Nate Warner MATH 434 Fall 2025

Problem set 2 - Due: Wednesday, October 15

1.7.10. Let

$$A = \begin{bmatrix} 2 & 1 & -1 & 3 \\ -2 & 0 & 0 & 0 \\ 4 & 1 & -2 & 6 \\ -6 & -1 & 2 & -3 \end{bmatrix}, \qquad b = \begin{bmatrix} 13 \\ -2 \\ 24 \\ -14 \end{bmatrix}.$$

(a) Calculate the appropriate (four) determinants to show that A can be transformed to (nonsingular) upper-triangular form by operations of type 1 only. (By the way, this is strictly an academic exercise. In practice one never calculates these determinants in advance.)

(b) Carry out the row operations of type 1 to transform the system $Ax = b$ to an equivalent system $Ux = y$, where U is upper triangular. Save the multipliers for use in Exercise 1.7.18.

(c) Carry out the back substitution on the system $Ux = y$ to obtain the solution of $Ax = b$. Don't forget to check your work.

Remark. Let $A \in \mathbb{R}^{n\times n}$. A admits an LU factorization $A = LU$ where L is unit lower triangular and U is upper triangular if and only if all leading principal submatrices are nonsingular. ■

a.) So, we check that

$$\det\left(\begin{bmatrix} 2 \end{bmatrix}\right), \ \det\left(\begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}\right), \ \det\left(\begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & 0 \\ 4 & 1 & -2 \end{bmatrix}\right), \ \det\left(\begin{bmatrix} 2 & 1 & -1 & 3 \\ -2 & 0 & 0 & 0 \\ 4 & 1 & -2 & 6 \\ -6 & -1 & 2 & -3 \end{bmatrix}\right)$$

are all nonzero. We see that

$$\begin{aligned}
\det\left(\begin{bmatrix} 2 \end{bmatrix}\right) &= 2 \neq 0, \\
\det\left(\begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}\right) &= 2(0) - (1)(-2) = 2 \neq 0, \\
\det\left(\begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & 0 \\ 4 & 1 & -2 \end{bmatrix}\right) &= -1 \cdot -2(1(-2) - (-1)(1)) = -1 \cdot 2(-2+1) = -2 \neq 0, \\
\det\left(\begin{bmatrix} 2 & 1 & -1 & 3 \\ -2 & 0 & 0 & 0 \\ 4 & 1 & -2 & 6 \\ -6 & -1 & 2 & -3 \end{bmatrix}\right) &= -1 \cdot -2 \det\left(\begin{bmatrix} 1 & -1 & 3 \\ 1 & -2 & 6 \\ -1 & 2 & -3 \end{bmatrix}\right) \\
&= 2\left(1(-2(-3) - 6(2)) - (-1)(1(-3) - 6(-1)) + 3(1(2) - (-2)(-1))\right) \\
&= 2(-3) = -6 \neq 0.
\end{aligned}$$

Thus, all leading principal submatrices are nonsingular and A can be transformed to nonsingular upper-triangular form by operations of type 1 only.

b.) We use Gaussian Elimination on the augmented system $[A|b] \to [U|y]$. We have

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 3 & 13 \\ -2 & 0 & 0 & 0 & -2 \\ 4 & 1 & -2 & 6 & 24 \\ -6 & -1 & 2 & -3 & -14 \end{array}\right].$$

The opertions to get $a_{21} = a_{31} = a_{41} = 0$ are

$$\begin{aligned}
&-(-1)r_1 + r_2 \to r_2', \\
&-2r_1 + r_3 \to r_3', \\
&-(-3)r_1 + r_4 \to r_4'.
\end{aligned}$$

Thus, $m_{21} = -1$, $m_{31} = 2$, $m_{41} = -3$ and the system becomes

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 3 & 13 \\ 0 & 1 & -1 & 3 & 11 \\ 0 & -1 & 0 & 0 & -2 \\ 0 & 2 & -1 & 6 & 25 \end{array}\right].$$

