

**Exercise 2.6.20** Define  $T: \mathbb{R}^n \rightarrow \mathbb{R}$  by  $T(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$ . Show that  $T$  is a linear transformation and find its matrix.

**Exercise 2.6.21** Given  $c$  in  $\mathbb{R}$ , define  $T_c: \mathbb{R}^n \rightarrow \mathbb{R}$  by  $T_c(\mathbf{x}) = c\mathbf{x}$  for all  $\mathbf{x}$  in  $\mathbb{R}^n$ . Show that  $T_c$  is a linear transformation and find its matrix.

**Exercise 2.6.22** Given vectors  $\mathbf{w}$  and  $\mathbf{x}$  in  $\mathbb{R}^n$ , denote their dot product by  $\mathbf{w} \cdot \mathbf{x}$ .

- Given  $\mathbf{w}$  in  $\mathbb{R}^n$ , define  $T_{\mathbf{w}}: \mathbb{R}^n \rightarrow \mathbb{R}$  by  $T_{\mathbf{w}}(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x}$  for all  $\mathbf{x}$  in  $\mathbb{R}^n$ . Show that  $T_{\mathbf{w}}$  is a linear transformation.
- Show that *every* linear transformation  $T: \mathbb{R}^n \rightarrow \mathbb{R}$  is given as in (a); that is  $T = T_{\mathbf{w}}$  for some  $\mathbf{w}$  in  $\mathbb{R}^n$ .

**Exercise 2.6.23** If  $\mathbf{x} \neq \mathbf{0}$  and  $\mathbf{y}$  are vectors in  $\mathbb{R}^n$ , show that there is a linear transformation  $T$ :

$\mathbb{R}^n \rightarrow \mathbb{R}^n$  such that  $T(\mathbf{x}) = \mathbf{y}$ . [Hint: By Definition 2.5, find a matrix  $A$  such that  $A\mathbf{x} = \mathbf{y}$ .]

**Exercise 2.6.24** Let  $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k$  be two linear transformations. Show directly that  $S \circ T$  is linear. That is:

- Show that  $(S \circ T)(\mathbf{x} + \mathbf{y}) = (S \circ T)\mathbf{x} + (S \circ T)\mathbf{y}$  for all  $\mathbf{x}, \mathbf{y}$  in  $\mathbb{R}^n$ .
- Show that  $(S \circ T)(a\mathbf{x}) = a[(S \circ T)\mathbf{x}]$  for all  $\mathbf{x}$  in  $\mathbb{R}^n$  and all  $a$  in  $\mathbb{R}$ .

**Exercise 2.6.25** Let  $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k \xrightarrow{R} \mathbb{R}^k$  be linear transformations. Show that  $R \circ (S \circ T) = (R \circ S) \circ T$  by showing directly that  $[R \circ (S \circ T)](\mathbf{x}) = [(R \circ S) \circ T](\mathbf{x})$  holds for each vector  $\mathbf{x}$  in  $\mathbb{R}^n$ .

## 2.7 LU-Factorization<sup>15</sup>

The solution to a system  $A\mathbf{x} = \mathbf{b}$  of linear equations can be solved quickly if  $A$  can be factored as  $A = LU$  where  $L$  and  $U$  are of a particularly nice form. In this section we show that gaussian elimination can be used to find such factorizations.

### Triangular Matrices

As for square matrices, if  $A = [a_{ij}]$  is an  $m \times n$  matrix, the elements  $a_{11}, a_{22}, a_{33}, \dots$  form the **main diagonal** of  $A$ . Then  $A$  is called **upper triangular** if every entry below and to the left of the main diagonal is zero. Every row-echelon matrix is upper triangular, as are the matrices

$$\begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & -3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 & 0 & 5 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

By analogy, a matrix  $A$  is called **lower triangular** if its transpose is upper triangular, that is if each entry above and to the right of the main diagonal is zero. A matrix is called **triangular** if it is upper or lower triangular.

<sup>15</sup>This section is not used later and so may be omitted with no loss of continuity.

