Chapter 7

LU Factorization

If we want to solve a system

$$Ax = \boldsymbol{b},$$

Gaussian elimination is probably the best method. There may be some situations where we want other methods, though. For example, suppose we have several systems

$$\begin{aligned} A\boldsymbol{x} &= \boldsymbol{b_1} \\ A\boldsymbol{x} &= \boldsymbol{b_2} \\ &\vdots \\ A\boldsymbol{x} &= \boldsymbol{b_n} \end{aligned}$$

with the same matrix of coefficients. In this case, it might make more sense to find A^{-1} and solve the systems by $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. The flaw in this otherwise brilliant plan is that A may not be invertible. It may be useful to develop other methods.

In some cases, we may be able to factor the matrix A into simpler matrices, which make the systems easier to solve. Let's look at special types of matrices.

Recall, a matrix with the same number of rows as columns is called a ***square matrix***.

Definition. If A is a square matrix, the elements a_{11}, a_{22}, $\ldots$ are called the ***diagonal elements***.

$$\begin{pmatrix} \boldsymbol{a_{11}} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & \boldsymbol{a_{22}} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & \boldsymbol{a_{33}} & \ldots & a_{3n} \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{3n} & \ldots & \boldsymbol{a_{nn}} \end{pmatrix}$$

Definition. A ***diagonal matrix*** is a matrix in which all elements *not* on the diagonal are 0; i.e., a matrix of the form

$$\begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & 0 & \cdots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & d_{nn} \end{pmatrix}$$

The diagonal elements of a diagonal matrix could be zero or non-zero.

Example. The matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are all diagonal matrices. ☑

Related to the diagonal matrices are the triangular matrices.

Definition. An ***upper triangular matrix*** is a matrix in which all elements *below* the diagonal are 0; i.e., a matrix of the form

$$\begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1n} \\ 0 & t_{22} & t_{23} & \cdots & t_{2n} \\ 0 & 0 & t_{33} & \cdots & t_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & t_{nn} \end{pmatrix}$$

Example. The matrices

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 5 \\ 0 & -1 & 4 \\ 0 & 0 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are all upper triangular matrices. ☑

Definition. A ***lower triangular matrix*** is a matrix in which all elements *above* the diagonal are 0; i.e., a matrix of the form

$$\begin{pmatrix} t_{11} & 0 & 0 & \cdots & 0 \\ t_{21} & t_{22} & 0 & \cdots & 0 \\ t_{31} & t_{32} & t_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & t_{n3} & \cdots & t_{nn} \end{pmatrix}$$

Example. The matrices

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 4 & 5 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are all lower triangular matrices. ☑

Now a square matrix in row echelon form is upper triangular. Like row echelon matrices, if the matrix of coefficients of a system of equations is upper triangular then the system can be solved with back-substitution. Similarly, if the matrix of coefficients of a system of equations is lower triangular, then the system can be solved with front-substitution; namely, you use the first equation to solve for the first variable, then move to the next equation to solve for the next variable, etc.

Example. Let

$$L = \begin{pmatrix} 2 & 0 \\ 4 & 3 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

Then

$$L\boldsymbol{x} = \boldsymbol{b}$$

corresponds to

$$\begin{array}{rcrl} 2x & & & = 4 \\ 4x & + & 3y & = 5. \end{array}$$

The first equation $2x = 4$ gives us $x = 2$, the second equation $4x + 3y = 5$ gives us $y = (5 - 4x)/3 = (5 - 4 \cdot 2)/3 = -1$. ☑

So if the matrix of coefficients of a system is upper or lower triangular, the system can be solved fairly easily. As we shall see, if the matrix of coefficients is a product of triangular matrices then the system can still be solved fairly easily.

Suppose we want to solve the system

$$A\boldsymbol{x} = \boldsymbol{b},$$

where $A = LU$. Then the system can be rewritten

$$LU\boldsymbol{x} = \boldsymbol{b}$$

or

$$L\boldsymbol{y} = \boldsymbol{b} \text{ where } U\boldsymbol{x} = \boldsymbol{y}.$$

The system can then be solved in two relatively simple steps.

Example. ***Given that the matrix***

$$\boldsymbol{A} = \begin{pmatrix} \mathbf{3} & \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{3} & \mathbf{4} \\ \mathbf{3} & \mathbf{2} & \mathbf{6} \end{pmatrix}$$

can be written $\boldsymbol{A} = \boldsymbol{LU}$, where

$$\boldsymbol{L} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{2} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{3} \end{pmatrix}$$

and

$$\boldsymbol{U} = \begin{pmatrix} \mathbf{3} & \mathbf{1} & \mathbf{2} \\ \mathbf{0} & \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix},$$

solve the system

$$\begin{array}{rcrcrl}\mathbf{3}\boldsymbol{x} & + & \boldsymbol{y} & + & \mathbf{2}\boldsymbol{z} & = \mathbf{3}\\ \mathbf{3}\boldsymbol{x} & + & \mathbf{3}\boldsymbol{y} & + & \mathbf{4}\boldsymbol{z} & = \mathbf{5}\\ \mathbf{3}\boldsymbol{x} & + & \mathbf{2}\boldsymbol{y} & + & \mathbf{6}\boldsymbol{z} & = \mathbf{1.}\end{array}$$

