

1 LU Factorization

Exploration 1 Consider the equation:

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

This equation is unusual in that it involves two matrices on the left-hand side. If we multiply the matrices together, we get

$$\begin{bmatrix} 2 & -1 & 2 \\ 6 & 1 & 5 \\ -4 & 10 & -3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix}$$

Gaussian elimination yields:

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

We conclude that

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

Now that we know the answer, we will turn to the original question and consider the advantages of the format of the original problem. Observe that the two matrices have a distinct pattern. The matrix on the left has zeros above the main diagonal, while the matrix on the right has zeros below the main diagonal. Matrices that follow such a pattern are called lower triangular and upper triangular matrices, respectively.

Let $\mathbf{y} = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 4 & -1 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x}$. Then we can write the original equation as

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \left(\begin{bmatrix} 2 & -1 & 2 \\ 0 & 4 & -1 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x} \right) = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \mathbf{y} = \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix}$$

The equation

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \mathbf{y} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix}$$

corresponds to the system

$$\begin{array}{rrcr} y_1 & & & = 1 \\ 3y_1 & +y_2 & & = 8 \\ -2y_1 & +2y_2 & +y_3 & = 5 \end{array}$$

This system can be easily solved by substitution, giving us

$$y_1 = 1$$

$$3(1) + y_2 = 8 \rightsquigarrow y_2 = 5$$

$$-2(1) + 2(5) + y_3 = 5 \rightsquigarrow y_3 = -3$$

So,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}$$

Recall that we let $\mathbf{y} = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 4 & -1 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x}$. The equation

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

corresponds to the system

$$\begin{aligned} 2x_1 - x_2 + 2x_3 &= 1 \\ 4x_2 - x_3 &= 5 \\ 3x_3 &= -3 \end{aligned}$$

which can be easily solved by back-substitution:

$$x_3 = \dots\dots$$

$$x_2 = \dots\dots$$

$$x_1 = \dots\dots$$

Thus, $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$. Observe that this is the same answer that we obtained in the beginning of the problem using Gaussian elimination.

Given a matrix equation such as

$$\begin{bmatrix} 2 & -1 & 2 \\ 6 & 1 & 5 \\ -4 & 10 & -3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix}$$

of Exploration ??, often in practice we express the matrix on the left as a product of an upper triangular matrix U and a lower triangular matrix L and use substitution to solve the equation instead of using Gaussian elimination to find the solution. In Exploration ??,

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

The process of taking a matrix A and expressing it as a product $A = LU$ of a lower triangular matrix L and an upper triangular matrix U is called *LU factorization*. We also adopt the (common) convention that L is a matrix with a diagonal consisting entirely of 1's, which is often called a *lower unit triangular* matrix.

Not every matrix has an LU factorization, but we will see that we can correct for that.

Example 1. Consider the system

$$\begin{array}{rrcr} & 12x_2 & - & 18x_3 & = & 6 \\ x_1 & + & 2x_2 & - & 3x_3 & = & 1 \\ x_1 & - & 2x_2 & + & x_3 & = & 1 \end{array}$$

We can express this system as the matrix equation $A\mathbf{x} = \mathbf{b}$, and we get

$$\begin{bmatrix} 0 & 12 & -18 \\ 1 & 2 & -3 \\ 1 & -2 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 6 \\ 1 \\ 1 \end{bmatrix}.$$

Unfortunately, for reasons we will explain below, it is not possible to find an LU factorization of this coefficient matrix. However, if we simply swap the first two equations:

$$\begin{array}{rrcr} x_1 & + & 2x_2 & - & 3x_3 & = & 1 \\ & 12x_2 & - & 18x_3 & = & 6 \\ x_1 & - & 2x_2 & + & x_3 & = & 1 \end{array},$$

we are able to find an LU factorization of the coefficient matrix.

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 12 & -18 \\ 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 12 & -18 \\ 0 & 0 & -2 \end{bmatrix}$$

(How we actually come up with an LU factorization of A will be discussed after this example.)

To understand how an LU factorization can be used, it is helpful to think of this system of equations as a matrix equation.

$$A\mathbf{x} = \mathbf{b}$$

$$(LU)\mathbf{x} = \mathbf{b}$$

$$L(U\mathbf{x}) = \mathbf{b}$$

So if we let $\mathbf{y} = U\mathbf{x}$ and also write $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, then we are able to solve for y_1 , y_2 , and y_3 using *forward-substitution*.

