Appendix D

Partial Solutions and Answers to Selected Problems

Chapter 2

2.4 (iii) Let $A = (a_1, \ldots, a_n)$ and $x = (x_1, \ldots, x_n)^T$, $\text{null}(A) = 0$ if and only if $Ax = 0$ implies $x = 0$. But $Ax = a_1x_1 + \cdots + a_nx_n$. Thus $\text{null}(A) = 0$ if and only if $a_1x_1 + \cdots + a_nx_n = 0$ implies $x = 0$, i.e., the columns of A are linearly independent.

2.22 (a) Let $a = \frac{x}{\|x\|_2}$, $b = \frac{y}{\|y\|_2}$. Then $\|a\|_2 = 1,\ \|b\|_2 = 1$. Now,

$$|a^T b| \le \frac{1}{2}(a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2)$$
$$= \frac{1}{2}(\|a\|^2 + \|b\|^2) = 1.$$

Thus $\frac{|x^T y|}{\|x\|_2\|y\|_2} \le 1$, i.e., $|x^T y| \le \|x\|_2\|y\|_2$.

(b) The inequality holds for either $x = 0$ or $y = 0$. If $x \neq 0,\ y \neq 0$, then $x^T x,\ y^T y \neq 0$. Thus

$$\|xy^T\|_2 = \sqrt{\text{maximum eigenvalue of } (xy^T)^T(xy^T)}$$
$$= \sqrt{x^T x}\sqrt{\text{maximum eigenvalue of } (yy^T)} = \|x\|_2\|y\|_2.$$

(Note that $x^T x$ is a scalar.)

2.23

$$\|x + y\|_2^2 = (x + y)^T(x + y) = x^T x + 2x^T y + y^T y$$
$$= x^T x + y^T y \quad (\text{since } x^T y = 0)$$
$$= \|x\|_2^2 + \|y\|_2^2.$$

2.27 (a) See solution of 2.29 (b). Solution of 2.27 (a) follows as a special case of this solution.

2.29 (a)

$$\|QAP\|_F^2 = \text{trace}(P^T A^T Q^T QAP)$$
$$= \text{trace}(P^T A^T AP) = \text{trace}(A^T A) = \|A\|_F^2.$$

(b) $\|QAP\|_2 = \sqrt{\text{maximum eigenvalue of } (P^T A^T Q^T QAP)}$

$$= \sqrt{\text{maximum eigenvalue of } (A^T A)} = \|A\|_2.$$

(Note that $P^T = P^{-1}$, and the eigenvalues remain invariant under similarity transformations.)

2.35 (i) Let $B = (b_1, \ldots, b_r)$, where $b_1, \ldots, b_r$ are columns of B. Then

$$\begin{aligned}
\|AB\|_F^2 &= \|A(b_1, \ldots, b_r)\|_F^2 \\
&= b_1^T A^T A b_1 + \cdots + b_r^T A^T A b_r \quad \text{(using result of Exercise 2.34)} \\
&= \|Ab_1\|_2^2 + \cdots + \|Ab_r\|_2^2 \\
&\le \|A\|_F^2 (\|b_1\|_2^2 + \cdots + \|b_r\|_2^2) \quad \text{(because of the property } \|Ax\|_2 \le \|A\|_F \|x\|_2) \\
&= \|A\|_F^2 \|B\|_F^2.
\end{aligned}$$

(ii) Similarly, we can prove $\|AB\|_F \le \|A\|_2 \|B\|_F$.

2.39 (c) Take

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

and $\|\cdot\| = \|\cdot\|_2$, which is subordinate. Then $A^k = 0$ for $k \ge 2$. But $\|A\|_2 = 1$, and thus the norm test fails.

Chapter 3

3.3 $x = 0.9999 \times 10^{14}$, $y = 0.9815 \times 10^{11}$, $z = 0.9825 \times 10^{11}$, $t = 4$. $x(y+z) = 0.1899 \times 10^{26}$, $xy + yz = 0.1799 \times 10^{26}$.