Next, we set a_{22} as the pivot element, r_2 as the pivot row, and perform the following operations to get $a_{23} = a_{24} = 0$. The operations are

$$\begin{aligned}
r_3' &\leftarrow r_3 - (-1)r_2 \implies m_{32} = -1, \\
r_4' &\leftarrow r_4 - 2r_2 \implies m_{42} = 2.
\end{aligned}$$

After these operations, the system becomes

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 3 & 13 \\ 0 & 1 & -1 & 3 & 11 \\ 0 & 0 & -1 & 3 & 9 \\ 0 & 0 & 1 & 0 & 3 \end{array}\right].$$

Next, we set a_{33} as the pivot element, and r_3 as the pivot row, and perform the operation

$$r_4' \leftarrow r_4 - (-1)r_3 \implies m_{43} = -1.$$

After this operation, the system becomes

$$\left[\begin{array}{cccc|c} 2 & 1 & -1 & 3 & 13 \\ 0 & 1 & -1 & 3 & 11 \\ 0 & 0 & -1 & 3 & 9 \\ 0 & 0 & 0 & 3 & 12 \end{array}\right].$$

Thus, the system $Ux = y$ is

$$\begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 13 \\ 11 \\ 9 \\ 12 \end{bmatrix}.$$

(c) Using back substitution, we can solve the above system.

$$\begin{aligned}
3x_4 = 12 &\implies x_4 = 4, \\
-x_3 + 3x_4 = 9 &\implies x_3 = -1(9 - 3(4)) = 3, \\
x_2 - x_3 + 3x_4 = 11 &\implies x_2 = 11 - 3(4) + 3 = 2, \\
2x_1 + x_2 - x_3 + 3x_4 = 13 &\implies x_1 = \frac{13 - 3(4) + 3 - 2}{2} = 1.
\end{aligned}$$

So, the solution to $Ax = b$ is

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}.$$

We can verify this solution by computing Ax, and observing that it equals the given b. We see that

$$\begin{bmatrix} 2 & 1 & -1 & 3 \\ -2 & 0 & 0 & 0 \\ 4 & 1 & -2 & 6 \\ -6 & -1 & 2 & -3 \end{bmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 2+2-3+12 \\ -2+0+0+0 \\ 4+2-6+24 \\ -6-2+6-12 \end{pmatrix} = \begin{pmatrix} 13 \\ -2 \\ 24 \\ -14 \end{pmatrix}.$$

Thus, the solution is verified.

Note: We can assemble our multipliers to form L, we have

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

1.7.18. Solve the linear system $Ax = \hat{b}$, where A is as in Exercise 1.7.10 and

$$\hat{b} = \begin{bmatrix} 12 & -8 & 21 & -26 \end{bmatrix}^T.$$

Use the L and U that you calculated in Exercise 1.7.10.

From the previous exercise, we have that

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -3 & 2 & -1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

So, the system $Ax = \hat{b}$ is solved using our LU decomposition for A. We have

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -3 & 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 \\ -8 \\ 21 \\ -26 \end{pmatrix}.$$

Let $Ux = y$, and $Ly = b$. First, we solve $Ly = b$ for y using forward substitution.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -3 & 2 & -1 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 12 \\ -8 \\ 21 \\ -26 \end{pmatrix}$$

implies

$$\begin{aligned} y_1 &= 12, \\ -y_1 + y_2 &= -8 \implies y_2 = -8 + 12 = 4, \\ 2y_1 - y_2 + y_3 &= 21 \implies y_3 = 21 + 4 - 2(12) = 1, \\ -3y_1 + 2y_2 - y_3 + y_4 &= -26 \implies y_4 = -26 + 1 - 2(4) + 3(12) = 3. \end{aligned}$$

So,

$$y = \begin{pmatrix} 12 \\ 4 \\ 1 \\ 3 \end{pmatrix}.$$

Now, we solve $Ux = y$ with backward substitution. We have

$$\begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \\ 1 \\ 3 \end{pmatrix}.$$

Which, implies that

$$\begin{aligned} 3x_4 = 3 &\implies x_4 = 1, \\ -x_3 + 3x_4 = 1 &\implies x_3 = -1(1 - 3(1)) = 2, \\ x_2 - x_3 + 3x_4 = 4 &\implies x_2 = 4 - 3(1) + 2 = 3, \\ 2x_1 + x_2 - x_3 + 3x_4 = 12 &\implies x_1 = \frac{12 - 3(1) + 2 - 3}{2} = 4. \end{aligned}$$