**Example 2.7.1**

Solve the system

$$\begin{aligned}x_1 + 2x_2 - 3x_3 - x_4 + 5x_5 &= 3 \\5x_3 + x_4 + x_5 &= 8 \\2x_5 &= 6\end{aligned}$$

where the coefficient matrix is upper triangular.

**Solution.** As in gaussian elimination, let the “non-leading” variables be parameters:  $x_2 = s$  and  $x_4 = t$ . Then solve for  $x_5$ ,  $x_3$ , and  $x_1$  in that order as follows. The last equation gives

$$x_5 = \frac{6}{2} = 3$$

Substitution into the second last equation gives

$$x_3 = 1 - \frac{1}{5}t$$

Finally, substitution of both  $x_5$  and  $x_3$  into the first equation gives

$$x_1 = -9 - 2s + \frac{2}{5}t.$$

The method used in Example 2.7.1 is called **back substitution** because later variables are substituted into earlier equations. It works because the coefficient matrix is upper triangular. Similarly, if the coefficient matrix is lower triangular the system can be solved by **forward substitution** where earlier variables are substituted into later equations. As observed in Section 1.2, these procedures are more efficient than gaussian elimination.

Now consider a system  $A\mathbf{x} = \mathbf{b}$  where  $A$  can be factored as  $A = LU$  where  $L$  is lower triangular and  $U$  is upper triangular. Then the system  $A\mathbf{x} = \mathbf{b}$  can be solved in two stages as follows:

1. First solve  $L\mathbf{y} = \mathbf{b}$  for  $\mathbf{y}$  by forward substitution.
2. Then solve  $U\mathbf{x} = \mathbf{y}$  for  $\mathbf{x}$  by back substitution.

Then  $\mathbf{x}$  is a solution to  $A\mathbf{x} = \mathbf{b}$  because  $A\mathbf{x} = LU\mathbf{x} = L\mathbf{y} = \mathbf{b}$ . Moreover, every solution  $\mathbf{x}$  arises this way (take  $\mathbf{y} = U\mathbf{x}$ ). Furthermore the method adapts easily for use in a computer.

This focuses attention on efficiently obtaining such factorizations  $A = LU$ . The following result will be needed; the proof is straightforward and is left as Exercises 7 and 8.

**Lemma 2.7.1**

Let  $A$  and  $B$  denote matrices.

1. If  $A$  and  $B$  are both lower (upper) triangular, the same is true of  $AB$ .
2. If  $A$  is  $n \times n$  and lower (upper) triangular, then  $A$  is invertible if and only if every main diagonal entry is nonzero. In this case  $A^{-1}$  is also lower (upper) triangular.

## LU-Factorization

Let  $A$  be an  $m \times n$  matrix. Then  $A$  can be carried to a row-echelon matrix  $U$  (that is, upper triangular). As in Section 2.5, the reduction is

$$A \rightarrow E_1 A \rightarrow E_2 E_1 A \rightarrow E_3 E_2 E_1 A \rightarrow \cdots \rightarrow E_k E_{k-1} \cdots E_2 E_1 A = U$$

where  $E_1, E_2, \dots, E_k$  are elementary matrices corresponding to the row operations used. Hence

$$A = LU$$

where  $L = (E_k E_{k-1} \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_{k-1}^{-1} E_k^{-1}$ . If we do not insist that  $U$  is reduced then, except for row interchanges, none of these row operations involve adding a row to a row *above* it. Thus, if no row interchanges are used, all the  $E_i$  are *lower* triangular, and so  $L$  is lower triangular (and invertible) by Lemma 2.7.1. This proves the following theorem. For convenience, let us say that  $A$  can be **lower reduced** if it can be carried to row-echelon form using no row interchanges.

### Theorem 2.7.1

If  $A$  can be lower reduced to a row-echelon matrix  $U$ , then

$$A = LU$$

where  $L$  is lower triangular and invertible and  $U$  is upper triangular and row-echelon.

### Definition 2.14

A factorization  $A = LU$  as in Theorem 2.7.1 is called an **LU-factorization** of  $A$ .

