Chapter 6

SECTION 1

2. If A is triangular then $A - a_{ii}I$ will be a triangular matrix with a zero entry in the (i, i) position. Since the determinant of a triangular matrix is the product of its diagonal elements it follows that

$$\det(A - a_{ii}I) = 0$$

Thus the eigenvalues of A are $a_{11}, a_{22}, \ldots, a_{nn}$.

3. A is singular if and only if $\det(A) = 0$. The scalar 0 is an eigenvalue if and only if

$$\det(A - 0I) = \det(A) = 0$$

Thus A is singular if and only if one of its eigenvalues is 0.

4. If A is a nonsingular matrix and λ is an eigenvalue of A, then there exists a nonzero vector $\mathbf{x}$ such that

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ A^{-1}A\mathbf{x} &= \lambda A^{-1}\mathbf{x} \end{aligned}$$

It follows from Exercise 3 that $\lambda \neq 0$. Therefore

$$A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x} \quad (\mathbf{x} \neq \mathbf{0})$$

and hence $1/\lambda$ is an eigenvalue of A^{-1}.

5. The proof is by induction. In the case where $m = 1$, $\lambda^1 = \lambda$ is an eigenvalue of A with eigenvector $\mathbf{x}$. Suppose λ^k is an eigenvalue of A^k and $\mathbf{x}$ is an eigenvector belonging to λ^k.

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

Thus λ^{k+1} is an eigenvalue of A^{k+1} and $\mathbf{x}$ is an eigenvector belonging to λ^{k+1}. It follows by induction that if λ an eigenvalue of A then λ^m is an eigenvalue of A^m, for $m = 1, 2, \ldots$.

6. If A is idempotent and λ is an eigenvalue of A with eigenvector $\mathbf{x}$, then

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ A^2\mathbf{x} &= \lambda A\mathbf{x} = \lambda^2\mathbf{x} \end{aligned}$$

and

$$A^2\mathbf{x} = A\mathbf{x} = \lambda\mathbf{x}$$

Therefore

$$(\lambda^2 - \lambda)\mathbf{x} = \mathbf{0}$$

Since $\mathbf{x} \neq \mathbf{0}$ it follows that

$$\begin{aligned} &\lambda^2 - \lambda = 0 \\ &\lambda = 0 \quad \text{or} \quad \lambda = 1 \end{aligned}$$

7. If λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k (Exercise 5). If $A^k = O$, then all of its eigenvalues are 0. Thus $\lambda^k = 0$ and hence $\lambda = 0$.

9. $\det(A-\lambda I) = \det((A-\lambda I)^T) = \det(A^T - \lambda I)$. Thus A and A^T have the same characteristic polynomials and consequently must have the same eigenvalues. The eigenspaces however will not be the same. For example

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

both have eigenvalues

$$\lambda_1 = \lambda_2 = 1$$

The eigenspace of A corresponding to $\lambda = 1$ is spanned by $(1,\ 0)^T$ while the eigenspace of A^T is spanned by $(0,\ 1)^T$. Exercise 27 shows how the eigenvectors of A and A^T are related.

10. $\det(A - \lambda I) = \lambda^2 - (2\cos\theta)\lambda + 1$. The discriminant will be negative unless θ is a multiple of π. The matrix A has the effect of rotating a real vector $\mathbf{x}$ about the origin by an angle of θ. Thus $A\mathbf{x}$ will be a scalar multiple of $\mathbf{x}$ if and only if θ is a multiple of π.

12. Since $\operatorname{tr}(A)$ equals the sum of the eigenvalues the result follows by solving

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$$

for λ_j.

13. $$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{21}a_{12})$$
$$= \lambda^2 - (\operatorname{tr} A)\lambda + \det(A)$$

14. If $\mathbf{x}$ is an eigenvector of A belonging to λ, then any nonzero multiple of $\mathbf{x}$ is also an eigenvector of A belonging to λ. By Exercise 5 we know that $A^m\mathbf{x} = \lambda^m\mathbf{x}$, so $A^m\mathbf{x}$ must be an eigenvector of A belonging to λ. Alternatively we could have proved the result by noting that

$$A^m\mathbf{x} = \lambda^m\mathbf{x} \neq \mathbf{0}$$

and

$$A(A^m\mathbf{x}) = A^{m+1}\mathbf{x} = A^m(A\mathbf{x}) = A^m(\lambda\mathbf{x}) = \lambda(A^m\mathbf{x})$$

15. If $A - \lambda_0 I$ has rank k then $N(A - \lambda_0 I)$ will have dimension $n - k$.

16. The subspace spanned by $\mathbf{x}$ and $A\mathbf{x}$ will have dimension 1 if and only if $\mathbf{x}$ and $A\mathbf{x}$ are linearly dependent and $\mathbf{x} \neq \mathbf{0}$. If $\mathbf{x} \neq \mathbf{0}$ then the vectors $\mathbf{x}$ and $A\mathbf{x}$ will be linearly dependent if and only if $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ.

17. (a) If $\alpha = a + bi$ and $\beta = c + di$, then

$$\overline{\alpha + \beta} = \overline{(a+c) + (b+d)i} = (a+c) - (b+d)i$$

and

$$\overline{\alpha} + \overline{\beta} = (a - bi) + (c - di) = (a+c) - (b+d)i$$

Therefore $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$.

Next we show that the conjugate of the product of two numbers is the product of the conjugates.

$$\overline{\alpha\beta} = \overline{(ac - bd) + (ad + bc)i} = (ac - bd) - (ad + bc)i$$
$$\overline{\alpha}\,\overline{\beta} = (a - bi)(c - di) = (ac - bd) - (ad + bc)i$$

Therefore $\overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$.

(b) If $A \in R^{m\times n}$ and $B \in R^{n\times r}$, then the (i,j) entry of $\overline{AB}$ is given by

$$\overline{a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}} = \overline{a_{i1}}\,\overline{b_{1j}} + \overline{a_{i2}}\,\overline{b_{2j}} + \cdots + \overline{a_{in}}\,\overline{b_{nj}}$$

The expression on the right is the (i,j) entry of $\overline{A}\,\overline{B}$. Therefore

$$\overline{AB} = \overline{A}\,\overline{B}$$

18. **(a)** If λ is an eigenvalue of an orthogonal matrix Q and $\mathbf{x}$ is a unit eigenvector belonging to λ then

$$|\lambda| = |\lambda|\,\|\mathbf{x}\| = \|\lambda\mathbf{x}\| = \|Q\mathbf{x}\| = \|\mathbf{x}\| = 1$$

(b) Since the eigenvalues of Q all have absolute value equal to 1, it follows that

$$|\det(Q)| = |\lambda_1\lambda_2\cdots\lambda_n| = 1$$

19. If Q is an orthogonal matrix with eigenvalue $\lambda = 1$ and $\mathbf{x}$ is an eigenvector belonging to $\lambda = 1$, then $Q\mathbf{x} = \mathbf{x}$ and since $Q^T = Q^{-1}$ we have

$$Q^T\mathbf{x} = Q^TQ\mathbf{x} = I\mathbf{x} = \mathbf{x}$$

Therefore $\mathbf{x}$ is an eigenvector of Q^T belonging to the eigenvector $\lambda = 1$.

20. **(a)** Each eigenvalue has absolute value 1 and the product of the eigenvalues is equal to 1. So if the eigenvalues are real and are ordered so that $\lambda_1 \geq \lambda_2 \geq \lambda_3$, then the only possible triples of eigenvalues are: $(1, 1, 1)$ and $(1, -1, -1)$.

(b) The complex eigenvalues must be of the form $\lambda_2 = \cos\theta + i\sin\theta$ and $\lambda_3 = \cos\theta - i\sin\theta$. It follows then that

$$\lambda_1\lambda_2\lambda_3 = \lambda_1(\cos\theta + i\sin\theta)(\cos\theta - i\sin\theta) = \lambda_1(\cos^2\theta + \sin^2\theta) = \lambda_1$$

Therefore

$$\lambda_1 = \lambda_1\lambda_2\lambda_3 = \det(A) = 1$$

(c) If the eigenvalues of Q are all real then by part (a) at least one of the eigenvalues must equal 1. If the eigenvalues are not all real then Q must have one pair of complex conjugate eigenvalues and one real eigenvalue. By part (b) the real eigenvalue must be equal to 1. Therefore if Q is a 3×3 orthogonal matrix with $\det(Q) = 1$, then $\lambda = 1$ must be an eigenvalue.

21. If $\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_r\mathbf{x}_r$ is an element of S, then

$$A\mathbf{x} = (c_1\lambda_1)\mathbf{x}_1 + (c_2\lambda_2)\mathbf{x}_2 + \cdots + (c_r\lambda_r)\mathbf{x}_r$$

Thus $A\mathbf{x}$ is also an element of S.

22. Since $\mathbf{x} \neq \mathbf{0}$ and S is nonsingular it follows that $S\mathbf{x} \neq \mathbf{0}$. If $B = S^{-1}AS$, then $AS = SB$ and it follows that

$$A(S\mathbf{x}) = (AS)\mathbf{x} = SB\mathbf{x} = S(\lambda\mathbf{x}) = \lambda(S\mathbf{x})$$

Therefore $S\mathbf{x}$ is an eigenvector of A belonging to λ.

23. If $\mathbf{x}$ is an eigenvector of A belonging to the eigenvalue λ and $\mathbf{x}$ is also an eigenvector of B corresponding to the eigenvalue μ, then

$$(\alpha A + \beta B)\mathbf{x} = \alpha A\mathbf{x} + \beta B\mathbf{x} = \alpha\lambda\mathbf{x} + \beta\mu\mathbf{x} = (\alpha\lambda + \beta\mu)\mathbf{x}$$

Therefore $\mathbf{x}$ is an eigenvector of $\alpha A + \beta B$ belonging to $\alpha\lambda + \beta\mu$.

24. If $\lambda \neq 0$ and $\mathbf{x}$ is an eigenvector belonging to λ, then

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ \mathbf{x} &= \frac{1}{\lambda}A\mathbf{x} \end{aligned}$$

Since $A\mathbf{x}$ is in $R(A)$ it follows that $\frac{1}{\lambda}A\mathbf{x}$ is in $R(A)$.

25. If

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T$$

then for $i = 1, \ldots, n$

$$A\mathbf{u}_i = \lambda_1\mathbf{u}_1\mathbf{u}_1^T\mathbf{u}_i + \lambda_2\mathbf{u}_2\mathbf{u}_2^T\mathbf{u}_i + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T\mathbf{u}_i$$

Since $\mathbf{u}_j^T\mathbf{u}_i = 0$ unless $j = i$, it follows that

$$A\mathbf{u}_i = \lambda_i\mathbf{u}_i\mathbf{u}_i^T\mathbf{u}_i = \lambda_i\mathbf{u}_i$$

and hence λ_i is an eigenvalue of A with eigenvector $\mathbf{u}_i$. The matrix A is symmetric since each $c_i\mathbf{u}_i\mathbf{u}_i^T$ is symmetric and any sum of symmetric matrices is symmetric.

26. If the columns of A each add up to a fixed constant δ then the row vectors of $A-\delta I$ all add up to $(0, 0, \ldots, 0)$. Thus the row vectors of $A-\delta I$ are linearly dependent and hence $A-\delta I$ is singular. Therefore δ is an eigenvalue of A.

27. Since $\mathbf{y}$ is an eigenvector of A^T belonging to λ_2 it follows that

$$\mathbf{x}^T A^T \mathbf{y} = \lambda_2 \mathbf{x}^T \mathbf{y}$$

The expression $\mathbf{x}^T A^T \mathbf{y}$ can also be written in the form $(A\mathbf{x})^T\mathbf{y}$. Since $\mathbf{x}$ is an eigenvector of A belonging to λ_1, it follows that

$$\mathbf{x}^T A^T \mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \lambda_1 \mathbf{x}^T \mathbf{y}$$

Therefore

$$(\lambda_1 - \lambda_2)\mathbf{x}^T\mathbf{y} = 0$$

and since $\lambda_1 \neq \lambda_2$, the vectors $\mathbf{x}$ and $\mathbf{y}$ must be orthogonal.

28. (a) If λ is a nonzero eigenvalue of AB with eigenvector $\mathbf{x}$, then let $\mathbf{y} = B\mathbf{x}$. Since

$$A\mathbf{y} = AB\mathbf{x} = \lambda\mathbf{x} \neq \mathbf{0}$$

it follows that $\mathbf{y} \neq \mathbf{0}$ and

$$BA\mathbf{y} = BA(B\mathbf{x}) = B(AB\mathbf{x}) = B\lambda\mathbf{x} = \lambda\mathbf{y}$$

Thus λ is also an eigenvalue of BA with eigenvector $\mathbf{y}$.

(b) If $\lambda = 0$ is an eigenvalue of AB, then AB must be singular. Since

$$\det(BA) = \det(B)\det(A) = \det(A)\det(B) = \det(AB) = 0$$

it follows that BA is also singular. Therefore $\lambda = 0$ is an eigenvalue of BA.

29. If $AB - BA = I$, then $BA = AB - I$. If the eigenvalues of AB are $\lambda_1, \lambda_2, \ldots, \lambda_n$, then it follows from Exercise 8 that the eigenvalues of BA are $\lambda_1 - 1$, $\lambda_2 - 1, \ldots, \lambda_n - 1$. This contradicts the result proved in Exercise 28 that AB and BA have the same eigenvalues.

30. (a) If λ_i is a root of $p(\lambda)$, then

$$\lambda_i^n = a_{n-1}\lambda_i^{n-1} + \cdots + a_1\lambda_i + a_0$$

Thus if $\mathbf{x} = (\lambda_i^{n-1}, \lambda_i^{n-2}, \ldots, \lambda_i, 1)^T$, then

$$C\mathbf{x} = (\lambda_i^n, \lambda_i^{n-1}, \ldots, \lambda_i^2, \lambda_i)^T = \lambda_i\mathbf{x}$$

and hence λ_i is an eigenvalue of C with eigenvector $\mathbf{x}$.

(b) If $\lambda_1, \ldots, \lambda_n$ are the roots of $p(\lambda)$, then

$$p(\lambda) = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

If $\lambda_1, \ldots, \lambda_n$ are all distinct then by part (a) they are the eigenvalues of C. Since the characteristic polynomial of C has lead coefficient $(-1)^n$ and roots $\lambda_1, \ldots, \lambda_n$, it must equal $p(\lambda)$.

31. Let

$$D_m(\lambda) = \begin{pmatrix} a_m & a_{m-1} & \cdots & a_1 & a_0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 1 & -\lambda \end{pmatrix}$$

It can be proved by induction on m that

$$\det(D_m(\lambda)) = (-1)^m(a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0)$$

If $\det(C - \lambda I)$ is expanded by cofactors along the first column one obtains

$$\begin{aligned} \det(C - \lambda I) &= (a_{n-1} - \lambda)(-\lambda)^{n-1} - \det(D_{n-2}) \\ &= (-1)^n(\lambda^n - a_{n-1}\lambda^{n-1}) - (-1)^{n-2}(a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0) \\ &= (-1)^n[(\lambda^n - a_{n-1}\lambda^{n-1}) - (a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0)] \\ &= (-1)^n[\lambda^n - a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - \cdots - a_1\lambda - a_0] \\ &= p(\lambda) \end{aligned}$$

SECTION 2

3. (a) If

$$\mathbf{Y}(t) = c_1e^{\lambda_1 t}\mathbf{x}_1 + c_2e^{\lambda_2 t}\mathbf{x}_2 + \cdots + c_ne^{\lambda_n t}\mathbf{x}_n$$

then

$$\mathbf{Y}_0 = \mathbf{Y}(0) = c_1\mathbf{x}_2 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

(b) It follows from part (a) that

$$\mathbf{Y}_0 = X\mathbf{c}$$

If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent then X is nonsingular and we can solve for $\mathbf{c}$

$$\mathbf{c} = X^{-1}\mathbf{Y}_0$$

7. It follows from the initial condition that

$$\begin{aligned} x_1'(0) &= a_1\sigma = 2 \\ x_2'(0) &= a_2\sigma = 2 \end{aligned}$$

and hence

$$a_1 = a_2 = 2/\sigma$$

Substituting for x_1 and x_2 in the system

$$\begin{aligned} x_1'' &= -2x_1 + x_2 \\ x_2'' &= x_1 - 2x_2 \end{aligned}$$

yields

$$\begin{aligned} -a_1\sigma^2\sin\sigma t &= -2a_1\sin\sigma t + a_2\sin\sigma t \\ -a_2\sigma^2\sin\sigma t &= a_1\sin\sigma t - 2a_2\sin\sigma t \end{aligned}$$

Replacing a_1 and a_2 by $2/\sigma$ we get

$$\sigma^2 = 1$$

Using either $\sigma = -1$, $a_1 = a_2 = -2$ or $\sigma = 1$, $a_1 = a_2 = 2$ we obtain the solution

$$\begin{aligned} x_1(t) &= 2\sin t \\ x_2(t) &= 2\sin t \end{aligned}$$

9. $m_1 y_1'' = k_1 y_1 - k_2(y_2 - y_1) - m_1 g$
$m_2 y_2'' = k_2(y_2 - y_1) - m_2 g$

11. If

$$y^{(n)} = a_0 y + a_1 y' + \cdots + a_{n-1} y^{(n-1)}$$

and we set

$$y_1 = y,\ y_2 = y_1' = y'',\ y_3 = y_2' = y''', \ldots, y_n = y_{n-1}' = y^n$$

then the nth order equation can be written as a system of first order equations of the form $\mathbf{Y}' = A\mathbf{Y}$ where

$$A = \begin{Bmatrix} 0 & y_2 & 0 & \cdots & 0 \\ 0 & 0 & y_3 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & y_n \\ a_0 & a_1 & a_2 & \cdots & a_{n-1} \end{Bmatrix}$$

SECTION 3

1. The factorization XDX^{-1} is not unique. However the diagonal elements of D must be eigenvalues of A and if λ_i is the ith diagonal element of D, then $\mathbf{x}_i$ must be an eigenvector belonging to λ_i

(a) $\det(A - \lambda I) = \lambda^2 - 1$ and hence the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$. $\mathbf{x}_1 = (1,\ 1)^T$ and $\mathbf{x}_2 = (-1,\ 1)^T$ are eigenvectors belonging to λ_1 and λ_2, respectively. Setting

$$X = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

we have

$$A = XDX^{-1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

(b) The eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 1$. If we take $\mathbf{x}_1 = (-2,\ 1)^T$ and $\mathbf{x}_2 = (-3,\ 2)^T$, then

$$A = XDX^{-1} = \begin{pmatrix} -2 & -3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & -3 \\ 1 & 2 \end{pmatrix}$$

(c) $\lambda_1 = 0$, $\lambda_2 = -2$. If we take $\mathbf{x}_1 = (4,\ 1)^T$ and $\mathbf{x}_2 = (2,\ 1)^T$, then

$$A = XDX^{-1} = \begin{pmatrix} 4 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1/2 & -1 \\ -1/2 & 2 \end{pmatrix}$$

(d) The eigenvalues are the diagonal entries of A. The eigenvectors corresponding to $\lambda_1 = 2$ are all multiples of $(1,\ 0,\ 0)^T$. The eigenvectors belonging to $\lambda_2 = 1$ are all multiples of $(2,\ -1,\ 0)$ and the eigenvectors corresponding to $\lambda_3 = -1$ are multiples $(1,\ -3,\ 3)^T$.

$$A = XDX^{-1} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -3 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & \frac{5}{3} \\ 0 & -1 & -1 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

(e) $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = -2$
$\mathbf{x}_1 = (3,\ 1,\ 2)^T$, $\mathbf{x}_2 = (0,\ 3,\ 1)^T$, $\mathbf{x}_3 = (0,\ -1,\ 1)^T$

$$A = XDX^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & -1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{5}{12} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

(f) $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = 0$, $\mathbf{x}_1 = (1,\ 2,\ 3)^T$, $\mathbf{x}_2 = (1,\ 0,\ 1)^T$, $\mathbf{x}_3 = (-2,\ 1,\ 0)^T$

$$A = XDX^{-1} = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{3}{2} & -3 & \frac{5}{2} \\ -1 & -1 & 1 \end{pmatrix}$$

2. If $A = XDX^{-1}$, then $A^6 = XD^6X^{-1}$.

(a) $D^6 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^6 = I$

$A^6 = XD^6X^{-1} = XX^{-1} = I$

(b) $A^6 = \begin{pmatrix} -2 & -3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^6 \begin{pmatrix} -2 & -3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 253 & 378 \\ -126 & -190 \end{pmatrix}$

(c) $A^6 = \begin{pmatrix} 4 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}^6 \begin{pmatrix} 1/2 & -1 \\ -1/2 & 2 \end{pmatrix} = \begin{pmatrix} -64 & 256 \\ -32 & 128 \end{pmatrix}$

(d) $$A^6 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -3 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}^6 \begin{pmatrix} 1 & 2 & 5/3 \\ 0 & -1 & -1 \\ 0 & 0 & 1/3 \end{pmatrix}$$

$$= \begin{pmatrix} 64 & 126 & 105 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(e) $$A^6 = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & -1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}^6 \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{5}{12} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -21 & 64 & 0 \\ -42 & 0 & 64 \end{pmatrix}$$

(f) $$A^6 = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^6 \begin{pmatrix} \frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{3}{2} & -3 & \frac{5}{2} \\ -1 & -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 32 & 64 & -32 \\ 64 & 128 & -64 \\ 96 & 192 & -96 \end{pmatrix}$$

3. If $A = XDX^{-1}$ is nonsingular, then $A^{-1} = XD^{-1}X^{-1}$

(a) $A^{-1} = XD^{-1}X^{-1} = XDX^{-1} = A$

(b) $$A^{-1} = \begin{pmatrix} -2 & -3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & -3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & -3 \\ 1 & \frac{5}{2} \end{pmatrix}$$

(d) $$A^{-1} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -3 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & \frac{5}{3} \\ 0 & -1 & -1 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & -1 & -\frac{3}{2} \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$$

(e) $$A^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & -1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{5}{12} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} & -\frac{3}{4} \end{pmatrix}$$

4. (a) The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 0$

$$A = XDX^{-1}$$

Since $D^2 = D$ it follows that

$$A^2 = XD^2X^{-1} = XDX^{-1} = A$$

(b) $$A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$B = XD^{1/2}X^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

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

5. If X diagonalizes A, then

$$X^{-1}AX = D$$

where D is a diagonal matrix. It follows that

$$D = D^T = X^TA^T(X^{-1})^T = Y^{-1}A^TY$$

Therefore Y diagonalizes A^T.