So,

$$x = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}.$$

We verify the result by computing Ax, and comparing it against $\hat{b}$. We have

$$\begin{bmatrix} 2 & 1 & -1 & 3 \\ -2 & 0 & 0 & 0 \\ 4 & 1 & -2 & 6 \\ -6 & -1 & 2 & -3 \end{bmatrix} \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2(4) + 1(3) - 1(2) + 3(1) \\ -2(4) + 0 + 0 + 0 \\ 4(4) + 1(3) - 2(2) + 6(1) \\ -6(4) - 1(3) + 2(2) - 3(1) \end{pmatrix} = \begin{pmatrix} 12 \\ -8 \\ 21 \\ -26 \end{pmatrix}.$$

The result is verified.

1.7.26. Use the inner-product formulation to calculate the LU decomposition of the matrix A in Exercise 1.7.10

Remark. The inner-product formulas to compute the LU decomposition are

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} \ell_{ik} u_{kj} \quad j = i, i+1, ..., n, \tag{1}$$

$$\ell_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} \ell_{ik} u_{kj}}{u_{jj}} \quad i = j+1, j+2, ..., n. \tag{2}$$

To use these formulas to find each u_{ij} we first need to plug $i = 1$ into (1), then after we get the first row of U, we can plug in $j = 1$ into (2) to get the first column of L, and so on. ■

Recall that the matrix A is given as

$$\begin{bmatrix} 2 & 1 & -1 & 3 \\ -2 & 0 & 0 & 0 \\ 4 & 1 & -2 & 6 \\ -6 & -1 & 2 & -3 \end{bmatrix}.$$

So, we first find the first row of U (set $i = 1$), we have

$$\begin{aligned} u_{11} &= a_{11} = 2, \\ u_{12} &= a_{12} = 1, \\ u_{13} &= a_{13} = -1, \\ u_{14} &= a_{14} = 3. \end{aligned}$$

Next, we find the first column of L (set $j = 1$),

$$\begin{aligned} \ell_{11} &= 1, \\ \ell_{21} &= \frac{a_{21}}{u_{11}} = \frac{-2}{2} = -1, \\ \ell_{31} &= \frac{a_{31}}{u_{11}} = \frac{4}{2} = 2, \\ \ell_{41} &= \frac{a_{41}}{u_{11}} = -\frac{6}{2} = -3. \end{aligned}$$

For the second row of U $(i = 2)$,

$$\begin{aligned} u_{22} &= a_{22} - \sum_{k=1}^{1} \ell_{2k} u_{k2} = 0 - (-1)(1) = 1, \\ u_{23} &= a_{23} - \sum_{k=1}^{1} \ell_{2k} u_{k3} = 0 - (-1)(-1) = -1, \\ u_{24} &= a_{24} - \sum_{k=1}^{1} \ell_{2k} u_{k4} = 0 - (-1)(3) = 3. \end{aligned}$$

For the second column of L $(j = 2)$,

$$\begin{aligned} \ell_{22} &= 1, \\ \ell_{32} &= \frac{a_{32} - \sum_{k=1}^{1} \ell_{3k} u_{k2}}{u_{22}} = \frac{1 - 2(1)}{1} = -1, \\ \ell_{42} &= \frac{a_{42} - \sum_{k=1}^{1} \ell_{4k} u_{k2}}{u_{22}} = \frac{-1 - (-3)(1)}{1} = 2. \end{aligned}$$

For the third row of U $(i = 3)$,

$$\begin{aligned} u_{33} &= a_{33} - \sum_{k=1}^{2} \ell_{3k} u_{k3} = a_{33} - (\ell_{31} u_{13} + \ell_{32} u_{23}) = -2 - (2(-1) + (-1)(-1)) = -1, \\ u_{34} &= a_{34} - \sum_{k=1}^{2} \ell_{3k} u_{k4} = a_{34} - (\ell_{31} u_{14} + \ell_{32} u_{24}) = 6 - (2(3) + (-1)(3)) = 3. \end{aligned}$$

