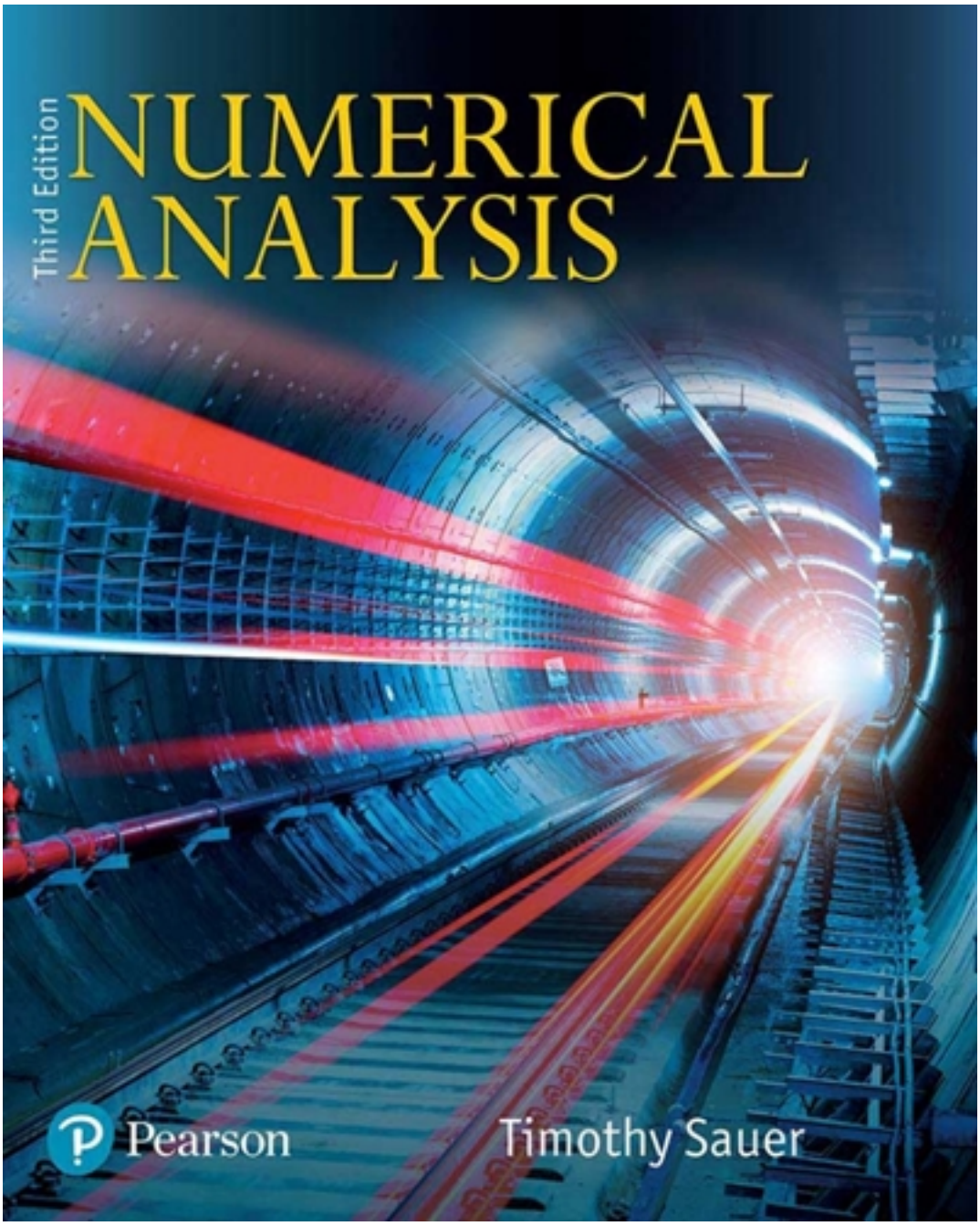
Solutions for Numerical Analysis 3rd Edition by Sauer

CLICK HERE TO ACCESS COMPLETE Solutions



Solutions

CHAPTER 2
Systems of Equations

EXERCISES 2.1 Gaussian Elimination

1 (a) Subtracting $\frac{5}{2}$ times the first equation from the second equation yields $\frac{3}{2}y = 3$, or $y = 2$. Substituting $y = 2$ into the first equation gives $2x - 3(2) = 2$, or $x = 4$.
1 (b) Subtracting 2 times the first equation from the second equation yields $-y = 3$, or $y = -3$. Substituting into the first equation gives $x - 6 = -1$, or $x = 5$.
1 (c) Subtracting -3 times the first equation from the second yields $7y = 21$, or $y = 3$. Substituting into the first equation gives $-x + 3 = 2$, or $x = 1$.

2 (a) $[1, 1, 2]$
2 (b) $[1, 1, 1]$
2 (c) $[-1, 3, 2]$

3 (a) $5z = 5$ implies $z = 1$; $3y - 4(1) = -1$ implies $y = 1$; $3x - 4(1) + 5(1) = 2$ implies $x = \frac{1}{3}$.
3 (b) $-3z = 3$ implies $z = -1$; $4y - 3(-1) = 1$ implies $y = -\frac{1}{2}$; $x - 2(-\frac{1}{2}) + (-1) = 2$ implies $x = 2$.

4 (a) $[1, 1/2, -1]$
4 (b) $[2, 1, 3]$

5 If n increases to $3n$, the approximate operation count changed from $2n^3/3$ to $2(3n)^3/3 = 54n^3/3$, which will take 27 times as long.

6 Approximately 17 seconds.

7 It is given that $(4000)^2$ operations require 0.002 seconds, corresponding to $500(4000)^2$ operations per second. Using the operation count $2n^3/3$, it will take about $(2(9000)^3/3)/(500(4000)^2) \approx 61$ seconds, to solve a general 9000×9000 system.

8 400

COMPUTER PROBLEMS 2.1

1 (a) Putting together the code fragments from the text gives the program

```
for j=1:n-1
  for i=j+1:n
    if abs(a(j,j))<eps; error('zero pivot encountered'); end
    mult = a(i,j)/a(j,j);
    for k = j+1:n
      a(i,k) = a(i,k) - mult*a(j,k);
    end
    b(i) = b(i) - mult*b(j);
  end
end
for i = n:-1:1
  for j = i+1:n
    b(i) = b(i) - a(i,j)*x(j);
  end
  x(i) = b(i)/a(i,i);
end
```

Define the coefficient matrix `a=[2 -2 -1;4 1 -2;-2 1 -1]`, `b=[-2;1;-3]` and apply the preceding MATLAB program. The result is $x = [1, 1, 2]$.

1 (b) Proceed as in (a); the result of the MATLAB program is $x = [1, 1, 1]$.

1 (c) Proceed as in (a); the result is $x = [-1, 3, 2]$.

2 (a) $[-2, 6]$

2 (b) $[5, -120, 630, -1120, 630]$

2 (c) $[-10, 990, -23760, 240240, -1261260, 3783780, -6726720, 7001280, -3938220, 923780]$

EXERCISES 2.2 The LU Factorization

1 (a) Subtracting 3 times the top row from the second row yields the upper triangular matrix $U = \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}$. The matrix of multipliers is $L = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$. Check by multiplication:

$$LU = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

1 (b) Subtracting 2 times the top row from the second row yields

$$U = \begin{bmatrix} 1 & 3 \\ 0 & -4 \end{bmatrix} \text{ and } L = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

Check by multiplication:

$$LU = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}.$$

1 (c) Subtracting $-\dfrac{5}{3}$ times the top row from the second row yields

$$U = \begin{bmatrix} 3 & -4 \\ 0 & -\frac{14}{3} \end{bmatrix} \text{ and } L = \begin{bmatrix} 1 & 0 \\ -\frac{5}{3} & 1 \end{bmatrix}.$$

Check by multiplication:

$$LU = \begin{bmatrix} 1 & 0 \\ -\frac{5}{3} & 1 \end{bmatrix} \begin{bmatrix} 3 & -4 \\ 0 & -\frac{14}{3} \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ -5 & 2 \end{bmatrix}.$$

2 (a) $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

2 (b) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

2 (c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

3 (a) Subtracting 2 times the top row from the second row gives the factorization

$$LU = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 7 \\ 0 & 13 \end{bmatrix}.$$

Solving $Lc = b$, or

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -11 \end{bmatrix},$$

yields $c_1 = 1$ and $2(1) + c_2 = -11$, or $c_2 = -13$. Solving $Ux = c$, or

$$\begin{bmatrix} 3 & 7 \\ 0 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -13 \end{bmatrix},$$

gives $x_2 = 1$ and $3x_1 + 7 = 1$, or $x_1 = -2$. Thus $x = [-2, 1]$.