6. If $A = XDX^{-1}$ where D is a diagonal matrix whose diagonal elements are all either 1 or -1, then $D^{-1} = D$ and

$$A^{-1} = XD^{-1}X^{-1} = XDX^{-1} = A$$

7. If $\mathbf{x}$ is an eigenvector belonging to the eigenvalue a, then

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & b-a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and it follows that

$$x_2 = x_3 = 0$$

Thus the eigenspace corresponding to $\lambda_1 = \lambda_2 = a$ has dimension 1 and is spanned by $(1,\ 0,\ 0)^T$. The matrix is defective since a is a double eigenvalue and its eigenspace only has dimension 1.

8. (a) The characteristic polynomial of the matrix factors as follows.

$$p(\lambda) = \lambda(2-\lambda)(\alpha-\lambda)$$

Thus the only way that the matrix can have a multiple eigenvalue is if $\alpha = 0$ or $\alpha = 2$. In the case $\alpha = 0$, we have that $\lambda = 0$ is an eigenvalue of multiplicity 2 and the corresponding eigenspace is spanned by the vectors $\mathbf{x}_1 = (-1, 1, 0)^T$ and $\mathbf{x}_2 = \mathbf{e}_3$. Since $\lambda = 0$ has two linearly independent eigenvectors, the matrix is not defective. Similarly in the case $\alpha = 2$ the matrix will not be defective since the eigenvalue $\lambda = 2$ possesses two linearly independent eigenvectors $\mathbf{x}_1 = (1, 1, 0)^T$ and $\mathbf{x}_2 = \mathbf{e}_3$.

9. If $A - \lambda I$ has rank 1, then

$$\dim N(A - \lambda I) = 4 - 1 = 3$$

Since λ has multiplicity 3 the matrix is not defective.

10. (a) The proof is by induction. In the case $m = 1$,

$$A\mathbf{x} = \sum_{i=1}^{n} \alpha_i A\mathbf{x}_i = \sum_{i=1}^{n} \alpha_i \lambda_i \mathbf{x}_i$$

If

$$A^k\mathbf{x} = \sum_{i=1}^{n} \alpha_i \lambda_i^k \mathbf{x}_i$$

then

$$A^{k+1}\mathbf{x} = A(A^k\mathbf{x}) = A\Big(\sum_{i=1}^{n} \alpha_i \lambda_i^k \mathbf{x}_i\Big) = \sum_{i=1}^{n} \alpha_i \lambda_i^k A\mathbf{x}_i = \sum_{i=1}^{n} \alpha_i \lambda_i^{k+1} \mathbf{x}_i$$

(b) If $\lambda_1 = 1$, then

$$A^m\mathbf{x} = \alpha_1\mathbf{x}_1 + \sum_{i=2}^{n} \alpha_i \lambda_i^m \mathbf{x}_i$$

Since $0 < \lambda_i < 1$ for $i = 2, \ldots, n$, it follows that $\lambda_i^m \to 0$ as $m \to \infty$. Hence

$$\lim_{m\to\infty} A^m\mathbf{x} = \alpha_1\mathbf{x}_1$$

11. If A is an $n \times n$ matrix and λ is an eigenvalue of multiplicity n then A is diagonalizable if and only if

$$\dim N(A - \lambda I) = n$$

or equivalently

$$\operatorname{rank}(A - \lambda I) = 0$$

The only way the rank can be 0 is if

$$\begin{aligned} A - \lambda I &= O \\ A &= \lambda I \end{aligned}$$

12. If A is nilpotent, then 0 is an eigenvalue of multiplicity n. It follows from Exercise 11 that A is diagonalizable if and only if $A = O$.

13. Let A be a diagonalizable $n \times n$ matrix. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the nonzero eigenvalues of A. The remaining eigenvalues are all 0.

$$\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_n = 0$$

If $\mathbf{x}_i$ is an eigenvector belonging to λ_i, then

$$\begin{aligned} A\mathbf{x}_i &= \lambda_i\mathbf{x}_i \qquad & i &= 1, \ldots, k \\ A\mathbf{x}_i &= \mathbf{0} \qquad & i &= k+1, \ldots, n \end{aligned}$$

Since A is diagonalizable we can choose eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ which form a basis for R^n. Given any vector $\mathbf{x} \in R^n$ we can write

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

It follows that

$$A\mathbf{x} = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_k\lambda_k\mathbf{x}_k$$

Thus $\mathbf{x}_1, \ldots, \mathbf{x}_k$ span the column space of A and since they are linearly independent they form a basis for the column space.

14. The matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has rank 1 even though all of its eigenvalues are 0.

15. (a) For $i = 1, \ldots, k$

$$\mathbf{b}_i = B\mathbf{e}_i = X^{-1}AX\mathbf{e}_i = X^{-1}A\mathbf{x}_i = \lambda X^{-1}\mathbf{x}_i = \lambda\mathbf{e}_i$$

Thus the first k columns of B will have λ's on the diagonal and 0's in the off diagonal positions.

(b) Clearly λ is an eigenvalue of B whose multiplicity is at least k. Since A and B are similar they have the same characteristic polynomial. Thus λ is an eigenvalue of A with multiplicity at least k.

16. (a) If $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors in R^n and $A = \mathbf{x}\mathbf{y}^T$, then A has rank 1. Thus

$$\dim N(A) = n - 1$$

It follows from Exercise 15 that $\lambda = 0$ is an eigenvalue with multiplicity greater than or equal to $n - 1$.

(b) By part (a)

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = 0$$

The sum of the eigenvalues is the trace of A which equals $\mathbf{x}^T\mathbf{y}$. Thus

$$\lambda_n = \sum_{i=1}^{n} \lambda_i = \operatorname{tr} A = \mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$$

Furthermore

$$A\mathbf{x} = \mathbf{x}\mathbf{y}^T\mathbf{x} = \lambda_n\mathbf{x}$$

so $\mathbf{x}$ is an eigenvector belonging to λ_n.

(c) Since $\dim N(A) = n-1$, it follows that $\lambda = 0$ has $n-1$ linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}$. If $\lambda_n \neq 0$ and $\mathbf{x}_n$ is an eigenvector belonging to λ_n, then $\mathbf{x}_n$ will be independent of $\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}$ and hence A will have n linearly independent eigenvectors.

17. If A is diagonalizable, then

$$A = XDX^{-1}$$

where D is a diagonal matrix. If B is similar to A, then there exists a nonsingular matrix S such that $B = S^{-1}AS$. It follows that

$$\begin{aligned} B &= S^{-1}(XDX^{-1})S \\ &= (S^{-1}X)D(S^{-1}X)^{-1} \end{aligned}$$

Therefore B is diagonalizable with diagonalizing matrix $S^{-1}X$.

18. If $A = XD_1X^{-1}$ and $B = XD_2X^{-1}$, where D_1 and D_2 are diagonal matrices, then

$$\begin{aligned} AB &= (XD_1X^{-1})(XD_2X^{-1}) \\ &= XD_1D_2X^{-1} \end{aligned}$$

$$\begin{aligned} &= XD_2D_1X^{-1} \\ &= (XD_2X^{-1})(XD_1X^{-1}) \\ &= BA \end{aligned}$$

19. If $\mathbf{r}_j$ is an eigenvector belonging $\lambda_j = t_{jj}$ then we claim that

$$r_{j+1,j} = r_{j+2,j} = \cdots + = r_{nj} = 0$$

The eigenvector $\mathbf{r}_j$ is a nontrivial solution to $(T - t_{jj}I)\mathbf{x} = \mathbf{0}$. The augmented matrix for this system is $(T - t_{jj}I \,|\, \mathbf{0})$. The equations corresponding to the last $n-j$ rows of the augmented matrix do not involve the variables $x_1, x_2, \ldots, x_j$. These last $n-j$ rows form a homogeneous system that is in strict triangular form with respect to the unknowns $x_{j+1}, x_{j+2}, \ldots, x_n$. The solution to this strictly triangular system is

$$x_{j+1} = x_{j+2} = \cdots = x_n = 0$$

Thus the last $n-j$ entries of the eigenvector $\mathbf{r}_j$ are all equal to 0. If we set $R = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ then R is upper triangular and R diagonalizes T.

23. If A is stochastic then the entries of each of its column vectors will all add up to 1, so the entries of each of the row vectors of A^T will all add up to 1 and consequently $A^T\mathbf{e} = \mathbf{e}$. Therefore $\lambda = 1$ is an eigenvalue of A^T. Since A and A^T have the same eigenvalues, it follows that $\lambda = 1$ is an eigenvalue of A.

24. Since the rows of a doubly stochastic matrix A all add up to 1 it follows that $\mathbf{e}$ is an eigenvector of A belonging to the eigenvalue $\lambda = 1$. If $\lambda = 1$ is the dominant eigenvalue then for any starting probability vector $\mathbf{x}_0$, the Markov chain will converge to a steady-state vector $\mathbf{x} = c\mathbf{e}$. Since the steady-state vector must be a probability vector we have

$$1 = x_1 + x_2 + \cdots + x_n = c + c + \cdots + c = nc$$

and hence $c = \frac{1}{n}$.

25. Let

$$\mathbf{w}_k = M\mathbf{x}_k \qquad \text{and} \qquad \alpha_k = \frac{\mathbf{e}^T\mathbf{x}_k}{n}$$

It follows from equation (5) in the textbook that

$$\mathbf{x}_{k+1} = A\mathbf{x}_k = pM\mathbf{x}_k + \frac{1-p}{n}\mathbf{e}\mathbf{e}^T\mathbf{x}_k = p\mathbf{w}_k + (1-p)\alpha_k\mathbf{e}$$

26. (a) Since $A^2 = O$, it follows that

$$e^A = I + A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$$

(c) Since

$$A^k = \begin{pmatrix} 1 & 0 & -k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad k = 1, 2, \ldots$$

it follows that

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

27. (b) $\begin{pmatrix} 2e - \frac{1}{e} & 2e - \frac{2}{e} \\ \\ -e + \frac{1}{e} & -e + \frac{2}{e} \end{pmatrix}$

28. (d) The matrix A is defective, so e^{At} must be computed using the definition of the matrix exponential. Since

$$A^2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{pmatrix} \quad \text{and} \quad A^3 = O$$

it follows that

$$\begin{aligned} e^{At} &= I + tA + \frac{t^2}{2}A^2 \\ &= \begin{pmatrix} 1 + t + \frac{1}{2}t^2 & t & t + \frac{1}{2}t^2 \\ t & 1 & t \\ -t - \frac{1}{2}t^2 & -t & 1 - t - \frac{1}{2}t^2 \end{pmatrix} \end{aligned}$$

The solution to the initial value problem is

$$\mathbf{Y} = e^{At}\mathbf{Y}_0 = \begin{pmatrix} 1+t \\ 1 \\ -1-t \end{pmatrix}$$

29. If λ is an eigenvalue of A and $\mathbf{x}$ is an eigenvector belonging to λ then

$$\begin{aligned} e^A\mathbf{x} &= \left(I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots\right)\mathbf{x} \\ &= \mathbf{x} + A\mathbf{x} + \frac{1}{2!}A^2\mathbf{x} + \frac{1}{3!}A^3\mathbf{x} + \cdots \\ &= \mathbf{x} + \lambda\mathbf{x} + \frac{1}{2!}\lambda^2\mathbf{x} + \frac{1}{3!}\lambda^3\mathbf{x} + \cdots \\ &= \left(1 + \lambda + \frac{1}{2!}\lambda^2 + \frac{1}{3!}\lambda^3 + \cdots\right)\mathbf{x} \\ &= e^\lambda\mathbf{x} \end{aligned}$$

30. If A is diagonalizable with linearly independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ then, by Exercise 29, $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are eigenvectors of e^A. Furthermore, if $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are eigenvectors corresponding to the eigenvalue λ of A and the eigenvalue e^λ of e^A, then these eigenvalues must have multiplicity at least k (see Exercise 15). Thus if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then $e^{\lambda_1}, \ldots, e^{\lambda_n}$ are the eigenvalues of e^A. Since the eigenvalues of e^A are all nonzero, e^A is nonsingular.

31. (a) Let A be a diagonalizable matrix with characteristic polynomial

$$p(\lambda) = a_1\lambda^n + a_2\lambda^{n-1} + \cdots + a_n\lambda + a_{n+1}$$

and let D be a diagonal matrix whose diagonal entries $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. The matrix

$$p(D) = a_1 D^n + a_2 D^{n-1} + \cdots + a_n D + a_{n+1} I$$

is diagonal since it is a sum of diagonal matrices. Furthermore the jth diagonal entry of $p(D)$ is

$$a_1 \lambda_j^n + a_2 \lambda_j^{n-1} + \cdots + a_n \lambda_j + a_{n+1} = p(\lambda_j) = 0$$

Therefore $p(D) = O$.

(b) If $A = XDX^{-1}$, then

$$\begin{aligned} p(A) &= a_1 A^n + a_2 A^{n-1} + \cdots + a_n A + a_{n+1} I \\ &= a_1 X D^n X^{-1} + a_2 X D^{n-1} X^{-1} + \cdots + a_n X D X^{-1} + a_{n+1} X I X^{-1} \\ &= X(a_1 D^n + a_2 D^{n-1} + \cdots + a_n D + a_{n+1}) X^{-1} \\ &= X p(D) X^{-1} \\ &= O \end{aligned}$$

(c) In part (b) we showed that

$$p(A) = a_1 A^n + a_2 A^{n-1} + \cdots + a_n A + a_{n+1} I = O$$

If $a_{n+1} \neq 0$, then we can solve for I.

$$I = c_1 A^n + c_2 A^{n-1} + \cdots + c_n A$$

where $c_j = -\dfrac{a_j}{a_{n+1}}$ for $j = 1, \ldots, n$. Thus if we set

$$q(A) = c_1 A^{n-1} + c_2 A^{n-2} + \cdots + c_{n-1} A + c_n I$$

then

$$I = Aq(A)$$

and it follows that A is nonsingular and

$$A^{-1} = q(A)$$

SECTION 4

2. (a) $\mathbf{z}_2^H \mathbf{z}_1 = \begin{pmatrix} \dfrac{-i}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \dfrac{1+i}{2} \\ \\ \dfrac{1-i}{2} \end{pmatrix} = 0$

$\mathbf{z}_1^H \mathbf{z}_1 = \begin{pmatrix} \dfrac{1-i}{2} & \dfrac{1+i}{2} \end{pmatrix} \begin{pmatrix} \dfrac{1+i}{2} \\ \\ \dfrac{1-i}{2} \end{pmatrix} = 1$

$$\mathbf{z}_2^H\mathbf{z}_2 = \begin{pmatrix} \frac{-i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{i}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = 1$$

5. There will not be a unique unitary diagonalizing matrix for a given Hermitian matrix A, however, the column vectors of any unitary diagonalizing matrix must be unit eigenvectors of A.

(a) $\lambda_1 = 3$ has a unit eigenvector $\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right)^T$

$\lambda_2 = 1$ has a unit eigenvector $\left(\frac{1}{\sqrt{2}},\ -\frac{1}{\sqrt{2}}\right)^T$.

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

(b) $\lambda_1 = 6$ has a unit eigenvector $\left(\frac{2}{\sqrt{14}},\ \frac{3-i}{\sqrt{14}}\right)^T$

$\lambda_2 = 1$ has a unit eigenvector $\left(\frac{-5}{\sqrt{35}},\ \frac{3-i}{\sqrt{25}}\right)^T$

$$Q = \begin{pmatrix} \frac{2}{\sqrt{14}} & -\frac{5}{\sqrt{35}} \\ \frac{3-i}{\sqrt{14}} & \frac{3-i}{\sqrt{35}} \end{pmatrix}$$

(c) $\lambda_1 = 3$ has a unit eigenvector $\left(-\frac{1}{\sqrt{2}},\ \frac{i}{\sqrt{2}},\ 0\right)^T$

$\lambda_2 = 2$ has a unit eigenvector $(0,\ 0,\ 1)^T$

$\lambda_3 = 1$ has a unit eigenvector $\left(\frac{1}{\sqrt{2}},\ \frac{i}{\sqrt{2}},\ 0\right)^T$

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 0 & 1 \\ i & 0 & i \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

(d) $\lambda_1 = 5$ has a unit eigenvector $\left(0,\ \frac{1}{\sqrt{2}},\ -\frac{1}{\sqrt{2}}\right)^T$

$\lambda_2 = 3$ has a unit eigenvector $\left(\frac{2}{\sqrt{6}},\ \frac{1}{\sqrt{6}},\ \frac{1}{\sqrt{6}}\right)^T$

$\lambda_3 = 0$ has a unit eigenvector $\left(-\frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right)^T$

$$Q = \frac{1}{\sqrt{6}}\begin{pmatrix} 0 & 2 & -\sqrt{2} \\ \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \end{pmatrix}$$

(e) The eigenvalue $\lambda_1 = -1$ has unit eigenvector $\frac{1}{\sqrt{2}}(-1,\ 0,\ 1)^T$.

The eigenvalues $\lambda_2 = \lambda_3 = 1$ have unit eigenvectors $\frac{1}{\sqrt{2}}(1,\ 0,\ 1)^T$ and $(0,\ 1,\ 0)^T$. The three vectors form an orthonormal set. Thus

$$Q = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

is an orthogonal diagonalizing matrix.

(f) $\lambda_1 = 3$ has a unit eigenvector $\mathbf{q}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)^T$.

$\lambda_2 = \lambda_3 = 0$. The eigenspace corresponding to $\lambda = 0$ has dimension 2. It consists of all vectors $\mathbf{x}$ such that

$$x_1 + x_2 + x_3 = 0$$

In this case we must choose a basis for the eigenspace consisting of orthogonal unit vectors. If we take $\mathbf{q}_2 = \frac{1}{\sqrt{2}}(-1,\ 0,\ 1)^T$ and $\mathbf{q}_3 = \frac{1}{\sqrt{6}}(-1,\ 2,\ -1)^T$ then

$$Q = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & -\sqrt{3} & -1 \\ \sqrt{2} & 0 & 2 \\ \sqrt{2} & \sqrt{3} & -1 \end{pmatrix}$$

(g) $\lambda_1 = 6$ has unit eigenvector $\frac{1}{\sqrt{6}}(-2,\ -1,\ 1)^T$, $\lambda_2 = \lambda_3 = 0$. The vectors $\mathbf{x}_1 = (1,\ 0,\ 2)^T$ and $\mathbf{x}_2 = (-1,\ 2,\ 0)^T$ form a basis for the eigenspace corresponding to $\lambda = 0$. The Gram–Schmidt process can be used to construct an orthonormal basis.

$$\begin{aligned} r_{11} &= \|\mathbf{x}_1\| = \sqrt{5} \\ \mathbf{q}_1 &= \frac{1}{\sqrt{5}}\mathbf{x}_1 = \frac{1}{\sqrt{5}}(1,\ 0,\ 2)^T \\ \mathbf{p}_1 &= (\mathbf{x}_2^T\mathbf{q}_1)\mathbf{q}_1 = -\frac{1}{\sqrt{5}}\mathbf{q}_1 = -\frac{1}{5}(1,\ 0,\ 2)^T \\ \mathbf{x}_2 - \mathbf{p}_1 &= \left(-\frac{4}{5},\ 2,\ \frac{2}{5} \right)^T \\ r_{22} &= \|\mathbf{x}_2 - \mathbf{p}_1\| = \frac{2\sqrt{30}}{5} \\ \mathbf{q}_2 &= \frac{1}{\sqrt{30}}(-2,\ 5,\ 1)^T \end{aligned}$$

Thus

$$Q = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{30}} & -\frac{2}{\sqrt{6}} \\ 0 & \frac{5}{\sqrt{30}} & -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

6. If A is Hermitian, then $A^H = A$. Comparing the diagonal entries of A^H and A we see that

$$\overline{a}_{ii} = a_{ii} \quad \text{for} \quad i = 1, \ldots, n$$

Thus if A is Hermitian, then its diagonal entries must be real.