Such a factorization may not exist (Exercise 4) because  $A$  cannot be carried to row-echelon form using no row interchange. A procedure for dealing with this situation will be outlined later. However, if an LU-factorization  $A = LU$  does exist, then the gaussian algorithm gives  $U$  and also leads to a procedure for finding  $L$ .

Example 2.7.2 provides an illustration. For convenience, the first nonzero column from the left in a matrix  $A$  is called the **leading column** of  $A$ .

### Example 2.7.2

Find an LU-factorization of  $A = \begin{bmatrix} 0 & 2 & -6 & -2 & 4 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{bmatrix}$ .

**Solution.** We lower reduce  $A$  to row-echelon form as follows:

$$A = \begin{bmatrix} 0 & 2 & -6 & -2 & 4 \\ 0 & -1 & 3 & 3 & 2 \\ 0 & -1 & 3 & 7 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 6 & 12 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = U$$

The circled columns are determined as follows: The first is the leading column of  $A$ , and is used (by lower reduction) to create the first leading 1 and create zeros below it. This completes the work on row 1, and we repeat the procedure on the matrix consisting of the remaining rows. Thus the second circled column is the leading column of this smaller matrix, which we use to create the second leading 1 and the zeros below it. As the remaining row is zero here, we are finished. Then  $A = LU$  where

$$L = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 2 & 0 \\ -1 & 6 & 1 \end{bmatrix}.$$

This matrix  $L$  is obtained from  $I_3$  by replacing the bottom of the first two columns by the circled columns in the reduction. Note that the rank of  $A$  is 2 here, and this is the number of circled columns.

The calculation in Example 2.7.2 works in general. There is no need to calculate the elementary matrices  $E_i$ , and the method is suitable for use in a computer because the circled columns can be stored in memory as they are created. The procedure can be formally stated as follows:

### LU-Algorithm

*Let  $A$  be an  $m \times n$  matrix of rank  $r$ , and suppose that  $A$  can be lower reduced to a row-echelon matrix  $U$ . Then  $A = LU$  where the lower triangular, invertible matrix  $L$  is constructed as follows:*

1. *If  $A = 0$ , take  $L = I_m$  and  $U = 0$ .*
2. *If  $A \neq 0$ , write  $A_1 = A$  and let  $\mathbf{c}_1$  be the leading column of  $A_1$ . Use  $\mathbf{c}_1$  to create the first leading 1 and create zeros below it (using lower reduction). When this is completed, let  $A_2$  denote the matrix consisting of rows 2 to  $m$  of the matrix just created.*
3. *If  $A_2 \neq 0$ , let  $\mathbf{c}_2$  be the leading column of  $A_2$  and repeat Step 2 on  $A_2$  to create  $A_3$ .*
4. *Continue in this way until  $U$  is reached, where all rows below the last leading 1 consist of zeros. This will happen after  $r$  steps.*
5. *Create  $L$  by placing  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r$  at the bottom of the first  $r$  columns of  $I_m$ .*

A proof of the LU-algorithm is given at the end of this section.

LU-factorization is particularly important if, as often happens in business and industry, a series of equations  $A\mathbf{x} = B_1, A\mathbf{x} = B_2, \dots, A\mathbf{x} = B_k$ , must be solved, each with the same coefficient matrix  $A$ . It is very efficient to solve the first system by gaussian elimination, simultaneously creating an LU-factorization of  $A$ , and then using the factorization to solve the remaining systems by forward and back substitution.