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1/3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 1 \end{bmatrix}$$

See if you can do it!

$$y_1 = \dots\dots, \quad y_2 = \dots\dots, \quad y_3 = \dots\dots$$

Once we have the values of $\mathbf{y}$, since $\mathbf{y} = U\mathbf{x}$, all that remains is to solve the following matrix equation using back-substitution.

$$U\mathbf{x} = \mathbf{y}$$

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 12 & -18 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$$

By now you are quite used to solving this kind of problem...

$$x_1 = \dots\dots, \quad x_2 = \dots\dots, \quad x_3 = \dots\dots$$

Finding an LU factorization by Inspection

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

One way to find the LU factorization is to simply look for it directly. We need

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 3 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ x & 1 & 0 \\ y & z & 1 \end{bmatrix} \begin{bmatrix} a & d & h & j \\ 0 & b & e & i \\ 0 & 0 & c & f \end{bmatrix}$$

By multiplying the latter matrices we get

$$\begin{bmatrix} a & d & h & j \\ xa & xd + b & xh + e & xj + i \\ ya & yd + zb & yh + ze + c & yj + iz + f \end{bmatrix}$$

and from this, it is a simple exercise to determine each of the unknowns. For example, from the first column, we get $a = 1$ and then $x = 1, y = 2$.

See if you can continue to determine the entire LU factorization of A .

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 3 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & \dots\dots & 1 \end{bmatrix} \begin{bmatrix} 1 & \dots\dots & \dots\dots & \dots\dots \\ 0 & \dots\dots & \dots\dots & \dots\dots \\ 0 & 0 & \dots\dots & \dots\dots \end{bmatrix}$$

Finding an LU factorization by the Multiplier Method

In the following example we demonstrate a method for computing an LU factorization known as the multiplier method.

Example 3. Find an LU factorization for

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

Explanation. Write the matrix as the following product.

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

In the matrix on the right, we begin with the left row and zero out the entries below the top using the row operation which involves adding a multiple of a row to another row. You do this and also update the matrix on the left so that the product will be unchanged.

Here is the first step. We take -2 times the top row and add to the second. Then take 2 times the top row and add to the second in the matrix on the left. This gives us

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

The next step is to take 2 times the top row and add to the bottom in the matrix on the right. To ensure that the product is unchanged, we place a -2 in the bottom left corner in the matrix on the left. Thus the second step yields

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 7 & 4 \end{bmatrix}$$

For our final step, we take 7 times the middle row on right and add to bottom row. Updating the matrix on the left in the manner we did earlier, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & -7 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & -31 \end{bmatrix}$$

At this point, we can stop. We have an LU factorization of the original matrix. We can always multiply the matrices to check, if we wish.

Notice that when we perform the multiplier method, we are making repeated use of a certain type of elementary row operation, namely, we are adding a scalar multiple of one row to a row below it. The reason this helps to create an LU factorization depends upon the fact that the elementary matrices corresponding to such operations are lower triangular. To understand how this works, we begin with a lemma.

Lemma 1. Let L be a lower (upper) triangular matrix $m \times m$ which has ones down the main diagonal. Then L^{-1} also is a lower (upper) triangular matrix which has ones down the main diagonal. In the case that L is of the form

$$L = \begin{bmatrix} 1 & & & \\ a_1 & 1 & & \\ \vdots & & \ddots & \\ a_n & & & 1 \end{bmatrix}, \quad (1)$$

where all entries are zero except for the left column and main diagonal, it is also the case that L^{-1} is obtained from L by simply multiplying each entry below the main diagonal in L by -1 . The same is true if the single nonzero column is in another position.

Proof Consider the usual setup for finding the inverse $\left[L \mid I \right]$. Each row operation used on L to transform this matrix to reduced row echelon form changes only the entries in I below the main diagonal. Whether we have the special case of L given in ?? where the nonzero nondiagonal entries are in the left column, or if the single nonzero column is in another position, it is clear that multiplication by -1 as described in the lemma gives us L^{-1} . ■

For a simple illustration of the lemma, observe:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & a & 1 & 0 & 0 & 1 \end{array} \right] \rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -a & 1 \end{array} \right]$$

Now let A be an $m \times n$ matrix, say

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

and assume A can be row reduced to an upper triangular form using only row operation 3. Thus, in particular, $a_{11} \neq 0$. Multiply on the left by

$$E_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{m1}}{a_{11}} & 0 & \cdots & 1 \end{bmatrix}$$

This is the product of elementary matrices which make modifications in the first column only. It is equivalent to taking $-a_{21}/a_{11}$ times the first row and adding to the second. Then taking $-a_{31}/a_{11}$ times the first row and adding to the third and so forth. The quotients in the first column of the above matrix are the multipliers. Thus the result is of the form

$$E_1 A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{m2} & \cdots & a'_{mn} \end{bmatrix}$$

By assumption, $a'_{22} \neq 0$ and so it is possible to use this entry to zero out all the entries below it in the matrix on the right by multiplication by a matrix of the form

$$E_2 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & E \end{bmatrix}$$

where E is an $(m-1) \times (m-1)$ matrix of the form

$$E = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a'_{32}}{a'_{22}} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a'_{m2}}{a'_{22}} & 0 & \cdots & 1 \end{bmatrix}$$

Again, the entries in the first column below the 1 are the multipliers. Continuing this way, zeroing out the entries below the diagonal entries, finally leads us to

$$E_{m-1}E_{m-2} \cdots E_1 A = U$$

where U is upper triangular. Each E_j has all ones down the main diagonal and is lower triangular. Now we can multiply both sides by the inverses of the E_j in the reverse order. This yields

$$A = E_1^{-1}E_2^{-1} \cdots E_{m-1}^{-1}U$$

By Lemma ??, this implies that the product of those E_j^{-1} is a lower triangular matrix having all ones down the main diagonal.