7. (a)

$$(A^H)^H = \overline{\left(\overline{A}^T\right)}^T = \left(\overline{\overline{A}}^T\right)^T = A$$

(b)

$$(\alpha A + \beta C)^H = \overline{\alpha A + \beta C}^T = (\overline{\alpha}\,\overline{A} + \overline{\beta}\,\overline{C})^T = \overline{\alpha}\,\overline{A}^T + \overline{\beta}\,\overline{C}^T = \overline{\alpha} A^H + \overline{\beta} C^H$$

(c) In general

$$\overline{AB} = \overline{A}\,\overline{B}$$

(See Exercise 17 of Section 1.) Using this we have

$$(AB)^H = (\overline{AB})^T = (\overline{A}\,\overline{B})^T = \overline{B}^T \overline{A}^T = B^H A^H$$

8. (i) $\langle \mathbf{z}, \mathbf{z} \rangle = \mathbf{z}^H \mathbf{z} = \Sigma |z_i|^2 \geq 0$ with equality if and only if $\mathbf{z} = \mathbf{0}$

(ii) $\overline{\langle \mathbf{w}, \mathbf{z} \rangle} = \overline{\mathbf{z}^H \mathbf{w}} = \mathbf{z}^T \overline{\mathbf{w}} = \overline{\mathbf{w}}^T \mathbf{z} = \mathbf{w}^H \mathbf{z} = \langle \mathbf{z}, \mathbf{w} \rangle$

(iii)
$$\begin{aligned} \langle \alpha\mathbf{z} + \beta\mathbf{w}, \mathbf{u} \rangle &= \mathbf{u}^H(\alpha\mathbf{z} + \beta\mathbf{w}) \\ &= \alpha\mathbf{u}^H\mathbf{z} + \beta\mathbf{u}^H\mathbf{w} \\ &= \alpha\langle \mathbf{z}, \mathbf{u} \rangle + \beta\langle \mathbf{w}, \mathbf{u} \rangle \end{aligned}$$

9.

$$\begin{aligned} \langle \mathbf{z}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle &= \overline{\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle} \\ &= \overline{\alpha\ \langle \mathbf{x}, \mathbf{z} \rangle + \beta\ \langle \mathbf{y}, \mathbf{z} \rangle} \\ &= \overline{\alpha}\,\overline{\langle \mathbf{x}, \mathbf{z} \rangle} + \overline{\beta}\,\overline{\langle \mathbf{y}, \mathbf{z} \rangle} \\ &= \overline{\alpha}\,\langle \mathbf{z}, \mathbf{x} \rangle + \overline{\beta}\,\langle \mathbf{z}, \mathbf{y} \rangle \end{aligned}$$

10. For $j = 1, \ldots, n$

$$\langle \mathbf{z}, \mathbf{u}_j \rangle = \langle a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n, \mathbf{u}_j \rangle = a_1\,\langle \mathbf{u}_1, \mathbf{u}_j \rangle + \cdots + a_n\,\langle \mathbf{u}_n, \mathbf{u}_j \rangle = a_j$$

Using the result from Exercise 9 we have

$$\begin{aligned} \langle \mathbf{z}, \mathbf{w} \rangle &= \langle \mathbf{z}, b_1\mathbf{u}_1 + \cdots + b_n\mathbf{u}_n \rangle \\ &= \overline{b_1}\,\langle \mathbf{z}, \mathbf{u}_1 \rangle + \cdots + \overline{b_n}\,\langle \mathbf{z}, \mathbf{u}_n \rangle \\ &= \overline{b_1}\,a_1 + \cdots + \overline{b_n}\,a_n \end{aligned}$$

11. The matrix A can be factored into a product $A = QDQ^H$ where

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & i & -i \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Let

$$E = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Note that $E^H E = D$. If we set

$$B = EQ^H = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -i & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

then

$$B^H B = (EQ^H)^H (EQ^H) = QE^H EQ^H = QDQ^H = A$$

12. (a) $U^H U = I = UU^H$

(c) If $\mathbf{x}$ is an eigenvector belonging to λ then

$$\|\mathbf{x}\| = \|U\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$$

Therefore $|\lambda|$ must equal 1.

14. Let U be a matrix that is both unitary and Hermitian. If λ is an eigenvalue of U and $\mathbf{z}$ is an eigenvector belonging to λ, then

$$U^2\mathbf{z} = U^H U\mathbf{z} = I\mathbf{z} = \mathbf{z}$$

and

$$U^2\mathbf{z} = U(U\mathbf{z}) = U(\lambda\mathbf{z}) = \lambda(U\mathbf{z}) = \lambda^2\mathbf{z}$$

Therefore

$$\begin{aligned} \mathbf{z} &= \lambda^2\mathbf{z} \\ (1 - \lambda^2)\mathbf{z} &= \mathbf{0} \end{aligned}$$

Since $\mathbf{z} \neq \mathbf{0}$ it follows that $\lambda^2 = 1$.

15. (a) A and T are similar and hence have the same eigenvalues. Since T is triangular, its eigenvalues are t_{11} and t_{22}.

(b) It follows from the Schur decomposition of A that

$$AU = UT$$

where U is unitary. Comparing the first columns of each side of this equation we see that

$$A\mathbf{u}_1 = U\mathbf{t}_1 = t_{11}\mathbf{u}_1$$

Hence $\mathbf{u}_1$ is an eigenvector belonging to t_{11}.

(c) Comparing the second column of $AU = UT$, we see that

$$\begin{aligned} A\mathbf{u}_2 &= U\mathbf{t}_2 \\ &= t_{12}\mathbf{u}_1 + t_{22}\mathbf{u}_2 \\ &\neq t_{22}\mathbf{u}_2 \end{aligned}$$

since $t_{12}\mathbf{u}_1 \neq \mathbf{0}$.

16. If A has Schur decomposition UTU^H and the diagonal entries of T are all distinct then by Exercise 19 in Section 3 there is an upper triangular matrix R that diagonalizes T. Thus we can factor T into a product RDR^{-1} where D is a diagonal matrix. It follows that

$$A = UTU^H = U(RDR^{-1})U^H = (UR)D(R^{-1}U^H)$$

and hence the matrix $X = UR$ diagonalizes A.

17. $M^H = (A - iB)^T = A^T - iB^T$
$\quad -M = -A - iB$
Therefore $M^H = -M$ if and only if $A^T = -A$ and $B^T = B$.

18. If A is skew Hermitian, then $A^H = -A$. Let λ be any eigenvalue of A and let $\mathbf{z}$ be a unit eigenvector belonging to λ. It follows that

$$\mathbf{z}^H A\mathbf{z} = \lambda\mathbf{z}^H\mathbf{z} = \lambda\|\mathbf{z}\|^2 = \lambda$$

and hence

$$\overline{\lambda} = \lambda^H = (\mathbf{z}^H A\mathbf{z})^H = \mathbf{z}^H A^H\mathbf{z} = -\mathbf{z}^H A\mathbf{z} = -\lambda$$

This implies that λ is purely imaginary.

19. If A is normal then there exists a unitary matrix U that diagonalizes A. If D is the diagonal matrix whose diagonal entries are the eigenvalues of A then $A = UDU^H$. The column vectors of U are orthonormal eigenvectors of A.

(a) Since $A^H = (UDU^H)^H = UD^HU^H$ and the matrix D^H is diagonal, we have that U diagonalizes A^H. Therefore A^H has a complete orthonormal set of eigenvectors and hence it is a normal matrix.

(b) $I + A = I + UDU^H = UIU^H + +UDU^H = +U(I + D)U^H$.
The matrix $I + D$ is diagonal, so U diagonalizes $I + A$. Therefore $I + A$ has a complete orthonormal set of eigenvectors and hence it is a normal matrix.

(c) $A^2 = UD^2U^H$.
The matrix D^2 is diagonal, so U diagonalizes A^2. Therefore A^2 has a complete orthonormal set of eigenvectors and hence it is a normal matrix.

20. $B = SAS^{-1} = \begin{pmatrix} a_{11} & \sqrt{a_{12}a_{21}} \\ \sqrt{a_{12}a_{21}} & a_{22} \end{pmatrix}$

Since B is symmetric it has real eigenvalues and an orthonormal set of eigenvectors. The matrix A is similar to B, so it has the same eigenvalues. Indeed, A is similar to the diagonal matrix D whose diagonal entries are the eigenvalues of B. Therefore A is diagonalizable and hence it has two linearly independent eigenvectors.

21. (a) $A^{-1} = \begin{pmatrix} 1 & 1-c & -1-c \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

$$A^{-1}CA = \begin{pmatrix} 0 & 1 & 0 \\ 1 & c+1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

(b) Let $B = A^{-1}CA$. Since B and C are similar they have the same eigenvalues. The eigenvalues of C are the roots of $p(x)$. Thus the roots of $p(x)$ are the eigenvalues of B. We saw in part (a) that B is symmetric. Thus all of the eigenvalues of B are real.

22. If A is Hermitian, then there is a unitary U that diagonalizes A. Thus

$$\begin{aligned} A &= UDU^H \\ &= (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_n^H \end{pmatrix} \\ &= (\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \ldots, \lambda_n\mathbf{u}_n) \begin{pmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_n^H \end{pmatrix} \\ &= \lambda_1\mathbf{u}_1\mathbf{u}_1^H + \lambda_2\mathbf{u}_2\mathbf{u}_2^H + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^H \end{aligned}$$

24. (a) Since the eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ form an orthonormal basis for C^n, the coordinates of $\mathbf{x}$ with respect to this basis are $c_i = \mathbf{u}_i^H\mathbf{x}_i$ for $i = 1, \ldots, n$. It follows then that

$$\begin{aligned} \mathbf{x} &= c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n \\ A\mathbf{x} &= c_1A\mathbf{u}_1 + c_2A\mathbf{u}_2 + \cdots + c_nA\mathbf{u}_n \\ &= \lambda_1c_1\mathbf{u}_1 + \lambda_2c_2\mathbf{u}_2 + \cdots + \lambda_nc_n\mathbf{u}_n \\ \mathbf{x}^HA\mathbf{x} &= \lambda_1c_1\mathbf{x}^H\mathbf{u}_1 + \lambda_2c_2\mathbf{x}^H\mathbf{u}_2 + \cdots + \lambda_nc_n\mathbf{x}^H\mathbf{u}_n \\ &= \lambda_1c_1\bar{c}_1 + \lambda_2c_2\bar{c}_2 + \cdots + \lambda_nc_n\bar{c}_n \\ &= \lambda_1|c_1|^2 + \lambda_2|c_2|^2 + \cdots + \lambda_n|c_n|^2 \end{aligned}$$

By Parseval's formula

$$\mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|^2 = \|\mathbf{c}\|^2$$

Thus

$$\begin{aligned}\rho(\mathbf{x}) &= \frac{\mathbf{x}^H A\mathbf{x}}{\mathbf{x}^H\mathbf{x}} \\ &= \frac{\lambda_1|c_1|^2 + \lambda_2|c_2|^2 + \cdots + \lambda_n|c_n|^2}{\|\mathbf{c}\|^2}\end{aligned}$$

(b) It follows from part (a) that

$$\frac{\lambda_{\min}\sum_{i=1}^{n}|c_i|^2}{\|\mathbf{c}\|^2} \le \rho(\mathbf{x}) \le \frac{\lambda_{\max}\sum_{i=1}^{n}|c_i|^2}{\|\mathbf{c}\|^2}$$

$$\lambda_{\min} \le \rho(\mathbf{x}) \le \lambda_{\max}$$

SECTION 5

1. If A has singular value decomposition $U\Sigma V^T$, then A^T has singular value decomposition $V\Sigma^T U^T$. The matrices Σ and Σ^T will have the same nonzero diagonal elements. Thus A and A^T have the same nonzero singular values.

3. If A is a matrix with singular value decomposition $U\Sigma V^T$, then the rank of A is the number of nonzero singular values it possesses, the 2-norm is equal to its largest singular value, and the closest matrix of rank 1 is $\sigma_1\mathbf{u}_1\mathbf{v}_1^T$.

(a) The rank of A is 1 and $\|A\|_2 = \sqrt{10}$. The closest matrix of rank 1 is A itself.

(c) The rank of A is 2 and $\|A\|_2 = 4$. The closest matrix of rank 1 is given by

$$4\mathbf{u}_1\mathbf{v}_1 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

(d) The rank of A is 3 and $\|A\|_2 = 3$. The closest matrix of rank 1 is given by

$$3\mathbf{u}_1\mathbf{v}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

5. (b) Basis for $R(A)$: $\mathbf{u}_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)^T$, $\mathbf{u}_2 = \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)^T$

Basis for $N(A^T)$: $\mathbf{u}_3 = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)$, $\mathbf{u}_4 = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)^T$

6. If A is symmetric then $A^TA = A^2$. Thus the eigenvalues of A^TA are $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$. The singular values of A are the positive square roots of the eigenvalues of A^TA.

7. The vectors $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ are all eigenvectors belonging to $\lambda = 0$. Hence these vectors are all in $N(A)$ and since dim $N(A) = n - r$, they form a basis for $N(A)$. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are all vectors in $N(A)^\perp = R(A^T)$. Since dim $R(A^T) = r$, it follows that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ form an orthonormal basis for $R(A^T)$.

8. If A is an $n \times n$ matrix with singular value decomposition $A = U\Sigma V^T$, then

$$A^TA = V\Sigma^2V^T \qquad \text{and} \qquad AA^T = U\Sigma^2U^T$$

If we set $X = VU^T$ then X is nonsingular and

$$X^{-1}(A^TA)X = UV^TV\Sigma^2V^TVU^T = U\Sigma^2U^T = AA^T$$

Therefore A^TA and AA^T are similar.

9. If σ is a singular value of A, then σ^2 is an eigenvalue of A^TA. Let $\mathbf{x}$ be an eigenvector of A^TA belonging to σ^2. It follows that

$$A^TA\mathbf{x} = \sigma^2\mathbf{x}$$

$$\mathbf{x}^TA^TA\mathbf{x} = \sigma^2\mathbf{x}^T\mathbf{x}$$

$$\|A\mathbf{x}\|_2^2 = \sigma^2\|\mathbf{x}\|_2^2$$

$$\sigma = \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

10. $$\begin{aligned} A^TA\hat{\mathbf{x}} &= A^TAA^+\mathbf{b} \\ &= V\Sigma^TU^TU\Sigma V^TV\Sigma^+U^T\mathbf{b} \\ &= V\Sigma^T\Sigma\Sigma^+U^T\mathbf{b} \end{aligned}$$

For any vector $\mathbf{y} \in R^m$

$$\Sigma^T\Sigma\Sigma^+\mathbf{y} = (\sigma_1y_1, \sigma_2y_2, \ldots, \sigma_ny_n)^T = \Sigma^T\mathbf{y}$$

Thus

$$A^TA\hat{\mathbf{x}} = V\Sigma^T\Sigma\Sigma^+(U^T\mathbf{b}) = V\Sigma^TU^T\mathbf{b} = A^T\mathbf{b}$$

11. $P = AA^+ = U\Sigma V^TV\Sigma^+U^T = U\Sigma\Sigma^+U^T$

The matrix $\Sigma\Sigma^+$ is an $m \times m$ diagonal matrix whose diagonal entries are all 0's and 1's. Thus we have

$$(\Sigma\Sigma^+)^T = \Sigma\Sigma^+ \quad \text{and} \quad (\Sigma\Sigma^+)^2 = \Sigma\Sigma^+$$

and it follows that

$$\begin{aligned} P^2 &= U(\Sigma\Sigma^+)^2U^T = U\Sigma^+\Sigma U^T = P \\ P^T &= U(\Sigma\Sigma^+)^TU^T = U\Sigma^+\Sigma U^T = P \end{aligned}$$

SECTION 6

1. (c) $\begin{pmatrix} 1 & 1/2 & -1 \\ 1/2 & 2 & 3/2 \\ -1 & 3/2 & 1 \end{pmatrix}$

2. $\lambda_1 = 4$, $\lambda_2 - 2$

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

If we set

$$\begin{pmatrix} x \\ y \end{pmatrix} = Q \begin{pmatrix} x' \\ y' \end{pmatrix}$$

then

$$(x \;\; y)A \begin{pmatrix} x \\ y \end{pmatrix} = (x' \;\; y')Q^T AQ \begin{pmatrix} x' \\ y' \end{pmatrix}$$

It follows that

$$Q^T AQ = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

and the equation of the conic can be written in the form

$$4(x')^2 + 2(y')^2 = 8$$

$$\frac{(x')^2}{2} + \frac{(y')^2}{4} = 1$$

The positive x' axis will be in the first quadrant in the direction of

$$\mathbf{q}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$$

The positive y' axis will be in the second quadrant in the direction of

$$\mathbf{q}_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$$

The graph will be exactly the same as Figure 6.6.3 except for the labeling of the axes.

3. (b) $A = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}$. The eigenvalues are $\lambda_1 = 7$, $\lambda_2 = -1$ with orthonormal eigenvectors

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T \quad \text{and} \quad \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T \quad \text{respectively.}$$

Let

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = Q^T \begin{pmatrix} x \\ y \end{pmatrix}$$

The equation simplifies to

$$7(x')^2 - (y')^2 = -28$$

$$\frac{(y')^2}{28} - \frac{(x')^2}{4} = 1$$

which is in standard form with respect to the $x'y'$ axis system.

(c) $A = \begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}$.

The eigenvalues are $\lambda_1 = 6$, $\lambda_2 = -4$ with orthonormal eigenvectors

$$\left(\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right)^T \quad \text{and} \quad \left(-\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)^T, \quad \text{respectively.}$$

Let

$$Q = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = Q^T \begin{pmatrix} x \\ y \end{pmatrix}$$

The equation simplifies to

$$6(x')^2 - 4(y')^2 = 24$$

$$\frac{(x')^2}{4} - \frac{(y')^2}{6} = 1$$

4. Using a suitable rotation of axes, the equation translates to

$$\lambda_1(x')^2 + \lambda_2(y')^2 = 1$$

Since λ_1 and λ_2 differ in sign, the graph will be an hyperbola.

5. The equation can be transformed into the form

$$\lambda_1(x')^2 + \lambda_2(y')^2 = \alpha$$

If either λ_1 and λ_2 is 0, then the graph is a pair of lines. Thus the conic section will be nondegenerate if and only if the eigenvalues of A are nonzero. The eigenvalues of A will be nonzero if and only if A is nonsingular.

6. (c) The eigenvalues are $\lambda_1 = 5$, $\lambda_2 = 2$. Therefore the matrix is positive definite.

(f) The eigenvalues are $\lambda_1 = 8$, $\lambda_2 = 2$, $\lambda_3 = 2$. Since all of the eigenvalues are positive, the matrix is positive definite.

7. (d) The Hessian of f is at $(1, 1)$ is

$$\begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

Its eigenvalues are $\lambda_1 = 9$, $\lambda_2 = 3$. Since both are positive, the matrix is positive definite and hence $(1, 1)$ is a local minimum.

(e) The Hessian of f at $(1,\ 0,\ 0)$ is

$$\begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Its eigenvalues are $\lambda_1 = 6$, $\lambda_2 = 1 + \sqrt{2}$, $\lambda_3 = 1 - \sqrt{2}$. Since they differ in sign, $(1,\ 0,\ 0)$ is a saddle point.

8. If A is symmetric positive definite, then all of its eigenvalues are positive. It follows that

$$\det(A) = \lambda_1\lambda_2\cdots\lambda_n > 0$$

The converse is not true. For example if I is the 2×2 identity matrix and we set $A = -I$ then $\det(A) = (-1)\cdot(-1) = 1$, however, A is not positive definite.

9. If A is symmetric positive definite, then all of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A are positive. Since 0 is not an eigenvalue, A is nonsingular. The eigenvalues of A^{-1} are $1/\lambda_1,\ 1/\lambda_2, \ldots, 1/\lambda_n$. Thus A^{-1} has positive eigenvalues and hence is positive definite.

10. A^TA is positive semidefinite since

$$\mathbf{x}^TA^TA\mathbf{x} = \|A\mathbf{x}\|^2 \geq 0$$

If A is singular then there exists a nonzero vector $\mathbf{x}$ such that

$$A\mathbf{x} = \mathbf{0}$$

It follows that

$$\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^TA^T\mathbf{0} = 0$$

and hence A^TA is not positive definite.