### Example 2.7.3

Find an LU-factorization for  $A = \begin{bmatrix} 5 & -5 & 10 & 0 & 5 \\ -3 & 3 & 2 & 2 & 1 \\ -2 & 2 & 0 & -1 & 0 \\ 1 & -1 & 10 & 2 & 5 \end{bmatrix}.$

**Solution.** The reduction to row-echelon form is

$$\begin{aligned}
 \begin{bmatrix} 5 & -5 & 10 & 0 & 5 \\ -3 & 3 & 2 & 2 & 1 \\ -2 & 2 & 0 & -1 & 0 \\ 1 & -1 & 10 & 2 & 5 \end{bmatrix} &\rightarrow \begin{bmatrix} 1 & -1 & 2 & 0 & 1 \\ 0 & 0 & 8 & 2 & 4 \\ 0 & 0 & 4 & -1 & 2 \\ 0 & 0 & 8 & 2 & 4 \end{bmatrix} \\
 &\rightarrow \begin{bmatrix} 1 & -1 & 2 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
 &\rightarrow \begin{bmatrix} 1 & -1 & 2 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = U
 \end{aligned}$$

If  $U$  denotes this row-echelon matrix, then  $A = LU$ , where

$$L = \begin{bmatrix} 5 & 0 & 0 & 0 \\ -3 & 8 & 0 & 0 \\ -2 & 4 & -2 & 0 \\ 1 & 8 & 0 & 1 \end{bmatrix}$$

The next example deals with a case where no row of zeros is present in  $U$  (in fact,  $A$  is invertible).

#### Example 2.7.4

Find an LU-factorization for  $A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix}$ .

**Solution.** The reduction to row-echelon form is

$$\begin{bmatrix} 2 & 4 & 2 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = U$$

Hence  $A = LU$  where  $L = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 2 & 5 \end{bmatrix}$ .

There are matrices (for example  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ ) that have no LU-factorization and so require at least one row interchange when being carried to row-echelon form via the gaussian algorithm. However, it turns out that, if all the row interchanges encountered in the algorithm are carried out first, the resulting matrix requires no interchanges and so has an LU-factorization. Here is the precise result.

### Theorem 2.7.2

Suppose an  $m \times n$  matrix  $A$  is carried to a row-echelon matrix  $U$  via the gaussian algorithm. Let  $P_1, P_2, \dots, P_s$  be the elementary matrices corresponding (in order) to the row interchanges used, and write  $P = P_s \cdots P_2 P_1$ . (If no interchanges are used take  $P = I_m$ .) Then:

1.  $PA$  is the matrix obtained from  $A$  by doing these interchanges (in order) to  $A$ .
2.  $PA$  has an LU-factorization.

The proof is given at the end of this section.

A matrix  $P$  that is the product of elementary matrices corresponding to row interchanges is called a **permutation matrix**. Such a matrix is obtained from the identity matrix by arranging the rows in a different order, so it has exactly one 1 in each row and each column, and has zeros elsewhere. We regard the identity matrix as a permutation matrix. The elementary permutation matrices are those obtained from  $I$  by a single row interchange, and every permutation matrix is a product of elementary ones.

### Example 2.7.5

If  $A = \begin{bmatrix} 0 & 0 & -1 & 2 \\ -1 & -1 & 1 & 2 \\ 2 & 1 & -3 & 6 \\ 0 & 1 & -1 & 4 \end{bmatrix}$ , find a permutation matrix  $P$  such that  $PA$  has an LU-factorization, and then find the factorization.

**Solution.** Apply the gaussian algorithm to  $A$ :

$$\begin{aligned}
 A \xrightarrow{*} \begin{bmatrix} -1 & -1 & 1 & 2 \\ 0 & 0 & -1 & 2 \\ 2 & 1 & -3 & 6 \\ 0 & 1 & -1 & 4 \end{bmatrix} &\rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 0 & -1 & 2 \\ 0 & -1 & -1 & 10 \\ 0 & 1 & -1 & 4 \end{bmatrix} \xrightarrow{*} \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & -1 & -1 & 10 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & -1 & 4 \end{bmatrix} \\
 &\rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 1 & 1 & -10 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -2 & 14 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 1 & 1 & -10 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 10 \end{bmatrix}
 \end{aligned}$$

Two row interchanges were needed (marked with \*), first rows 1 and 2 and then rows 2 and 3. Hence, as in Theorem 2.7.2,

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

If we do these interchanges (in order) to  $A$ , the result is  $PA$ . Now apply the LU-algorithm to  $PA$ :

$$PA = \begin{bmatrix} -1 & -1 & 1 & 2 \\ 2 & 1 & -3 & 6 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & -1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & -1 & -1 & 10 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & -1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 1 & 1 & -10 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -2 & 14 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 1 & 1 & -10 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 1 & 1 & -10 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix} = U$$

Hence,  $PA = LU$ , where  $L = \begin{bmatrix} 1 & 1 & -1 & -2 \\ 0 & 1 & 1 & -10 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$  and  $U = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -2 & 10 \end{bmatrix}$ .