3.6 (a) $S_{1000} = \sum_{k=1}^{1000} \frac{1}{k}$, where each term $\frac{1}{k}$ is rounded to four significant digits. When added in ascending order the sum is 7.486, the same as the answer in four-digit arithmetic. When done in descending order it is 7.449. Verify this yourself.

3.12 For $t = 4$,

$$\text{fl}(A^T A) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix};$$

for $t = 9$,

$$\text{fl}(A^T A) = \begin{pmatrix} 1 + 10^{-8} & 1 \\ 1 & 1 + 10^{-8} \end{pmatrix}.$$

3.13 (a) $e^x - x - 1 = \dfrac{1}{1 + y + \frac{y^2}{2!} + \cdots} + y - 1$, where $x = -y$, $y > 0$.

(b) $\sqrt{x^4+1}-x^2 = \dfrac{1}{\sqrt{x^4+1}+x^2}$.

(c) $\dfrac{1}{x} - \dfrac{1}{1+x} = \dfrac{1}{x(x+1)}$.

(d) and (e) Use series formulas for $\sin x$ and $\cos x$, respectively.

3.16

$$m = max \begin{pmatrix} |a| \\ |b| \end{pmatrix},$$

$$y_1 = \frac{a}{m}, \; y_2 = \frac{b}{m},$$

$$e = m\sqrt{y_1^2 + y_2^2}.$$

3.18 (a) $-b$ is about the same size as $\sqrt{b^2-4ac}$. A catastrophic cancellation will occur. Take

$$x_1 = \frac{-b - \operatorname{sign}(b)\sqrt{b^2-4ac}}{2a} = 10^6,$$

$$x_2 = \frac{c}{ax_1} = 10^{-6}.$$

Chapter 4

4.1 (a) $\mathrm{fl}(xy) = xy(1+\delta) = xy'$. Take $x' = x,\; y' = y(1+\delta)$. Since $|\delta| \le \mu$, and y' is close to y, the computation is backward stable.

(b) The backward stability of the inner product follows from Theorems 3.16 and 3.17. As of the outer product, note that the matrix of the computed outer product will very likely not have rank 1. Thus it cannot in general be expressed as the outer product of two nearby vectors: $(x+\delta x)(y+\delta y)^T$.

4.3 (a) Change the coefficient 3 to 2.99, then the computed of the perturbed polynomial roots are $1.6327 \pm 1.0493i$ and -0.2655.

4.5 (a) Multiplication of two lower triangular matrices:

Input: Two lower triangular matrices A and B.
Output: Matrix C, $C = A * B$.

For $i = 1, \ldots, n$ do
  For $j = 1, \ldots, i$ do
    For $k = j, \ldots, i$ do
      $c_{ij} = c_{ij} + a_{ik} * b_{kj}$

(c) Multiplication of two tridiagonal matrices:

Input: Two tridiagonal matrices A and B.
Output: Matrix C, $C = A * B$.

For $i = 1, \dots, n$ do
  For $j = \max(1, i-2), \dots, \min(i+2, n)$ do
    For $k = \max(i-1, 1), \dots, \min(i+1, n)$ do
      $c_{ij} = c_{ij} + a_{ik} * b_{kj}$

(d) Multiplication of an arbitrary matrix by an upper Hessenberg matrix:

Input: Matrices A and B.
Output: $C = A * B$.

For $i = 1, \dots, n$ do
  For $j = 1, \dots, n$ do
    For $k = 1, \dots, \min(j+1, n)$ do
      $c_{ij} = c_{ij} + a_{ik} * b_{kj}$

(e) Computation of $[I + xy^T] * A$:

Input: $n \times n$ matrix A and vectors x and y.

For $j = 1, \dots, n$ do
  $\alpha = 0$
    For $i = 1, \dots, n$ do
      $\alpha = \alpha + y_i * a_{ij}$
        For $i = 1, \dots, n$ do
          $a_{ij} = a_{ij} + \alpha * x_i$

4.7 Addition of two $n \times n$ symmetric matrices A and B:

Input: Symmetric matrices A and B.
Output: $C = A + B$. The algorithm overwrites B with C.