11. Let X be an orthogonal diagonalizing matrix for A. If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are the column vectors of X then by the remarks following Corollary 6.4.5 we can write

$$A\mathbf{x} = \lambda_1(\mathbf{x}^T\mathbf{x}_1)\mathbf{x}_1 + \lambda_2(\mathbf{x}^T\mathbf{x}_2)\mathbf{x}_2 + \cdots + \lambda_n(\mathbf{x}^T\mathbf{x}_n)\mathbf{x}_n$$

Thus

$$\mathbf{x}^TA\mathbf{x} = \lambda_1(\mathbf{x}^T\mathbf{x}_1)^2 + \lambda_2(\mathbf{x}^T\mathbf{x}_2)^2 + \cdots + \lambda_n(\mathbf{x}^T\mathbf{x}_n)^2$$

12. If A is positive definite, then

$$\mathbf{e}_i^TA\mathbf{e}_i > 0 \quad \text{for} \quad i = 1, \ldots, n$$

but

$$\mathbf{e}_i^TA\mathbf{e}_i = \mathbf{e}_i^T\mathbf{a}_i = a_{ii}$$

13. Let $\mathbf{x}$ be any nonzero vector in R^n and let $\mathbf{y} = S\mathbf{x}$. Since S is nonsingular, $\mathbf{y}$ is nonzero and

$$\mathbf{x}^TS^TAS\mathbf{x} = \mathbf{y}^TA\mathbf{y} > 0$$

Therefore S^TAS is positive definite.

14. If A is symmetric, then by Corollary 6.4.5 there is an orthogonal matrix U that diagonalizes A.

$$A = UDU^T$$

Since A is positive definite, the diagonal elements of D are all positive. If we set

$$Q = UD^{1/2}$$

then the columns of Q are mutually orthogonal and

$$\begin{aligned} A &= (UD^{1/2})((D^{1/2})^TU^T) \\ &= QQ^T \end{aligned}$$

SECTION 7

3. (a)

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{3} & 1 & 0 \\ 0 & 0 & -\frac{3}{4} & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & 0 \\ 0 & 0 & \frac{4}{3} & -1 \\ 0 & 0 & 0 & \frac{5}{4} \end{pmatrix}$$

(b) Since the diagonal entries of U are all positive it follows that A can be reduced to upper triangular form using only row operation III and the pivot elements are all positive. Therefore A must be positive definite.

6. A is symmetric positive definite

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$$

(i) $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^T A \mathbf{x} > 0 \quad (\mathbf{x} \neq \mathbf{0})$
since A is positive definite.

(ii) $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = (A\mathbf{x})^T \mathbf{y} = \mathbf{y}^T A \mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle$

(iii) $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = (\alpha\mathbf{x} + \beta\mathbf{y})^T A \mathbf{z}$

$$= \alpha\mathbf{x}^T A \mathbf{z} + \beta\mathbf{y}^T A \mathbf{z}$$
$$= \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$$

7. If $L_1 D_1 U_1 = L_2 D_2 U_2$, then

$$D_2^{-1} L_2^{-1} L_1 D_1 = U_2 U_1^{-1}$$

The left hand side represents a lower triangular matrix and the right hand side represents an upper triangular matrix. Therefore both matrices must be diagonal. Since the matrix U_1 can be transformed into the identity matrix using only row operation III it follows that the diagonal entries of U_1^{-1} must all be 1. Thus

$$U_2 U_1^{-1} = I$$

and hence

$$L_2^{-1} L_1 = D_2 D_1^{-1}$$

Therefore $L_2^{-1} L_1$ is a diagonal matrix and since its diagonal entries must also be 1's we have

$$U_2 U_1^{-1} = I = L_2^{-1} L_1 = D_2 D_1^{-1}$$

or equivalently

$$U_1 = U_2, \qquad L_1 = L_2, \qquad D_1 = D_2$$

8. If A is a positive definite symmetric matrix then A can be factored into a product $A = QDQ^T$ where Q is orthogonal and D is a diagonal matrix whose diagonal elements are all positive. Let E be a diagonal matrix with $e_{ii} = \sqrt{d_{ii}}$ for $i = 1, \ldots, n$. Since $E^T E = E^2 = D$ it follows that

$$A = QE^T EQ^T = (EQ^T)^T (EQ^T) = B^T B$$

where $B = EQ^T$.

9. If B is an $m \times n$ matrix of rank n and $\mathbf{x} \neq \mathbf{0}$, then $B\mathbf{x} \neq \mathbf{0}$. It follows that

$$\mathbf{x}^T B^T B\mathbf{x} = \|B\mathbf{x}\|^2 > 0$$

Therefore $B^T B$ is positive definite.

10. If A is symmetric, then its eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all real and there is an orthogonal matrix Q that diagonalizes A. It follows that

$$A = QDQ^T \qquad \text{and} \qquad e^A = Qe^D Q^T$$

The matrix e^A is symmetric since

$$(e^A)^T = Q(e^D)^T Q^T = Qe^D Q^T = e^A$$

The eigenvalues of e^A are the diagonal entries of e^D

$$\mu_1 = e^{\lambda_1}, \mu_2 = e^{\lambda_2}, \ldots, \mu_n = e^{\lambda_n}$$

Since e^A is symmetric and its eigenvalues are all positive, it follows that e^A is positive definite.

11. Since B is symmetric

$$B^2 = B^T B$$

Since B is also nonsingular, it follows from Theorem 6.7.1 that B^2 is positive definite.

12. (a) A is positive definite since A is symmetric and its eigenvalues $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{3}{2}$ are both positive. If $\mathbf{x} \in R^2$, then

$$\mathbf{x}^T A\mathbf{x} = x_1^2 - x_1 x_2 + x_2^2 = \mathbf{x}^T B\mathbf{x}$$

(b) If $\mathbf{x} \neq \mathbf{0}$, then

$$\mathbf{x}^T B\mathbf{x} = \mathbf{x}^T A\mathbf{x} > 0$$

since A is positive definite. Therefore B is also positive definite. However,

$$B^2 = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

is not positive definite. Indeed if $\mathbf{x} = (1, 1)^T$, then

$$\mathbf{x}^T B^2 \mathbf{x} = 0$$

13. (a) If A is an symmetric negative definite matrix, then its eigenvalues are all negative. Since the determinant of A is the product of the eigenvalues, it follows that $\det(A)$ will be positive if n is even and negative if n is odd.

(b) Let A_k denote the leading principal submatrix of A of order k and let $\mathbf{x}_1$ be a nonzero vector in R^k. If we set

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{pmatrix} \qquad \mathbf{x} \in R^n$$

then

$$\mathbf{x}_1^T A_k \mathbf{x}_1 = \mathbf{x}^T A\mathbf{x} < 0$$

Therefore the leading principal submatrices are all negative definite.

(c) The result in part (c) follows as an immediate consequence of the results from parts (a) and (b).

14. (a) Since $L_{k+1}L_{k+1}^T = A_{k+1}$, we have

$$\begin{pmatrix} L_k & \mathbf{0} \\ \mathbf{x}_k^T & \alpha_k \end{pmatrix} \begin{pmatrix} L_k^T & \mathbf{x}_k \\ \mathbf{0}^T & \alpha_k \end{pmatrix} = \begin{pmatrix} A_k & \mathbf{y}_k \\ \mathbf{y}_k^T & \beta_k \end{pmatrix}$$

$$\begin{pmatrix} L_kL_k^T & L_k\mathbf{x}_k \\ \mathbf{x}_k^TL_k^T & \mathbf{x}_k^T\mathbf{x}_k + \alpha_k^2 \end{pmatrix} = \begin{pmatrix} A_k & \mathbf{y}_k \\ \mathbf{y}_k^T & \beta_k \end{pmatrix}$$

Thus

$$L_k\mathbf{x}_k = \mathbf{y}_k$$

and hence

$$\mathbf{x}_k = L_k^{-1}\mathbf{y}_k$$

Once $\mathbf{x}_k$ has been computed one can solve for α_k.

$$\begin{aligned} \mathbf{x}_k^T\mathbf{x}_k + \alpha_k^2 &= \beta_k \\ \alpha_k &= (\beta_k - \mathbf{x}_k^T\mathbf{x}_k)^{1/2} \end{aligned}$$

(b) **Cholesky Factorization Algorithm**

Set $L_1 = (\sqrt{a_{11}})$

For $k = 1, \ldots, n-1$

(1) Let $\mathbf{y}_k$ be the vector consisting of the first k entries of $\mathbf{a}_{k+1}$ and let β_k be the $(k+1)$st entry of $\mathbf{a}_{k+1}$.

(2) Solve the lower triangular system $L_k\mathbf{x}_k = \mathbf{y}_k$ for $\mathbf{x}_k$.

(3) Set $\alpha_k = (\beta_k - \mathbf{x}_k^T\mathbf{x}_k)^{1/2}$

(4) Set

$$L_{k+1} = \begin{pmatrix} L_k & 0 \\ \mathbf{x}_k^T & \alpha_k \end{pmatrix}$$

End (For Loop)

$L = L_n$

The Cholesky decomposition of A is LL^T.

SECTION 8

7. (b)

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(c)

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

8. It follows from Theorem 6.8.2 that the other two eigenvalues must be

$$\lambda_2 = 2\exp\left(\frac{2\pi i}{3}\right) = -1 + i\sqrt{3}$$

and

$$\lambda_3 = 2\exp\left(\frac{4\pi i}{3}\right) = -1 - i\sqrt{3}$$

9. (a) $A\hat{\mathbf{x}} = \begin{pmatrix} B & O \\ O & C \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} B\mathbf{x} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \lambda\mathbf{x} \\ \mathbf{0} \end{pmatrix} = \lambda\hat{\mathbf{x}}$

(b) Since B is a positive matrix it has a positive eigenvalue r_1 satisfying the three conditions in Perron's Theorem. Similarly C has a positive eigenvalue r_2 satisfying the conditions of Perron's Theorem. Let $r = \max(r_1, r_2)$. By part (a), r is an eigenvalue of A and condition (iii) of Perron's Theorem implies its multiplicity can be at most 2. (It would have multiplicity 2 in the case that $r_1 = r_2$.) If r_1 has a positive eigenvector $\mathbf{x}$ and r_2 has a positive eigenvector $\mathbf{y}$ then r will have an eigenvector that is either of the form

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \quad \text{or of the form} \quad \begin{pmatrix} \mathbf{0} \\ \mathbf{y} \end{pmatrix}$$

(c) The eigenvalues of A are the eigenvalues of B and C. If $B = C$, then

$$r = r_1 = r_2 \quad \text{(from part (b))}$$

is an eigenvalue of multiplicity 2. If $\mathbf{x}$ is a positive eigenvector of B belonging to r then let

$$\mathbf{z} = \begin{pmatrix} \mathbf{x} \\ \mathbf{x} \end{pmatrix}$$

It follows that

$$A\mathbf{z} = \begin{pmatrix} B & O \\ O & B \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} B\mathbf{x} \\ B\mathbf{x} \end{pmatrix} = \begin{pmatrix} r\mathbf{x} \\ r\mathbf{x} \end{pmatrix} = r\mathbf{z}$$

Thus $\mathbf{z}$ is a positive eigenvector belonging to r.

10. There are only two possible partitions of the index set $\{1,\ 2\}$. If $I_1 = \{1\}$ and $I_2 = \{2\}$ then A will be reducible provided $a_{12} = 0$. If $I_1 = \{2\}$ and $I_2 = \{1\}$ then A will be reducible provided $a_{21} = 0$. Thus A is reducible if and only if $a_{12}a_{21} = 0$.

11. If A is an irreducible nonnegative 2×2 matrix then it follows from Exercise 10 that $a_{12}a_{21} > 0$. The characteristic polynomial of A

$$p(\lambda) = \lambda^2 - (a_{11} + a_{12})\lambda + (a_{11}a_{22} - a_{12}a_{21})$$

has roots

$$\frac{(a_{11}+a_{22}) \pm \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2}$$

The discriminant can be simplified to

$$(a_{11} - a_{22})^2 + 4a_{12}a_{21}.$$

Thus both roots are real. The larger root r_1 is obtained using the + sign.

$$\begin{aligned} r_1 &= \frac{(a_{11}+a_{22}) + \sqrt{(a_{11}-a_{22})^2 + 4a_{12}a_{21}}}{2} \\ &> \frac{a_{11}+a_{22}+|a_{11}-a_{22}|}{2} \\ &= \max(a_{11}, a_{22}) \\ &\geq 0 \end{aligned}$$

Finally r_1 has a positive eigenvector

$$\mathbf{x} = \begin{pmatrix} a_{12} \\ r_1 - a_{11} \end{pmatrix}$$

The case where A has two eigenvalues of equal modulus can only occur when

$$a_{11} = a_{22} = 0$$

In this case $\lambda_1 = \sqrt{a_{21}a_{12}}$ and $\lambda_2 = -\sqrt{a_{21}a_{12}}$.

12. The eigenvalues of A^k are $\lambda_1^k = 1, \lambda_2^k, \ldots, \lambda_n^k$. Clearly $|\lambda_j^k| \leq 1$ for $j = 2, \ldots, n$. However, A^k is a positive matrix and therefore by Perron's theorem $\lambda = 1$ is the dominant eigenvalue and it is a simple root of the characteristic equation for A^k. Therefore $|\lambda_j^k| < 1$ for $j = 2, \ldots, n$ and hence $|\lambda_j| < 1$ for $j = 2, \ldots, n$.

13. (a) It follows from Exercise 12 that $\lambda_1 = 1$ is the dominant eigenvector of A. By Perron's theorem it has a positive eigenvector $\mathbf{x}_1$.

(b) Each $\mathbf{y}_j$ in the chain is a probability vector and hence the coordinates of each vector are nonnegative numbers adding up to 1. Therefore

$$\|\mathbf{y}_j\|_1 = 1 \quad j = 1, 2, \ldots$$

(c) If

$$\mathbf{y}_0 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

then

$$\mathbf{y}_k = c_1\mathbf{x}_1 + c_2\lambda_2^k\mathbf{x}_2 + \cdots + c_n\lambda_n^k\mathbf{x}_n$$

and since $\|\mathbf{y}_k\| = 1$ for each k and

$$c_2\lambda_2^k\mathbf{x}_2 + \cdots + c_n\lambda_n^k\mathbf{x}_n \to 0 \quad k \to \infty$$

it follow that $c_1 \neq 0$.

(d) Since

$$\mathbf{y}_k = c_1\mathbf{x}_1 + c_2\lambda_2^k\mathbf{x}_2 + \cdots + c_n\lambda_n^k\mathbf{x}_n$$

and $|\lambda_j| < 1$ for $j = 2, \ldots, n$ it follows that

$$\lim_{k\to\infty} \mathbf{y}_k = c_1\mathbf{x}_1$$

$c_1\mathbf{x}_1$ is the steady-state vector.

(e) Each $\mathbf{y}_k$ is a probability vector and hence the limit vector $c_1\mathbf{x}_1$ must also be a probability vector. Since $\mathbf{x}_1$ is positive it follows that $c_1 > 0$. Thus we have

$$\|c_1\mathbf{x}_1\|_\infty = 1$$

and hence

$$c_1 = \frac{1}{\|\mathbf{x}_1\|_\infty}$$

14. In general if the matrix is nonnegative then there is no guarantee that it has a dominant eigenvalue with a positive eigenvector. So the results from parts (c) and (d) of Exercise 13 would not hold in this case. On the other hand if A^k is a positive matrix for some k, then by Exercise 12, $\lambda_1 = 1$ is the dominant eigenvalue of A and it has a positive eigenvector $\mathbf{x}_1$. Therefore the results from Exercise 13 will be valid in this case.

MATLAB EXERCISES

1. Initially $\mathbf{x} = \mathbf{e}_1$, the standard basis vector, and

$$A\mathbf{x} = \frac{5}{4}\mathbf{e}_1 = \frac{5}{4}\mathbf{x}$$

is in the same direction as $\mathbf{x}$. So $\mathbf{x}_1 = \mathbf{e}_1$ is an eigenvector of A belonging to the eigenvalue $\lambda_1 = \frac{5}{4}$. When the initial vector is rotated so that $\mathbf{x} = \mathbf{e}_2$ the image will be

$$A\mathbf{x} = \frac{3}{4}\mathbf{e}_2 = \frac{3}{4}\mathbf{x}$$

so $\mathbf{x}_2 = \mathbf{e}_2$ is an eigenvector of A belonging to the eigenvalue $\lambda_2 = \frac{3}{4}$. The second diagonal matrix has the same first eigenvalue-eigenvector pair and the second eigenvector is again $\mathbf{x}_2 = \mathbf{e}_2$, however, this time the eigenvalue is negative since $\mathbf{x}_2$ and $A\mathbf{x}_2$ are in opposite directions. In general for any 2×2 diagonal matrix D, the eigenvalues will be d_{11} and d_{22} and the corresponding eigenvectors will be $\mathbf{e}_1$ and $\mathbf{e}_2$.

2. For the identity matrix the eigenvalues are the diagonal entries so $\lambda_1 = \lambda_2 = 1$. In this case not only are $\mathbf{e}_1$ and $\mathbf{e}_2$ eigenvectors, but any vector $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ is an eigenvector.

3. In this case $\mathbf{x}$ and $A\mathbf{x}$ are equal when $\mathbf{x}$ makes an angle of 45° with the x axis. So $\lambda_1 = 1$ is an eigenvalue with eigenvector

$$\mathbf{x}_1 = \left(\cos\frac{\pi}{4}, \sin\frac{\pi}{4}\right)^T = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$$

The vectors $\mathbf{x}$ and $A\mathbf{x}$ are unit vectors in opposite directions when $\mathbf{x}$ makes an angle of 135° with the x axis. So $\lambda_2 = -1$ is an eigenvalue and the corresponding eigenvector is

$$\mathbf{x}_2 = \left(\cos\frac{3\pi}{4}, \sin\frac{3\pi}{4}\right)^T = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$$

4. In this case $\mathbf{x}$ and $A\mathbf{x}$ are never parallel so A cannot have any real eigenvalues. Therefore the two eigenvalues of A must be complex numbers.

6. For the ninth matrix the vectors $\mathbf{x}$ and $A\mathbf{x}$ are never parallel so A must have complex eigenvalues.

7. The tenth matrix is singular, so one of its eigenvalues is 0. To find the eigenvector using the `eigshow` utility you most rotate $\mathbf{x}$ until $A\mathbf{x}$ coincides with the zero vector. The other eigenvalue of this matrix is $\lambda_2 = 1.5$. Since the eigenvalues are distinct their corresponding eigenvectors must be linearly independent. The next two matrices both have multiple eigenvalues and both are defective. Thus for either matrix any pair of eigenvectors would be linearly dependent.

8. The characteristic polynomial of a 2×2 matrix is a quadratic polynomial and its graph will be a parabola. The eigenvalues will be equal when the graph of the parabola corresponding to the characteristic polynomial has its vertex on the x axis. For a random 2×2 matrix the probability that this will happen should be 0.

11. (a) $A - I$ is a rank one matrix. Therefore the dimension of the eigenspace corresponding to $\lambda = 1$ is 9, the nullity of $A - I$. Thus $\lambda = 1$ has multiplicity at least 9. Since the trace is 20, the remaining eigenvalue $\lambda_{10} = 11$. For symmetric matrices, eigenvalue computations should be quite accurate. Thus one would expect to get nearly full machine accuracy in the computed eigenvalues of A.

(b) The roots of a tenth degree polynomial are quite sensitive, i.e., any small roundoff errors in either the data or in the computations are liable to lead to significant errors in the computed roots. In particular if $p(\lambda)$ has multiple roots, the computed eigenvalues are liable to be complex.

12. (a) When $t = 4$, the eigenvalues change from real to complex. The matrix C corresponding to $t = 4$ has eigenvalues $\lambda_1 = \lambda_2 = 2$. The matrix X of eigenvectors is singular. Thus C does not have two linearly independent eigenvectors and hence must be defective.

(b) The eigenvalues of A correspond to the two points where the graph crosses the x-axis. For each t the graph of the characteristic polynomial will be a parabola. The vertices of these parabolas rise as t increases. When $t = 4$ the vertex will be tangent to the x-axis at $x = 2$. This corresponds to a double eigenvalue. When $t > 4$ the vertex will be above the x-axis. In this case there are no real roots and hence the eigenvalues must be complex.

13. If the rank of B is 2, then its nullity is $4 - 2 = 2$. Thus 0 is an eigenvalue of B and its eigenspace has dimension 2.

14. The reduced row echelon form of C has three lead 1's. Therefore the rank of C is 3 and its nullity is 1. Since $C^4 = O$, all of the eigenvalues of C must be 0. Thus $\lambda = 0$ is an eigenvalue of multiplicity 4 and its eigenspace only has dimension 1. Hence C is defective.