Theorem 2.7.2 provides an important general factorization theorem for matrices. If  $A$  is any  $m \times n$  matrix, it asserts that there exists a permutation matrix  $P$  and an LU-factorization  $PA = LU$ . Moreover, it shows that either  $P = I$  or  $P = P_s \cdots P_2 P_1$ , where  $P_1, P_2, \dots, P_s$  are the elementary permutation matrices arising in the reduction of  $A$  to row-echelon form. Now observe that  $P_i^{-1} = P_i$  for each  $i$  (they are elementary row interchanges). Thus,  $P^{-1} = P_1 P_2 \cdots P_s$ , so the matrix  $A$  can be factored as

$$A = P^{-1}LU$$

where  $P^{-1}$  is a permutation matrix,  $L$  is lower triangular and invertible, and  $U$  is a row-echelon matrix. This is called a **PLU-factorization** of  $A$ .

The LU-factorization in Theorem 2.7.1 is not unique. For example,

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

However, it is necessary here that the row-echelon matrix has a row of zeros. Recall that the rank of a matrix  $A$  is the number of nonzero rows in any row-echelon matrix  $U$  to which  $A$  can be carried by row operations. Thus, if  $A$  is  $m \times n$ , the matrix  $U$  has no row of zeros if and only if  $A$  has rank  $m$ .

### Theorem 2.7.3

Let  $A$  be an  $m \times n$  matrix that has an LU-factorization

$$A = LU$$

If  $A$  has rank  $m$  (that is,  $U$  has no row of zeros), then  $L$  and  $U$  are uniquely determined by  $A$ .

**Proof.** Suppose  $A = MV$  is another LU-factorization of  $A$ , so  $M$  is lower triangular and invertible and  $V$  is row-echelon. Hence  $LU = MV$ , and we must show that  $L = M$  and  $U = V$ . We write  $N = M^{-1}L$ . Then  $N$  is lower triangular and invertible (Lemma 2.7.1) and  $NU = V$ , so it suffices to prove that  $N = I$ . If  $N$  is  $m \times$

$m$ , we use induction on  $m$ . The case  $m = 1$  is left to the reader. If  $m > 1$ , observe first that column 1 of  $V$  is  $N$  times column 1 of  $U$ . Thus if either column is zero, so is the other ( $N$  is invertible). Hence, we can assume (by deleting zero columns) that the  $(1, 1)$ -entry is 1 in both  $U$  and  $V$ .

Now we write  $N = \begin{bmatrix} a & 0 \\ X & N_1 \end{bmatrix}$ ,  $U = \begin{bmatrix} 1 & Y \\ 0 & U_1 \end{bmatrix}$ , and  $V = \begin{bmatrix} 1 & Z \\ 0 & V_1 \end{bmatrix}$  in block form. Then  $NU = V$  becomes  $\begin{bmatrix} a & aY \\ X & XY + N_1U_1 \end{bmatrix} = \begin{bmatrix} 1 & Z \\ 0 & V_1 \end{bmatrix}$ . Hence  $a = 1$ ,  $Y = Z$ ,  $X = 0$ , and  $N_1U_1 = V_1$ . But  $N_1U_1 = V_1$  implies  $N_1 = I$  by induction, whence  $N = I$ .  $\square$

If  $A$  is an  $m \times m$  invertible matrix, then  $A$  has rank  $m$  by Theorem 2.4.5. Hence, we get the following important special case of Theorem 2.7.3.