For $i = 1, \dots, n$ do
  For $j = i, \dots, n$ do
    $b_{ij} = a_{ij} + b_{ij}$
      If $i \neq j$
        $b(j, i) = b(i, j)$

4.16 (b) (i) $\text{Cond}(A) = \|A\| \|A^{-1}\| \geq \|AA^{-1}\| = \|I\| \geq 1$.

(ii) $\text{Cond}_2(A^T A) = \|A^T A\|_2 \|(A^T A)^{-1}\|_2$. Now $\|A^T A\|_2 = \|A\|_2^2$, $\|(A^T A)^{-1}\|_2 = \|A^{-1}\|_2^2$. So, $\text{Cond}_2(A^T A) = \text{Cond}_2^2(A)$.

(iii) $\text{Cond}(cA) = \|cA\| \|(cA)^{-1}\| = c\|A\| \dfrac{\|(A)^{-1}\|}{c} = \|A\| \|(A)^{-1}\| = \text{Cond}(A)$.

4.17 (a) $\text{Cond}_2(A) = \|A\|_2 \|(A)^{-1}\|_2 = \|A\|_2 \|A^T\|_2 = 1 \times 1 = 1$ (since A is orthogonal).

(b) Let $O = \alpha A$; $\text{Cond}_2(O) = \text{Cond}_2(A) = 1$. If $\text{Cond}_2(O)$ is 1, then $\|(O)^{-1}\|_2$ has to be reciprocal of $\|O\|_2$.

4.18 Since U is upper triangular, $\|U\|_\infty \geq \max |u_{ii}|$ and $\|U^{-1}\|_\infty \geq \dfrac{1}{\min |u_{ii}|}$. Take

$$U = \begin{pmatrix} 0.0001 & 2 & 3 \\ 0 & 1 & 12 \\ 0 & 0 & 12 \end{pmatrix}, \quad \text{Cond}_\infty(U) = 617,500.$$

4.21 (a) When $a \to 1$.

(b) $A^{-1} = \dfrac{1}{1-a^2} \begin{pmatrix} 1 & -a \\ -a & 1 \end{pmatrix}$.

Chapter 5

5.1 (b) $(I + me_k^T)(I - me_k^T) = I - me_k^T + me_k^T - me_k^T me_k^T = I$ (since the last term is the zero matrix).

5.2 (a) $M = 10^5 \begin{pmatrix} 10^{-5} & 0 \\ -1 & 10^{-5} \end{pmatrix}$.

(b) $U = \begin{pmatrix} 10^{-5} & 1 \\ 0 & -0.999 \times 10^5 \end{pmatrix}$,

$L = \begin{pmatrix} 1 & 0 \\ 10^5 & 1 \end{pmatrix}$.

5.15 (a) *Hint:* Prove by induction that $|u_{ij}| \leq 2^{i-1} \max_{k \leq i} |a_{kj}|$.

(b) $A = \begin{pmatrix} \epsilon & 1 \\ 1 & 0 \end{pmatrix}$. The ratio is of order $\dfrac{1}{\epsilon}$.

5.16 (a) *(Without Pivoting.)* Let $r = \min\{m-1,\ n\}$ and U overwrites A.
For $k = 1, 2, \ldots, r$ do

$$a_{ik} \equiv m_{ik} = -\frac{a_{ik}}{a_{kk}} \quad (i = k+1, \ldots, m),$$

$$a_{ij} \equiv a_{ij} + m_{ik}a_{kj}$$
$$= a_{ij} - a_{ik}a_{kj} \quad (i = k+1, \ldots, m,\ j = k+1, \ldots, m).$$

The matrix L can be formed out of the multipliers m_{ik}, as shown in the book.

Chapter 6

6.1 Let x, y, z, be the required cubic yards. Then $0.6x + 0.4y + 0.2z = 5000$; $0.2x + 0.4y + 0.3z = 5500$; $0.2x + 0.2y + 0.5y = 6000$.