15. In theory A and B should have the same eigenvalues. However for a defective matrix it is difficult to compute the eigenvalues accurately. Thus even though B would be defective if computed in exact arithmetic, the matrix computed using floating point arithmetic may have distinct eigenvalues and the computed matrix X of eigenvectors may turn out to be nonsingular. If, however, `rcond` is very small, this would indicate that the column vectors of X are nearly dependent and hence that B may be defective.

16. (a) Both $A - I$ and $A + I$ have rank 3, so the eigenspaces corresponding to $\lambda_1 = 1$ and $\lambda_2 = -1$ should both have dimension 1.

(b) Since $\lambda_1 + \lambda_2 = 0$ and the sum of all four eigenvalues is 0, it follows that

$$\lambda_3 + \lambda_4 = 0$$

Since $\lambda_1\lambda_2 = -1$ and the product of all four eigenvalues is 1, it follows that

$$\lambda_3\lambda_4 = -1$$

Solving these two equations, we get $\lambda_3 = 1$ and $\lambda_4 = -1$. Thus 1 and -1 are both double eigenvalues. Since their eigenspaces each have dimension 1, the matrix A must be defective.

(d) The computed eigenvectors are linearly independent, but the computed matrix of eigenvectors does not diagonalize A.

17. Since

$$x(2)^2 = \frac{9}{10,000}$$

it follows that $x(2) = 0.03$. This proportion should remain constant in future generations. The proportion of genes for color-blindness in the male population should approach 0.03 as the number of generations increases. Thus in the long run 3% of the male population should be color-blind. Since $x(2)^2 = 0.0009$, one would expect that 0.09% of the female population will be color-blind in future generations.

18. (a) By construction S has integer entries and $\det(S) = 1$. It follows that $S^{-1} = \operatorname{adj} S$ will also have integer entries.

19. (a) By construction the matrix A is Hermitian. Therefore its eigenvalues should be real and the matrix X of eigenvectors should be unitary.

(b) The matrix B should be normal. Thus in exact arithmetic $B^H B$ and BB^H should be equal.

20. (a) If $A = USV^T$ then

$$AV = USV^T V = US$$

(b)

$$AV = (A\mathbf{v}_1, A\mathbf{v}_2) \qquad \text{and} \qquad US = (s_1\mathbf{u}_1, s_2\mathbf{u}_2)$$

Since $AV = US$ their corresponding column vectors must be equal. Thus we have

$$A\mathbf{v}_1 = s_1\mathbf{u}_1 \qquad \text{and} \qquad A\mathbf{v}_2 = s_2\mathbf{u}_2$$

(c) V and U are orthogonal matrices so $\mathbf{v}_1$, $\mathbf{v}_2$ are orthonormal vectors in R^n and $\mathbf{u}_1$, $\mathbf{u}_2$ are orthonormal vectors in R^m. The images $A\mathbf{v}_1$ and $A\mathbf{v}_2$ are orthogonal since

$$(A\mathbf{v}_1)^T A\mathbf{v}_2 = s_1 s_2 \mathbf{u}_1^T \mathbf{u}_2 = 0$$

(d) $\|A\mathbf{v}_1\| = \|s_1\mathbf{u}_1\| = s_1$ and $\|A\mathbf{v}_2\| = \|s_2\mathbf{u}_2\| = s_2$

21. If s_1, s_2 are the singular values of A, $\mathbf{v}_1$, $\mathbf{v}_2$ are the right singular vectors and $\mathbf{u}_1$, $\mathbf{u}_2$, are the corresponding left singular vectors, then the vectors $A\mathbf{x}$ and $A\mathbf{y}$ will be orthogonal when $\mathbf{x} = \mathbf{v}_1$ and $\mathbf{y} = \mathbf{v}_2$. When this happens

$$A\mathbf{x} = A\mathbf{v}_1 = s_1\mathbf{u}_1 \qquad \text{and} \qquad A\mathbf{y} = A\mathbf{v}_2 = s_2\mathbf{u}_2$$

Thus the image $A\mathbf{x}$ is a vector in the direction of $\mathbf{u}_1$ with length s_1 and the image $A\mathbf{y}$ is a vector in the direction of $\mathbf{u}_2$ with length s_2.
If you rotate the axes a full 360° the image vectors will trace out an ellipse. The major axis of the ellipse will be the line corresponding to the span of $\mathbf{u}_1$ and the diameter of the ellipse along its major axis will be $2s_1$ The minor axis of the ellipse will be the line corresponding to the span of $\mathbf{u}_2$ and the diameter of the ellipse along its minor axis will be $2s_2$.

22. The stationary points of the Hessian are $(-\frac{1}{4}, 0)$ and $(-\frac{71}{4}, 4)$. If the stationary values are substituted into the Hessian, then in each case we can compute the eigenvalues using the MATLAB's *eig* command. If we use the *double* command to view the eigenvalues in numeric format, the displayed values should be 7.6041 and −2.1041 for the first stationary point and −7.6041, 2.1041 for the second stationary point. Thus both stationary points are saddle points.

23. (a) The matrix C is symmetric and hence cannot be defective. The matrix X of eigenvectors should be an orthogonal matrix. The rank of $C - 7I$ is 1 and hence its nullity is 5. Therefore the dimension of the eigenspace corresponding to $\lambda = 7$ is 5.
(b) The matrix C is clearly symmetric and all of its eigenvalues are positive. Therefore C must be positive definite.
(c) In theory R and W should be equal. To see how close the computed matrices actually are, use MATLAB to compute the difference $R - W$.

24. In the $k \times k$ case, U and L will both be bidiagonal. All of the superdiagonal entries of U will be −1 and the diagonal entries will be

$$u_{11} = 2,\ u_{22} = \frac{3}{2},\ u_{33} = \frac{4}{3}, \ldots, u_{kk} = \frac{k+1}{k}$$

L will have 1's on the main diagonal and the subdiagonal entries will be

$$l_{21} = -\frac{1}{2},\ l_{32} = -\frac{2}{3},\ l_{43} = -\frac{3}{4}, \ldots, l_{k,k-1} = -\frac{k-1}{k}$$

Since A can be reduced to upper triangular form U using only row operation III and the diagonal entries of U are all positive, it follows that A must be positive definite.

25. (a) If you subtract 1 from the (6,6) entry of P, the resulting matrix will be singular.

(c) The matrix P is symmetric. The leading principal submatrices of P are all Pascal matrices. If all have determinant equal to 1, then all have positive determinants. Therefore P should be positive definite. The Cholesky factor R is a unit upper triangular matrix. Therefore

$$\det(P) = \det(R^T)\det(R) = 1$$

(d) If one sets $r_{88} = 0$, then R becomes singular. It follows that Q must also be singular since

$$\det(Q) = \det(R^T)\det(R) = 0$$

Since R is upper triangular, when one sets $r_{88} = 0$ it will only affect the $(8,8)$ entry of the product $R^T R$. Since R has 1's on the diagonal, changing r_{88} from 1 to 0 will have the effect of decreasing the $(8,8)$ entry of $R^T R$ by 1.

CHAPTER TEST A

1. The statement is true. If A were singular then we would have

$$\det(A - 0I) = \det(A) = 0$$

so $\lambda = 0$ would have to be an eigenvalue. Therefore if all of the eigenvalues are nonzero, then A cannot be singular.

One could also show that the statement is true by noting that if the eigenvalues of A are all nonzero then

$$\det(A) = \lambda_1\lambda_2\cdots\lambda_n \neq 0$$

and therefore A must be nonsingular.

2. The statement is false in general. A and A^T have the same eigenvalues but generally do not have the same eigenvectors. For example if

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

then $A\mathbf{e}_1 = \mathbf{e}_1$ so $\mathbf{e}_1$ is an eigenvector of A. However $\mathbf{e}_1$ is not an eigenvector of A^T since $A^T\mathbf{e}_1$ is not a multiple of $\mathbf{e}_1$.

3. The statement is false in general. The 2×2 identity matrix has eigenvalues $\lambda_1 = \lambda_2 = 1$, but it is not defective.

4. The statement is false. If A is a 4×4 matrix of rank 3, then the nullity of A is 1. Since $\lambda = 0$ is an eigenvalue of multiplicity 3 and the eigenspace has dimension 1, the matrix must be defective.

5. The statement is false. If A is a 4×4 matrix of rank 1, then the nullity of A is 3. Since $\lambda = 0$ is an eigenvalue of multiplicity 3 and the dimension of the eigenspace is also 3, the matrix is diagonalizable.

6. The statement is false in general. The matrix

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

has rank 1 even though all of its eigenvalues are 0.

7. The statement is true. If A has singular value decomposition $U\Sigma V^T$, then since U and V are orthogonal matrices, it follows that A and Σ have the same rank. The rank of the diagonal matrix Σ is equal to the number of nonzero singular values.

8. The statement is true. A and T are similar so they have the same eigenvalues. Since T is upper triangular its eigenvalues are its diagonal entries.

9. The statement is true. If A is symmetric positive definite then its eigenvalues are all positive and its determinant is positive. So A must be nonsingular. The inverse of a symmetric matrix is symmetric and the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A. It follows from Theorem 6.6.2 that A^{-1} must be positive definite.

10. The statement is false in general. For example let

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Although $\det(A) > 0$, the matrix is not positive definite since $\mathbf{x}^T A\mathbf{x} = -2$.

CHAPTER TEST B

1. **(a)** The eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = -1$, and $\lambda_3 = 0$,

(b) Each eigenspace has dimension 1. The vectors that form bases for the eigenspaces are $\mathbf{x}_1 = (1, 1, 1)^T, \mathbf{x}_2 = (0, 1, 2)^T, \mathbf{x}_3 = (0, 1, 1)^T$

(c)

$$\begin{aligned} A &= XDX^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix} \\ A^7 &= XD^7X^{-1} = XDX^{-1} = A \end{aligned}$$

2. Since A has real entries $\lambda_2 = 3 - 2i$ must be an eigenvalue and since A is singular the third eigenvalue is $\lambda_3 = 0$. We can find the last eigenvalue if we make use of the result that the trace of A is equal to the sum of its eigenvalues. Thus we have

$$\operatorname{tr}(A) = 4 = (3 + 2i) + (3 - 2i) + 0 + \lambda_4 = 6 + \lambda_4$$

and hence $\lambda_4 = -2$.

3. **(a)** $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$. If A is nonsingular then $\det(A) \neq 0$ and hence all of the eigenvalues of A must be nonzero.

(b) If λ is an eigenvalue of A then there exists a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. Multiplying both sides of this equation by A^{-1} we get

$$\begin{aligned} A^{-1}A\mathbf{x} &= A^{-1}(\lambda\mathbf{x}) \\ \mathbf{x} &= \lambda A^{-1}\mathbf{x} \\ \frac{1}{\lambda}\mathbf{x} &= A^{-1}\mathbf{x} \end{aligned}$$

and hence $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

4. The scalar a is a triple eigenvalue of A. The vector space $N(A-aI)$ consists of all vectors whose third entry is 0. The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ form a basis for this eigenspace and hence the dimension of the eigenspace is 2. Since the dimension of the eigenspace is less than the multiplicity of the eigenvalue, the matrix must be defective.

5. **(a)**

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 10 & 10 \\ 2 & 10 & 14 \end{pmatrix} \to \begin{pmatrix} 4 & 2 & 2 \\ 0 & 9 & 9 \\ 0 & 9 & 13 \end{pmatrix} \to \begin{pmatrix} 4 & 2 & 2 \\ 0 & 9 & 9 \\ 0 & 0 & 4 \end{pmatrix}$$

Since we were able to reduce A to upper triangular form U using only row operation III and the diagonal entries of U are all positive, it follows that A is positive definite.

(b)

$$\begin{aligned} U &= DL^T = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ A &= LDL^T = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

(c)

$$\begin{aligned} L_1 &= LD^{\frac{1}{2}} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 3 & 2 \end{pmatrix} \\ A &= L_1L_1^T = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 3 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned}$$

6. The first partials of F are

$$f_x = 3x^2y + 2x - 2 \qquad \text{and} \qquad f_y = x^3 + 2y - 1$$

At $(1,0)$ we have $f_x(1,0) = 0$ and $f_y(1,0) = 0$. So $(1,0)$ is a stationary point. The second partials of f are

$$f_{xx} = 6xy + 2, \quad f_{xy} = f_{yx} = 3x^2, \quad f_{yy} = 2$$

At the point $(1,0)$ the Hessian is

$$H = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

The eigenvalues of H are $\lambda_1 = 5$ and $\lambda_2 = -1$. Since the eigenvalues differ in sign it follows that H is indefinite and hence the stationary point $(1,0)$ is a saddle point.

7. The eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = -2$ and the corresponding eigenvectors are $\mathbf{x}_1 = (1,1)^T$ and $\mathbf{x}_2 = (2,3)^T$. The matrix $X = (\mathbf{x}_1, \mathbf{x}_2)$ diagonalizes A and $e^{tA} = Xe^{tD}X^{-1}$. The solution to the initial value problem is

$$\begin{aligned} \mathbf{Y}(t) = e^{tA}\mathbf{Y}_0 &= Xe^{tD}X^{-1}\mathbf{Y}_0 \\ &= \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{pmatrix} \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} e^{-t} + 2e^{-2t} \\ e^{-t} + 3e^{-2t} \end{pmatrix} \end{aligned}$$

8. (a) Since A is symmetric there is an orthogonal matrix that diagonalizes A. So A cannot be defective and hence the eigenspace corresponding to the triple eigenvalue $\lambda = 0$ (that is, the nullspace of A) must have dimension 3.

(b) Since λ_1 is distinct from the other eigenvalues, the eigenvector $\mathbf{x}_1$ will be orthogonal to $\mathbf{x}_2$, $\mathbf{x}_3$, and $\mathbf{x}_4$.

(c) To construct an orthogonal matrix that diagonalizes A, set $\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1$. The vectors $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$ form a basis for $N(A)$. Use the Gram-Schmidt process to transform this basis into an orthonormal basis $\{\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$. Since the vector $\mathbf{u}_1$ is in $N(A)^\perp$, it follows that $U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$ is an orthogonal matrix and U diagonalizes A.

(d) Since A is symmetric it can be factored into a product $A = QDQ^T$ where Q is orthogonal and D is diagonal. It follows that $e^A = Qe^DQ^T$. The matrix e^A is symmetric since

$$(e^A)^T = Q(e^D)^TQ^T = Qe^DQ^T = e^A$$

The eigenvalues of e^A are $\lambda_1 = e$ and $\lambda_2 = \lambda_3 = \lambda_4 = 1$. Since e^A is symmetric and its eigenvalues are all positive, it follows that e^A is positive definite.

9. (a) $\mathbf{u}_1^H\mathbf{z} = 5 - 7i$ and $\mathbf{z}^H\mathbf{u}_1 = 5 + 7i$.
$c_2 = \mathbf{u}_2^H\mathbf{z} == 1 - 5i$.

(b)

$$\begin{aligned} \|\mathbf{z}\|^2 = |c_1|^2 + |c_2|^2 &= (5-7i)(5+7i) + (1-5i)(1+5i) \\ &= 25 + 49 + 1 + 25 \end{aligned}$$

$$= 100$$

Therefore $\|\mathbf{z}\| = 10$.

10. (a) The matrix B is symmetric so it eigenvalues are all real. Furthermore, if $\mathbf{x} \neq \mathbf{0}$, then

$$\mathbf{x}^T B\mathbf{x} = \mathbf{x}^T A^T A\mathbf{x} = \|A\mathbf{x}\|^2 \geq 0$$

So B is positive semidefinite and hence its eigenvalues are all nonnegative. Furthermore $N(A)$ has dimension 2, so $\lambda = 0$ is an eigenvalue of multiplicity 2. In summary B is a symmetric positive semidefinite matrix with a double eigenvalue $\lambda = 0$.

(b) The matrix B can be factored into a product QDQ^T where Q is an orthogonal matrix and D is diagonal. It follows that $C = Qe^D Q^T$. So C is symmetric and its eigenvalues are the diagonal entries of e^D which are all positive. Therefore C is a symmetric positive definite matrix.

11. (a) If A has Schur decomposition UTU^H, then U is unitary and T is upper triangular. The matrices A and T are similar so they have the same eigenvalues. Since T is upper triangular it follows that $t_{11}, t_{22}, \ldots, t_{nn}$ are the eigenvalues of both T and A.

(b) If B is Hermitian with Schur decomposition WSW^H, then W is unitary and S is diagonal. The eigenvalues of B are the diagonal entries of S and the column vectors of W are the corresponding eigenvectors.

12. (a) Since A has 3 nonzero singular values, its rank is 3.

(b) If U is the matrix on the left in the given factorization then its first 3 columns, $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ form an orthonormal basis for $R(A)$.

(c) The matrix on the right in the factorization is V^T. The nullity of A is 1 and the vector $\mathbf{v}_4 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2})^T$ forms a basis for $N(A)$.

(d)

$$B = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T = 100 \begin{pmatrix} \frac{2}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \\ 30 & 30 & 30 & 30 \end{pmatrix}$$

(e) $\|B - A\|_F = \sqrt{10^2 + 10^2} = 10\sqrt{2}$.

CHAPTER 7

SECTION 1

The answers to all of the exercises in this section are included in the text.

SECTION 2

4. (a) (i) $n(mr + mn + n)$ multiplications and $(n-1)m(n+r)$ additions.
(ii) $(mn + nr + mr)$ multiplications and $(n-1)(m+r)$ additions.
(iii) $mn(r+2)$ multiplications and $m(n-1)(r+1)$ additions.

5. (a) The matrix $\mathbf{e}_k\mathbf{e}_i^T$ will have a 1 in the (k,i) position and 0's in all other positions. Thus if $B = I - \alpha\mathbf{e}_k\mathbf{e}_i^T$, then

$$b_{ki} = -\alpha \quad \text{and} \quad b_{sj} = \delta_{sj} \qquad (s,j) \neq (k,i)$$

Therefore $B = E_{ki}$

(b)
$$\begin{aligned} E_{ji}E_{ki} &= (I - \beta\mathbf{e}_j\mathbf{e}_i^T)(I - \alpha\mathbf{e}_k\mathbf{e}_i^T) \\ &= I - \alpha\mathbf{e}_k\mathbf{e}_i^T - \beta\mathbf{e}_j\mathbf{e}_i^T + \alpha\beta\mathbf{e}_j\mathbf{e}_i^T\mathbf{e}_k\mathbf{e}_i^T \\ &= I - (\alpha\mathbf{e}_k + \beta\mathbf{e}_j)e_i^T \end{aligned}$$

(c)
$$\begin{aligned} (I + \alpha\mathbf{e}_k\mathbf{e}_i^T)E_{ki} &= (I + \alpha\mathbf{e}_k\mathbf{e}_i^T)(I - \alpha\mathbf{e}_k\mathbf{e}_i^T) \\ &= I - \alpha^2\mathbf{e}_k\mathbf{e}_i^T\mathbf{e}_k\mathbf{e}_i^T \\ &= I - \alpha^2(\mathbf{e}_i^T\mathbf{e}_k)\mathbf{e}_k\mathbf{e}_i^T \\ &= I \qquad (\text{since } \mathbf{e}_i^T\mathbf{e}_k = 0) \end{aligned}$$

Therefore

$$E_{ki}^{-1} = I + \alpha\mathbf{e}_k\mathbf{e}_i^T$$

6. $\det(A) = \det(L)\det(U) = 1 \cdot \det(U) = u_{11}u_{22}\cdots u_{nn}$

7. Algorithm for solving $LDL^T\mathbf{x} = \mathbf{b}$

For $k = 1, \ldots, n$

Set $y_k = b_k - \sum_{i=1}^{k-1} \ell_{ki} y_i$

Set $z_k = y_k / d_{ii}$

End (For Loop)

For $k = n-1, \ldots, 1$

Set $x_k = z_k - \sum_{j=k+1}^{n} \ell_{jk} x_j$

End (For Loop)

8. (a) Algorithm for solving tridiagonal systems using diagonal pivots

For $k = 1, \ldots, n-1$

Set $m_k := c_k / a_k$

$a_{k+1} := a_{k+1} - m_k b_k$

$d_{k+1} := d_{k+1} - m_k d_k$

End (For Loop)

Set $x_n := d_n / a_n$

For $k = n-1, n-2, \ldots, 1$

Set $x_k := (d_k - b_k x_{k+1}) / a_k$

End (For Loop)

. 9 (b) To solve $A\mathbf{x} = \mathbf{e}_j$, one must first solve $L\mathbf{y} = \mathbf{e}_j$ using forward substitution. From part (a) it follows that this requires $[(n-j)(n-j+1)]/2$ multiplications and $[(n-j-1)(n-j)]/2$ additions. One must then perform back substitution to solve $U\mathbf{x} = \mathbf{y}$. This requires n divisions, $n(n-1)/2$ multiplications and $n(n-1)/2$ additions. Thus altogether, given the LU factorization of A, the number of operations to solve $A\mathbf{x} = \mathbf{e}_j$ is

$$\frac{(n-j)(n-j+1) + n^2 + n}{2} \quad \text{multiplications/divisions}$$

and

$$\frac{(n-j-1)(n-j) + n^2 - n}{2} \quad \text{additions/subtractions}$$

10. Given A^{-1} and $\mathbf{b}$, the multiplication $A^{-1}\mathbf{b}$ requires n^2 scalar multiplications and $n(n-1)$ scalar additions. The same number of operations is required in order to solve $LU\mathbf{x} = \mathbf{b}$ using Algorithm 7.2.2. Thus it is not really worthwhile to calculate A^{-1}, since this calculation requires three times the amount of work it would take to determine L and U.

