

# The LU and PLU factorizations

Nicholas Hu · Last updated on 2024-03-24

## The LU factorization

Let  $\mathbb{F}$  be a field and  $A \in \mathbb{F}^{n \times n}$ . Given a  $b \in \mathbb{F}^n$ , recall that the linear system  $Ax = b$  can be solved by performing **Gaussian elimination** (also known as **row reduction**) on the augmented matrix  $[A \ b]$  to transform it into an augmented matrix of the form  $[U \ y]$ , where  $U \in \mathbb{F}^{n \times n}$  is upper triangular and  $y \in \mathbb{F}^n$ , which represents an equivalent linear system  $Ux = y$ . Such a system can then be solved readily by backward substitution. In fact, this process (under certain conditions) amounts to computing an **LU factorization** of  $A$  as  $LU$ , where  $L \in \mathbb{F}^{n \times n}$  is *unit* lower triangular and  $U \in \mathbb{F}^{n \times n}$  is upper triangular.

To illustrate this, suppose that  $\mathbb{F} = \mathbb{R}$  and

$$\underbrace{\begin{bmatrix} 1 & 5 & 1 & 2 \\ 2 & 6 & 3 & 4 \\ 3 & 7 & 2 & 4 \\ 4 & 8 & 1 & 3 \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_x = \underbrace{\begin{bmatrix} 7 \\ 8 \\ 7 \\ 8 \end{bmatrix}}_b.$$

Then the variable  $x_1$  can be eliminated from the second equation  $2x_1 + 6x_2 + 3x_3 + 4x_4 = 8$  by multiplying the first equation  $1x_1 + 5x_2 + 1x_3 + 2x_4 = 7$  by  $\frac{2}{1}$  and subtracting the resulting equation from the second. The quotient  $\frac{2}{1} = 2$  is called a **multiplier** and the divisor 1 is called a **pivot**. In other words (ignoring the vector  $b$  for now),

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

Similarly,  $x_1$  can be eliminated from the third and fourth equations, yielding the factorization

$$\begin{bmatrix} 1 & 5 & 1 & 2 \\ 2 & 6 & 3 & 4 \\ 3 & 7 & 2 & 4 \\ 4 & 8 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 \begin{bmatrix} 1 & 5 & 1 & 2 \end{bmatrix} \\ 2 \begin{bmatrix} 1 & 5 & 1 & 2 \end{bmatrix} + 1 \begin{bmatrix} 0 & -4 & 1 & 0 \end{bmatrix} \\ 3 \begin{bmatrix} 1 & 5 & 1 & 2 \end{bmatrix} + 1 \begin{bmatrix} 0 & -8 & -1 & -2 \end{bmatrix} \\ 4 \begin{bmatrix} 1 & 5 & 1 & 2 \end{bmatrix} + 1 \begin{bmatrix} 0 & -12 & -3 & -5 \end{bmatrix} \end{bmatrix} \\ = \begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 3 & & 1 & \\ 4 & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 & 1 & 2 \\ & -4 & 1 & 0 \\ & -8 & -1 & -2 \\ & -12 & -3 & -5 \end{bmatrix}$$

(where blanks denote zeroes).

The variable  $x_2$  can then be eliminated analogously from the new third and fourth equations. For instance,

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

In view of the previous step, this means that

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

Thus, after the second step of the factorization, we have

$$\begin{bmatrix} 1 & 5 & 1 & 2 \\ 2 & 6 & 3 & 4 \\ 3 & 7 & 2 & 4 \\ 4 & 8 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 3 & 2 & 1 & \\ 4 & 3 & & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 & 1 & 2 \\ & -4 & 1 & 0 \\ & & -3 & -2 \\ & & & -6 & -5 \end{bmatrix}.$$

Finally, an entirely analogous third step completes the LU factorization:

$$\underbrace{\begin{bmatrix} 1 & 5 & 1 & 2 \\ 2 & 6 & 3 & 4 \\ 3 & 7 & 2 & 4 \\ 4 & 8 & 1 & 3 \end{bmatrix}}_A = \underbrace{\begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 3 & 2 & 1 & \\ 4 & 3 & 2 & 1 \end{bmatrix}}_L \underbrace{\begin{bmatrix} 1 & 5 & 1 & 2 \\ & -4 & 1 & 0 \\ & & -3 & -2 \\ & & & -1 \end{bmatrix}}_U.$$