The system is

$$A\boldsymbol{x} = \boldsymbol{b},$$

where

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}.$$

Since this can be written

$$LU\boldsymbol{x} = \boldsymbol{b},$$

it is enough to solve

$$L\boldsymbol{y} = \boldsymbol{b}$$

and

$$U\boldsymbol{x} = \boldsymbol{y},$$

where

$$\boldsymbol{y} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}.$$

Now $L\boldsymbol{y} = \boldsymbol{b}$ represents the system

$$\begin{array}{rcrcrl} u & & & & & = 3\\ u & + & 2v & & & = 5\\ u & + & v & + & 3w & = 1,\end{array}$$

which has solution

$$\begin{aligned} u &= 3\\ v &= (5-u)/2 = (5-3)/2 = 1\\ w &= (1-u-v)/3 = (1-3-1)/3 = -1.\end{aligned}$$

Next, $U\boldsymbol{x} = \boldsymbol{y}$ represents

$$\begin{array}{rcrcrl} 3x & + & y & + & 2z & = 3\\ & & y & + & z & = 1\\ & & & & z & = -1,\end{array}$$

which has solution

$$\begin{aligned}
z &= -1 \\
y &= 1 - z = 1 - (-1) = 2 \\
x &= (3 - y - 2z)/3 = (3 - 2 - 2(-1))/3 = 1.
\end{aligned}$$

So the solution is $x = 1$, $y = 2$, $z = -1$. ☑

Definition. If a square matrix A can be written as the product of a lower triangular matrix L and an upper triangular matrix U,

$$A = LU,$$

then LU is an ***LU decomposition*** of A. □

Not every square matrix has an LU decomposition, but will if we allow the rows to be rearranged beforehand. Given a square matrix A, there will be a permutation matrix P (which is merely the identity matrix with the rows rearranged, or permuted) such that PA has an LU decomposition. We will only concern ourselves with finding LU decompisition of matrices which have them for now.

Even when a matrix has an LU decomposition, it isn't necessarily unique.

Example. For

$$A = \begin{pmatrix} 2 & 2 \\ 6 & 4 \end{pmatrix},$$

both

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

and

$$A = \begin{pmatrix} 2 & 0 \\ 6 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

are LU decompositions of A. ☑

Given a matrix, we still have to find an LU factorization for it, if such a factorization is possible. The key observation is that when performing Gaussian elimination to turn a matrix into an upper triangular matrix, the elementary row operations that are used are:

1. multiply a row by a non-zero number
2. add a multiple of one row to a later row
3. switch rows.

(The row operations which aren't used are adding a multiply of a row to an earlier row.) The row operations listed above, except row switching, correspond to lower triangular elementary matrices. We start with the following proposition.

Proposition. *If $E = E_1 E_2 \ldots E_n$ is a product of elementary matrices, where no elementary matrix corresponds to switching rows or adding a multiply of a row to a previous row, then both E and $E^{-1} = E_n^{-1} E_{n-1}^{-1} \ldots E_1^{-1}$ are lower triangular.* □

Recall that when finding the inverse of a product of matrices, the order of the factors is reversed.

Given a matrix A, suppose we can use Gaussian elimination (without row switching) to put A in upper triangular form; we can then reduce

$$(A|B)$$

to

$$(U|E)\,,$$

and $EA = U$. We then get $A = E^{-1}U$, or

$$A = LU,$$

where $L = E^{-1}$ is lower triangular. If only a few row operations are needed, then rather than augmenting the matrix A and then inverting E at the end, it is often easier to keep track of the row operations as they're performed, and at the end find $L = E^{-1} = E_n^{-1} E_{n-1}^{-1} \ldots E_1^{-1}$ by applying the opposites of the row operations in reverse order. (It's easy to find the opposite of a row operation.) Let's do an example.

Example. ***Find an LU factorization of***

$$\boldsymbol{A = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 2 & 6 \\ 2 & -10 & 2 \end{pmatrix}.}$$

Starting with

$$A = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 2 & 6 \\ 2 & -10 & 2 \end{pmatrix}.$$

we can

1. Subtract 2 times row 1 from row 3:
$$\begin{pmatrix} 1 & -3 & 0 \\ 0 & 2 & 6 \\ 0 & -4 & 2 \end{pmatrix}.$$
2. Divide row 2 by 2:
$$\begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & -4 & 2 \end{pmatrix}.$$
(Note that this step isn't necessary, since we only want an upper triangular matrix, not necessarily a row echelon matrix. But this step doesn't hurt ....)
3. Add 4 times row 2 to row 3:
$$\begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 14 \end{pmatrix}.$$

This last matrix is the U part of our LU decomposition. To find the L part, we apply the inverse of the row operations in opposite order (starting with the identity matrix). The inverses of the row operations are what you'd expect: replace addition by subtraction, multiplication by division, etc. So we get

1. Subtract 4 times row 2 from row 3:
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix}.$$
2. Multiply row 2 by 2:
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & -4 & 1 \end{pmatrix}.$$

3. Add 2 times row 1 to row 3:

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

This last matrix is our L. So our LU decomposition is $A = LU$, where

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 2 & -4 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 14 \end{pmatrix}.$$

So if A can be put in upper triangular form using elementary row operations (but no row switches), then we know how to find an LU factorization. If A cannot be put in upper triangular form without row switches, then A has no LU factorization.