For the third column of L $(j = 3)$,

$$\begin{aligned} \ell_{33} &= 1, \\ \ell_{43} &= \frac{a_{43} - \sum_{k=1}^{2} \ell_{4k} u_{k3}}{u_{33}} = \frac{a_{43} - (\ell_{41} u_{13} + \ell_{42} u_{23})}{u_{33}} = \frac{2 - ((-3)(-1) + (2)(-1))}{-1} = -1. \end{aligned}$$

For the fourth row of U $(i = 4)$,

$$u_{44} = a_{44} - \sum_{k=1}^{3} \ell_{4k} u_{k4} = a_{44} - (\ell_{41} u_{14} + \ell_{42} u_{24} + \ell_{43} u_{34}) = -3 - ((-3)(3) + 2(3) + (-1)(3)) = -3.$$

For the fourth column of L ($j = 4$),

$$\ell_{44} = 1.$$

So, the LU decomposition according to the inner-product formulas is

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -3 & 2 & -1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Which is exactly the same decomposition that we got with Gaussian Elimination.

1.7.34. In this exercise you will show that performing an elementary row operation of type 1 is equivalent to left multiplication by a matrix of a special type. Suppose $\tilde{A}$ is obtained from A by adding m times the jth row to the ith row.

(a) Show that $\tilde{A} = MA$, where M is the triangular matrix obtained from the identity matrix by replacing the zero by an m in the (i, j) position. For example, when $i > j$, M has the form

$$M = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & m & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}.$$

Notice that this is the matrix obtained by applying the type 1 row operation directly to the identity matrix. We call M an *elementary matrix of type 1.*

(b) Show that $\det(M) = 1$ and $\det(\tilde{A}) = \det(A)$. Thus we see (again) that $\tilde{A}$ is nonsingular if and only if A is.

(c) Show that M^{-1} differs from M only in that it has $-m$ instead of m in the (i, j) position. M^{-1} is also an elementary matrix of type 1. To which elementary operation does it correspond?

Remark. Let $A \in \mathbb{R}^{n\times n}$, let e_i be the i^{th} standard basis vector in $\mathbb{R}^n$. That is, a vector of size n with a one in the i^{th} position, and zeros everywhere else. Then,

$$\text{col}_j(A) = Ae_j.$$

If we take the transpose of both sides,

$$(\text{col}_j(A))^T = e_j^T A^T,$$

which implies that

$$\left(\text{col}_j(A^T)\right)^T = e_j^T A, .$$

But, we know that $\text{row}_j(A) = (\ \text{col}_j(A^T))^T$, and $\text{col}_j(A) = (\text{row}_j(A^T))^T$. Thus,

$$\begin{aligned} &\left(\text{col}_j(A^T)\right)^T = e_j^T A \\ \implies &\ \text{row}_j(A) = e_j^T A. \end{aligned}$$

a.) In $\tilde{A}$, we have

$$\begin{aligned}\text{If } k \neq i,\ \text{row}_k(\tilde{A}) &= \ \text{row}_k(A),\\ \text{If } k = i,\ \text{row}_k(\tilde{A}) &= \ \text{row}_k(A) + m \cdot \text{row}_j(A).\end{aligned}$$

Let E_{ij} be the zero matrix except for a one at e_{ij}. Thus,

$$M = I + mE_{ij}.$$

Observe that $E_{ij} = e_i e_j^T$, so

$$M = I + me_i e_j^T.$$

From this fact, we have

$$MA = (I + me_i e_j^T)A = A + me_i(e_j^T A).$$

Recall that $e_j^T A$ is the j^{th} row of A, so

$$MA = A + me_i \cdot \text{row}_j(A).$$

Further observe that $e_i \cdot \text{row}_j(A)$ is a matrix of size $n \times n$, where the i^{th} row is $\text{row}_j(A)$, and all other rows are zero.

