$

Since $U_r U_r^T \mathbf{y}$ is the orthogonal projection of **y** onto $\text{Col}\, U_r$ by Theorem 10 in Section 6.3, and since $\text{Col}\, U_r = \text{Col}\, A$ by (5) in Example 6 of Section 7.4, $AA^+\mathbf{y}$ is the orthogonal projection of **y** onto Col A.

b. Because the columns of U_r are orthonormal,

$$A^+A\mathbf{x} = (V_r D^{-1} U_r^T)(U_r D V_r^T)\mathbf{x} = (V_r D^{-1} D V_r^T)\mathbf{x} = V_r V_r^T \mathbf{x}$$

Since $V_r V_r^T \mathbf{x}$ is the orthogonal projection of **x** onto $\text{Col}\, V_r$ by Theorem 10 in Section 6.3, and since $\text{Col}\, V_r = \text{Row}\, A$ by (8) in Example 6 of Section 7.4, $A^+A\mathbf{x}$ is the orthogonal projection of **x** onto Row A.

c. Using the reduced singular value decomposition, the definition of A^+, and the associativity of matrix multiplication gives:

$$\begin{aligned} AA^+A &= (U_r D V_r^T)(V_r D^{-1} U_r^T)(U_r D V_r^T) = (U_r D D^{-1} U_r^T)(U_r D V_r^T) \\ &= U_r D D^{-1} D V_r^T = U_r D V_r^T = A \\ A^+AA^+ &= (V_r D^{-1} U_r^T)(U_r D V_r^T)(V_r D^{-1} U_r^T) = (V_r D^{-1} D V_r^T)(V_r D^{-1} U_r^T) \\ &= V_r D^{-1} D D^{-1} U_r^T = V_r D^{-1} U_r^T = A^+ \end{aligned}$$

13. **a**. If $\mathbf{b} = A\mathbf{x}$, then $\mathbf{x}^+ = A^+\mathbf{b} = A^+A\mathbf{x}$. By Exercise 12(a), $\mathbf{x}^+$ is the orthogonal projection of **x** onto Row A.

b. From part (a) and Exercise 12(c), $A\mathbf{x}^+ = A(A^+A\mathbf{x}) = (AA^+A)\mathbf{x} = A\mathbf{x} = \mathbf{b}$.

c. Let $A\mathbf{u} = \mathbf{b}$. Since $\mathbf{x}^+$ is the orthogonal projection of **x** onto Row A, the Pythagorean Theorem shows that $\|\mathbf{u}\|^2 = \|\mathbf{x}^+\|^2 + \|\mathbf{u} - \mathbf{x}^+\|^2 \ge \|\mathbf{x}^+\|^2$, with equality only if $\mathbf{u} = \mathbf{x}^+$.

14. The least-squares solutions of $A\mathbf{x} = \mathbf{b}$ are precisely the solutions of $A\mathbf{x} = \hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is the orthogonal projection of **b** onto Col A. From Exercise 13, the minimum length solution of $A\mathbf{x} = \hat{\mathbf{b}}$ is $A^+\hat{\mathbf{b}}$, so $A^+\hat{\mathbf{b}}$ is the minimum length least-squares solution of $A\mathbf{x} = \mathbf{b}$. However, $\hat{\mathbf{b}} = AA^+\mathbf{b}$ by Exercise 12(a) and hence $A^+\hat{\mathbf{b}} = A^+AA^+\mathbf{b} = A^+\mathbf{b}$ by Exercise 12(c). Thus $A^+\mathbf{b}$ is the minimum length least-squares solution of $A\mathbf{x} = \mathbf{b}$.