3 (b) Subtracting 2 times the top row from the second row gives the factorization

$$LU = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}.$$

Solving $Lc = b$, or

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

yields $c_1 = 1$ and $2(1) + c_2 = 3$, or $c_2 = 1$. Solving $Ux = c$, or

$$\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

gives $x_2 = 1$ and $2x_1 + 3(1) = 1$, or $x_1 = -1$. Thus $x = [-1, 1]$.

4 (a) $[-1, 1, 1]$
4 (b) $[1, -1, 2]$

5 Since A is already factored as LU, only the back substitution is needed. Solving $Lc = b$, or

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 4 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix},$$

yields $c_1 = 1, c_2 = 1, c_3 = -2$, and $c_4 = -1$. Solving $Ux = c$, or

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ -1 \end{bmatrix},$$

yields $x_4 = -1, x_3 = 1, x_2 = -1$, and $x_1 = 1$. Thus $x = [1, -1, 1, -1]$.

6 34 seconds

7 To solve 1000 upper-triangular 500×500 systems requires 1000 back substitutions, or approximately $1000(500)^2$ operations. To solve one full 5000×5000 system requires approximately $2(5000)^3/3 + 2(5000)^2 \approx 2(5000)^3/3$ operations. The number of seconds to solve the latter is the ratio

$$\frac{2(5000)^3/3}{1000(500)^2} = \frac{1000}{3} \approx 333 \text{ seconds},$$

or 5 minutes, 33 seconds (or 5 minutes, 34 seconds if the $2(5000)^2$ term is not neglected).

8 10 min., 40 sec.

9 The first problem $Ax = b_0$ requires approximately $2n^3/3$ multiplications, while the 100 subsequent problems require $2n^2$ each. Setting the two equal gives the equation $\dfrac{2n^3}{3} = 200n^2$, or $n = 300$.

COMPUTER PROBLEMS 2.2

1 The elimination part of the code must be supplemented by filling in the entries of L and U. The diagonal entries of L are ones, and the remaining entries are the multipliers from `mult`. It is also necessary to change the `k` loop to go from `j` to `n`, in order to place a zero in the eliminated location of U. MATLAB code follows:

```
l=diag(ones(n,1));
for j = 1:n-1
  for i = j+1:n
    if abs(a(j,j))<eps; error('zero pivot encountered'); end
    mult = a(i,j)/a(j,j);l(i,j)=mult;
    for k = j:n
      a(i,k) = a(i,k) - mult*a(j,k);
    end
  end
end
l
u=a
```

EXERCISES 2.3 Sources of Error

1 (a) The matrix infinity norm is the maximum of the absolute row sums, in this case the maximum of 3 and 7. So $||A||_\infty = 7$.

1 (b) The maximum of the absolute row sums is $|1| + |-7| + |0| = 8$.

2 (a) 21

2 (b) 2403

2 (c) does not exist

3 (a) The solution of the system

$$\begin{bmatrix} 1 & 1 \\ 1.0001 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2.0001 \end{bmatrix},$$

is $[1, 1]$. The forward error is $||[1, 1] - [-1, 3]||_\infty = 2$. The backward error is the infinity norm of

$$b - Ax_c = \begin{bmatrix} 2 \\ 2.0001 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1.0001 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.0002 \end{bmatrix},$$

which is 0.0002. The error magnification factor is the ratio of the relative forward and backward errors, or $(2/1)/(0.0002/2.0001) = 20001$.

3 (b) The forward error is $||[1, 1] - [0, 2]|| = 1$. The backward error is the infinity norm of

$$b - Ax_c = \begin{bmatrix} 2 \\ 2.0001 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.0001 \end{bmatrix},$$

which is 0.0001. The error magnification factor is the ratio of the relative forward and backward errors, or $(1/1)/(0.0001/2.0001) = 20001$.

3 (c) The calculation is similar to (a) and (b). The forward error is $||[1,1]-[2,2]|| = 1$, the backward error is $||[-2,-2.0001]|| = 2.0001$, and the error magnification factor is $(1/1)/(2.0001/2.0001) = 1$.

3 (d) Forward error is $||[1,1]-[-2,4]|| = 3$, the backward error is $||[0,0.0003]|| = 0.0003$, and the error magnification factor is $(3/1)/(0.0003/2.0001) = 20001$.

3 (e) Forward error is $||[1,1]-[-2,4.0001]|| = 3.0001$, the backward error is $||[0.0001,0.0002]|| = 0.0002$, and the error magnification factor is $(3.0001/1)/(0.0002/2.0001) = 30002.5$.

4 (a) FE $= 2$, BE $= 0.01$, EMF $= 400$

4 (b) FE $= 2$, BE $= 0.01$, EMF $= 400$

4 (c) FE $= 1$, BE $= 0.005$, EMF $= 400$

5 (a) The solution of the system

$$\begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix},$$

is $[1,-1]$. The forward error is $||[1,-1]-[-2,-4]||_\infty = 3$, and the relative forward error is $3/||[1,-1]|| = 3$. The backward error is the infinity norm of

$$b - Ax_c = \begin{bmatrix} 3 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} -2 \\ -4 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix},$$

and the relative backward error is is $||[-3,-3||/||[3,7]|| = 3/7$. The error magnification factor is the ratio of the relative forward and backward errors, or $3/\frac{3}{7} = 7$.

5 (b) The forward error is $||[1,-1]-[-2,-3]||_\infty = 3$, and the relative forward error is $3/1 = 3$. The backward error is the infinity norm of

$$b - Ax_c = \begin{bmatrix} 3 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} -2 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and the relative backward error is is $||[0,1||/7 = 1/7$. The error magnification factor is the ratio $3/\frac{1}{7} = 21$.

5 (c) The forward error is $||[1,-1]-[0,-2]||_\infty = 1$, and the relative forward error is $1/1 = 1$. The backward error is the infinity norm of $[-1,-1]$, and the relative backward error is is $1/7$. The error magnification factor is the ratio of the relative forward and backward errors, or $1/\frac{1}{7} = 7$.

5 (d) The forward error is $||[1,-1]-[-1,-1]||_\infty = 2$, and the relative forward error is $2/1 = 2$. The backward error is $||[3,7]-[1,1]||_\infty = 6$, and the relative backward error is is $6/7$. The error magnification factor is $2/\frac{6}{7} = 7/3$.

5 (e) The inverse matrix is

$$A^{-1} = \begin{bmatrix} -2 & 1 \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix}.$$

The condition number of A is $||A|| \cdot ||A^{-1}|| = 7 \cdot 3 = 21$.

6 (a) FE $= 11$, BE $= 0.324$, EMF $= 33.9$
6 (b) FE $= 101$, BE $= 0.418$, EMF $= 241.8$
6 (c) FE $= 601$, BE $= 1/6.01$, EMF $= 3612.01$
6 (d) FE $= 600$, BE $= 0.499168$, EMF $= 1202$
6 (e) 3612.01

7 The maximum row of the 5×5 Hilbert matrix is the top row $[1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}]$, and $||H||_\infty = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} = \frac{137}{60}$.

8 (a) $4/\delta + 4 + \delta$
8 (b) $4/\delta + 4 + \delta$

9 (a) The inverse is

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

The condition number is $||A||_\infty \cdot ||A^{-1}||_\infty = 1 \cdot 1 = 1$.

9 (b) The inverse D^{-1} is a diagonal matrix with entries $d_1^{-1}, \ldots, d_n^{-1}$. The condition number is $||D||_\infty \cdot ||D^{-1}||_\infty = \max |d_i| \cdot \max 1/|d_i| = \max |d_i| / \min |d_i|$.

10 (a) $\kappa(A) = 36012.001$.
10 (b) RFE $= 6001$. RBE $= 1/6.001$. EMF $= 36012.001$.