So, we see that

$$\begin{aligned}&\text{If } k \neq i,\ \text{row}_k(E_{ij}A) = 0, \ \ \text{so row}_k(MA) = \ \text{row}_k(A),\\ &\text{If } k = i,\ \text{row}_k(E_{ij}A) = \text{row}_j(A), \ \ \text{so row}_k(MA) = \text{row}_i(A) + m \cdot \text{row}_j(A).\end{aligned}$$

Thus, $\tilde{A} = MA$

b.) Since M is triangular, the determinant is

$$\det(M) = \prod_{i=1}^{n} m_{ii}.$$

But, $m_{ii} = 1$ for $i = 1, 2, ..., n$. Thus, $\det(M) = 1$. The determinant of $\tilde{A}$ is

$$\det(\tilde{A}) = \det(MA) = \det(M)\det(A) = 1 \det(A) = \det(A).$$

1.7.36. Suppose $\tilde{A}$ is obtained from A by multiplying the ith row by the nonzero constant c.

1. Find the form of the matrix M (an *elementary matrix of type 3*) such that $\tilde{A} = MA$.
2. Find M^{-1} and state its function as an elementary matrix.
3. Find $\det(M)$ and determine the relationship between $\det(\tilde{A})$ and $\det(A)$. Deduce that $\tilde{A}$ is nonsingular if and only if A is.

1.8.4. Let

$$A = \begin{bmatrix} 2 & 2 & -4 \\ 1 & 1 & 5 \\ 1 & 3 & 6 \end{bmatrix}, \qquad b = \begin{bmatrix} 10 \\ -2 \\ -5 \end{bmatrix}.$$

Use Gaussian elimination with partial pivoting (by hand) to find matrices L and U such that U is upper triangular, L is unit lower triangular with $|l_{ij}| \leqslant 1$ for all $i > j$, and $LU = \tilde{A}$, where $\tilde{A}$ can be obtained from A by making row interchanges. Use your LU decomposition to solve the system $Ax = b$.

We begin by initializing our permutation matrix P as the identity matrix I. That is, $P = I$. We set row one as the pivot row, and a_{11} as the pivot element. We look to the first column of A and see that a_{11} has the maximum absolute value, so no partial pivoting at this stage.

Let r_i denote the i^{th} row of A. We perform the operations

$$\begin{aligned} r_2 &\leftarrow r_2 - m_{21} r_1, \quad m_{21} = \frac{1}{2}, \\ r_3 &\leftarrow r_3 - m_{31} r_1, \quad m_{31} = \frac{1}{2}. \end{aligned}$$

Thus,

$$\left[\begin{array}{ccc|c} 2 & 2 & -4 & 10 \\ 1 & 1 & 5 & -2 \\ 1 & 3 & 6 & -5 \end{array}\right] \sim \left[\begin{array}{ccc|c} 2 & 2 & -4 & 10 \\ \boxed{\frac{1}{2}} & 0 & 7 & -7 \\ \boxed{\frac{1}{2}} & 2 & 8 & -10 \end{array}\right].$$

Note that the boxed numbers are entries of L. Next, row two is the pivot row, and a_{22} is the pivot element. Using partial pivoting, we swap rows two and three. We make the same swap in P. So,

$$\left[\begin{array}{ccc|c} 2 & 2 & -4 & 10 \\ \boxed{\frac{1}{2}} & 0 & 7 & -7 \\ \boxed{\frac{1}{2}} & 2 & 8 & -10 \end{array}\right] \sim \left[\begin{array}{ccc|c} 2 & 2 & -4 & 10 \\ \boxed{\frac{1}{2}} & 2 & 8 & -10 \\ \boxed{\frac{1}{2}} & 0 & 7 & -7 \end{array}\right], \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Next, we perform the operation

$$r_3 \leftarrow r_3 - 0 r_2.$$

So,

$$\left[\begin{array}{ccc|c} 2 & 2 & -4 & 10 \\ \boxed{\frac{1}{2}} & 2 & 8 & -10 \\ \boxed{\frac{1}{2}} & 0 & 7 & -7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 2 & 2 & -4 & 10 \\ \boxed{\frac{1}{2}} & 2 & 8 & -10 \\ \boxed{\frac{1}{2}} & \boxed{0} & 7 & -7 \end{array}\right].$$