11. If

$$A(E_1E_2E_3) = L$$

then

$$A = L(E_1E_2E_3)^{-1} = LU$$

The elementary matrices E_1^{-1}, E_2^{-1}, E_3^{-1} will each be upper triangular with ones on the diagonal. Indeed,

$$E_1^{-1} = \begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad E_2^{-1} = \begin{pmatrix} 1 & 0 & \frac{a_{13}}{a_{11}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{a_{23}}{a_{22}^{(1)}} \\ 0 & 0 & 1 \end{pmatrix}$$

where $a_{22}^{(1)} = a_{22} - \frac{a_{12}}{a_{11}}$. If we let

$$u_{12} = \frac{a_{12}}{a_{11}}, \quad u_{13} = \frac{a_{13}}{a_{11}}, \quad u_{23} = \frac{a_{23}}{a_{22}^{(1)}}$$

then

$$U = E_3^{-1}E_2^{-1}E_1^{-1} = \begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

SECTION 3

6. (a) $\left(\begin{array}{rrr|r} 5 & 4 & 7 & 2 \\ 2 & -4 & 3 & -5 \\ 2 & 8 & 6 & 4 \end{array}\right) \to \left(\begin{array}{rrr|r} 4 & 0 & 4 & 0 \\ 3 & 0 & 6 & -3 \\ 2 & 8 & 6 & 4 \end{array}\right)$

$\to \left(\begin{array}{rrr|r} 2 & 0 & 0 & 2 \\ 3 & 0 & 6 & -3 \\ 2 & 8 & 6 & 4 \end{array}\right)$

$$\begin{array}{ll} 2x_1 = 2 & x_1 = 1 \\ 3 + 6x_3 = -3 & x_3 = -1 \\ 2 + 8x_2 - 6 = 4 & x_2 = 1 \end{array}$$

$\mathbf{x} = (1,\ 1,\ -1)^T$

(b) The pivot rows were 3, 2, 1 and the pivot columns were 2, 3, 1. Therefore

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Rearranging the rows and columns of the reduced matrix from part (a), we get

$$U = \begin{pmatrix} 8 & 6 & 2 \\ 0 & 6 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

The matrix L is formed using the multipliers $-\frac{1}{2}, \frac{1}{2}, \frac{2}{3}$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{2}{3} & 1 \end{pmatrix}$$

(c) The system can be solved in 3 steps.

(1) Solve $L\mathbf{y} = P\mathbf{c}$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ -\frac{1}{2} & 1 & 0 & -4 \\ \frac{1}{2} & \frac{2}{3} & 1 & 5 \end{array}\right) \qquad \begin{array}{l} y_1 = 2 \\ y_2 = -3 \\ y_3 = 6 \end{array}$$

(2) Solve $U\mathbf{z} = \mathbf{y}$

$$\left(\begin{array}{ccc|c} 8 & 6 & 2 & 2 \\ 0 & 6 & 3 & -3 \\ 0 & 0 & 2 & 6 \end{array}\right) \qquad \begin{array}{l} z_1 = 1 \\ z_2 = -2 \\ z_3 = 3 \end{array}$$

(3) Set $\mathbf{x} = Q\mathbf{z}$

$$\mathbf{x} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$$

SECTION 4

3. Let $\mathbf{x}$ be a nonzero vector in R^2

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{|x_1|}{\sqrt{x_1^2 + x_2^2}} \le 1$$

Therefore

$$\|A\|_2 = \max_{\mathbf{x} \ne \mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le 1$$

On the other hand

$$\|A\|_2 \ge \frac{\|A\mathbf{e}_1\|_2}{\|\mathbf{e}_1\|_2} = 1$$

Therefore $\|A\|_2 = 1$.

4. (a) D has singular value decomposition $U\Sigma V^T$ where the diagonal entries of Σ are $\sigma_1 = 5$, $\sigma_2 = 4$, $\sigma_3 = 3$, $\sigma_4 = 2$ and

$$U = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(b) $\|D\|_2 = \sigma_1 = 5$

5. If D is diagonal then its singular values are the square roots of the eigenvalues of $D^TD = D^2$. The eigenvalues of D^2 are $d_{11}^2, d_{22}^2, \ldots, d_{nn}^2$ and hence it follows that

$$\|D\|_2 = \sigma_1 = \max_{1\le i\le n} |d_{ii}|$$

6. It follows from Theorem 7.4.2 that

$$\|D\|_1 = \|D\|_\infty = \max_{1\le i\le n} |d_{ii}|$$

and it follows from Exercise 5 that this is also the value of $\|D\|_2$. Thus for a diagonal matrix all 3 norms are equal.

8. (a) If $\|\cdot\|_M$ and $\|\cdot\|_V$ are compatible, then for any nonzero vector $\mathbf{x}$,

$$\|\mathbf{x}\|_V = \|I\mathbf{x}\|_V \le \|V\|_M \|\mathbf{x}\|_V$$

Dividing by $\|\mathbf{x}\|_V$ we get

$$1 \le \|I\|_M$$

(b) If $\|\cdot\|_M$ is subordinate to $\|\cdot\|_V$, then

$$\frac{\|I\mathbf{x}\|_V}{\|\mathbf{x}\|_V} = 1$$

for all nonzero vectors $\mathbf{x}$ and it follows that

$$\|I\|_M = \max_{\mathbf{x}\ne 0} \frac{\|I\mathbf{x}\|_V}{\|\mathbf{x}\|_V} = 1$$

9. (a) $\|X\|_\infty = \|\mathbf{x}\|_\infty$ since the ith row sum is just $|x_i|$ for each i.

(b) The 1-norm of a matrix is equal to the maximum of the 1-norm of its column vectors. Since X only has one column its 1-norm is equal to the 1-norm of that column vector.

11. Let $\mathbf{x}$ be a nonzero vector in R^n

$$\begin{aligned}
\frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} &= \frac{\displaystyle\max_{1\le i\le m} |\sum_{j=1}^n a_{ij}x_j|}{\displaystyle\max_{1\le j\le n} |x_j|} \\
&\le \frac{\displaystyle\max_{1\le j\le n} |x_j| \max_{1\le i\le m} |\sum_{j=1}^n a_{ij}|}{\displaystyle\max_{1\le j\le n} |x_j|} \\
&= \max_{1\le i\le m} \left|\sum_{j=1}^n a_{ij}\right| \\
&\le \max_{1\le i\le m} \sum_{j=1}^n |a_{ij}|
\end{aligned}$$

Therefore

$$\|A\|_\infty = \max_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$$

Let k be the index of the row of A for which $\sum_{j=1}^{n} |a_{ij}|$ is a maximum. Define $x_j = \operatorname{sgn} a_{kj}$ for $j = 1, \ldots, n$ and let $\mathbf{x} = (x_1, \ldots, x_n)^T$. Note that $\|\mathbf{x}\|_\infty = 1$ and $a_{kj}x_j = |a_{kj}|$ for $j = 1, \ldots, n$. Thus

$$\|A\|_\infty \geq \|A\mathbf{x}\|_\infty = \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \geq \sum_{j=1}^{n} |a_{kj}|$$

Therefore

$$\|A\|_\infty = \max_{1\leq i\leq m} \left(\sum_{j=1}^{n} |a_{ij}| \right)$$

12. $\|A\|_F = \left(\sum_j \sum_i a_{ij}^2 \right)^{1/2} = \left(\sum_i \sum_j a_{ij}^2 \right)^{1/2} = \|A^T\|_F$

13. $\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ji}| = \|A\|_1$

14. $\|A\|_2 = \sigma_1 = 5$ and

$$\|A\|_F = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_5^2)^{\frac{1}{2}} = 6$$

15. (a) Let $k = \min(m, n)$.

$$\|A\|_2 = \sigma_1 \leq (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2)^{\frac{1}{2}} = \|A\|_F \tag{6}$$

(b) Equality will hold in (6) if $\sigma_2^2 = \cdots = \sigma_k^2 = 0$. It follows then that $\|A\|_2 = \|A\|_F$ if and only if the matrix A has rank 1.

16. Since

$$\{\mathbf{x} \mid \ \|\mathbf{x}\| = 1\} = \{\mathbf{x} \mid \ \mathbf{x} = \frac{1}{\|\mathbf{y}\|}\mathbf{y}, \quad \mathbf{y} \in R^n \ \text{ and } \ \mathbf{y} \neq \mathbf{0}\}$$

$$\begin{aligned} \|A\|_M &= \max_{\mathbf{y}\neq\mathbf{0}} \frac{\|A\mathbf{y}\|}{\|\mathbf{y}\|} \\ &= \max_{\mathbf{y}\neq\mathbf{0}} \left\| A\left(\frac{1}{\|\mathbf{y}\|}\mathbf{y}\right) \right\| \\ &= \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \end{aligned}$$

17. If $\mathbf{x}$ is a unit eigenvector belonging to the eigenvalue λ, then

$$|\lambda| = \|\lambda\mathbf{x}\| = \|A\mathbf{x}\| \leq \|A\|_M \|\mathbf{x}\| = \|A\|_M$$

18. If A is a stochastic matrix then $\|A\|_1 = 1$. It follows from Exercise 17 that if λ is an eigenvalue of A then

$$|\lambda| \le \|A\|_1 = 1$$

19. (b) $\|A\mathbf{x}\|_2 \le n^{1/2}\|A\mathbf{x}\|_\infty \le n^{1/2}\|A\|_\infty\|\mathbf{x}\|_\infty \le n^{1/2}\|A\|_\infty\|\mathbf{x}\|_2$

(c) Let $\mathbf{x}$ be any nonzero vector in R^n. It follows from part (a) that

$$\frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \le n^{1/2}\|A\|_2$$

and it follows from part (b) that

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le n^{1/2}\|A\|_\infty$$

Consequently

$$\|A\|_\infty = \max_{\mathbf{x}\ne\mathbf{0}} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \le n^{1/2}\|A\|_2$$

and

$$\|A\|_2 = \max_{\mathbf{x}\ne\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le n^{1/2}\|A\|_\infty$$

Thus

$$n^{-1/2}\|A\|_2 \le \|A\|_\infty \le n^{1/2}\|A\|_2$$

20. Let A be a symmetric matrix with orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. If $\mathbf{x} \in R^n$ then by Theorem 5.5.2

$$\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$$

where $c_i = \mathbf{u}_i^T\mathbf{x}$, $i = 1, \ldots, n$.

(a) $A\mathbf{x} = c_1A\mathbf{u}_1 + c_2A\mathbf{u}_2 + \cdots + c_nA\mathbf{u}_n$
$= c_1\lambda_1\mathbf{u}_1 + c_2\lambda_2\mathbf{u}_2 + \cdots + c_n\lambda_n\mathbf{u}_n.$

It follows from Parseval's formula that

$$\|A\mathbf{x}\|_2^2 = \sum_{i=1}^n (\lambda_i c_i)^2$$

(b) It follows from part (a) that

$$\min_{1\le i\le n} |\lambda_i| \left(\sum_{i=1}^n c_i^2\right)^{1/2} \le \|A\mathbf{x}\|_2 \le \max_{1\le i\le n} |\lambda_i| \left(\sum_{i=1}^n c_i^2\right)^{1/2}$$

Using Parseval's formula we see that

$$\left(\sum_{j=1}^n c_i^2\right)^{1/2} = \|\mathbf{x}\|_2$$

and hence for any nonzero vector $\mathbf{x}$ we have

$$\min_{1\le i\le n} |\lambda_i| \le \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \max_{1\le i\le n} |\lambda_i|$$

(c) If

$$|\lambda_k| = \max_{1 \le i \le n} |\lambda_i|$$

and $\mathbf{x}_k$ is an eigenvector belonging to λ_k, then

$$\frac{\|A\mathbf{x}_k\|_2}{\|\mathbf{x}_k\|_2} = |\lambda_k| = \max_{1 \le i \le n} |\lambda_i|$$

and hence it follows from part (b) that

$$\|A\|_2 = \max_{1 \le i \le n} |\lambda_i|$$

21.

$$A^{-1} = \begin{pmatrix} 100 & 99 \\ 100 & 100 \end{pmatrix}$$

22. Let A be the coefficient matrix of the first system and A' be the coefficient matrix of the second system. If $\mathbf{x}$ is the solution to the first system and $\mathbf{x}'$ is the solution to the second system then

$$\frac{\|\mathbf{x} - \mathbf{x}'\|_\infty}{\|\mathbf{x}\|_\infty} \approx 3.03$$

while

$$\frac{\|A - A'\|_\infty}{\|A\|_\infty} \approx 0.014$$

The systems are ill-conditioned in the sense that a relative change of 0.014 in the coefficient matrix results in a relative change of 3.03 in the solution.

24. $\operatorname{cond}(A) = \|A\|_M \|A^{-1}\|_M \ge \|AA^{-1}\|_M = \|I\|_M = 1.$

26. The given conditions allow us to determine the singular values of the matrix. Indeed, $\sigma_1 = \|A\|_2 = 8$ and since

$$\frac{\sigma_1}{\sigma_3} = \operatorname{cond}_2(A) = 2$$

it follows that $\sigma_3 = 4$. Finally

$$\begin{aligned} \sigma_1^2 + \sigma_2^2 + \sigma_3^2 &= \|A\|_F^2 \\ 64 + \sigma_2^2 + 16 &= 144 \end{aligned}$$

and hence $\sigma_2 = 8$.

27. (c) $\displaystyle \frac{1}{\operatorname{cond}_\infty(A)} \frac{\|\mathbf{r}\|_\infty}{\|\mathbf{b}\|_\infty} \le \frac{\|\mathbf{x} - \mathbf{x}'\|_\infty}{\|\mathbf{x}\|_\infty} \le \operatorname{cond}_\infty(A) \frac{\|\mathbf{r}\|_\infty}{\|\mathbf{b}\|_\infty}$

$0.0006 = \frac{1}{20}(0.012) \le \frac{\|\mathbf{x}-\mathbf{x}'\|_\infty}{\|\mathbf{x}\|_\infty} \le 20(0.012) = 0.24$

31. $\operatorname{cond}(AB) = \|AB\| \, \|(AB)^{-1}\| \le \|A\| \, \|B\| \, \|B^{-1}\| \, \|A^{-1}\| = \operatorname{cond}(A) \operatorname{cond}(B)$

32. It follows from Exercises 5 and 6 that

$$\|D\|_1 = \|D\|_2 = \|D\|_\infty = d_{\max}$$

and

$$\|D^{-1}\|_1 = \|D^{-1}\|_2 = \|D^{-1}\|_\infty = \frac{1}{d_{\min}}$$

Therefore the condition number of D will be $\dfrac{d_{\max}}{d_{\min}}$ no matter which of the 3 norms is used.

33. (a) For any vector $\mathbf{x}$

$$\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$$

Thus if $\mathbf{x}$ is nonzero, then

$$\frac{\|Q\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = 1$$

and hence

$$\|Q\|_2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|Q\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = 1$$

(b) The matrix $Q^{-1} = Q^T$ is also orthogonal and hence by part (a) we have

$$\|Q^{-1}\|_2 = 1$$

Therefore

$$\operatorname{cond}_2(Q) = 1$$

(c) $$\frac{1}{\operatorname{cond}_2(Q)}\frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2} \le \frac{\|\mathbf{e}\|_2}{\|\mathbf{x}\|_2} \le \operatorname{cond}_2(Q)\frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2}$$

Since $\operatorname{cond}_2(Q) = 1$, it follows that

$$\frac{\|\mathbf{e}\|_2}{\|\mathbf{x}\|_2} = \frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2}$$

34. (a) If $\mathbf{x}$ is any vector in R^r, then $A\mathbf{x}$ is a vector in R^n and

$$\|QA\mathbf{x}\|_2 = \|A\mathbf{x}\|_2$$

Thus for any nonzero vector $\mathbf{x}$

$$\frac{\|QA\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

and hence

$$\begin{aligned}\|QA\|_2 &= \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|QA\mathbf{x}\|_2}{\|\mathbf{x}\|_2}\\ &= \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}\\ &= \|A\|_2\end{aligned}$$

(b) For each nonzero vector $\mathbf{x}$ in R^n set $\mathbf{y} = V\mathbf{x}$. Since V is nonsingular it follows that $\mathbf{y}$ is nonzero. Furthermore

$$\{\mathbf{y}\,|\,\mathbf{y} = V\mathbf{x} \text{ and } \mathbf{x} \neq \mathbf{0}\} = \{\mathbf{x}\,|\,\mathbf{x} \neq \mathbf{0}\}$$

since any nonzero $\mathbf{y}$ can be written as

$$\mathbf{y} = V\mathbf{x} \qquad \text{where } \mathbf{x} = V^T\mathbf{y}$$

It follows that if $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} = V\mathbf{x}$, then

$$\frac{\|AV\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{\|AV\mathbf{x}\|_2}{\|V\mathbf{x}\|_2} = \frac{\|A\mathbf{y}\|_2}{\|\mathbf{y}\|_2}$$

and hence

$$\|AV\|_2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|AV\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_{\mathbf{y}\neq\mathbf{0}} \frac{\|A\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \|A\|_2$$

(c) It follows from parts (a) and (b) that

$$\|QAV\|_2 = \|Q(AV)\|_2 = \|AV\|_2 = \|A\|_2$$

35. (a) If A has singular value decomposition $U\Sigma V^T$, then it follows from the Cauchy-Schwarz inequality that

$$|\mathbf{x}^T A\mathbf{y}| \leq \|\mathbf{x}\|_2\|A\mathbf{y}\|_2 \leq \|\mathbf{x}\|_2\|\mathbf{y}\|_2\|A\|_2 = \sigma_1\|\mathbf{x}\|_2\|\mathbf{y}\|_2$$

Thus if $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors, then

$$\frac{|\mathbf{x}^T A\mathbf{y}|}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2} \leq \sigma_1$$

(b) If we set $\mathbf{x}_1 = \mathbf{u}_1$ and $\mathbf{y}_1 = \mathbf{v}_1$, then

$$\|\mathbf{x}_1\|_2 = \|\mathbf{u}_1\|_2 = 1 \qquad \text{and} \qquad \|\mathbf{y}_1\|_2 = \|\mathbf{v}_1\|_2 = 1$$

and

$$A\mathbf{y}_1 = A\mathbf{v}_1 = \sigma_1\mathbf{u}_1$$

Thus

$$\mathbf{x}_1^T A\mathbf{y}_1 = \mathbf{u}_1^T(\sigma_1\mathbf{u}_1) = \sigma_1$$

and hence

$$\frac{|\mathbf{x}_1^T A\mathbf{y}_1|}{\|\mathbf{x}_1\|_2\|\mathbf{y}_1\|_2} = \sigma_1$$

Combining this with the result from part (a) we have

$$\max_{\mathbf{x}\neq\mathbf{0},\mathbf{y}\neq\mathbf{0}} \frac{|\mathbf{x}^T A\mathbf{y}|}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2} = \sigma_1$$

36. For each nonzero vector $\mathbf{x}$ in R^n

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{\|U\Sigma V^T\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{\|\Sigma V^T\mathbf{x}\|_2}{\|V^T\mathbf{x}\|_2} = \frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2}$$

where $\mathbf{y} = V^T\mathbf{x}$. Thus

$$\min_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \min_{\mathbf{y}\neq 0} \frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2}$$

For any nonzero vector $\mathbf{y} \in R^n$

$$\frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \frac{\left(\sum_{i=1}^{n} \sigma_i^2 y_i^2\right)^{1/2}}{\left(\sum_{i=1}^{n} y_i^2\right)^{1/2}} \geq \frac{\sigma_n\|\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \sigma_n$$

Thus

$$\min \frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \geq \sigma_n$$

On the other hand

$$\min_{\mathbf{y}\neq 0} \frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \leq \frac{\|\Sigma\mathbf{e}_n\|_2}{\|\mathbf{e}_n\|_2} = \sigma_n$$

Therefore

$$\min_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \min_{\mathbf{y}\neq 0} \frac{\|\Sigma\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \sigma_n$$

37. For any nonzero vector $\mathbf{x}$

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \|A\|_2 = \sigma_1$$

It follows from Exercise 33 that

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \geq \sigma_n$$

Thus if $\mathbf{x} \neq \mathbf{0}$, then

$$\sigma_n\|\mathbf{x}\|_2 \leq \|A\mathbf{x}\|_2 \leq \sigma_1\|\mathbf{x}\|_2$$

Clearly this inequality is also valid if $\mathbf{x} = \mathbf{0}$.