11 (a) The three properties that define a vector norm must be checked.
(i) $||x||_\infty \geq 0$ is guaranteed by the definition $||x||_\infty = \max |x_i|$, and if $||x||_\infty = 0$, then all components x_i must be zero.
(ii) For a scalar α,

$$\begin{aligned} ||\alpha x||_\infty &= \max\{|\alpha x_1|, \ldots, |\alpha x_n|\} \\ &= \max\{|\alpha||x_1|, \ldots, |\alpha||x_n|\} \\ &= |\alpha| \max\{|x_1|, \ldots, |x_n|\} \\ &= |\alpha| \cdot ||x||_\infty. \end{aligned}$$

(iii)

$$\begin{aligned} ||x + y||_\infty &= \max\{|x_1 + y_1|, \ldots, |x_n + y_n|\} \\ &\leq \max\{|x_1| + |y_1|, \ldots, |x_n| + |y_n|\} \\ &\leq \max\{|x_1|, \ldots, |x_n|\} + \max\{|y_1|, \ldots, |y_n|\} \\ &= ||x||_\infty + ||y||_\infty \end{aligned}$$

11 (b) The three properties:
(i) $||x||_1 \geq 0$ is guaranteed by the definition $||x||_1 = |x_1| + \ldots + |x_n|$, and if $||x||_1 = 0$, all components x_i must be zero.
(ii) For a scalar α,

$$\begin{aligned} ||\alpha x||_1 &= |\alpha x_1| + \ldots + |\alpha x_n| \\ &= |\alpha|(|x_1| + \ldots + |x_n|) \\ &= |\alpha| \cdot ||x||_1 \end{aligned}$$

(iii) The triangle inequality is

$$\begin{aligned} ||x + y||_1 &= |x_1 + y_1| + \ldots + |x_n + y_n| \\ &\leq |x_1| + |y_1| + \ldots + |x_n| + |y_n| \\ &= ||x||_1 + ||y||_1 \end{aligned}$$

13 For a matrix A, the operator norm of the vector infinity norm is

$$\max \frac{||Ax||_\infty}{||x||_\infty}$$

where the maximum is taken over all vectors x. By property (ii) of vector norms, this is equal to the maximum $||Ax||_\infty$, where the maximum is taken over all unit vectors x in the infinity norm, or

$$\max ||Ax||_\infty = \max\{a_{11}x_1 + \ldots + a_{1n}x_n, \ldots, a_{n1}x_1 + \ldots + a_{nn}x_n\}$$

where $|x_1|, \ldots, |x_n| \leq 1$. In fact, the maximum is reached when all x_i are $+1$ or -1, where the sign of x_i is chosen to match the sign of a_{i1}. Here i denotes the largest row in the sense of the infinity vector norm. Therefore

$$||A||_\infty = \text{ maximum absolute row sum of } A = \max_{||x||=1} ||Ax|| = \max_x \frac{||Ax||_\infty}{||x||_\infty}.$$

15 (a) The unit vector that maximizes $||Ax||_\infty$ is $x = [1, 1]$, so that

$$Ax = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}.$$

Note that $||x||_\infty = 1, ||Ax||_\infty = 7$, and $||A||_\infty = 7$. Any scalar multiple of x will work as well.

15 (b) The unit vector that maximizes $||Ax||_\infty$ is $x = [1, -1, 1]$. The signs are chosen to maximize row 3 of Ax. Since

$$Ax = \begin{bmatrix} 1 & 5 & 1 \\ -1 & 2 & -3 \\ 1 & -7 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \\ 8 \end{bmatrix},$$

we have $||x||_\infty = 1, ||Ax||_\infty = 8$, and $||A||_\infty = 8$. Any scalar multiple of x will also work.

16 (a) $[0, 1]$
16 (b) $[0, 1, 0]$

17 Applying Gaussian elimination yields the matrices

$$\begin{bmatrix} 10 & 20 & 1 \\ 1 & 1.99 & 6 \\ 0 & 50 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 10 & 20 & 1 \\ 0 & -0.01 & 5.9 \\ 0 & 50 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 10 & 20 & 1 \\ 0 & -0.01 & 5.9 \\ 0 & 0 & 29501 \end{bmatrix}$$

where the last multiplier is $l_{32} = -5000$. The LU-factorization is

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ 0.1 & 1 & 0 \\ 0 & -5000 & 1 \end{bmatrix} \begin{bmatrix} 10 & 20 & 1 \\ 0 & -0.01 & 5.9 \\ 0 & 0 & 29501 \end{bmatrix},$$

and the largest magnitude multiplier is -5000.

COMPUTER PROBLEMS 2.3

1 (a) Since the answers depend on rounding errors, they will vary slightly with the exact sequence of operations used. For example, using the naive Gaussian elimination code of Computer Problem 2.1.1 gives the forward error $||x - x_c||_\infty \approx 6.6 \times 10^{-10}$ and error magnification factor $\approx 4.6 \times 10^6$, while the MATLAB backslash command, a more sophisticated algorithm, returns forward error $\approx 5.4 \times 10^{-10}$ and error magnification factor $\approx 3.7 \times 10^6$. The condition number of A is approximately 7×10^7.

1 (b) The MATLAB code of Computer Problem 2.1.1 gives forward error $||x - x_c||_\infty \approx 1.5 \times 10^{-3}$ and error magnification factor $\approx 6.2 \times 10^{12}$, while the MATLAB backslash command returns forward error $\approx 1.1 \times 10^{-3}$ and error magnification factor $\approx 9.1 \times 10^{12}$. The condition number of A is approximately 1.3×10^{14}.

2

	n	FE	EMF	cond(A)
(a)	6	8.88×10^{-16}	5.83	8.61
(b)	10	1.11×10^{-15}	9.33	11.26

3 Using naive Gaussian elimination as in Computer Problem 2.1.1, the results are:

n	FE	EMF	cond (A)
100	5.3×10^{-11}	1.2×10^{3}	1.0×10^{4}
200	5.8×10^{-10}	6.3×10^{3}	4.0×10^{4}
300	3.0×10^{-9}	8.7×10^{3}	9.0×10^{4}
400	4.5×10^{-9}	7.0×10^{3}	1.6×10^{5}
500	9.6×10^{-9}	4.8×10^{4}	2.5×10^{5}

The MATLAB backslash command is slightly more efficient, yielding the results:

n	FE	EMF	cond (A)
100	5.7×10^{-12}	6.3×10^{3}	1.0×10^{4}
200	3.4×10^{-11}	1.9×10^{4}	4.0×10^{4}
300	6.2×10^{-11}	3.2×10^{4}	9.0×10^{4}
400	1.8×10^{-10}	9.6×10^{4}	1.6×10^{5}
500	2.6×10^{-10}	1.1×10^{5}	2.5×10^{5}

4

n	FE	EMF	cond (A)
100	4.29×10^{-9}	5.87×10^{6}	6.18×10^{7}
200	7.61×10^{-7}	4.18×10^{8}	1.29×10^{10}
300	1.61×10^{-4}	8.30×10^{10}	6.20×10^{11}
400	0.00830	3.04×10^{12}	1.48×10^{13}
500	0.0578	2.07×10^{13}	2.28×10^{14}

5 The exact n depends slightly on the code, as in Computer Problem 1. Using the naive Gaussian elimination code of Computer Problem 2.1.1, the solution for $n = 11$ rounds to the correct solution $x = [1.0, \ldots, 1.0]$ within one correct decimal place, while the solution for $n = 12$ does not. If the MATLAB backslash command is used, the $n = 12$ solution rounds correctly to one decimal place and $n = 13$ does not.

EXERCISES 2.4 The PA=LU Factorization

1 (a)

$$\begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} \longrightarrow \begin{matrix} P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \text{exchange rows 1 and 2} \end{matrix} \longrightarrow \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix} \longrightarrow \begin{matrix} \text{sub } \frac{1}{2} \text{ x row 1} \\ \text{from row 2} \end{matrix} \longrightarrow \begin{bmatrix} 2 & 3 \\ \textcircled{\frac{1}{2}} & \frac{3}{2} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} = PA = LU = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & \frac{3}{2} \end{bmatrix}$$

1 (b)

$$\begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix} \longrightarrow \begin{array}{c}\text{subtract } \frac{1}{2} \text{ x row 1} \\ \text{from row 2}\end{array} \longrightarrow \begin{bmatrix} 2 & 4 \\ \textcircled{\frac{1}{2}} & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix} = PA = LU = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$$

1 (c)

$$\begin{bmatrix} 1 & 5 \\ 5 & 12 \end{bmatrix} \longrightarrow \begin{array}{c} P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \text{exchange rows 1 and 2}\end{array} \longrightarrow \begin{bmatrix} 5 & 12 \\ 1 & 5 \end{bmatrix} \longrightarrow \begin{array}{c}\text{sub } \frac{1}{5} \text{ x row 1} \\ \text{from row 2}\end{array} \longrightarrow \begin{bmatrix} 5 & 12 \\ \textcircled{\frac{1}{5}} & \frac{13}{5} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 5 \\ 5 & 12 \end{bmatrix} = PA = LU = \begin{bmatrix} 1 & 0 \\ \frac{1}{5} & 1 \end{bmatrix}\begin{bmatrix} 5 & 12 \\ 0 & \frac{13}{5} \end{bmatrix}$$