Thus,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}, \; U = \begin{bmatrix} 2 & 2 & -4 \\ 0 & 2 & 8 \\ 0 & 0 & 7 \end{bmatrix}, \; y = \begin{pmatrix} 10 \\ -10 \\ -7 \end{pmatrix}, \; P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

and

$$Pb = \begin{pmatrix} 10 \\ -5 \\ -2 \end{pmatrix}.$$

We solve the system $Ax = b$, by splitting into two triangular systems

$$\begin{aligned} Ax = b \implies PAx = Pb \implies \tilde{A} = Pb \implies LUx = Pb \\ \implies \begin{cases} Ly & = Pb \\ Ux & = y \end{cases}. \end{aligned}$$

So, we solve $Ux = y$ with backward substitution

$$Ux = y \implies \begin{bmatrix} 2 & 2 & -4 \\ 0 & 2 & 8 \\ 0 & 0 & 7 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 10 \\ -10 \\ -7 \end{pmatrix}.$$

Which, implies that

$$\begin{aligned} 7x_3 = -7 &\implies x_3 = -1, \\ 2x_2 + 8x_3 = -10 &\implies x_2 = \frac{-10 - 2(-1)}{2} = -1, \\ 2x_1 + 2x_2 - 4x_3 = 10 &\implies x_1 = \frac{10 + 4(-1) - 2(-1)}{2} = 4. \end{aligned}$$

So,

$$x = \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix}.$$

We can verify this solution by checking that $Ax = b$, we have

$$\begin{bmatrix} 2 & 2 & -4 \\ 1 & 1 & 5 \\ 1 & 3 & 6 \end{bmatrix} \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 8 - 2 + 4 \\ 4 - 1 - 5 \\ 4 - 3 - 6 \end{pmatrix} = \begin{pmatrix} 10 \\ -2 \\ -5 \end{pmatrix}.$$

Therefore the solution is verified.

1.8.9. Let A be the matrix in Exercise 1.8.4. Determine matrices P, L, and U with the properties stated in Theorem 1.8.8, such that $A = P^T LU$

The matrices P, L, U are precisely the matrices obtained in the previous exercise, since A is unchanged. The matrices are

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \ L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}, \ U = \begin{bmatrix} 2 & 2 & -4 \\ 0 & 2 & 8 \\ 0 & 0 & 7 \end{bmatrix}.$$

We have

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & -4 \\ 0 & 2 & 8 \\ 0 & 0 & 7 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & -4 \\ \frac{1}{2}(2) & \frac{1}{2}(2) + 2 & \frac{1}{2}(-4) + 8 \\ \frac{1}{2}(2) & \frac{1}{2}(2) & \frac{1}{2}(-4) + 7 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & -4 \\ 1 & 3 & 6 \\ 1 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -4 \\ 1 & 1 & 5 \\ 1 & 3 & 6 \end{bmatrix} = A. \end{aligned}$$

1.8.12. Write an algorithm that implements Gaussian elimination with partial pivoting. Store L and U over A, and save a record of the row interchanges.

```
proc row_swap(A, i, k)
    if (i = k) return
    for ℓ = 1, ..., n
        tmp = a_{iℓ}
        a_{iℓ} = a_{kℓ}
        a_{kℓ} = tmp
    end
endproc

proc partial_pivot(A, P, K)
    max = |a_{kk}|, max_i = k
    for i = k + 1, ..., n
        if (|a_{ik}| > max)
            max = |a_{ik}|, max_i = i
        end
    end

    if (max = 0) set error flag, exit

    row_swap(A, k, max_i)
    row_swap(P, k, max_i)
endproc

proc gaussian(A)
    P = I
    for k = 1, ..., n
        partial_pivot(A, P, k)
        for i = k + 1, ..., n
            m = a_{ik}/a_{kk}
            for j = k, ..., n
                a_{ij} = a_{ij} − m · a_{kj}
            end
            a_{ik} = m
        end
    end

    return P
endproc
```

2.1.10. Prove that the 1-norm is a norm.

Remark. $\|\cdot\|$ is a norm if and only if the following properties are satisfied

1. $\|x\| \geqslant 0$ and $\|x\| = 0 \iff x = 0$
2. $\|\alpha x\| = |\alpha| \, \|x\|$
3. $\|x + y\| \leqslant \|x\| + \|y\|$ (triangle inequality)

The 1-norm for a vector $x \in \mathbb{R}^n$ is $\|x\|_1 = \sum_{i=1}^{n} |x_i|$.