### Corollary 2.7.1

*If an invertible matrix  $A$  has an LU-factorization  $A = LU$ , then  $L$  and  $U$  are uniquely determined by  $A$ .*

Of course, in this case  $U$  is an upper triangular matrix with 1s along the main diagonal.

## Proofs of Theorems

### Proof of the LU-Algorithm

If  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r$  are columns of lengths  $m, m-1, \dots, m-r+1$ , respectively, write  $L^{(m)}(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r)$  for the lower triangular  $m \times m$  matrix obtained from  $I_m$  by placing  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r$  at the bottom of the first  $r$  columns of  $I_m$ .

Proceed by induction on  $n$ . If  $A = 0$  or  $n = 1$ , it is left to the reader. If  $n > 1$ , let  $\mathbf{c}_1$  denote the leading column of  $A$  and let  $\mathbf{k}_1$  denote the first column of the  $m \times m$  identity matrix. There exist elementary matrices  $E_1, \dots, E_k$  such that, in block form,

$$(E_k \cdots E_2 E_1)A = \left[ 0 \mid \mathbf{k}_1 \mid \begin{array}{c} X_1 \\ A_1 \end{array} \right] \quad \text{where } (E_k \cdots E_2 E_1)\mathbf{c}_1 = \mathbf{k}_1.$$

Moreover, each  $E_j$  can be taken to be lower triangular (by assumption). Write

$$G = (E_k \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$

Then  $G$  is lower triangular, and  $G\mathbf{k}_1 = \mathbf{c}_1$ . Also, each  $E_j$  (and so each  $E_j^{-1}$ ) is the result of either multiplying row 1 of  $I_m$  by a constant or adding a multiple of row 1 to another row. Hence,

$$G = (E_1^{-1} E_2^{-1} \cdots E_k^{-1})I_m = \left[ \mathbf{c}_1 \mid \begin{array}{c} 0 \\ I_{m-1} \end{array} \right]$$

in block form. Now, by induction, let  $A_1 = L_1 U_1$  be an LU-factorization of  $A_1$ , where  $L_1 = L^{(m-1)}[\mathbf{c}_2, \dots, \mathbf{c}_r]$  and  $U_1$  is row-echelon. Then block multiplication gives

$$G^{-1}A = \left[ 0 \mid \mathbf{k}_1 \mid \begin{array}{c} X_1 \\ L_1 U_1 \end{array} \right] = \left[ \begin{array}{c|c} 1 & 0 \\ 0 & L_1 \end{array} \right] \left[ \begin{array}{c|c|c} 0 & 1 & X_1 \\ 0 & 0 & U_1 \end{array} \right]$$



Hence  $A = LU$ , where  $U = \left[ \begin{array}{c|c|c} 0 & 1 & X_1 \\ \hline 0 & 0 & U_1 \end{array} \right]$  is row-echelon and

$$L = \left[ \begin{array}{c|c} \mathbf{c}_1 & 0 \\ \hline & I_{m-1} \end{array} \right] \left[ \begin{array}{c|c} 1 & 0 \\ \hline 0 & L_1 \end{array} \right] = \left[ \begin{array}{c|c} \mathbf{c}_1 & 0 \\ \hline & L \end{array} \right] = L^{(m)}[\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r].$$

This completes the proof. □

### Proof of Theorem 2.7.2

Let  $A$  be a nonzero  $m \times n$  matrix and let  $\mathbf{k}_j$  denote column  $j$  of  $I_m$ . There is a permutation matrix  $P_1$  (where either  $P_1$  is elementary or  $P_1 = I_m$ ) such that the first nonzero column  $\mathbf{c}_1$  of  $P_1 A$  has a nonzero entry on top. Hence, as in the LU-algorithm,

$$L^{(m)}[\mathbf{c}_1]^{-1} \cdot P_1 \cdot A = \left[ \begin{array}{c|c|c} 0 & 1 & X_1 \\ \hline 0 & 0 & A_1 \end{array} \right]$$

in block form. Then let  $P_2$  be a permutation matrix (either elementary or  $I_m$ ) such that