The above discussion and lemma explain how the multiplier method works. The expressions

$$-\frac{a_{21}}{a_{11}}, -\frac{a_{31}}{a_{11}}, \dots, -\frac{a_{m1}}{a_{11}},$$

which we obtained in building E_1 , and which we denote respectively by $M_{21}, \dots, M_{m1}$ to save notation, are the multipliers. Therefore, according to the lemma, to find E_1^{-1} we simply write

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -M_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -M_{m1} & 0 & \cdots & 1 \end{bmatrix}$$

Similar considerations apply to the other E_j^{-1} . Thus L is a product of the form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -M_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -M_{m1} & 0 & \cdots & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & -M_{m(m-1)} & 1 \end{bmatrix},$$

where each factor has at most one nonzero column, the position of which moves from left to right as we scan the above product of matrices from left to right. It follows from what we know about the effect of multiplying on the left by an elementary matrix that the above product is of the form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -M_{21} & 1 & \cdots & 0 & 0 \\ \vdots & -M_{32} & \ddots & \vdots & \vdots \\ -M_{(M-1)1} & \vdots & \cdots & 1 & 0 \\ -M_{M1} & -M_{M2} & \cdots & -M_{MM-1} & 1 \end{bmatrix}$$

To sum up the procedure, beginning at the left column and moving toward the right, you simply insert, into the corresponding position in the identity matrix, -1 times the multiplier which was used to zero out an entry in that position below the main diagonal in A , while retaining the main diagonal which consists entirely of ones. This will give us L as we create U using row operation 3.

We now return to Example ??, to understand why we could not use the multiplier method to find an LU factorization of the coefficient matrix. Suppose we had written

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 12 & -18 \\ 1 & 2 & -3 \\ 1 & -2 & 1 \end{bmatrix}.$$

Our first step of Gaussian elimination would require a row swap to get a nonzero entry into the top left corner (we

swapped rows 1 and 2 in Example ??). Unfortunately, the elementary matrix that accomplishes this, $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

is NOT lower triangular. So we cannot apply Lemma ?? to generate a lower triangular L .

If the elementary matrices used to reduce our matrix to row-echelon form are all lower triangular, then we can find an LU factorization. But what about the general case?

Theorem 1. Suppose an $m \times n$ matrix A is transformed to a row-echelon matrix U using Gaussian elimination. 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.

For a proof of the above theorem, the reader is referred to [Nicholson].

The matrix P in the above theorem is called a *permutation matrix*. These matrices have other important applications, as we will see later.

Practice Problems

Problem 2 Use the given LU factorization of the coefficient matrix to solve the system of equations.

$$\begin{array}{rrcr} x & + & 2y & + & 3z & = & 5 \\ 2x & + & 3y & + & z & = & 6 \\ 3x & + & 5y & + & 4z & = & 11 \end{array}$$

Observe that an LU factorization of the coefficient matrix is

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

Problems ??-??

Find the LU factorization of the coefficient matrix and use it to solve the system of equations.

Problem 3

$$\begin{aligned}x + 2y &= 5 \\2x + 3y &= 6\end{aligned}$$
$$L = \begin{bmatrix} \dots & 0 \\ \dots & \dots \end{bmatrix}, \quad U = \begin{bmatrix} \dots & \dots \\ 0 & \dots \end{bmatrix}$$
$$x = \dots, y = \dots$$

Problem 4

$$\begin{aligned}x + 2y + z &= 1 \\y + 3z &= 2 \\2x + 3y &= 6\end{aligned}$$
$$L = \begin{bmatrix} \dots & 0 & 0 \\ \dots & \dots & 0 \\ \dots & \dots & \dots \end{bmatrix}, \quad U = \begin{bmatrix} \dots & \dots & \dots \\ 0 & \dots & \dots \\ 0 & 0 & \dots \end{bmatrix}$$
$$x = \dots, y = \dots, z = \dots$$

Problem 5 *Is there only one LU factorization for a given matrix?*

Problem 6 *Can you show that every permutation matrix is invertible? If so, What does the inverse of a permutation matrix look like? (Recall that a permutation matrix is matrix P of Theorem ??.)*

Text Source

The text in this section is an adaptation of Section 2.2 of Ken Kuttler's *A First Course in Linear Algebra*. (CC-BY)
Ken Kuttler, *A First Course in Linear Algebra*, Lyryx 2017, Open Edition, p. 99-106.

Bibliography

[Nicholson] W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, pp. 123–127.