1.) Suppose $\|x\|_1 = 0$, then

$$\begin{aligned} \sum_{i=1}^{n} |x_i| = 0 &\implies |x_1| + |x_2| + ... + |x_n| = 0 \\ &\implies x_1 = x_2 = ... = x_n = 0. \end{aligned}$$

Suppose $x = 0$, then $x_1 = x_2 = ... = x_n = 0$, and

$$\sum_{i=1}^{n} |x_i| = \sum_{i=1}^{n} 0 = 0 + 0 + ... + 0 = 0.$$

2.)

$$\begin{aligned} \|\alpha x\|_1 &= \sum_{i=1}^{n} |\alpha x_i| = |\alpha x_1| + |\alpha x_2| + ... + |\alpha x_n| \\ &= |\alpha| |x_1| + |\alpha| |x_2| + ... + |\alpha| |x_n| \\ &= |\alpha| (|x_1| + |x_2| + ... + |x_n|) = |\alpha| \|x\|_1 . \end{aligned}$$

3.) We can use the triangular inequality for absolute value,

$$\begin{aligned} \|x + y\|_1 &= \sum_{i=1}^{n} |x_i + y_i| \\ &= |x_1 + y_1| + |x_2 + y_2| + ... + |x_n + y_n| \\ &\leqslant |x_1| + |y_1| + |x_2| + |y_2| + ... + |x_n| + |y_n| \\ &= |x_1| + |x_2| + ... + |x_n| + |y_1| + |y_2| + ... + |y_n| \\ &= \|x\|_1 + \|y\|_1 . \end{aligned}$$

■

2.1.13. Prove that the ∞-norm is a norm.

The ∞-norm for a vector $x \in \mathbb{R}$ is $\|x\|_\infty = \max_{i=1}^{n} |x_i|$

1.) Suppose $\|x\| = 0$, then

$$\max_{i=1}^{n} |x_i| = 0$$

implies that $x_1 = x_2 = ... = x_n = 0$, since $|x_i| \geqslant 0$ for all i. Next, suppose that $x = 0$, then

$$\max_{i=1}^{n} 0 = \max\{0, 0, ..., 0\} = 0.$$

Thus $\|x\| = 0 \iff x = 0$ holds for the ∞-norm.

2.)

$$\|\alpha x\| = \max\{|\alpha x_1| , |\alpha x_2| , ..., |\alpha x_n|\} = |\alpha| \max\{|x_1| , |x_2| , ..., |x_n|\}.$$

This follows from the fact that if $|\alpha x_\ell| \geqslant |\alpha x_i|$, for all $i \neq \ell$, then

$$\begin{aligned} |\alpha| |x_\ell| &\geqslant |\alpha| |x_i| \\ \implies |x_\ell| &\geqslant |x_i| . \end{aligned}$$

3.) If

$$\|x+y\| = |x_\ell + y_\ell|\,.$$

Then, it follows from the triangle inequality for absolute value that

$$|x_\ell + y_\ell| \leqslant |x_\ell| + |y_\ell| = \max_{i=1}^{n} |x_i| + \max_{i=1}^{n} |y_i| = \|x\| + \|y\|\,.$$

Therefore, the ∞-norm is a norm. ■

2.1.17.

1. Let A be a positive definite matrix, and let R be its Cholesky factor, so that $A = R^T R$. Verify that for all $x \in \mathbb{R}^n$,
$$\|x\|_A = \|Rx\|_2.$$
2. Using the fact that the 2-norm is indeed a norm on $\mathbb{R}^n$, prove that the A-norm is a norm on $\mathbb{R}^n$.

2.2.6.