1 (d)

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \longrightarrow \begin{array}{c} P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \text{exchange rows 1 and 2}\end{array} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = PA = LU = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2 (a) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & -1 \\ -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & -\frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix}$

2 (b) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 3 \\ 2 & 1 & 1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix}$

2 (c) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & 2 \\ -1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & -4 \end{bmatrix}$

2 (d) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

3 (a)

$$\begin{bmatrix} 3 & 7 \\ 6 & 1 \end{bmatrix} \longrightarrow \begin{matrix} P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \text{exchange rows 1 and 2} \end{matrix} \longrightarrow \begin{bmatrix} 6 & 1 \\ 3 & 7 \end{bmatrix} \longrightarrow \begin{matrix} \text{sub } \frac{1}{2} \text{ x row 1} \\ \text{from row 2} \end{matrix} \longrightarrow \begin{bmatrix} 6 & 1 \\ \textcircled{\frac{1}{2}} & \frac{13}{2} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 7 \\ 6 & 1 \end{bmatrix} = PA = LU = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 6 & 1 \\ 0 & \frac{13}{2} \end{bmatrix}$$

$Lc = Pb$:

$$\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -11 \end{bmatrix} = \begin{bmatrix} -11 \\ 1 \end{bmatrix}$$

Solving from the top,

$$c_1 = -11$$
$$\frac{1}{2}(-11) + c_2 = 1 \Rightarrow c_2 = \frac{13}{2}$$

$Ux = c$:

$$\begin{bmatrix} 6 & 1 \\ 0 & \frac{13}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -11 \\ \frac{13}{2} \end{bmatrix}$$

From the bottom,

$$\frac{13}{2}x_2 = \frac{13}{2} \Rightarrow x_2 = 1$$
$$6x_1 + 1(1) = -11 \Rightarrow x_1 = -2$$

The solution is $x = [-2, 1]$.

3 (b)

$$\begin{bmatrix} 3 & 1 & 2 \\ 6 & 3 & 4 \\ 3 & 1 & 5 \end{bmatrix} \longrightarrow \begin{matrix} P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \text{exchange rows 1 and 2} \end{matrix} \longrightarrow \begin{bmatrix} 6 & 3 & 4 \\ 3 & 1 & 2 \\ 3 & 1 & 5 \end{bmatrix}$$

$$\longrightarrow \begin{matrix} \text{subtract } \frac{1}{2} \text{ x row 1} \\ \text{from row 2} \end{matrix} \longrightarrow \begin{bmatrix} 6 & 3 & 4 \\ \textcircled{\frac{1}{2}} & -\frac{1}{2} & 0 \\ 3 & 1 & 5 \end{bmatrix}$$

$$\longrightarrow \begin{matrix}\text{subtract } \frac{1}{2} \text{ x row 1} \\ \text{from row 3}\end{matrix} \longrightarrow \begin{bmatrix} 6 & 3 & 4 \\ \textcircled{\frac{1}{2}} & -\frac{1}{2} & 0 \\ \textcircled{\frac{1}{2}} & -\frac{1}{2} & 3 \end{bmatrix} \longrightarrow \begin{matrix}\text{subtract 1 x row 2} \\ \text{from row 3}\end{matrix} \longrightarrow \begin{bmatrix} 6 & 3 & 4 \\ \textcircled{\frac{1}{2}} & -\frac{1}{2} & 0 \\ \textcircled{\frac{1}{2}} & \textcircled{1} & 3 \end{bmatrix}$$

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

Solve $Lc = Pb$:

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$$

Starting at the top,

$$\begin{aligned} c_1 &= 1 \\ \frac{1}{2}(1) + c_2 = 0 &\Rightarrow c_2 = -\frac{1}{2} \\ \frac{1}{2}(1) + 1(-\frac{1}{2}) + c_3 = 3 &\Rightarrow c_3 = 3 \end{aligned}$$

Solve $Ux = c$:

$$\begin{bmatrix} 6 & 3 & 4 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{2} \\ 3 \end{bmatrix}$$

Starting at the bottom,

$$\begin{aligned} 3x_3 = 3 \quad &\Rightarrow \quad x_3 = 1 \\ -\frac{1}{2}x_2 = -\frac{1}{2} \quad &\Rightarrow \quad x_2 = 1 \\ 6x_1 + 3(1) + 4(1) = 1 \quad &\Rightarrow \quad x_1 = -1 \end{aligned}$$

Therefore the solution is $x = [-1, 1, 1]$.

4 (a) $[1, -1, 2]$
4 (b) $[5, 4, 3]$

5 According to Theorem 2.8, simply exchange rows 2 and 5 of the identity matrix.

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

6 (a) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

6 (b) The second and fourth columns will be exchanged.

7 The matrix has been changed by moving row 1 to row 4, row 4 to row 3, and row 3 to row 1. According to Theorem 2.8, this can be done by multiplying on the left with a permutation matrix constructed by applying the same changes to the identity matrix. Therefore the leftmost matrix is

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

8 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 10 & 20 & 1 \\ 1 & 1.99 & 6 \\ 0 & 50 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/10 & -\frac{1}{5000} & 1 \end{bmatrix} \begin{bmatrix} 10 & 20 & 1 \\ 0 & 50 & 1 \\ 0 & 0 & 5.9002 \end{bmatrix}.$

Largest multiplier is $1/10$.

9 (a)

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ \textcircled{-1} & 1 & 0 & 2 \\ \textcircled{-1} & -1 & 1 & 2 \\ \textcircled{-1} & -1 & -1 & 2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ \textcircled{-1} & 1 & 0 & 2 \\ \textcircled{-1} & \textcircled{-1} & 1 & 4 \\ \textcircled{-1} & \textcircled{-1} & -1 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ \textcircled{-1} & 1 & 0 & 2 \\ \textcircled{-1} & \textcircled{-1} & 1 & 4 \\ \textcircled{-1} & \textcircled{-1} & \textcircled{-1} & 8 \end{bmatrix}$$

The PA=LU factorization is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

9 (b) Following the pattern in (a), $P = I$, since partial pivoting results in no row exchanges. The entries of L are 1 on the main diagonal, and -1 in all lower triangular locations. The matrix U is the identity matrix except for column n, which is $[2^0, 2^1, 2^2, \ldots, 2^{n-1}]^T$.

10 (a) Under partial pivoting, all multipliers are less than one in absolute value. During the elimination of column k, each entry a_{ij} is changed by the addition of at most the largest entry of the matrix A. Therefore the largest entry of the matrix can at most double in absolute value during elimination of each column k. There are $n-1$ columns to eliminate, so the largest entry of U is at most 2^{n-1}.

10 (b) The analogous fact is that the ratio $\max\{|u_{ij}|\}/\max\{|a_{ij}|\} \leq 2^{n-1}$.

EXERCISES 2.5 Iterative Methods

1 (a) The Jacobi equations are

$$u_{k+1} = \frac{5+v_k}{3}$$
$$v_{k+1} = \frac{4+u_k}{2}$$

Starting with $[u_0, v_0] = [0, 0]$, the first two steps are

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ 2 \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{7}{3} \\ \frac{17}{6} \end{bmatrix}.$$

The Gauss-Seidel equations are

$$u_{k+1} = \frac{5+v_k}{3}$$
$$v_{k+1} = \frac{4+u_{k+1}}{2}$$

Starting with $[u_0, v_0] = [0, 0]$, the first two steps are

$$\begin{bmatrix} u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 5/3 \\ 17/6 \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{47}{18} \\ \frac{119}{36} \end{bmatrix}.$$

1 (b) The Jacobi equations are

$$u_{k+1} = \frac{v_k}{2}$$
$$v_{k+1} = \frac{u_k + w_k + 2}{2}$$
$$w_{k+1} = \frac{v_k}{2}$$

Starting with $[u_0, v_0, w_0] = [0, 0, 0]$, the first two steps are

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix}.$$

The Gauss-Seidel equations are

$$\begin{aligned} u_{k+1} &= \frac{v_k}{2} \\ v_{k+1} &= \frac{u_{k+1} + w_k + 2}{2} \\ w_{k+1} &= \frac{v_{k+1}}{2} \end{aligned}$$

Starting with $[u_0, v_0, w_0] = [0, 0, 0]$, the first two steps are

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 3/2 \\ 3/4 \end{bmatrix}.$$