$$P_2 \cdot L^{(m)}[\mathbf{c}_1]^{-1} \cdot P_1 \cdot A = \left[ \begin{array}{c|c|c} 0 & 1 & X_1 \\ \hline 0 & 0 & A'_1 \end{array} \right]$$

and the first nonzero column  $\mathbf{c}_2$  of  $A'_1$  has a nonzero entry on top. Thus,

$$L^{(m)}[\mathbf{k}_1, \mathbf{c}_1]^{-1} \cdot P_2 \cdot L^{(m)}[\mathbf{c}_1]^{-1} \cdot P_1 \cdot A = \left[ \begin{array}{c|c|c} 0 & 1 & X_1 \\ \hline 0 & 0 & \begin{array}{c|c|c} 0 & 1 & X_2 \\ \hline 0 & 0 & A_2 \end{array} \end{array} \right]$$

in block form. Continue to obtain elementary permutation matrices  $P_1, P_2, \dots, P_r$  and columns  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r$  of lengths  $m, m-1, \dots$ , such that

$$(L_r P_r L_{r-1} P_{r-1} \cdots L_2 P_2 L_1 P_1) A = U$$

where  $U$  is a row-echelon matrix and  $L_j = L^{(m)}[\mathbf{k}_1, \dots, \mathbf{k}_{j-1}, \mathbf{c}_j]^{-1}$  for each  $j$ , where the notation means the first  $j-1$  columns are those of  $I_m$ . It is not hard to verify that each  $L_j$  has the form  $L_j = L^{(m)}[\mathbf{k}_1, \dots, \mathbf{k}_{j-1}, \mathbf{c}'_j]$  where  $\mathbf{c}'_j$  is a column of length  $m-j+1$ . We now claim that each permutation matrix  $P_k$  can be “moved past” each matrix  $L_j$  to the right of it, in the sense that

$$P_k L_j = L'_j P_k$$

where  $L'_j = L^{(m)}[\mathbf{k}_1, \dots, \mathbf{k}_{j-1}, \mathbf{c}''_j]$  for some column  $\mathbf{c}''_j$  of length  $m-j+1$ . Given that this is true, we obtain a factorization of the form

$$(L_r L'_{r-1} \cdots L'_2 L'_1) (P_r P_{r-1} \cdots P_2 P_1) A = U$$

If we write  $P = P_r P_{r-1} \cdots P_2 P_1$ , this shows that  $PA$  has an LU-factorization because  $L_r L'_{r-1} \cdots L'_2 L'_1$  is lower triangular and invertible. All that remains is to prove the following rather technical result. □

**Lemma 2.7.2**

Let  $P_k$  result from interchanging row  $k$  of  $I_m$  with a row below it. If  $j < k$ , let  $\mathbf{c}_j$  be a column of length  $m - j + 1$ . Then there is another column  $\mathbf{c}'_j$  of length  $m - j + 1$  such that

$$P_k \cdot L^{(m)} [\mathbf{k}_1, \dots, \mathbf{k}_{j-1}, \mathbf{c}_j] = L^{(m)} [\mathbf{k}_1, \dots, \mathbf{k}_{j-1}, \mathbf{c}'_j] \cdot P_k$$

The proof is left as Exercise 11.

## Exercises for 2.7

**Exercise 2.7.1** Find an LU-factorization of the following matrices.

a. 
$$\begin{bmatrix} 2 & 6 & -2 & 0 & 2 \\ 3 & 9 & -3 & 3 & 1 \\ -1 & -3 & 1 & -3 & 1 \end{bmatrix}$$

b. 
$$\begin{bmatrix} 2 & 4 & 2 \\ 1 & -1 & 3 \\ -1 & 7 & -7 \end{bmatrix}$$

c. 
$$\begin{bmatrix} 2 & 6 & -2 & 0 & 2 \\ 1 & 5 & -1 & 2 & 5 \\ 3 & 7 & -3 & -2 & 5 \\ -1 & -1 & 1 & 2 & 3 \end{bmatrix}$$