(a) Show that $\kappa(A) = \kappa(A^{-1})$

(b) Show that for any nonzero scalar c, $\kappa(cA) = \kappa(A)$

a.)

$$\kappa(A) = \|A\| \left\|A^{-1}\right\|,$$
$$\kappa(A^{-1}) = \left\|A^{-1}\right\| \left\|\left(A^{-1}\right)^{-1}\right\| = \left\|A^{-1}\right\| \|A\| = \|A\| \left\|A^{-1}\right\| = \kappa(A).$$

2.2.15. Let us take another look at the ill-conditioned matrices

$$A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix}$$

from Example 2.2.8. Notice that

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1999 \\ 1997 \end{bmatrix}. \tag{2.2.16}$$

If we use the ∞-norm to measure lengths, the magnification factor

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

is 1999, which equals $\|A\|_\infty$. Thus $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a vector that is magnified maximally by A.

Since the amount by which a vector is magnified depends only on its direction and not on its length, we say that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is in a *direction of maximum magnification* by A.

Equivalently we can say that $\begin{bmatrix} 1999 \\ 1997 \end{bmatrix}$ lies in a *direction of minimum magnification.*

Looking now at A^{-1}, we note that

$$A^{-1}\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1997\\-1999\end{bmatrix}.$$

The magnification factor $\frac{\|A^{-1}x\|_\infty}{\|x\|_\infty}$ is 1999, which equals $\|A^{-1}\|_\infty$, so $\begin{bmatrix}-1\\1\end{bmatrix}$ is in a direction of maximum magnification by A^{-1}. Equivalently

$$A\begin{bmatrix}1997\\-1999\end{bmatrix}=\begin{bmatrix}-1\\1\end{bmatrix}, \tag{2.2.17}$$

and $\begin{bmatrix}1997\\-1999\end{bmatrix}$ is in a direction of minimum magnification by A.

We will use the vectors in (2.2.16) and (2.2.17) to construct a spectacular example. Suppose we wish to solve the system

$$\begin{bmatrix}1000 & 999\\999 & 998\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}1999\\1997\end{bmatrix}. \tag{2.2.18}$$

That is, $Ax=b$, where $b=\begin{bmatrix}1999\\1997\end{bmatrix}$. Then by (2.2.16) the unique solution is

$$x=\begin{bmatrix}1\\1\end{bmatrix}.$$

Now suppose that we solve instead the slightly perturbed system

$$\begin{bmatrix}1000 & 999\\999 & 998\end{bmatrix}\begin{bmatrix}\hat{x}_1\\\hat{x}_2\end{bmatrix}=\begin{bmatrix}1998.99\\1997.01\end{bmatrix}. \tag{2.2.19}$$

This is $\hat{A}x=b+\delta b$, where $\delta b=\begin{bmatrix}-0.01\\0.01\end{bmatrix}=0.01\begin{bmatrix}-1\\1\end{bmatrix}$, which is in a direction of maximum magnification by A^{-1}. By (2.2.17), $A\delta x=\delta b$, where $\delta x=\begin{bmatrix}19.97\\-19.99\end{bmatrix}$. Therefore $\hat{x}=x+\delta x=\begin{bmatrix}20.97\\-18.99\end{bmatrix}$. Thus the nearly identical problems (2.2.18) and (2.2.19) have very different solutions.

Repeat the proof of Theorem 2.3.3.

Remark. Theorem 2.3.3. Let A be nonsingular, let $b\neq 0$, and let x and $\hat{x}=x+\delta x$ be solutions of $Ax=b$ and $(A+\delta A)\hat{x}=b$, respectively. Then,

$$\frac{\|\delta x\|}{\|\hat{x}\|}\leqslant\kappa(A)\frac{\|\delta A\|}{\|A\|}. \tag{2.3.4}$$

Proof. Rewriting the equation $(A+\delta A)\hat{x}=b$ as $Ax+A\delta x+\delta A\hat{x}=b$, using the equation $Ax=b$, and reorganizing the resulting equation, we obtain

$$\delta x=-A^{-1}\delta A\hat{x}.$$

Thus

$$\|\delta x\|\leqslant\|A^{-1}\|\,\|\delta A\|\,\|\hat{x}\|. \tag{2.3.5}$$

Dividing through by $\|\hat{x}\|$ and using the definition $\kappa(A)=\|A\|\,\|A^{-1}\|$, we obtain the desired result.