1 (c) The Jacobi equations are

$$\begin{aligned} u_{k+1} &= \frac{6 - v_k - w_k}{3} \\ v_{k+1} &= \frac{3 - u_k - w_k}{3} \\ w_{k+1} &= \frac{5 - u_k - v_k}{3} \end{aligned}$$

Starting with $[u_0, v_0, w_0] = [0, 0, 0]$, the first two steps are

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ \frac{5}{3} \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} 10/9 \\ -2/9 \\ 2/3 \end{bmatrix}.$$

The Gauss-Seidel equations are

$$\begin{aligned} u_{k+1} &= \frac{6 - v_k - w_k}{3} \\ v_{k+1} &= \frac{3 - u_{k+1} - w_k}{3} \\ w_{k+1} &= \frac{5 - u_{k+1} - v_{k+1}}{3} \end{aligned}$$

Starting with $[u_0, v_0, w_0] = [0, 0, 0]$, the first two steps are

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{1}{3} \\ \frac{8}{9} \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{43}{27} \\ \frac{14}{81} \\ \frac{262}{243} \end{bmatrix}.$$

2 (a) Jacobi $[u_2, v_2] = [22/15, -11/15]$ Gauss-Seidel $[u_2, v_2] = [134/75, -209/225]$
2 (b) Jacobi $[u_2, v_2, w_2] = [-39/40, -49/120, 23/24]$
Gauss-Seidel $[u_2, v_2, w_2] = [-191/180, -361/720, 89/80]$
2 (c) Jacobi $[u_2, v_2, w_2] = [-3/4, 5/4, 3/8]$ Gauss-Seidel $[u_2, v_2, w_2] = [-9/32, 169/128, 87/256]$

3 (a) The SOR equations are

$$u_{k+1} = (1-\omega)u_k + \omega\frac{5+v_k}{3}$$
$$v_{k+1} = (1-\omega)v_k + \omega\frac{4+u_{k+1}}{2}$$

where $\omega = 1.5$. Starting with $[u_0, v_0] = [0, 0]$, the first two steps are

$$u_1 = -\frac{1}{2}u_0 + \frac{3(5+v_0)}{6} = \frac{5}{2}$$
$$v_1 = -\frac{1}{2}v_0 + \frac{3(4+u_1)}{4} = \frac{39}{8}$$

and

$$u_2 = -\frac{1}{2}u_1 + \frac{3(5+v_1)}{6} = \frac{59}{16}$$
$$v_2 = -\frac{1}{2}v_1 + \frac{3(4+u_2)}{4} = \frac{213}{64}$$

3 (b) The SOR equations are

$$u_{k+1} = (1-\omega)u_k + \omega\frac{v_k}{2}$$
$$v_{k+1} = (1-\omega)v_k + \omega\frac{u_{k+1}+w_k+2}{2}$$
$$w_{k+1} = (1-\omega)w_k + \omega\frac{v_{k+1}}{2}$$

where $\omega = 1.5$. Starting with $[u_0, v_0, w_0] = [0, 0, 0]$, the first two steps are $[u_1, v_1, w_1] = [0, \frac{3}{2}, \frac{9}{8}]$ and $[u_2, v_2, w_2] = [\frac{9}{8}, \frac{39}{16}, \frac{81}{64}]$.

3 (c) The SOR equations are

$$u_{k+1} = (1-\omega)u_k + \omega\frac{6 - v_k - w_k}{3}$$
$$v_{k+1} = (1-\omega)v_k + \omega\frac{3 - u_{k+1} - w_k}{3}$$
$$w_{k+1} = (1-\omega)w_k + \omega\frac{5 - u_{k+1} - v_{k+1}}{3}$$

where $\omega = 1.5$. Starting with $[u_0, v_0, w_0] = [0, 0, 0]$, the first two steps are $[u_1, v_1, w_1] = [3, 0, 1]$ and $[u_2, v_2, w_2] = [1, \frac{1}{2}, \frac{5}{4}]$

4 (a) $\omega = 1$, $[u_2, v_2] = [134/75, -209/225]$; $\omega = 1.2$, $[u_2, v_2] = [2.089, -1.040]$
4 (b) $\omega = 1$, $[u_2, v_2, w_2] = [-191/180, -361/720, 89/80]$;
$\omega = 1.2$, $[u_2, v_2, w_2] = [-1.235, -0.646, 1.168]$
4 (c) $\omega = 1$, $[u_2, v_2, w_2] = [-9/32, 169/128, 87/256]$; $\omega = 1.2$, $[u_2, v_2, w_2] = [-0.27, 1.281, 0.371]$

5 (a) By dividing an eigenvector v associated to λ by its largest magnitude entry, we can find an eigenvector whose largest magnitude entry v_m is exactly 1. The mth row of the eigenvalue equation $Av = \lambda v$ is therefore

$$A_{m1}v_1 + \ldots + A_{m,m-1}v_{m-1} + A_{mm} + A_{m,m+1}v_{m+1} + \ldots + A_{mn}v_n = \lambda.$$

Since $|v_i| \leq 1$ for all $1 \leq i \leq n$, it follows that

$$\begin{aligned} |A_{mm} - \lambda| &= |A_{m1}v_1 + \ldots + A_{m,m-1}v_{m-1} + A_{m,m+1}v_{m+1} + \ldots + A_{mn}v_n| \\ &\leq \sum_{j \neq m} |A_{mj}|. \end{aligned}$$

5 (b) If $\lambda = 0$ is an eigenvalue of A, then by the Gerschgorin Circle Theorem there exists an m such that $|A_{mm}| \leq \sum_{j \neq m} |A_{mj}|$, which contradicts strict diagonal dominance.

COMPUTER PROBLEMS 2.5

1 The MATLAB program `jacobi.m` can be used to solve the system after defining A and b with an altered version of `sparsesetup.m`. The initial vector is set to $[0, \ldots, 0]$. By checking the infinity norm error of the solution x, say by the command `norm(x-1,inf)`, the Jacobi method can be iterated until the error is less than 0.5×10^{-6}. For $n = 100$, 36 Jacobi steps are required. For $n = 100000$, 36 steps are required. In both cases, the backward error is approximately 4.6×10^{-7}.

2 16209 steps, BE $= 4.84 \times 10^{-7}$

3 The Gauss-Seidel method can be coded in MATLAB as follows:

```
% Gauss-Seidel
% Inputs: sparse matrix a, r.h.s b,
%         d = diagonal of a, r = rest of a,
%         numsteps = number of Jacobi iterations
% Output: solution x
function x = gaussseidel(a,b,k)
n=length(b);          % find n
d=diag(diag(a)); u=triu(a,1);l=tril(a,-1);
x=zeros(n,1);         % Initialize vector x
for j=1:k             % loop for GS iteration
  b1=b-u*x;
  for i=1:n
      x(i)=(b1(i)-l(i,:)*x)/d(i,i);
  end
end
```

5 Using the code from Computer Problem 3, 21 steps of Gauss-Seidel iteration, starting from $x = [0, \ldots, 0]$, are needed to converge to the correct solution within 6 decimal places. Using the code from Computer Problem 4, 16 steps of SOR with $\omega = 1.2$, starting from $x = [0, \ldots, 0]$, are needed to converge to the correct solution within 6 decimal places.

6 (a) 8110 steps, BE $= 4.82 \times 10^{-7}$
6 (b) 2699 steps, BE $= 4.86 \times 10^{-7}$

7 The results will depend on the computer. Using the sparse matrix capability of MATLAB and the Gauss-Seidel code from Computer Problem 3, typical results for one second of computation are given in the table.

n	steps	forward error
400	50	1.1×10^{-8}
800	15	1.7×10^{-3}
1200	7	2.5×10^{-2}

EXERCISES 2.6 Methods for Symmetric Positive-Definite Matrices

1 (a) For $x = [x_1, x_2] \neq 0$,

$$x^T Ax = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 3x_2^2 > 0.$$

1 (b) For $x = [x_1, x_2] \neq 0$,

$$x^T Ax = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 6x_1x_2 + 10x_2^2 = (x_1 + 3x_2)^2 + x_2^2 > 0.$$

1 (c) For $x = [x_1, x_2, x_3] \neq 0$,

$$x^T Ax = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1^2 + 2x_2^2 + 3x_3^2 > 0.$$

2 (a) $[1, 1]$
2 (b) $[2, -1]$
2 (c) $[0, 1]$
2 (d) $[0, 1, 0]$

3 (a) Clearly $R^T R = \begin{bmatrix} \sqrt{1} & 0 \\ 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{1} & 0 \\ 0 & \sqrt{3} \end{bmatrix}$ is a Cholesky factorization; alternatively, R can be chosen to be the negative of this matrix.