6.3
$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 10 & -10 & 0 & -15 & -5 \\ 5 & -10 & 0 & -20 & 0 & 0 \end{pmatrix} \begin{pmatrix} i_{12} \\ i_{52} \\ i_{32} \\ i_{65} \\ i_{54} \\ i_{43} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 150 \end{pmatrix}.$$

6.4 node(1, 1) : $-4T_{11} + T_{21} + T_{12} = -150$; node(2, 1) : $T_{11} - 4T_{21} + T_{31} + T_{22} = -50$; node(3, 1) : $T_{21} - 4T_{31} + T_{32} = -125$; node(1, 2) : $T_{11} - 4T_{12} + T_{22} + T_{13} = -100$; node(2, 2) : $-4T_{22} + T_{32} + T_{23} + T_{12} + T_{21} = 0$; node(3, 2) : $-4T_{32} + T_{33} + T_{22} + T_{31} = -75$; node(1, 3) : $-4T_{13} + T_{23} + T_{12} = -100$; node(2, 3) : $-4T_{23} + T_{33} + T_{13} + T_{22} = 0$; node(3, 3) : $-4T_{33} + T_{23} + T_{32} = -75$.

6.19 (b) Proof follows from Theorem 4.19 by taking $\delta b = -r$.

6.27 (a) (i) The Cholesky factor

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0.0316 & 0 \\ 1 & 0.0316 & 0.9995 \end{pmatrix}.$$

(b) The growth factor $\rho = \dfrac{\max(2, 1, 1)}{2} = 1$.

6.28 (a) The leading principal minors obtained from the U matrix of the LU factorization of A using Gaussian elimination without pivoting are 4, 15, 56, 192. A is, therefore, positive definite. To see it from the Cholesky factorization of A, compute Cholesky factor H by using the MATLAB command **chol**(A) and see that all the diagonal entries of the Cholesky factor are positive.

(b) Using the computed Cholesky factor H from 6.27(a), we have $\text{Cond}_2(A) = \text{Cond}_2(HH^T) = \text{Cond}_2(H).\ \text{Cond}_2(H^T) = 3$.

6.35 (a) For $k = 1$,

$$\sum_{\substack{i=2 \\ i \neq j}}^{n} |a_{ij}^{(1)}| = \sum \left| a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}} \right| \leq \sum |a_{ij}| + \frac{|a_{1j}|}{|a_{11}|} \sum |a_{i1}|$$

$$\leq \sum |a_{ij}| + \frac{|a_{1j}|}{|a_{11}|}(|a_{11}| - |a_{j1}|) \text{ (since } A \text{ is column-diagonally dominant)}$$

$$= \sum |a_{ij}| + |a_{1j}| - \frac{|a_{1j}|}{|a_{11}|}|a_{j1}| \text{ (because of column dominance of } A\text{)}$$

$$\leq \left| a_{jj} - \frac{a_{j1}a_{1j}}{a_{11}} \right| = |a_{jj}^{(1)}|.$$

6.38 (a) (i) Let $H = (h_{ij})$ be $n \times n$ upper Hessenberg. For each value of k from 1 to $n - 1$ perform the following steps:

1. Interchange h_{kj} and $h_{k+1,j}$ if $|h_{kk}| < |h_{k+1,k}|$, $j = k, \ldots, n$.

2. Compute $m = \dfrac{-h_{k+1,k}}{h_{k,k}}$.
3. Update $h_{k+1,j} \equiv h_{k+1,j} + m h_{kj}$, $j = k+1, \ldots, n$.

6.39 Let ΔL and ΔU be the error matrices associated with the matrices L and U, respectively. Denote by E' the error matrix associated with the LU factorization. Then

$$\begin{aligned}|E| &\le |E' + L\Delta U + ALU + \Delta L AU|\\ &\le |E'| + |L||AU| + |\Delta L||U| + |\Delta L||AU|\\ &\le n\mu|L||U| + n\mu|L||U| + n\mu|L||U| + n^2\mu^2|L||U|\\ &\approx 3n\mu|L||U| \text{ (ignoring the last term, which is small).}\end{aligned}$$

To prove the bound for $\|E\|_\infty$, use $\|L\|_\infty \le n$ and $\|U\|_\infty \le n\rho\|A\|_\infty$.