15. **[M]** The reduced SVD of A is $A = U_r D V_r^T$, where

$$U_r = \begin{bmatrix} .966641 & .253758 & -.034804 \\ .185205 & -.786338 & -.589382 \\ .125107 & -.398296 & .570709 \\ .125107 & -.398296 & .570709 \end{bmatrix}, D = \begin{bmatrix} 9.84443 & 0 & 0 \\ 0 & 2.62466 & 0 \\ 0 & 0 & 1.09467 \end{bmatrix},$$

$$\text{and } V_r = \begin{bmatrix} -.313388 & .009549 & .633795 \\ -.313388 & .009549 & .633795 \\ -.633380 & .023005 & -.313529 \\ .633380 & -.023005 & .313529 \\ .035148 & .999379 & .002322 \end{bmatrix}$$

So the pseudoinverse $A^+ = V_r D^{-1} U_r^T$ may be calculated, as well as the solution $\hat{\mathbf{x}} = A^+\mathbf{b}$ for the system $A\mathbf{x} = \mathbf{b}$:

$$A^+ = \begin{bmatrix} -.05 & -.35 & .325 & .325 \\ -.05 & -.35 & .325 & .325 \\ -.05 & .15 & -.175 & -.175 \\ .05 & -.15 & .175 & .175 \\ .10 & -.30 & -.150 & -.150 \end{bmatrix}, \hat{\mathbf{x}} = \begin{bmatrix} .7 \\ .7 \\ -.8 \\ .8 \\ .6 \end{bmatrix}$$

Row reducing the augmented matrix for the system $A^T\mathbf{z} = \hat{\mathbf{x}}$ shows that this system has a solution, so $\hat{\mathbf{x}}$ is in $\text{Col}\, A^T = \text{Row}\, A$. A basis for Nul A is $\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$, and an arbitrary element of Nul A is $\mathbf{u} = c\mathbf{a}_1 + d\mathbf{a}_2$. One computes that $\|\hat{\mathbf{x}}\| = \sqrt{131/50}$, while $\|\hat{\mathbf{x}} + \mathbf{u}\| = \sqrt{(131/50) + 2c^2 + 2d^2}$. Thus if $\mathbf{u} \neq \mathbf{0}$, $\|\hat{\mathbf{x}}\| < \|\hat{\mathbf{x}} + \mathbf{u}\|$, which confirms that $\hat{\mathbf{x}}$ is the minimum length solution to $A\mathbf{x} = \mathbf{b}$.

16. [M] The reduced SVD of A is $A = U_r D V_r^T$, where

$$U_r = \begin{bmatrix} -.337977 & .936307 & .095396 \\ .591763 & .290230 & -.752053 \\ -.231428 & -.062526 & -.206232 \\ -.694283 & -.187578 & -.618696 \end{bmatrix}, D = \begin{bmatrix} 12.9536 & 0 & 0 \\ 0 & 1.44553 & 0 \\ 0 & 0 & .337763 \end{bmatrix},$$

$$\text{and } V_r = \begin{bmatrix} -.690099 & .721920 & .050939 \\ 0 & 0 & 0 \\ .341800 & .387156 & -.856320 \\ .637916 & .573534 & .513928 \\ 0 & 0 & 0 \end{bmatrix}$$

So the pseudoinverse $A^+ = V_r D^{-1} U_r^T$ may be calculated, as well as the solution $\hat{\mathbf{x}} = A^+\mathbf{b}$ for the system $A\mathbf{x} = \mathbf{b}$:

$$A^+ = \begin{bmatrix} .5 & 0 & -.05 & -.15 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & .5 & 1.5 \\ .5 & -1 & -.35 & -1.05 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \hat{\mathbf{x}} = \begin{bmatrix} 2.3 \\ 0 \\ 5.0 \\ -.9 \\ 0 \end{bmatrix}$$

Row reducing the augmented matrix for the system $A^T\mathbf{z} = \hat{\mathbf{x}}$ shows that this system has a solution, so $\hat{\mathbf{x}}$ is in $\text{Col}\, A^T = \text{Row}\, A$. A basis for Nul A is $\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, and an arbitrary element of Nul A is $\mathbf{u} = c\mathbf{a}_1 + d\mathbf{a}_2$. One computes that $\|\hat{\mathbf{x}}\| = \sqrt{311/10}$, while $\|\hat{\mathbf{x}} + \mathbf{u}\| = \sqrt{(311/10) + c^2 + d^2}$. Thus if $\mathbf{u} \neq \mathbf{0}$, $\|\hat{\mathbf{x}}\| < \|\hat{\mathbf{x}} + \mathbf{u}\|$, which confirms that $\hat{\mathbf{x}}$ is the minimum length solution to $A\mathbf{x} = \mathbf{b}$.