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

e. 
$$\begin{bmatrix} 2 & 2 & 4 & 6 & 0 & 2 \\ 1 & -1 & 2 & 1 & 3 & 1 \\ -2 & 2 & -4 & -1 & 1 & 6 \\ 0 & 2 & 0 & 3 & 4 & 8 \\ -2 & 4 & -4 & 1 & -2 & 6 \end{bmatrix}$$

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

a. 
$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & 4 \\ 3 & 5 & 1 \end{bmatrix}$$

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

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

d. 
$$\begin{bmatrix} -1 & -2 & 3 & 0 \\ 2 & 4 & -6 & 5 \\ 1 & 1 & -1 & 3 \\ 2 & 5 & -10 & 1 \end{bmatrix}$$

**Exercise 2.7.3** In each case use the given LU-decomposition of  $A$  to solve the system  $A\mathbf{x} = \mathbf{b}$  by finding  $\mathbf{y}$  such that  $L\mathbf{y} = \mathbf{b}$ , and then  $\mathbf{x}$  such that  $U\mathbf{x} = \mathbf{y}$ :

a. 
$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix};$$

$$\mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

**Exercise 2.7.2** Find a permutation matrix  $P$  and an LU-factorization of  $PA$  if  $A$  is:

$$\text{b. } A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$$\mathbf{b} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{c. } A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & 1 & -4 \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix};$$

$$\mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$$

$$\text{d. } A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -1 & 1 & 2 & 0 \\ 3 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$$\mathbf{b} = \begin{bmatrix} 4 \\ -6 \\ 4 \\ 5 \end{bmatrix}$$

**Exercise 2.7.4** Show that  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = LU$  is impossible where  $L$  is lower triangular and  $U$  is upper triangular.

**Exercise 2.7.5** Show that we can accomplish any row interchange by using only row operations of other types.

### Exercise 2.7.6

- a. Let  $L$  and  $L_1$  be invertible lower triangular matrices, and let  $U$  and  $U_1$  be invertible upper triangular matrices. Show that  $LU = L_1U_1$  if

and only if there exists an invertible diagonal matrix  $D$  such that  $L_1 = LD$  and  $U_1 = D^{-1}U$ . [Hint: Scrutinize  $L^{-1}L_1 = UU_1^{-1}$ .]

- b. Use part (a) to prove Theorem 2.7.3 in the case that  $A$  is invertible.

**Exercise 2.7.7** Prove Lemma 2.7.1(1). [Hint: Use block multiplication and induction.]

**Exercise 2.7.8** Prove Lemma 2.7.1(2). [Hint: Use block multiplication and induction.]

**Exercise 2.7.9** A triangular matrix is called **unit triangular** if it is square and every main diagonal element is a 1.

- a. If  $A$  can be carried by the gaussian algorithm to row-echelon form using no row interchanges, show that  $A = LU$  where  $L$  is unit lower triangular and  $U$  is upper triangular.
- b. Show that the factorization in (a) is unique.

**Exercise 2.7.10** Let  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r$  be columns of lengths  $m, m-1, \dots, m-r+1$ . If  $\mathbf{k}_j$  denotes column  $j$  of  $I_m$ , show that  $L^{(m)}[\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r] = L^{(m)}[\mathbf{c}_1] L^{(m)}[\mathbf{k}_1, \mathbf{c}_2] L^{(m)}[\mathbf{k}_1, \mathbf{k}_2, \mathbf{c}_3] \cdots L^{(m)}[\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_{r-1}, \mathbf{c}_r]$ . The notation is as in the proof of Theorem 2.7.2. [Hint: Use induction on  $m$  and block multiplication.]

**Exercise 2.7.11** Prove Lemma 2.7.2. [Hint:  $P_k^{-1} = P_k$ . Write  $P_k = \begin{bmatrix} I_k & 0 \\ 0 & P_0 \end{bmatrix}$  in block form where  $P_0$  is an  $(m-k) \times (m-k)$  permutation matrix.]