Chapter 7

7.3 (c) The following algorithm computes AJ, given i and k. AJ is stored over A:

For $j = 1, \ldots, n$ do
  $\alpha = a_{ji}$,
  $\beta = a_{jk}$,
  $a_{ji} = c\alpha - s\beta$,
  $a_{jk} = s\alpha + c\beta$.

Flops: 4 flops for each j.
Total flops $= 4n$.

7.9 For $k = 1, 2, \ldots, n$ do
  For $i = 1, 2, \ldots, k-1$ do
    $r_{ik} = q_i^T a_k,\ a_k \equiv a_k - r_{ik} q_i$
$\hat{q}_k = a_k,\ r_{kk} = \|\hat{q}_k\|_2$
$q_k = \dfrac{\hat{q}_k}{r_{kk}}$

7.16 (c) Since A is symmetric, $A^T A = A^2$. Thus, the eigenvalues of $A^T A$ are $\lambda_1^2, \ldots, \lambda_n^2$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Because the singular values of A are the nonnegative square roots of the eigenvalues of $A^T A$, the singular values of A are then $|\lambda_1|, \ldots, |\lambda_n|$.

7.17 Take $\tilde{U} = UP$ and $\tilde{V} = VP$, where P permutes row i and row $i+1$ of Σ.

7.19 (i) Let $A = U\Sigma V^T$. Then $A^T A = V\Sigma^T\Sigma V^T$, and $AA^T = U\Sigma\Sigma^T U^T$. Because $\Sigma^T\Sigma$ and $\Sigma\Sigma^T$ have the same number of nonzero singular values, rank $(\Sigma^T\Sigma) =$ rank $(\Sigma\Sigma^T)$, and hence $A^T A$ and AA^T have the same rank.

7.21 (i) $\|AU\|_2 = \|P\Sigma QU\|_2 = \|P\Sigma Q'\|_2 = \|\Sigma\|_2 = \|A\|_2$.
(ii) Similar to 7.21 (i).

(iii) *Hint:*

$$\max_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} \leq \sigma_1. \qquad \text{Again, } \frac{\|Av_1\|_2}{\|v_1\|_2} = \sigma_1.$$

7.24 Let $E = B - A$. Let $\sigma_i,\ i = 1, \dots, n$, and $\tilde{\sigma}_i,\ i = 1, \dots, n$, be, respectively, the singular values of A and B. By Theorem 10.10 (see Chapter 10), $|\tilde{\sigma}_i - \sigma_i| \leq \|E\|_2 < \sigma_r$ for each i. Thus $\tilde{\sigma}_i$ cannot be zero for $i = 1, \dots, r$.

7.35 First we show that $R(I - P_S) = S^{\perp}$. Let $y = x - P_S x \in R(I - P_S)$ for some $x \in R^n$. Then for all $z \in S$, $z^T y = z^T x - z^T P_S x = z^T x - z^T x = 0$. Thus $R(I - P_S) \subseteq S^{\perp}$. Conversely, for all $y \in S^{\perp}$, $P_S y = 0$. Thus $y = y - P_S y = (I - P_S)y \in R(I - P_S)$, i.e., $S^{\perp} \subseteq R(I - P_S)$. Also, $(I - P_S)^T = I - P_S$, $(I - P_S)^2 = I - P_S$.

7.36 (i) It is obvious that $R(A(A^TA)^{-1}A^T) \subseteq R(A)$. Let $x \in R(A)$; then there exists y, such that $x = Ay$. So $A(A^TA)^{-1}A^T(Ay) = Ay = x$. Thus $x \in R(A(A^TA)^{-1}A^T)$. Hence $R(A(A^TA)^{-1}A^T) = R(A)$. Also, it is easy to see that $A(A^TA)^{-1}A^T$ are idempotent and symmetric.