38. (a) It follows from Exercise 34 that

$$\|QA\|_2 = \|A\|_2 \quad \text{and} \quad \|A^{-1}Q^T\|_2 = \|A^{-1}\|_2$$

$$\|AQ\|_2 = \|A\|_2 \quad \text{and} \quad \|Q^TA^{-1}\|_2 = \|A^{-1}\|_2$$

Thus

$$\begin{aligned} \text{cond}_2(QA) &= \|QA\|_2\|A^{-1}Q^T\|_2 = \text{cond}_2(A) \\ \text{cond}_2(AQ) &= \|AQ\|_2\|Q^TA^{-1}\|_2 = \text{cond}_2(A) \end{aligned}$$

(b) It follows from Exercise 34 that

$$\|B\|_2 = \|A\|_2$$

and

$$\|B^{-1}\|_2 = \|Q^TA^{-1}Q\|_2 = \|A^{-1}\|_2$$

Therefore

$$\text{cond}_2(B) = \text{cond}_2(A)$$

39. If A is a symmetric $n \times n$ matrix, then there exists an orthogonal matrix Q that diagonalizes A.

$$Q^T A Q = D$$

The diagonal elements of D are the eigenvalues of A. Since A is symmetric and nonsingular its eigenvalues are all nonzero real numbers. It follows from Exercise 38 that

$$\operatorname{cond}_2(A) = \operatorname{cond}_2(D)$$

and it follows from Exercise 32 that

$$\operatorname{cond}_2(D) = \frac{\lambda_{\max}}{\lambda_{\min}}$$

SECTION 5

7. (b)

$$G = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$(GA \mid G\mathbf{b}) = \left(\begin{array}{cc|c} \sqrt{2} & 3\sqrt{2} & 3\sqrt{2} \\ 0 & \sqrt{2} & 2\sqrt{2} \end{array}\right), \qquad \mathbf{x} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

(c)

$$G = \begin{pmatrix} \frac{4}{5} & 0 & -\frac{3}{5} \\ 0 & 1 & 0 \\ -\frac{3}{5} & 0 & -\frac{4}{5} \end{pmatrix}$$

$$(GA \mid G\mathbf{b}) = \left(\begin{array}{ccc|c} 5 & -5 & 2 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & -2 \end{array}\right), \qquad \mathbf{x} = \begin{pmatrix} 9 \\ 8 \\ -2 \end{pmatrix}$$

12. (a) $\|\mathbf{x}-\mathbf{y}\|^2 = (\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y})$

$$\begin{aligned} &= \mathbf{x}^T\mathbf{x} - \mathbf{x}^T\mathbf{y} - \mathbf{y}^T\mathbf{x} + \mathbf{y}^T\mathbf{y} \\ &= 2\mathbf{x}^T\mathbf{x} - 2\mathbf{y}^T\mathbf{x} \\ &= 2(\mathbf{x}-\mathbf{y})^T\mathbf{x} \end{aligned}$$

(b) It follows from part (a) that

$$2\mathbf{u}^T\mathbf{x} = \frac{2}{\|\mathbf{x}-\mathbf{y}\|}(\mathbf{x}-\mathbf{y})^T\mathbf{x} = \|\mathbf{x}-\mathbf{y}\|$$

Thus

$$2\mathbf{u}\mathbf{u}^T\mathbf{x} = (2\mathbf{u}^T\mathbf{x})\mathbf{u} = \mathbf{x}-\mathbf{y}$$

and hence

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

13. (a) $Q\mathbf{u} = (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{u} = \mathbf{u} - 2(\mathbf{u}^T\mathbf{u})\mathbf{u} = -\mathbf{u}$
The eigenvalue is $\lambda = -1$.

(b) $Q\mathbf{z} = (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{z} = \mathbf{z} - 2(\mathbf{u}^T\mathbf{z})\mathbf{u} = \mathbf{z}$
Therefore $\mathbf{z}$ is an eigenvector belonging to the eigenvalue $\lambda = 1$.

(c) The eigenspace corresponding to $\lambda = 1$ is

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

The matrix $\mathbf{u}\mathbf{u}^T$ has rank 1 and hence its nullity must be $n - 1$. Thus the dimension of the eigenspace corresponding to $\lambda = 1$ is $n - 1$. Therefore the multiplicity of the eigenvalue must be at least $n-1$. Since we know that -1 is an eigenvalue, it follows that $\lambda = 1$ must have multiplicity $n - 1$. Since the determinant is equal to the product of the eigenvalues we have

$$\det(Q) = -1 \cdot (1)^n = -1$$

14. If R is a plane rotation then expanding its determine by cofactors we see that

$$\det(R) = \cos^2\theta + \sin^2\theta = 1$$

By Exercise 13(c) an elementary orthogonal matrix has determinant equal to -1, so it follows that a plane rotation cannot be an elementary orthogonal matrix.

15. (a) Let $Q = Q_1^T Q_2 = R_1 R_2^{-1}$. The matrix Q is orthogonal and upper triangular. Since Q is upper triangular, Q^{-1} must also be upper triangular. However

$$Q^{-1} = Q^T = (R_1 R_2^{-1})^T$$

which is lower triangular. Therefore Q must be diagonal.

(b) $R_1 = (Q_1^T Q_2)R_2 = QR_2$. Since

$$|q_{ii}| = \|Q\mathbf{e}_i\| = \|\mathbf{e}_i\| = 1$$

it follows that $q_{ii} = \pm 1$ and hence the ith row of R_1 is ± 1 times the ith row of R_2.

16. Since $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors, there exist Householder matrices H_1 and H_2 such that

$$H_1\mathbf{x} = \|\mathbf{x}\|\mathbf{e}_1^{(m)} \quad \text{and} \quad H_2\mathbf{y} = \|\mathbf{y}\|\mathbf{e}_2^{(n)}$$

where $\mathbf{e}_1^{(m)}$ and $\mathbf{e}_1^{(n)}$ denote the first column vectors of the $m \times m$ and $n \times n$ identity matrices. It follows that

$$\begin{aligned} H_1 A H_2 &= H_1 \mathbf{x}\mathbf{y}^T H_2 \\ &= (H_1\mathbf{x})(H_2\mathbf{y})^T \\ &= \|\mathbf{x}\|\,\|\mathbf{y}\|\mathbf{e}_1^{(m)}(\mathbf{e}_1^{(n)})^T \end{aligned}$$

Set

$$\Sigma = \|\mathbf{x}\|\,\|\mathbf{y}\|\mathbf{e}_1^{(m)}(\mathbf{e}_1^{(n)})^T$$

Σ is an $m \times n$ matrix whose entries are all zero except for the $(1,1)$ entry which equals $\|\mathbf{x}\|\,\|\mathbf{y}\|$. We have then

$$H_1AH_2 = \Sigma$$

Since H_1 and H_2 are both orthogonal and symmetric it follows that A has singular value decomposition $H_1\Sigma H_2$.

17. In constructing the Householder matrix we set

$$\beta = \alpha(\alpha - x_1) \qquad \text{and} \qquad \mathbf{v} = (x_1 - \alpha, x_2, \ldots, x_n)^T$$

In both computations we can avoid loss of significant digits by choosing α to have the opposite sign of x_1.

18.

$$\begin{aligned} ULU &= \begin{pmatrix} 1 & \frac{\cos\theta-1}{\sin\theta} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin\theta & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos\theta-1}{\sin\theta} \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \frac{\cos\theta-1}{\sin\theta} \\ \sin\theta & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos\theta-1}{\sin\theta} \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{aligned}$$

SECTION 6

3. (a) $\mathbf{v}_1 = A\mathbf{u}_0 = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \qquad \mathbf{u}_1 = \frac{1}{3}\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2/3 \end{pmatrix}$

$\mathbf{v}_2 = A\mathbf{u}_1 = \begin{pmatrix} -1/3 \\ -1/3 \end{pmatrix} \qquad \mathbf{u}_2 = -3\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\mathbf{v}_3 = A\mathbf{u}_2 = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \qquad \mathbf{u}_3 = \frac{1}{3}\mathbf{v}_3 = \begin{pmatrix} 1 \\ -2/3 \end{pmatrix}$

$\mathbf{v}_4 = A\mathbf{u}_3 = \begin{pmatrix} -1/3 \\ -1/3 \end{pmatrix} \qquad \mathbf{u}_4 = -3\mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

6. (a and b). Let $\mathbf{x}_j$ be an eigenvector of A belonging to λ_j.

$$B^{-1}\mathbf{x}_j = (A - \lambda I)\mathbf{x}_j = (\lambda_j - \lambda)\mathbf{x}_j = \frac{1}{\mu_j}\mathbf{x}_j$$

Multiplying through by $\mu_j B$ we obtain

$$B\mathbf{x}_j = \mu_j\mathbf{x}_j$$

Thus μ_j is an eigenvalue of B and $\mathbf{x}_j$ is an eigenvector belonging to μ_j.

(c) If λ_k is the eigenvalue of A that is closest to λ, then

$$|\mu_k| = \frac{1}{|\lambda_k - \lambda|} > \frac{1}{|\lambda_j - \lambda|} = |\mu_j|$$

for $j \neq k$. Therefore μ_k is the dominant eigenvalue of B. Thus when the power method is applied to B, it will converge to an eigenvector $\mathbf{x}_k$ of μ_k. By part (b), $\mathbf{x}_k$ will also be an eigenvector belonging to λ_k.

7. (a) Since $A\mathbf{x} = \lambda\mathbf{x}$, the ith coordinate of each side must be equal. Thus

$$\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i$$

(b) It follows from part (a) that

$$(\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j$$

Since $|x_1| = \|\mathbf{x}\|_\infty > 0$ it follows that

$$|\lambda - a_{ii}| = \left| \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{a_{ij}x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \left| \frac{x_j}{x_i} \right| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

8. (a) Let $B = X^{-1}(A+E)X$. Since $X^{-1}AX$ is a diagonal matrix whose diagonal entries are the eigenvalues of A we have

$$b_{ij} = \begin{cases} c_{ij} & \text{if } i \neq j \\ \lambda_i + c_{ii} & \text{if } i = j \end{cases}$$

It follows from Exercise 7 that

$$|\lambda - b_{ii}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |b_{ij}|$$

for some i. Thus

$$|\lambda - \lambda_i - c_{ii}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |c_{ij}|$$

Since

$$|\lambda - \lambda_i| - |c_{ii}| \leq |\lambda - \lambda_i - c_{ii}|$$

it follows that

$$|\lambda - \lambda_i| \leq \sum_{j=1}^{n} |c_{ij}|$$

(b) It follows from part (a) that

$$\begin{aligned}\min_{1\le j\le n}|\lambda-\lambda_j| &\le \max_{1\le i\le n}\left(\sum_{j=1}^{n}|c_{ij}|\right)\\ &= \|C\|_\infty\\ &\le \|X^{-1}\|_\infty\|E\|_\infty\|X\|_\infty\\ &= \mathrm{cond}_\infty(X)\|E\|_\infty\end{aligned}$$

9. The proof is by induction on k. In the case $k=1$

$$AP_1 = (Q_1R_1)Q_1 = Q_1(R_1Q_1) = P_1A_2$$

Assuming $P_mA_{m+1} = AP_m$ we will show that $P_{m+1}A_{m+2} = AP_{m+1}$.

$$\begin{aligned}AP_{m+1} &= AP_mQ_{m+1}\\ &= P_mA_{m+1}Q_{m+1}\\ &= P_mQ_{m+1}R_{m+1}Q_{m+1}\\ &= P_{m+1}A_{m+2}\end{aligned}$$

10. (a) The proof is by induction on k. In the case $k=1$

$$P_2U_2 = Q_1Q_2R_2R_1 = Q_1A_2R_1 = P_1A_2U_1$$

It follows from Exercise 9 that

$$P_1A_2U_1 = AP_1U_1$$

Thus

$$P_2U_2 = P_1A_2U_1 = AP_1U_1$$

If

$$P_{m+1}U_{m+1} = P_mA_{m+1}U_m = AP_mU_m$$

then

$$\begin{aligned}P_{m+2}U_{m+2} &= P_{m+1}Q_{m+2}R_{m+2}U_{m+1}\\ &= P_{m+1}A_{m+2}U_{m+1}\end{aligned}$$

Again by Exercise 9 we have

$$P_{m+1}A_{m+2} = AP_{m+1}$$

Thus

$$P_{m+2}U_{m+2} = P_{m+1}A_{m+2}U_{m+1} = AP_{m+1}U_{m+1}$$

(b) Prove: $P_kU_k = A^k$. The proof is by induction on k. In the case $k=1$

$$P_1U_1 = Q_1R_1 = A = A^1$$

If

$$P_mU_m = A^m$$

then it follows from part (a) that

$$P_{m+1}U_{m+1} = AP_mU_m = AA^m = A^{m+1}$$

11. To determine $\mathbf{x}_k$ and β, compare entries on both sides of the block multiplication for the equation $R_{k+1}U_{k+1} = U_{k+1}D_{k+1}$.

$$\begin{pmatrix} R_k & \mathbf{b}_k \\ \mathbf{0}^T & \beta_k \end{pmatrix} \begin{pmatrix} U_k & \mathbf{x}_k \\ \mathbf{0}^T & 1 \end{pmatrix} = \begin{pmatrix} U_k & \mathbf{x}_k \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} D_k & \mathbf{0} \\ \mathbf{0}^T & \beta \end{pmatrix}$$

$$\begin{pmatrix} R_kU_k & R_k\mathbf{x}_k + b_k \\ \mathbf{0}^T & \beta_k \end{pmatrix} = \begin{pmatrix} U_kD_k & \beta\mathbf{x}_k \\ \mathbf{0}^T & \beta \end{pmatrix}$$

By hypothesis, $R_kU_k = U_kD_k$, so if we set $\beta = \beta_k$, then the diagonal blocks of both sides will match up. Equating the $(1, 2)$ blocks of both sides we get

$$R_k\mathbf{x}_k + \mathbf{b}_k = \beta_k\mathbf{x}_k$$

$$(R_k - \beta_kI)\mathbf{x}_k = -\mathbf{b}_k$$

This is a $k\times k$ upper triangular system. The system has a unique solution since β_k is not an eigenvalue of R_k. The solution $\mathbf{x}_k$ can be determined by back substitution.

12. (a) Algorithm for computing eigenvectors of an $n \times n$ upper triangular matrix with no multiple eigenvalues.

Set $U_1 = (1)$
For $k = 1, \ldots, n-1$
Use back substitution to solve

$$(R_k - \beta_kI)\,\mathbf{x}_k = -\mathbf{b}_k$$

where

$$\beta_k = r_{k+1,k+1} \quad \text{and} \quad \mathbf{b}_k = (r_{1,k+1}, r_{2,k+1}, \ldots, r_{k,k+1})^T$$

Set

$$U_{k+1} = \begin{pmatrix} U_k & \mathbf{x}_k \\ \mathbf{0}^T & 1 \end{pmatrix}$$

End (For Loop)

The matrix U_n is upper triangular with 1's on the diagonal. Its column vectors are the eigenvectors of R.

(b) All of the arithmetic is done in solving the $n - 1$ systems

$$(R_k - \beta_kI)\mathbf{x}_k = -\mathbf{b}_k \qquad k = 1, \ldots, n-1$$

by back substitution. Solving the kth system requires

$$1 + 2 + \cdots + k = \frac{k(k+1)}{2} \quad \text{multiplications}$$

and k divisions. Thus the kth step of the loop requires $\frac{1}{2}k^2 + \frac{3}{2}k$ multiplications/divisions. The total algorithm requires

$$\frac{1}{2}\sum_{k=1}^{n-1}(k^2 + 3k) \;=\; \frac{1}{2}\left(\frac{n(2n-1)(n-1)}{6} + \frac{3n(n-1)}{2}\right)$$

$$= \frac{n^3}{6} + \frac{4n^2 - n - 4}{6} \text{ multiplications/divisions}$$

The dominant term is $n^3/6$.

SECTION 7

3. (a) $\alpha_1 = \|\mathbf{a}_1\| = 2,\ \beta_1 = \alpha_1(\alpha_1 - \alpha_{11}) = 2,\quad \mathbf{v}_1 = (-1,\ 1,\ 1,\ 1)^T$

$$H_1 = I - \frac{1}{\beta_1}\mathbf{v}_1\mathbf{v}_1^T$$

$$H_1A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \\ 0 & 1 \\ 0 & -2 \end{pmatrix} \qquad H_1\mathbf{b} = \begin{pmatrix} 8 \\ -1 \\ -8 \\ -5 \end{pmatrix}$$

$\alpha_2 = \|(2,\ 1,\ -2)^T\| = 3 \qquad \beta_2 = 3(3-2) = 3 \qquad \mathbf{v}_2 = (-1,\ 1,\ -2)^T$

$$H_2 = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & H_{22} \end{pmatrix} \text{ where } H_{22} = I - \frac{1}{\beta_2}\mathbf{v}_2\mathbf{v}_2^T$$

$$H_2H_1A = \begin{pmatrix} 2 & 3 \\ 0 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad H_2H_1\mathbf{b} = \begin{pmatrix} 8 \\ 0 \\ -9 \\ -3 \end{pmatrix}$$

5. Let A be an $m \times n$ matrix with nonzero singular values $\sigma_1, \ldots, \sigma_r$ and singular value decomposition $U\Sigma V^T$. We will show first that Σ^+ satisfies the four Penrose conditions. Note that the matrix $\Sigma\Sigma^+$ is an $m \times m$ diagonal matrix whose first r diagonal entries are all 1 and whose remaining diagonal entries are all 0. Since the only nonzero entries in the matrices Σ and Σ^+ occur in the first r diagonal positions it follows that

$$(\Sigma\Sigma^+)\Sigma = \Sigma \quad \text{and} \quad \Sigma^+(\Sigma\Sigma^+) = \Sigma^+$$

Thus Σ^+ satisfies the first two Penrose conditions. Since both $\Sigma\Sigma^+$ and $\Sigma^+\Sigma$ are square diagonal matrices they must be symmetric

$$\begin{aligned}(\Sigma\Sigma^+)^T &= \Sigma\Sigma^+ \\ (\Sigma^+\Sigma)^T &= \Sigma^+\Sigma\end{aligned}$$

Thus Σ^+ satisfies all four Penrose conditions. Using this result it is easy to show that $A^+ = V\Sigma^+U^T$ satisfies the four Penrose conditions.

(1) $AA^+A = U\Sigma V^TV\Sigma^+U^TU\Sigma V^T = U\Sigma\Sigma^+\Sigma V^T = U\Sigma V^T = A$

(2) $A^+AA^+ = V\Sigma^+U^TU\Sigma V^TV\Sigma^+U^T = V\Sigma^+\Sigma\Sigma^+U^T = V\Sigma^+U^T = A^+$

(3) $$\begin{aligned}(AA^+)^T &= (U\Sigma V^TV\Sigma^+U^T)^T \\ &= (U\Sigma\Sigma^+U^T)^T \\ &= U(\Sigma\Sigma^+)^TU^T\end{aligned}$$

$$
\begin{aligned}
&= U(\Sigma\Sigma^+)U^T \\
&= AA^+
\end{aligned}
$$

$$
\begin{aligned}
(4)\ \ (A^+A)^T &= (V\Sigma^+U^TU\Sigma V^T)^T \\
&= (V\Sigma^+\Sigma V^T)^T \\
&= V(\Sigma^+\Sigma)^T V^T \\
&= V(\Sigma^+\Sigma)V^T \\
&= A^+A
\end{aligned}
$$

6. Let B be a matrix satisfying Penrose condition (1) and (3), that is,

$$ABA = A \qquad \text{and} \qquad (AB)^T = AB$$

If $\mathbf{x} = B\mathbf{b}$, then

$$A^TA\mathbf{x} = A^TAB\mathbf{b} = A^T(AB)^T\mathbf{b} = (ABA)^T\mathbf{b} = A^T\mathbf{b}$$

7. If $X = \dfrac{1}{\|\mathbf{x}\|_2^2}\mathbf{x}^T$, then

$$X\mathbf{x} = \frac{1}{\|\mathbf{x}\|_2^2}\mathbf{x}^T\mathbf{x} = 1$$

Using this it is easy to verify that $\mathbf{x}$ and X satisfy the four Penrose conditions.

(1) $\mathbf{x}X\mathbf{x} = \mathbf{x}1 = \mathbf{x}$

(2) $X\mathbf{x}X = 1X = X$

(3) $(\mathbf{x}X)^T = X^T\mathbf{x} = \dfrac{1}{\|\mathbf{x}\|^2}\mathbf{x}\mathbf{x}^T = \mathbf{x}X$

(4) $(X\mathbf{x})^T = 1^T = 1 = X\mathbf{x}$

8. If A has singular value decomposition $U\Sigma V^T$ then

$$A^TA = V\Sigma^TU^TU\Sigma V^T = V\Sigma^T\Sigma V^T \tag{7}$$

The matrix $\Sigma^T\Sigma$ is an $n \times n$ diagonal matrix with diagonal entries $\sigma_1^2, \ldots, \sigma_n^2$. Since A has rank n its singular values are all nonzero and it follows that $\Sigma^T\Sigma$ is nonsingular. It follows from equation (7) that

$$
\begin{aligned}
(A^TA)^{-1}A^T &= (V(\Sigma^T\Sigma)^{-1}V^T)(V\Sigma^TU^T) \\
&= V(\Sigma^T\Sigma)^{-1}\Sigma^TU^T \\
&= V\Sigma^+U^T \\
&= A^+
\end{aligned}
$$

9. Let

$$\mathbf{b} = AA^+\mathbf{b} = A(A^+\mathbf{b})$$

since

$$R(A) = \{A\mathbf{x} \mid \mathbf{x} \in R^n\}$$

it follows that $\mathbf{b} \in R(A)$.