3 (b) The top row of R is $R_{11} = \sqrt{a_{11}} = 1$, followed by $R_{12} = u = \dfrac{[3]}{\sqrt{a_{11}}} = 3$. Subtracting the outer product uu^T from the lower principal submatrix $[10]$ leaves $10 - 3 = 1$. Repeating the factorization step for the remaining 1×1 matrix yields $R_{22} = \sqrt{1} = 1$. Therefore $R = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$ satisfies $R^T R = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 10 \end{bmatrix}$.

3 (c) Clearly $R^T R = \begin{bmatrix} \sqrt{1} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{1} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}$ is a Cholesky factorization; alternatively, R can be chosen to be the negative of this matrix.

5 (a) The top row of R is $R_{11} = \sqrt{a_{11}} = 1$, followed by $R_{12} = u = \dfrac{[2]}{\sqrt{a_{11}}} = 2$. Subtracting the outer product uu^T from the lower principal submatrix $[8]$ leaves $8 - 4 = 4$. Repeating the factorization step for the remaining 1×1 matrix yields $R_{22} = \sqrt{4} = 2$. Therefore $R = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$ satisfies $R^T R = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix}$.

5 (b) The top row of R is $R_{11} = \sqrt{a_{11}} = 2$, followed by $R_{12} = u = \dfrac{[-2]}{\sqrt{a_{11}}} = -1$. Subtracting the outer product uu^T from the lower principal submatrix $[5/4]$ leaves $5/4 - 1 = 1/4$. Repeating the factorization step for the remaining 1×1 matrix yields $R_{22} = \sqrt{1/4} = 1/2$. Therefore $R = \begin{bmatrix} 2 & -1 \\ 0 & 1/2 \end{bmatrix}$ satisfies $R^T R = \begin{bmatrix} 2 & 0 \\ -1 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 5/4 \end{bmatrix}$.

5 (c) The top row of R is $R_{11} = \sqrt{a_{11}} = 5$, followed by $R_{12} = u = \dfrac{[5]}{\sqrt{a_{11}}} = 1$. Subtracting the outer product uu^T from the lower principal submatrix $[26]$ leaves $26 - 1 = 25$. Repeat-

ing the factorization step for the remaining 1×1 matrix yields $R_{22} = \sqrt{25} = 5$. Therefore $R = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ satisfies $R^T R = \begin{bmatrix} 5 & 0 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 25 & 5 \\ 5 & 26 \end{bmatrix}$.

5 (d) The top row of R is $R_{11} = \sqrt{a_{11}} = 1$, followed by $R_{12} = u = \dfrac{[-2]}{\sqrt{a_{11}}} = -2$. Subtracting the outer product uu^T from the lower principal submatrix $[8]$ leaves $5 - (-2)(-2) = 1$. Repeating the factorization step for the remaining 1×1 matrix yields $R_{22} = \sqrt{1} = 1$. Therefore $R = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$ satisfies $R^T R = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$.

6 (a) $R = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$

6 (b) $R = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

6 (c) $R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

6 (d) $R = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

7 (a) The Cholesky factorization is $R^T R = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$.

The two-part back substitution is $R^T c = b$ followed by $Rx = c$.

The solution of $\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -7 \end{bmatrix}$ is $c = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$,

and the solution of $\begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ is $x = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

7 (b) The Cholesky factorization is $R^T R = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}$.

The two-part back substitution is $R^T c = b$ followed by $Rx = c$.

The solution of $\begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \end{bmatrix}$ is $c = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$,

and the solution of $\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ is $x = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

8 (a) $x = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

8 (b) $x = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

9 Multiply out $x^T Ax = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 4x_1x_2 + dx_2^2 > 0$ and complete the square as $(x_1 + 2x_2)^2 + (d-4)x_2^2$. If $d > 4$, then $x^T Ax$ is expressed as a sum of squares and is positive for all $x \neq 0$.

10 A is positive definite if $d > 4$.

11 We attempt to find the Cholesky factorization. The matrix A is positive-definite exactly when the diagonal entries of R are positive. The top row of R is $R_{11} = \sqrt{a_{11}} = 1$, followed by $[R_{12}\ R_{13}] = u = \dfrac{[-1\ 0]}{\sqrt{a_{11}}} = [-1\ 0]$. Subtracting the outer product $uu^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ from the lower principal submatrix $uu^T = \begin{bmatrix} 2 & 1 \\ 1 & d \end{bmatrix}$ leaves $uu^T = \begin{bmatrix} 1 & 1 \\ 1 & d \end{bmatrix}$. Repeating the factorization step for the remaining 2×2 matrix yields $R_{22} = \sqrt{1} = 1$, followed by $R_{23} = u = \dfrac{[1]}{\sqrt{a_{22}}} = 1$. Subtracting the outer product uu^T from the lower principal submatrix $[d]$ leaves $d - 1$. The R_{33} entry will be the square root of $d - 1$. The matrix is positive-definite if and only if $d > 1$.

13 (a) Following the Conjugate Gradient Method pseudocode:

$$x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, r_0 = d_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\alpha_0 = \frac{\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{2}{10} = \frac{1}{5}$$

$$x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 1/5 \end{bmatrix}$$

$$r_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix}$$

$$\beta_0 = \frac{r_1^T r_1}{r_0^T r_0} = 0.16$$

$$d_1 = 12\begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix} + 0.16\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.56 \\ -0.24 \end{bmatrix}$$

$$\alpha_1 = \frac{\begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix}^T \begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix}}{\begin{bmatrix} 0.56 \\ -0.24 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0.56 \\ -0.24 \end{bmatrix}} = 5$$

$$x_2 = \begin{bmatrix} 1/5 \\ 1/5 \end{bmatrix} + 5\begin{bmatrix} 0.56 \\ -0.24 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$r_2 = \begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix} - 5\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0.56 \\ -0.24 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

13 (b)

$$x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, r_0 = d_0 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\alpha_0 = \frac{\begin{bmatrix} 1 \\ 3 \end{bmatrix}^T \begin{bmatrix} 1 \\ 3 \end{bmatrix}}{\begin{bmatrix} 1 \\ 3 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}} = \frac{10}{58} = \frac{5}{29}$$

$$x_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \frac{5}{29}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 5/29 \\ 15/29 \end{bmatrix}$$

$$r_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{5}{29}\begin{bmatrix} 7 \\ 17 \end{bmatrix} = \begin{bmatrix} -6/29 \\ 2/29 \end{bmatrix}$$

$$\beta_0 = \frac{r_1^T r_1}{r_0^T r_0} = \frac{4}{(29)^2}$$

$$d_1 = 12\begin{bmatrix} -6/29 \\ 2/29 \end{bmatrix} + \frac{4}{(29)^2}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{170}{(29)^2} \\ \frac{70}{(29)^2} \end{bmatrix}$$

$$\alpha_1 = \frac{\begin{bmatrix} -6/29 \\ 2/29 \end{bmatrix}^T \begin{bmatrix} -6/29 \\ 2/29 \end{bmatrix}}{\frac{1}{(29)^4} \begin{bmatrix} -170 \\ 70 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -170 \\ 70 \end{bmatrix}} = 5.8$$

$$x_2 = \begin{bmatrix} 5/29 \\ 15/29 \end{bmatrix} + 5.8 \begin{bmatrix} -170/(29)^2 \\ 70/(29)^2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$r_2 = \begin{bmatrix} -6/29 \\ 2/29 \end{bmatrix} - 5.8 \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -170/(29)^2 \\ 70/(29)^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

14 (a) $[1, 1]$
14 (b) $[-1, 1]$

15 $\alpha_0 = 1/A, x_1 = b/A, r_1 = b - Ab/A = 0$

COMPUTER PROBLEMS 2.6

1 (a) The Conjugate Gradient loop written in pseudocode in the textbook can be coded as follows.

```
function x=cg(a,b,n)
% Inputs: symm. pos. def. matrix a, right-hand side b, number of steps n
% Output: solution x
x=zeros(n,1);
r=b-a*x;
d=r;d1(:,1)=d;r1(:,1)=r;
for i=1:n
    if max(abs(r))<eps break; end
    alf=d'*r/(d'*a*d);
    x=x+alf*d;
    rold=r;
    r=rold-alf*a*d;
    beta=r'*r/(rold'*rold);
    d=r+beta*d;d1(:,i+1)=d;r1(:,i+1)=r;
end
```

The test for r equal to zero uses `eps`, the machine epsilon. The MATLAB command

```
>> x=cg([1 0;0 2],[2;4],2)
```

returns the solution $x = [2, 2]$.