Chapter 8

8.1 (a) Let A have full rank. Then for any nonzero vector x, $Ax = y$ is nonzero. Thus, $x^T A^T A x = y^T y > 0$. Now reverse the argument to prove the converse.

(b) Let $z \in R(A)$. Then there is a y such that $Ay = z$, $z^T r = z^T(b - Ax) = z^T b - z^T A(A^TA)^{-1}A^T b = y^T A^T b - y^T A^T A(A^TA)^{-1}A^T b = 0$ (since $x = (A^TA)^{-1}A^T b$).

8.2 Let $b - Ax \in N(A^T)$. Then $A^T(b - x) = 0$, that is, $A^TAx = A^Tb$.

8.11 (a) $\|A\|_2^2 = \|A^TA\|_2$, $\|A^{\dagger}\|_2^2 = \|A^{\dagger}A^{\dagger^T}\|_2 = \|(A^TA)^{-1}\|_2$. Thus, $\text{Cond}_2^2(A) = \|A\|_2^2\|A^{\dagger}\|_2^2 = \text{Cond}_2(A^TA)$.

8.16 (b) The basic solution

$$x_B = P\begin{pmatrix} R_{11}^{-1}c \\ 0 \end{pmatrix}.$$

The minimum-norm solution is computed by taking y_2 as

$$\min_{y_2}\left\|\begin{pmatrix} R_{11}^{-1}R_{12} \\ -I_{n-r} \end{pmatrix} y_2 - \begin{pmatrix} R_{11}^{-1}c \\ 0 \end{pmatrix}\right\|_2.$$

Therefore, the basic solution x_B cannot be a minimum-norm solution unless $R_{12} = 0$.

8.18 (b) $$\|Ax - b\|^2 = \|Q\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T x - b\|^2$$

$$= \|Ry_1 - c\|^2 + \|d\|^2,$$

where

$$Q^T b = \begin{pmatrix} c \\ d \end{pmatrix} \quad \text{and} \quad V^T x = y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Therefore, all least-squares solutions are of the form

$$x = V \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

where y_1 is the unique solution of $Ry_1 = c$. The minimum-norm solution is obtained by setting $y_2 = 0$.

8.25 Since $X = (A^T A)^{-1} = R^{-1}(R^{-1})^T$, the diagonal entries of X are just the 2-norms squared of the rows of R^{-1}: $x_{ii} = \sum_{j=i}^{n} s_{ij}^2,\ i = 1, 2, \ldots, n$, where $S = (s_{ij}) = R^{-1}$.

Chapter 9

9.5 By Theorem 9.6, $|\lambda - a_{ii}| \leq r_i = \sum_{j \neq i}^{n} |a_{ij}|$. Since A is diagonally dominant, we have $|\lambda - a_{ii}| \leq \sum_{j=1}^{n} |a_{ij}| < |a_{ii}|$. This means that $\lambda \neq 0$.

9.7 The eigenvalues of a symmetric matrix A are real. Theorem 9.6, therefore, gives $-r_i \leq \lambda - a_{ii} \leq r_i$.

9.9 $(A - \sigma I)v_i = Av_i - \sigma v_i = \lambda_i v_i - \sigma v_i = (\lambda - \sigma)v_i$.

9.10 (a) $\lambda_2 = 2.9$ is close to $\lambda_1 = 3$.

9.18 (b) First, find a Householder matrix P_1 such that $P_1 b$ is a multiple of e_1. Form $H_1 = P_1 A P_1^T$. Transform H_1 to an upper Hessenberg matrix using Householder transformations, updating the new vector b every time a Householder transformation is applied in the process.

9.19 (a) $A^{(k)} = P_k A P_k^T$. $P_k A$ requires $4(n-k)^2$ flops. $(P_k A) P_k^T$ requires $4n(n-k)$ flops. Follow now the hint given in the problem.

(d) If A is symmetric, the reduced form is symmetric tridiagonal, and therefore only *about half* the entries need to be computed.