Conversely if $\mathbf{b} \in R(A)$, then $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x} \in R^n$. It follows that

$$\begin{aligned} A^+\mathbf{b} &= A^+A\mathbf{x} \\ AA^+\mathbf{b} &= AA^+A\mathbf{x} = A\mathbf{x} = \mathbf{b} \end{aligned}$$

10. A vector $\mathbf{x} \in R^n$ minimizes $\|\mathbf{b} - A\mathbf{x}\|_2$ if and only if $\mathbf{x}$ is a solution to the normal equations. It follows from Theorem 7.9.1 that $A^+\mathbf{b}$ is a particular solution. Since $A^+\mathbf{b}$ is a particular solution it follows that a vector $\mathbf{x}$ will be a solution if and only if

$$\mathbf{x} = A^+\mathbf{b} + \mathbf{z}$$

where $\mathbf{z} \in N(A^TA)$. However, $N(A^TA) = N(A)$. Since $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ form a basis for $N(A)$ (see Exercise 7, Section 7), it follows that $\mathbf{x}$ is a solution if and only if

$$\mathbf{x} = A^+\mathbf{b} + c_{r+1}\mathbf{v}_{r+1} + \cdots + c_n\mathbf{v}_n$$

13. (a) $(\Sigma^+)^+$ is an $m \times n$ matrix whose nonzero diagonal entries are the reciprocals of the nonzero diagonal entries of Σ^+. Thus $(\Sigma^+)^+ = \Sigma$. If $A = U\Sigma V^T$, then

$$(A^+)^+ = (V\Sigma^+U^T)^+ = U(\Sigma^+)^+V^T = U\Sigma V^T = A$$

(b) $\Sigma\Sigma^+$ is an $m \times m$ diagonal matrix whose diagonal entries are all 0's and 1's. Thus $(\Sigma\Sigma^+)^2 = \Sigma\Sigma^+$ and it follows that

$$\begin{aligned} (AA^+)^2 &= (U\Sigma V^TV\Sigma^+U^T)^2 = (U\Sigma\Sigma^+U^T)^2 = U(\Sigma\Sigma^+)^2U^T \\ &= U\Sigma\Sigma^+U^T = AA^+ \end{aligned}$$

(c) $\Sigma^+\Sigma$ is an $n \times n$ diagonal matrix whose diagonal entries are all 0's and 1's. Thus $(\Sigma^+\Sigma)^2 = \Sigma^+\Sigma$ and it follows that

$$\begin{aligned} (A^+A)^2 &= (V\Sigma^+U^TU\Sigma V^T)^2 = (V\Sigma^+\Sigma V^T)^2 = V(\Sigma^+\Sigma)^2V^T \\ &= V\Sigma^+\Sigma V^T = A^+A \end{aligned}$$

15. (1)
$$\begin{aligned} ABA &= XY^T[Y(Y^TY)^{-1}(X^TX)^{-1}X^T]XY^T \\ &= X(Y^TY)(Y^TY)^{-1}(X^TX)^{-1}(X^TX)Y^T \\ &= XY^T \\ &= A \end{aligned}$$

(2)
$$\begin{aligned} BAB &= [Y(Y^TY)^{-1}(X^TX)^{-1}X^T](XY^T)[Y(Y^TY)^{-1}(X^TX)^{-1}X^T] \\ &= Y(Y^TY)^{-1}(X^TX)^{-1}(X^TX)(Y^TY)(Y^TY)^{-1}(X^TX)^{-1}X^T \\ &= Y(Y^TY)^{-1}(X^TX)^{-1}X^T \\ &= B \end{aligned}$$

(3)
$$\begin{aligned} (AB)^T &= B^TA^T \\ &= [Y(Y^TY)^{-1}(X^TX)^{-1}X^t]^T(YX^T) \\ &= X(X^TX)^{-1}(Y^TY)^{-1}Y^TYX^T \\ &= X(X^TX)^{-1}X^T \\ &= X(Y^TY)(Y^TY)^{-1}(X^TX)^{-1}X^T \end{aligned}$$

$$
\begin{aligned}
&= (XY^T)[Y(Y^TY)^{-1}(X^TX)^{-1}X^T] \\
&= AB
\end{aligned}
$$

$$
\begin{aligned}
(4)\ \ (BA)^T &= A^TB^T \\
&= (YX^T)[Y(Y^TY)^{-1}(X^TX)^{-2}X^T]^T \\
&= YX^TX(X^TX)^{-1}(Y^TY)^{-1}Y^T \\
&= Y(Y^TY)^{-1}Y^T \\
&= Y(Y^TY)^{-1}(X^TX)^{-1}(X^TX)Y^T \\
&= [Y(Y^TY)^{-1}(X^TX)^{-1}X^T](XY^T) \\
&= BA
\end{aligned}
$$

MATLAB EXERCISES

1. The system is well conditioned since perturbations in the solutions are roughly the same size as the perturbations in A and $\mathbf{b}$.

2. (a) The entries of $\mathbf{b}$ and the entries of $V\mathbf{s}$ should both be equal to the row sums of V.

3. (a) Since L is lower triangular with 1's on the diagonal, it follows that $\det(L) = 1$ and

$$\det(C) = \det(L)\det(L^T) = 1$$

and hence $C^{-1} = \operatorname{adj}(C)$. Since C is an integer matrix its adjoint will also consist entirely of integers.

7. Since A is a magic square, the row sums of $A-tI$ will all be 0. Thus the row vectors of $A-tI$ must be linearly dependent. Therefore $A-tI$ is singular and hence t is an eigenvalue of A. Since the sum of all the eigenvalues is equal to the trace, the other eigenvalues must add up to 0. The condition number of X should be small, which indicates that the eigenvalue problem is well-conditioned.

8. Since A is upper triangular no computations are necessary to determine its eigenvalues. Thus MATLAB will give you the exact eigenvalues of A. However the eigenvalue problem is moderately ill-conditioned and consequently the eigenvalues of A and $A1$ will differ substantially.

9. (b) $\text{Cond}(X)$ should be on the order of 10^8, so the eigenvalue problem should be moderately ill-conditioned.

10. (b) $K\mathbf{e} = -H\mathbf{e}$.

12. (a) The graph has been rotated 45° in the counterclockwise direction.

(c) The graph should be the same as the graph from part (b). Reflecting about a line through the origin at an angle of $\frac{\pi}{8}$ is geometrically the same as reflecting about the x-axis and then rotating 45 degrees. The later pair of operations can be represented by the matrix product

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} c & s \\ s & -c \end{pmatrix}$$

where $c = \cos\frac{\pi}{4}$ and $s = \sin\frac{\pi}{4}$.

13. (b)

$$\mathbf{b}(1,:) = \mathbf{b}(2,:) = \mathbf{b}(3,:) = \mathbf{b}(4,:) = \frac{1}{2}(\mathbf{a}(2,:) + \mathbf{a}(3,:))$$

(c) Both A and B have the same largest singular value $s(1)$. Therefore

$$\|A\|_2 = s(1) = \|B\|_2$$

The matrix B is rank 1. Therefore $s(2) = s(3) = s(4) = 0$ and hence

$$\|B\|_F = \|\mathbf{s}\|_2 = s(1)$$

14. (b)

$$\|A\|_2 = s(1) = \|B\|_2$$

(c) To construct C, set

$$D(4,4) = 0 \qquad \text{and} \qquad C = U * D * V'$$

It follows that

$$\|C\|_2 = s(1) = \|A\|_2$$

and

$$\|C\|_F = \sqrt{s(1)^2 + s(2)^2 + s(3)^2} < \|\mathbf{s}\|_2 = \|A\|_F$$

15. (a) The rank of A should be 4. To determine $V1$ and $V2$ set

$$V1 = V(:, 1:4) \qquad\qquad V2 = V(:, 5:6)$$

P is the projection matrix onto $N(A)$. Therefore $\mathbf{r}$ must be in $N(A)$. Since $\mathbf{w} \in R(A^T) = N(A)^{\perp}$, we have

$$\mathbf{r}^T\mathbf{w} = \mathbf{0}$$

(b) Q is the projection matrix onto $N(A^T)$. Therefore $\mathbf{y}$ must be in $N(A^T)$. Since $\mathbf{z} \in R(A) = N(A^T)^{\perp}$, we have

$$\mathbf{y}^T\mathbf{z} = \mathbf{0}$$

(d) Both AX and $U1(U1)^T$ are projection matrices onto $R(A)$. Since the projection matrix onto a subspace is unique, it follows that

$$AX = U1(U1)^T$$

16. (b) The disk centered at 50 is disjoint from the other two disks, so it contains exactly one eigenvalue. The eigenvalue is real so it must lie in the interval $[46, 54]$. The matrix C is similar to B and hence must have the same eigenvalues. The disks of C centered at 3 and 7 are disjoint from the other disks. Therefore each of the two disks contains an eigenvalue. These eigenvalues are real and consequently must lie in the intervals $[2.7, 3.3]$ and $[6.7, 7.3]$. The matrix C^T has the same eigenvalues as C and B. Using the Gerschgorin disk corresponding to the third row of C^T we see that the dominant eigenvalue must lie in the interval $[49.6, 50.4]$. Thus without computing the eigenvalues of B we are able to obtain nice approximations to their actual locations.

CHAPTER TEST A

1. The statement is false in general. For example, if

$$a = 0.11 \times 10^0, \quad b = 0.32 \times 10^{-2}, \quad c = 0.33 \times 10^{-2}$$

and 2-digit decimal arithmetic is used, then

$$fl(fl(a + b) + c) = a = 0.11 \times 10^0$$

and

$$fl(a + fl(b + c)) = 0.12 \times 10^0$$

2. The statement is false in general. For example, if A and B are both 2×2 matrices and C is a 2×1 matrix, then the computation of $A(BC)$ requires 8 multiplications and 4 additions, while the computation of $(AB)C$ requires 12 multiplications and 6 additions.

3. The statement is false in general. It is possible to have a large relative error if the coefficient matrix is ill-conditioned. For example, the $n \times n$ Hilbert matrix H is defined by

$$h_{ij} = \frac{1}{i + j - 1}$$

For $n = 12$, the matrix H is nonsingular, but it is very ill-conditioned. If you tried to solve a nonhomogeneous linear system with this coefficient matrix you would not get an accurate solution.

4. The statement is true. For a symmetric matrix the eigenvalue problem is well conditioned. (See the remarks following Theorem 7.6.1.) If a stable algorithm is used then the computed eigenvalues should be the exact eigenvalues of a nearby matrix, i.e., a matrix of the form $A + E$ where $\|E\|$ is small. Since the problem is well conditioned the eigenvalues of nearby matrices will be good approximations to the eigenvalues of A.

5. The statement is false in general. If the matrix is nonsymmetric then the eigenvalue problem could be ill-conditioned. If so, then even a stable algorithm will not necessary guarantee accurate eigenvalues. In particular if A has an eigenvalue-eigenvector decomposition XDX^{-1} and X is very ill-conditioned, then the eigenvalue problem will be ill-conditioned and it will not be possible to compute the eigenvalues accurately.

6. The statement is false. If A^{-1} and the LU factorization are both available the it doesn't matter which you use since it takes the same number of arithmetic operations to solve $LU\mathbf{x} = \mathbf{b}$ using forward and back substitution as it does to multiply $A^{-1}\mathbf{b}$.

7. The statement is true. The 1-norm is computed by taking the sum of the absolute values on the entries in each column of A and then taking the maximum of the column sums. The infinity norm is computed by taking the sum of the absolute values on the entries in each row of A and then taking the maximum of the row sums. If A is symmetric then the row sums and column sums will be the same and hence the both norms will be equal.

8. The statement is false in general. For example if

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$$

then $\|A\|_2 = 4$ and $\|A\|_F = 5$.

9. The statement is false in general. If A has rank n, then the least squares problem will have a unique solution. However, if A is ill-conditioned the computed solution may not be a good approximation to the exact solution even though it produces a small residual vector.

10. The statement is false in general. For example, if

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 10^{-8} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

then A and B are close since $\|A-B\|_F = 10^{-8}$. However their pseudoinverses are not close. In fact, $\|A^+ - B^+\|_F = 10^8$

CHAPTER TEST B

1. If $\mathbf{y} = B\mathbf{x}$ then the computation of a single entry of $\mathbf{y}$ requires n multiplications and $n-1$ additions. Since $\mathbf{y}$ has n entries, the computation of the matrix-vector product $B\mathbf{x}$ requires n^2 multiplications and $n(n-1)$ additions. The computation $A(B\mathbf{x}) = A\mathbf{y}$ requires 2 matrix-vector multiplications. So the number of scalar multiplications and scalar additions that are necessary is $2n^2$ and $2n(n-1)$.

On the other hand if $C = AB$ then the computation of the jth column of C requires a matrix-vector multiplication $\mathbf{c}_j = A\mathbf{b}_j$ and hence the computation of C requires n matrix-vector multiplications. Therefore the computation $(AB)\mathbf{x} = C\mathbf{x}$ will require $n+1$ matrix-vector multiplications. The total number of arithmetic operations will be $(n+1)n^2$ scalar multiplications and $(n+1)n(n-1)$ scalar additions.

For $n > 1$ the computation $A(B\mathbf{x})$ is more efficient.

2. **(a)**

$$\left(\begin{array}{ccc|c} 2 & 3 & 6 & 3 \\ 4 & 4 & 8 & 0 \\ 1 & 3 & 4 & 4 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 4 & 4 & 8 & 0 \\ 2 & 3 & 6 & 3 \\ 1 & 3 & 4 & 4 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 4 & 4 & 8 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 2 & 4 \end{array}\right)$$

$$\rightarrow \left(\begin{array}{ccc|c} 4 & 4 & 8 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 1 & 2 & 3 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 4 & 4 & 8 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

The solution $\mathbf{x} = (-3, 1, 1)^T$ is obtained using back substitution.

(b)

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad PA = \begin{pmatrix} 4 & 4 & 8 \\ 1 & 3 & 4 \\ 2 & 3 & 6 \end{pmatrix}$$

and

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 4 & 4 & 8 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

(c) If we set $\mathbf{d} = P\overrightarrow{=}(8, 2, 1)^T$ and solve $L\mathbf{y} = \mathbf{d}$ by forward substitution

$$(L \ \ \mathbf{d}) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ \frac{1}{4} & 1 & 0 & 2 \\ \frac{1}{2} & \frac{1}{2} & 1 & 1 \end{array}\right)$$

then the solution is $\mathbf{y} = (8, 0, -3)^T$. To find the solution to the system $A\mathbf{x} = \mathbf{c}$, we solve $U\mathbf{x} = \mathbf{y}$ using back substitution.

$$(U \ \ \mathbf{y}) = \left(\begin{array}{ccc|c} 4 & 4 & 8 & 8 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & -3 \end{array}\right)$$

The solution is $\mathbf{x} = (5, 3, -3)^T$.

3. If Q is a 4×4 orthogonal matrix then for any nonzero $\mathbf{x}$ in R^4 we have $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ and hence

$$\|Q\|_2 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|Q\mathbf{x}\|}{\|\mathbf{x}\|} = 1$$

To determine the Frobenius norm of Q, note that

$$\|Q\|_F^2 = \|\mathbf{q}_1\|^2 + \|\mathbf{q}_2\|^2 + \|\mathbf{q}_3\|^2 + \|\mathbf{q}_4\|^2 = 4$$

and hence $\|Q\|_F = 2$.

4. (a) $\|H\|_1 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$

$\|H^{-1}\|_1 = \max(516,\ 5700,\ 13620,\ 8820) = 13620$

(b) From part (a) we have $\text{cond}_1(H) = \frac{25}{12} \cdot 13620 = 28375$ and hence

$$\frac{\|\mathbf{x} - \mathbf{x}'\|_1}{\|\mathbf{x}\|_1} \leq \text{cond}_1(H) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|_1} = 28375 \cdot \frac{0.36 \times 10^{-11}}{50} = 2.043 \times 10^{-9}$$

5. The relative error in the solution is bounded by

$$\text{cond}_\infty(A) \frac{\|\mathbf{r}\|_\infty}{\|\mathbf{b}\|_\infty} \approx 10^7 \epsilon$$

so it is possible that one could lose as many as 7 digits of accuracy.

6. **(a)** $\alpha = 3$, $\beta = 3(3-1) = 6$, $\mathbf{v} = (-2, 2, -2)^T$

$$H = I - \frac{1}{\beta}\mathbf{v}\mathbf{v}^T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

(b)

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

7. If A has QR-factorization $A = QR$ and $B = RQ$ then

$$Q^T AQ = Q^T QRQ = RQ = B$$

The matrices A and B are similar and consequently must have the same eigenvalues. Furthermore, if λ is an eigenvalue of B and $\mathbf{x}$ is an eigenvector belonging to λ then

$$Q^T AQ\mathbf{x} = B\mathbf{x} = \lambda\mathbf{x}$$

and hence

$$AQ\mathbf{x} = \lambda Q\mathbf{x}$$

So $Q\mathbf{x}$ is an eigenvector of A belonging to λ.

8. The estimate you get will depend upon your choice of a starting vector. If we start with $\mathbf{u}_0 = \mathbf{x}_0 = \mathbf{e}_1$, then

$$\begin{aligned}
&\mathbf{v}_1 = A\mathbf{e}_1 = \mathbf{a}_1, && \mathbf{u}_1 = \tfrac{1}{4}\mathbf{v}_1 = (0.25, 1)^T \\
&\mathbf{v}_2 = A\mathbf{u}_1 = (2.25, 4)^T, && \mathbf{u}_2 = \tfrac{1}{4}\mathbf{v}_2 = (0.5625, 1)^T \\
&\mathbf{v}_3 = A\mathbf{u}_2 = (2.5625, 5.25)^T, && \mathbf{u}_3 = \tfrac{1}{5.25}\mathbf{v}_3 = (0.548810, 1)^T \\
&\mathbf{v}_4 = A\mathbf{u}_3 = (2.48810, 4.95238)^T, && \mathbf{u}_4 = (0.502404, 1)^T \\
&\mathbf{v}_5 = A\mathbf{u}_4 = (2.50240, 5.00962)^T, && \mathbf{u}_5 == (0.499520, 1)^T \\
&\mathbf{v}_6 = A\mathbf{u}_5 = (2.49952, 4.99808)^T
\end{aligned}$$

Our computed eigenvalue is the second coordinate of $\mathbf{v}_6$, 4.99808 (rounded to 6 digits) and the computed eigenvector is $\mathbf{u}_5$. The actual dominant eigenvalue of A is $\lambda = 5$ and $\mathbf{x} = (0.5, 1)^T$ is an eigenvector belonging to λ.

9. The least squares solution with the smallest 2-norm is

$$\mathbf{x} = A^+\mathbf{b} = V\Sigma^+U^T\mathbf{b} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{2} \\ \frac{1}{12} \end{pmatrix}$$

10. If we set $\alpha_1 = \|\mathbf{a}_1\| = 2$, $\beta_1 = 2$, $\mathbf{v}_1 = (-1, 1, 1, 1)^T$ and $H_1 = I - \frac{1}{\beta_1}\mathbf{v}_1\mathbf{v}_1^T$ then $H_1\mathbf{a}_1 = 2\mathbf{e}_1$. If we multiply the augmented matrix $(A \;\; \mathbf{b})$ by H_1 we get

$$H_1(A \;|\, \mathbf{b}) = \left(\begin{array}{rr|r} 2 & 9 & 7 \\ 0 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & -2 & -2 \end{array}\right)$$

Next we construct a 3×3 Householder matrix H_2 to zero out the last 2 entries of the vector $(1, 2, -2)^T$. If we set $\alpha_2 = 3$, $\beta_2 = 6$ and $\mathbf{v}_2 = (-2, 2, -2)^T$, then $H_2 = I - \frac{1}{\beta}\mathbf{v}_2\mathbf{v}_2^T$. If we apply H_2 to the last 3 rows of $H_1(A \;|\, \mathbf{b})$ we end up with the matrix

$$\left(\begin{array}{rr|r} 2 & 9 & 7 \\ 0 & 3 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{array}\right)$$

The first two rows of this matrix form a triangular system. The solution $\mathbf{x} = (2, \frac{1}{3})^T$ to the triangular system is the solution to the least squares problem.