1 (b) Applying the code from part (a) returns the solution $x = [3, -1]$.

2 (a) $[1, 1, 1]$
2 (b) $[2, -1, 1]$

3 (a) The Conjugate Gradient code from Computer Problem 1 can be used with `a = hilb(4)` and `b = ones(4,1)` to yield the solution $x = [-4, 60, -180, 140]$ after 4 steps.

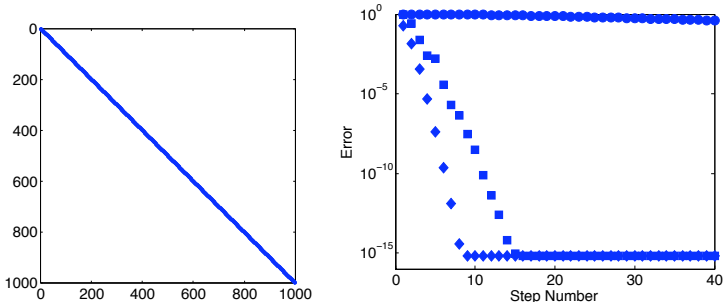
3 (b) The exact solution $x = [-8, 504, -7560, 46200, -138600, 216216, -168168, 51480]$. is approached after more than 20 steps of Conjugate Gradient.

4 (a) $[1, \ldots, 1]$

4 (b) $[1, \ldots, 1]$

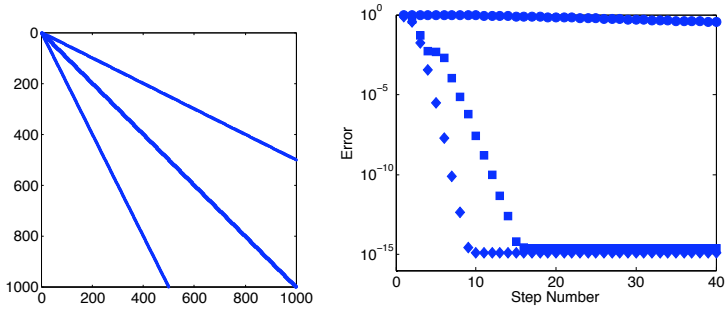
5 Use Program 2.1, `sparsesetup.m` to define the matrix a and right-hand side (b). For $n = 100$, Conjugate Gradient runs 34 steps before the residual r is smaller than machine epsilon in the infinity norm. The final residual is $r \approx 9.76 \times 10^{-17}$. For $n = 1000$, only 35 steps are needed to make the residual $r \approx 7.12 \times 10^{-17}$. For $n = 10000$, 35 steps are needed to make the residual $r \approx 7.17 \times 10^{-17}$.

6 Part (a) shows the output of MATLAB `spy` command on the matrix A. Part (b) shows the error as a function of step number for no preconditioner (circles), Jacobi preconditioner (squares), and Gauss-Seidel preconditioner (diamonds).



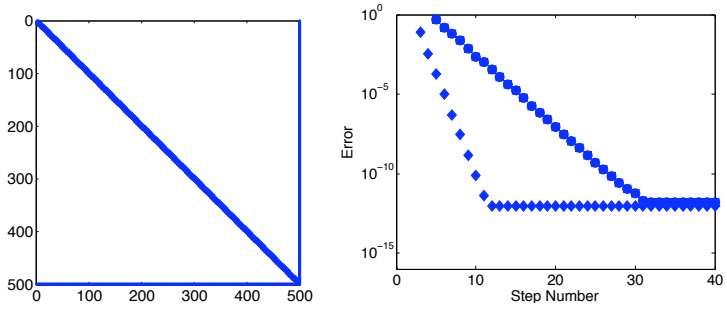
(a) (b)

7 Part (a) shows the output of MATLAB `spy` command on the matrix A. The code shown in the answer to Computer Problem 1(a) above can be slightly modified to carry out the Preconditioned Conjugate Gradient Method outlined in pseudocode in the textbook. Applying this code to the A and b defined in the problem result in Part (b), showing the error as a function of step number for no preconditioner (circles), Jacobi preconditioner (squares), and Gauss-Seidel preconditioner (diamonds).



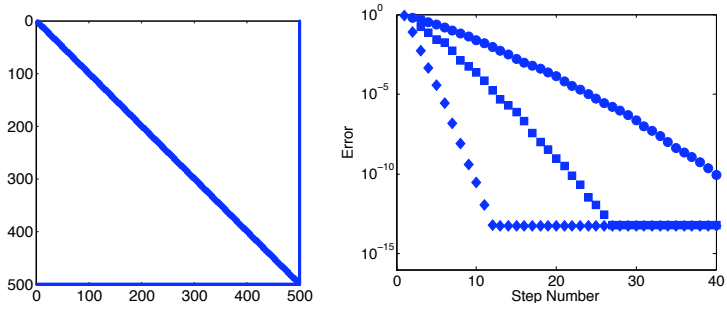
(a) (b)

8 Part (a) shows the output of MATLAB `spy` command on the matrix A. Part (b) shows the error as a function of step number for no preconditioner and Jacobi preconditioner (circles), and Gauss-Seidel preconditioner (diamonds).



(a) (b)

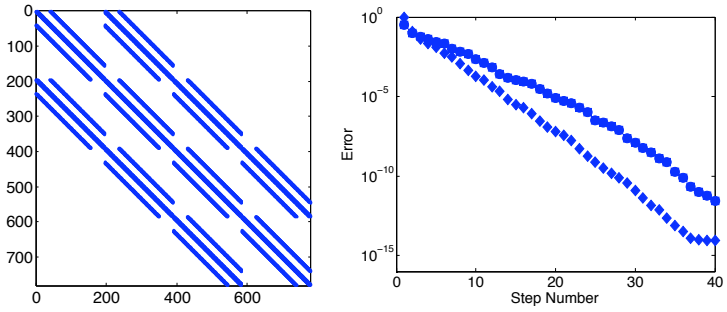
9 Part (a) shows the output of MATLAB `spy` command on the matrix A. Applying the code as described in the answer to Computer Problem 7 to the A and b defined in the problem result in Part (b), showing the error as a function of step number for no preconditioner (circles), Jacobi preconditioner (squares), and Gauss-Seidel preconditioner (diamonds).



(a) (b)

10 Part (a) shows the output of MATLAB `spy` command on the matrix A. Part (b) shows the

error as a function of step number for no preconditioner and Jacobi preconditioner (circles), and Gauss-Seidel preconditioner (diamonds).



(a) (b)

EXERCISES 2.7 Nonlinear Systems of Equations

1 (a) $DF(u,v) = \begin{bmatrix} 3u^2 & 0 \\ v^3 & 3uv^2 \end{bmatrix}$

1 (b) $DF(u,v) = \begin{bmatrix} v\cos uv & u\cos uv \\ ve^{uv} & ue^{uv} \end{bmatrix}$

1 (c) $DF(u,v) = \begin{bmatrix} 2u & 2v \\ 2(u-1) & 2v \end{bmatrix}$

1 (d) $DF(u,v,w) = \begin{bmatrix} 2u & 1 & -2w \\ vw\cos uvw & uw\cos uvw & uv\cos uvw \\ vw^4 & uw^4 & 4uvw^3 \end{bmatrix}$

2 (a) $\begin{bmatrix} 2+u+2v \\ u+v \end{bmatrix}$

2 (b) $\begin{bmatrix} 2+2(u-1)-(v-1) \\ 3+2(u-1)+(v-1) \end{bmatrix}$

3 (a) The curves are circles with radius 1 centered at $(u,v) = (0,0)$ and $(1,0)$, respectively. Solving the first equation for v^2 and substituting into the second yields $(u-1)^2+1-u^2 = 1$ or $-2u+1=0$, so $u = \frac{1}{2}$. The two solutions are $(u,v) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$.

3 (b) The curves are ellipses with semimajor axes 1 and 2 centered at zero and aligned with the x and y axes. Solving by substitution gives the four solutions $(u,v) = (\pm\frac{2}{\sqrt{5}}, \pm\frac{2}{\sqrt{5}})$.

3 (c) The curves are a hyperbola and a circle that intersects one half of the hyperbola in two points. Solving by substitution gives the two solutions $(u,v) = (\frac{4}{5}(1+\sqrt{6}), \pm\frac{1}{5}\sqrt{3+8\sqrt{6}})$.