9.20 (a) Take $X = (x_1, x_2, \ldots, x_n)$, $x_1 = (1, 0, \ldots, 0)^T = e_1^T$ and compute $x_i = H^{i-1} e_1, i = 2, 3, \ldots, n$. The matrix X is an upper triangular matrix with $h_{21}, h_{21}h_{32}, \ldots, h_{21} \ldots h_{n,n-1}$ as the diagonal entries. Since H is unreduced, X is nonsingular. Moreover $X^{-1}HX$ is a companion matrix in upper Hessenberg form.

(b) If the subdiagonal entries of H are very small, X will be ill-conditioned.

9.21 (c) Suppose $h_{21} = 0$, and all other subdiagonal entries are nonzero, then H can be written in the form

$$H = \begin{pmatrix} h_{11} & x \\ 0 & \hat{H} \end{pmatrix},$$

where $\hat{H}$ is an unreduced $(n-1) \times (n-1)$ upper Hessenberg matrix.

9.24 (e) Note that all s_j in this case are 1. The conclusion, therefore, follows from Theorem 9.41.

9.25 The eigenvalues are well-conditioned, because A is symmetric. But when ϵ is small, the eigenvalues $1+\epsilon$ and $1-\epsilon$ are close to each other. By Theorem 9.41, the eigenvectors corresponding to these eigenvalues are then ill-conditioned.

9.28 (a) $X = \begin{pmatrix} 1 & 1 \\ 0 & \epsilon \end{pmatrix}$.

If ϵ is small, X is ill-conditioned; then by the Bauer–Fike theorem (Theorem 9.37), the eigenvalues of A are ill-conditioned.

9.30 (a) The matrix Q is Hessenberg. The first column of QR is therefore $(q_{11}r_{11},\ q_{21}r_{11},\ 0,\dots,0)^T$. The first column of $H-\lambda I$ is $(h_{11}-\lambda,\ h_{21},0,\dots,0)^T$. Since they have to be the same, the first column of Q is $\frac{1}{r_{11}}$ times the first column of $H-\lambda I$.

(c) Each of the Givens rotations in the problem has 1 in the $(1,1)$ entry. Thus $\tilde{Q}$ has the same first column as P_0 and has the same first column as Q by part (b).

9.34 It has been proved in Section 9.8.4 that H_2 is similar to H_0. Thus, they have the same eigenvalues.

Chapter 10

10.2 (a) Apply Theorem 10.4 to show that $|\lambda| < 4$. The Sturm sequence evaluated at $\mu = 2$ gives $1,\ -1,\ -1,\ 1,\ 1$. There are two agreements and two variations of sign. Thus, there are two eigenvalues of A greater than 2 and two that are less than 2.

10.18 (a) Let $v \in R(A)$. Then there exists u such that $v = Au$. Using Property 1 of the pseudoinverse, we then have $AA^\dagger v = AA^\dagger Au = Au = v$.

(b) Let $x \in N(A^T)$. Then $A^\dagger x = A^\dagger AA^\dagger x = A^\dagger (AA^\dagger)^T x = A^\dagger (A^\dagger)^T A^T x = 0$.

10.19 (a) Let $X = (A^TA)^{-1}A^T$. Then X exists since A has full column rank. Since $AXA = A(A^TA)^{-1}A^TA = A$, $XAX = (A^TA)^{-1}A^TA(A^TA)^{-1}A^T = (A^TA)^{-1}A^T = X$, $(AX)^T = X^TA^T = A(A^TA)^{-1}A^T = AX$, $(XA)^T = A^TX^T = A^TA(A^TA)^{-1} = I$, and $XA = (A^TA)^{-1}A^TA = I$, we have $A^\dagger = (A^TA)^{-1}A^T$. (Note that $(XA)^T = XA = I$.)

10.21 (a) Without loss of generality, assume that B is 2×2. Assume further that none of the diagonal entries is zero. Let

$$B = \begin{pmatrix} \alpha_1 & \beta_1 \\ 0 & \alpha_2 \end{pmatrix}.$$

Then

$$B^TB = \begin{pmatrix} \alpha_1^2 & \alpha_1\beta_1 \\ \beta_1\alpha_1 & \beta_1^2+\alpha_2^2 \end{pmatrix}.$$

The characteristic polynomial $p(\lambda)$ of B^TB is $\lambda^2 - (\alpha_1^2+\alpha_2^2+\beta_1^2)\lambda + \alpha_1^2\alpha_2^2$. Thus B has a multiple singular value if and only if $p(\lambda)$ has a multiple zero. Since α_1 and α_2 are not zero, $p(\lambda)$ can have a multiple zero only if $\beta_1 = 0$.

