PREFACE

This solutions manual is designed to accompany the eighth edition of *Linear Algebra with Applications* by Steven J. Leon. The answers in this manual supplement those given in the answer key of the textbook. In addition this manual contains the complete solutions to all of the nonroutine exercises in the book.

At the end of each chapter of the textbook there are two chapter tests (A and B) and a section of computer exercises to be solved using MATLAB. The questions in each Chapter Test A are to be answered as either *true* or *false*. Although the true-false answers are given in the Answer Section of the textbook, students are required to explain or prove their answers. This manual includes explanations, proofs, and counterexamples for all Chapter Test A questions. The chapter tests labeled B contain problems similar to the exercises in the chapter. The answers to these problems are not given in the Answers to Selected Exercises Section of the textbook, however, they are provided in this manual. Complete solutions are given for all of the nonroutine Chapter Test B exercises.

In the MATLAB exercises most of the computations are straightforward. Consequently they have not been included in this solutions manual. On the other hand, the text also includes questions related to the computations. The purpose of the questions is to emphasize the significance of the computations. The solutions manual does provide the answers to most of these questions. There are some questions for which it is not possible to provide a single answer. For example, some exercises involve randomly generated matrices. In these cases the answers may depend on the particular random matrices that were generated.

Steven J. Leon
sleon@umassd.edu

Contents

Chapter 1

Matrices and Systems of Equations

1 SYSTEMS OF LINEAR EQUATIONS

2. (d) $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & -2 & 1 \\ 0 & 0 & 4 & 1 & -2 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

5. (a) $3x_1 + 2x_2 = 8$
$x_1 + 5x_2 = 7$

Therefore

$$(\alpha + \beta)f = \alpha f + \beta f$$

A7. For each x in $[a, b]$,

$$[(\alpha\beta)f](x) = \alpha\beta f(x) = \alpha[\beta f(x)] = [\alpha(\beta f)](x)$$

Therefore

$$(\alpha\beta)f = \alpha(\beta f)$$

A8. For each x in $[a, b]$

$$1f(x) = f(x)$$

Therefore

$$1f = f$$

6. The proof is exactly the same as in Exercise 5.

9. (a) If $\mathbf{y} = \beta\mathbf{0}$ then

$$\mathbf{y} + \mathbf{y} = \beta\mathbf{0} + \beta\mathbf{0} = \beta(\mathbf{0} + \mathbf{0}) = \beta\mathbf{0} = \mathbf{y}$$

and it follows that

$$\begin{aligned} &(\mathbf{y} + \mathbf{y}) + (-\mathbf{y}) = \mathbf{y} + (-\mathbf{y}) \\ &\mathbf{y} + [\mathbf{y} + (-\mathbf{y})] = \mathbf{0} \\ &\mathbf{y} + \mathbf{0} = \mathbf{0} \\ &\mathbf{y} = \mathbf{0} \end{aligned}$$

(b) If $\alpha\mathbf{x} = \mathbf{0}$ and $\alpha \neq 0$ then it follows from part (a), A7 and A8 that

$$\mathbf{0} = \frac{1}{\alpha}\mathbf{0} = \frac{1}{\alpha}(\alpha\mathbf{x}) = \left(\frac{1}{\alpha}\alpha\right)\mathbf{x} = 1\mathbf{x} = \mathbf{x}$$

10. Axiom 6 fails to hold.

$$\begin{aligned} (\alpha + \beta)\mathbf{x} &= ((\alpha + \beta)x_1,\ (\alpha + \beta)x_2) \\ \alpha\mathbf{x} + \beta\mathbf{x} &= ((\alpha + \beta)x_1, 0) \end{aligned}$$

12. A1. $x \oplus y = x \cdot y = y \cdot x = y \oplus x$

A2. $(x \oplus y) \oplus z = x \cdot y \cdot z = x \oplus (y \oplus z)$

A3. Since $x \oplus 1 = x \cdot 1 = x$ for all x, it follows that 1 is the zero vector.

A4. Let

$$-x = -1 \circ x = x^{-1} = \frac{1}{x}$$

It follows that

$$x \oplus (-x) = x \cdot \frac{1}{x} = 1 \quad \text{(the zero vector)}.$$

Therefore $\frac{1}{x}$ is the additive inverse of x for the operation $\oplus$.