4 (a) $x_2 = [1/2, 7/8]$
4 (b) $x_2 = [161/180, 161/180]$
4 (c) $x_2 = [2269/388, 651/388]$

5 (a) Given initial values $A_0 = I$ and $x_0 = [1,1]^T$, set $F\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u^2+v^2-1 \\ (u-1)^2+v^2-1 \end{bmatrix}$.

According to Broyden's Method,

$$\begin{aligned} x_1 &= x_0 - A_0^{-1}F(x_0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \Delta_1 &= F(x_1) - F(x_0) = F\begin{bmatrix} 0 \\ 1 \end{bmatrix} - F\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ \delta_1 &= x_1 - x_0 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A_1 &= A_0 + \frac{(\Delta_1 - A_0\delta_1)\delta_1^T}{\delta_1^T\delta_1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \end{bmatrix}}{\begin{bmatrix} -1 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ x_2 &= x_1 - A_1^{-1}F(x_1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

5 (b) Proceed as in (a), with $F\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u^2+4v^2-4 \\ 4u^2+v^2-4 \end{bmatrix}$. According to Broyden's Method,

$$\begin{aligned} x_1 &= x_0 - A_0^{-1}F(x_0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Delta_1 &= F(x_1) - F(x_0) = F\begin{bmatrix} 0 \\ 0 \end{bmatrix} - F\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -5 \end{bmatrix} \\ \delta_1 &= x_1 - x_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \\ A_1 &= A_0 + \frac{(\Delta_1 - A_0\delta_1)\delta_1^T}{\delta_1^T\delta_1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\begin{bmatrix} -4 \\ -4 \end{bmatrix}\begin{bmatrix} -1 & -1 \end{bmatrix}}{\begin{bmatrix} -1 & -1 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix}} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \\ x_2 &= x_1 - A_1^{-1}F(x_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}^{-1}\begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.8 \end{bmatrix} \end{aligned}$$

5 (c) Proceed as in (a), with $F\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u^2 - 4v^2 - 4 \\ (u-1)^2 + v^2 - 4 \end{bmatrix}$. According to Broyden's Method,

$$x_1 = x_0 - A_0^{-1}F(x_0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -7 \\ -3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\Delta_1 = F(x_1) - F(x_0) = F\begin{bmatrix} 8 \\ 4 \end{bmatrix} - F\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 61 \end{bmatrix} - \begin{bmatrix} -7 \\ -3 \end{bmatrix} = \begin{bmatrix} 3 \\ 64 \end{bmatrix}$$

$$\delta_1 = x_1 - x_0 = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$$

$$A_1 = A_0 + \frac{(\Delta_1 - A_0\delta_1)\delta_1^T}{\delta_1^T\delta_1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\begin{bmatrix} -4 \\ 61 \end{bmatrix}\begin{bmatrix} 7 & 3 \end{bmatrix}}{\begin{bmatrix} 7 & 3 \end{bmatrix}\begin{bmatrix} 7 \\ 3 \end{bmatrix}} = \begin{bmatrix} 0.5172 & -0.2069 \\ 7.3621 & 4.1552 \end{bmatrix}$$

$$x_2 = x_1 - A_1^{-1}F(x_1) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} - A_1^{-1}\begin{bmatrix} -4 \\ 61 \end{bmatrix} = \begin{bmatrix} 9.0892 \\ -12.6103 \end{bmatrix}$$

6 (a) $x_1 = [0, 1]$, $x_2 = [0, 0]$
6 (b) $x_1 = [0, 0]$, $x_2 = [0.8, 0.8]$
6 (c) $x_1 = [8, 4]$, $x_2 \approx [9.0892, -12.6103]$

COMPUTER PROBLEMS 2.7

1 (a) For the function $F\begin{pmatrix} u \\ v \end{pmatrix} = \begin{bmatrix} u^2 + v^2 - 1 \\ (u-1)^2 + v^2 - 1 \end{bmatrix}$,
the Jacobian is $DF(u, v) = \begin{bmatrix} 2u & 2v \\ 2(u-1) & 2v \end{bmatrix}$.
Multivariate Newton's Method from p. 131 of the textbook with appropriate initial vectors converges to the roots shown in the Exercise 3(a) solution above.

1 (b) Similar to (a); check solutions with Exercise 3(b).

1 (c) Similar to (a); check solutions with Exercise 3(c).

2 $[1, 1], [0.865939, 0.462168], [0.886809, -0.294007]$

3 Given the multivariate function $F\begin{pmatrix} u \\ v \end{pmatrix} = \begin{bmatrix} u^3 - v^3 + u \\ u^2 + v^2 - 1 \end{bmatrix}$,
the Jacobian is $DF(u, v) = \begin{bmatrix} 3u^2 + 1 & -3v^2 \\ 2u & 2v \end{bmatrix}$.

Using the initial vector $[1, 1]$, Newton's Method converges to $[0.50799200, 0.86136179]$. Using the initial vector $[-1, -1]$, Newton's Method converges to root $[-0.50799200, -0.86136179]$.

4 The two solutions are $[2, 1, -1]$ and $\approx [1.0960, -1.1592, -0.2611]$.

5 (a) Points that lie on all three spheres satisfy

$$\begin{aligned}(u-1)^2+(v-1)^2+w^2-1 &= 0\\ (u-1)^2+v^2+(w-1)^2-1 &= 0\\ u^2+(v-1)^2+(w-1)^2-1 &= 0.\end{aligned}$$

The Jacobian is $DF = \begin{bmatrix} 2(u-1) & 2(v-1) & 2w \\ 2(u-1) & 2v & 2(w-1) \\ 2u & 2(v-1) & 2(w-1) \end{bmatrix}$. Under the Newton iteration (2.51), initial guesses near each of the roots $[1, 1, 1]$ and $[1/3, 1/3, 1/3]$ converge to them.

5 (b) Points that lie on all three spheres satisfy

$$\begin{aligned}(u-1)^2+(v+2)^2+w^2-25 &= 0\\ (u+2)^2+(v-2)^2+(w+1)^2-25 &= 0\\ (u-4)^2+(v+2)^2+(w-3)^2-25 &= 0.\end{aligned}$$

The Jacobian is $DF = \begin{bmatrix} 2(u-1) & 2(v+2) & 2w \\ 2(u+2) & 2(v-2) & 2(w+1) \\ 2(u-4) & 2(v+2) & 2(w-3) \end{bmatrix}$. Under the Newton iteration (2.51), initial guesses near each of the roots $[17/9, 22/9, 19/9]$ and $[1, 2, 3]$ converge to them.

6 Newton's Method converges linearly to the double root $[1, 2, 3]$.

7 (a) Broyden I can be used to compute the root with initial vector $[1, 1]$. Convergence occurs within 15 decimal places to the root $(1/2, \sqrt{3}/2)$ after about 11 steps.

7 (b) Similar to (a). Broyden I converges within 15 decimal places to the root $(2/\sqrt{5}, 2/\sqrt{5})$ after about 13 steps.

7 (c) Similar to (a). Broyden I converges to the root $(4(1+\sqrt{6})/5, \sqrt{3+8\sqrt{6}}/5)$ within 15 decimal places after about 14 steps.

8 (a) The MATLAB code `broyden2` can be used to compute the root with initial vectors $[1, 1]$ and $[1, 2]$. Broyden II converges within 15 decimal places to the root $(1/2, \sqrt{3}/2)$ after about 11 steps.

8 (b) Similar to (a). Broyden II converges within 15 decimal places to the root $(2/\sqrt{5}, 2/\sqrt{5})$ after about 13 steps.

8 (c) Similar to (a). Broyden II converges to the root $(4(1+\sqrt{6})/5, \sqrt{3+8\sqrt{6}}/5)$ within 15 decimal places after about 14 steps.

9 (a) Applying Broyden I with initial matrix $A_0 = I$ and initial guesses near each of the roots $[1, 1, 1]$ and $[1/3, 1/3, 1/3]$ converge to them.
9 (b) Applying Broyden I with initial matrix $A_0 = I$ and initial guesses near each of the roots $[17/9, 22/9, 19/9]$ and $[1, 2, 3]$ converge to them.

10 Broyden I converges linearly to $[1, 2, 3]$.

11 (a) Applying Broyden II with initial matrix $B_0 = I$ and initial guesses near each of the roots $[1, 1, 1]$ and $[1/3, 1/3, 1/3]$ converge to them.
11 (b) Applying Broyden II with initial matrix $B_0 = I$ and initial guesses near each of the roots $[17/9, 22/9, 19/9]$ and $[1, 2, 3]$ converge to them.

12 Broyden II converges linearly to $[1, 2, 3]$.