Chapter 11

11.1 (ii) The eigenvalues are 1, ∞; since B is singular, det $(A - \lambda B)$ is a polynomial of degree 1.

11.2 Let x be the common nonzero null vector of A and B. Then $(A - \lambda B)x = 0$ for all λ. Since $x \neq 0$, $\det(A - \lambda B) = 0$ for any λ.

11.3 (ii) Let $B = P^{-1}\Lambda P$, where Λ is the Jordan canonical form of B. Then $p(\lambda) = \det(A - \lambda B) = \det(PAP^{-1} - \lambda\Lambda)$. Now B is nonsingular if and only if Λ is nonsingular, and nonsingularity of Λ implies that $p(\lambda)$ is of degree n.

(iv) If x is an eigenvector of $(A,\ B)$ corresponding to an eigenvalue λ, then $y = V^T x$ is an eigenvector of the orthogonally equivalent pair $(\tilde{A},\ \tilde{B})$ corresponding to λ, where $\tilde{B} = UBV$.

11.4 Q_1 is a Householder matrix that transforms the first column n_1 of the matrix N to a multiple of e_1. n_1 has three nonzero entries $x,\ y,\ z$ and it has been shown in Section 11.4.2 that these entries can be computed by inverting the 2×2 leading principal minor of B.

11.11 Step 1. Find the ordered (from smallest to largest) Shur decomposition of B: $V^T BV = D$, and let $L = VD^{\frac{1}{2}}$.
Step 2. Form $C = L^{-1}A(L^T)^{-1}$.
Step 3. Compute the eigenvalues and eigenvectors y_i, $i = 1, \dots, n$, of the symmetric matrix C using QR iteration.
Step 4. Compute the eigenvectors x_i of the pencil $(A,\ B)$ by solving $L^T x_i = y_i,\ i = 1, \dots, n$.
Give your own example now using the above algorithm.

11.12 Choose $Q_0 = I_{4\times 4}$. Then at the end of the 50th iteration, we have the eigenvectors for the pair $(A,\ B)$.

11.14 From $B^{-1}Ax = \lambda x$, we have

$$|\lambda| = \frac{\|B^{-1}Ax\|}{\|x\|} \le \|B^{-1}A\|.$$

Since B is symmetric positive definite, the eigenvalues of (A, B) are real and we have the result.

Chapter 14

14.5 (a) $L = \begin{pmatrix} 1 & 0 \\ 1.67 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 3 & 4 \\ 0 & -0.68 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 \\ 0.01 & 0 \end{pmatrix}$.

(b) $L = \begin{pmatrix} 1 & 0 \\ 0.04 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 0.25 & 0.79 \\ 0 & 0.09 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 \\ 0 & 0.0016 \end{pmatrix}$.

(c) $L = \begin{pmatrix} 1 & 0 \\ 0.8 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 10 & 9 \\ 0 & -2.2 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

(d) $L = \begin{pmatrix} 1 & 0 \\ 0.25 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 4 & 1 \\ 0 & 1.75 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

(e) $L = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 0.01 & 0.05 \\ 0 & -0.14 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

14.6 (a) $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $L = \begin{pmatrix} 1 & 0 \\ 0.6 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 5 & 6 \\ 0 & 0.4 \end{pmatrix}$,

$$E = LU - PA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(e) $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $L = \begin{pmatrix} 1 & 0 \\ 0.33 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 0.03 & 0.01 \\ 0 & 0.05 \end{pmatrix}$,

$$E = LU - PA = \begin{pmatrix} 0 & 0 \\ -0.0001 & 0.0033 \end{pmatrix}.$$