A5. $\alpha \circ (x \oplus y) = (x \oplus y)^\alpha = (x \cdot y)^\alpha = x^\alpha \cdot y^\alpha$

$\alpha \circ x \oplus \alpha \circ y = x^\alpha \oplus y^\alpha = x^\alpha \cdot y^\alpha$

A6. $(\alpha + \beta) \circ x = x^{(\alpha+\beta)} = x^\alpha \cdot x^\beta$

$\alpha \circ x \oplus \beta \circ x = x^\alpha \oplus x^\beta = x^\alpha \cdot x^\beta$

10. (a) By the Consistency Theorem $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in $R(A)$. We are given that $\mathbf{b}$ is in $N(A^T)$. So if the system is consistent then $\mathbf{b}$ would be in $R(A) \cap N(A^T) = \{\mathbf{0}\}$. Since $\mathbf{b} \neq \mathbf{0}$, the system must be inconsistent.

(b) If A has rank 3 then A^TA also has rank 3 (see Exercise 13 in Section 2). The normal equations are always consistent and in this case there will be 2 free variables. So the least squares problem will have infinitely many solutions.

11. (a) $P^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = P$

(b) Prove: $P^k = P$ for $k = 1, 2, \ldots$.
Proof: The proof is by mathematical induction. In the case $k = 1$ we have $P^1 = P$. If $P^m = P$ for some m then

$$P^{m+1} = PP^m = PP = P^2 = P$$

(c)
$$\begin{aligned} P^T &= [A(A^TA)^{-1}A^T]^T \\ &= (A^T)^T[(A^TA)^{-1}]^TA^T \\ &= A[(A^TA)^T]^{-1}A^T \\ &= A(A^TA)^{-1}A^T \\ &= P \end{aligned}$$

12. If

$$\begin{pmatrix} A & I \\ O & A^T \end{pmatrix} \begin{pmatrix} \hat{\mathbf{x}} \\ \mathbf{r} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}$$

then

$$\begin{aligned} A\hat{\mathbf{x}} + \mathbf{r} &= \mathbf{b} \\ A^T\mathbf{r} &= \mathbf{0} \end{aligned}$$

We have then that

$$\begin{aligned} \mathbf{r} &= \mathbf{b} - A\hat{\mathbf{x}} \\ A^T\mathbf{r} &= A^T\mathbf{b} - A^TA\hat{\mathbf{x}} = \mathbf{0} \end{aligned}$$

Therefore

$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$$

So $\hat{\mathbf{x}}$ is a solution to the normal equations and hence is a least squares solution to $A\mathbf{x} = \mathbf{b}$.

13. If $\hat{\mathbf{x}}$ is a solution to the least squares problem, then $\hat{\mathbf{x}}$ is a solution to the normal equations

$$A^TA\mathbf{x} = A^T\mathbf{b}$$

It follows that a vector $\mathbf{y} \in \mathbb{R}^n$ will be a solution if and only if

$$\mathbf{y} = \hat{\mathbf{x}} + \mathbf{z}$$

for some $\mathbf{z} \in N(A^TA)$. (See Exercise 26, Chapter 3, Section 2). Since

$$N(A^TA) = N(A)$$

18. (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 of A belonging to t_{11}.

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

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

Since $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent, $t_{12}\mathbf{u}_1 + t_{22}\mathbf{u}_2$ cannot not be equal to a scalar times $\mathbf{u}_2$. So $\mathbf{u}_2$ is not an eigenvector of A.

19. (a) If the eigenvalues are all real, then there will be five 1×1 blocks. The blocks can occur in any order depending on how the eigenvalues are ordered.

(b) If A has three real eigenvalues and one pair of complex conjugate eigenvalues, then there will be three 1×1 blocks corresponding to the real eigenvalues and one 2×2 block corresponding to the pair of complex conjugate eigenvalues. The blocks may appear in any order on the diagonal of the Schur form matrix T.

(c) If A has one real eigenvalue and two pairs of complex eigenvalues then there will be a single 1×1 block and two pairs of 2×2 blocks. The three blocks may appear in any order along the diagonal of the Schur form matrix T.

20. If A has Schur decomposition UTU^H and the diagonal entries of T are all distinct then by Exercise 20 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.

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

22. 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.

(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.