In general, we can regard LU factorization as partitioning a matrix  $A \in \mathbb{F}^{n \times n}$  as

$$A = \begin{bmatrix} \alpha & b^\top \\ c & D \end{bmatrix},$$

where  $\alpha \in \mathbb{F}$ ,  $b \in \mathbb{F}^{n-1}$ ,  $c \in \mathbb{F}^{n-1}$ , and  $D \in \mathbb{F}^{(n-1) \times (n-1)}$ . Provided that the pivot  $\alpha$  is nonzero, we can eliminate  $c$  by using the multipliers  $\frac{c}{\alpha}$  to write  $\begin{bmatrix} c & D \end{bmatrix} - \frac{c}{\alpha} \begin{bmatrix} \alpha & b^\top \end{bmatrix} = \begin{bmatrix} 0 & A' \end{bmatrix}$  for some  $A' \in \mathbb{F}^{(n-1) \times (n-1)}$ ; the matrix  $A' = D - \frac{c}{\alpha} b^\top$  is known as the **Schur complement** of  $[\alpha]$  in  $A$ . This yields the factorization

$$A = \begin{bmatrix} 1 & \\ \frac{c}{\alpha} & I \end{bmatrix} \begin{bmatrix} \alpha & b^\top \\ A' \end{bmatrix},$$

which can be recursively continued by computing an LU factorization  $L'U'$  of  $A'$ . If one exists, then  $\begin{bmatrix} c & D \end{bmatrix} - \frac{c}{\alpha} \begin{bmatrix} \alpha & b^\top \end{bmatrix} = L' \begin{bmatrix} 0 & U' \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$ , so

$$A = \underbrace{\begin{bmatrix} 1 & \\ \frac{c}{\alpha} & L' \end{bmatrix}}_L \underbrace{\begin{bmatrix} \alpha & b^\top \\ U' \end{bmatrix}}_U$$

is an LU factorization of  $A$ .

## Existence and uniqueness

Let  $a_{ij}^{(k)}$  denote the  $(i, j)$ -entry of  $A$  after the  $k^{\text{th}}$  step in the LU factorization. From the example above, it is clear that  $A$  will have an LU factorization provided that the pivots  $a_{11}^{(0)}, a_{22}^{(1)}, \dots, a_{n-1, n-1}^{(n-2)}$  are nonzero. Moreover, we see after  $k-1$  steps that the  $k^{\text{th}}$  leading principal minor of  $A$  is equal to  $a_{11}^{(0)} a_{22}^{(1)} \cdots a_{kk}^{(k-1)}$ , so  $A$  will have an LU factorization if its first  $n-1$  leading principal minors are nonzero.

If  $A$  is *nonsingular*, this condition is also *necessary* because the  $k^{\text{th}}$  leading principal minor of  $A$  must be equal to  $u_{11} u_{22} \cdots u_{kk}$ . Furthermore, the LU factorization is *unique* in this case, for if  $L_1 U_1$  and  $L_2 U_2$  are both LU factorizations of  $A$ , then  $L_2^{-1} L_1 = U_2 U_1^{-1}$  must be simultaneously unit lower triangular and upper triangular, and hence equal to the identity.

## Solving linear systems

If the same row operations that were applied to  $A$  are applied to  $b$  as well, evidently the resulting vector  $y$  will satisfy  $b = Ly$  (since  $[A \ b] = L[U \ y]$ ). Therefore the solutions of  $Ax = b$  will indeed be those of  $Ux = y$ . Notably, once an LU factorization of  $A$  has been computed, it is possible to solve  $Ax = b$  for any given  $b \in \mathbb{F}^n$  by solving  $Ly = b$  for  $y$  and then  $Ux = y$  for  $x$ , reusing the computed factors  $L$  and  $U$ ! (The former system can be solved readily by forward substitution and the latter by backward substitution.)

## The PLU factorization

It is possible to produce an LU-like factorization for *any*  $A \in \mathbb{F}^{n \times n}$  by allowing for *row interchanges* in addition to the elementary row operation above. The resulting **PLU factorization** consists of a permutation matrix  $P \in \mathbb{F}^{n \times n}$  along with matrices  $L$  and  $U$  as above such that  $PA = LU$  (or equivalently,  $A = P^\top LU$ ).

As an illustration, consider the matrix

$$\underbrace{\begin{bmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 2 & 2 \\ 4 & 5 & 5 & 6 \\ 2 & 4 & 4 & 4 \end{bmatrix}}_A.$$

Following one elimination step as before, we arrive at

$$A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 2 & 2 \\ 4 & 5 & 5 & 6 \\ 2 & 4 & 4 & 4 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ \frac{1}{4} & 1 & & \\ 1 & & 1 & \\ \frac{1}{2} & & & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 & 4 & 4 \\ 0 & 1 & 1 & \\ 1 & 1 & 2 & \\ 2 & 2 & 2 & \end{bmatrix}.$$

Since the (2, 2)-entry of the factor on the right is zero, it is impossible to introduce zeroes in the second column using this entry as a pivot. However, by interchanging the second row with a row below it whose entry in the second column is *nonzero*, we can obtain a nonzero pivot to perform the elimination.<sup>1</sup> Also, note that had such a row not existed, it could only have been because all subdiagonal entries in the second column were zero, in which case we could have immediately continued to the next elimination step!

For example, suppose we wish to use the (4, 2)-entry, 2, as a pivot. We have

$$\begin{cases} [1 & 1 & 2 & 2] = \frac{1}{4} [4 & 4 & 4 & 4] + [0 & 0 & 1 & 1] \\ [2 & 4 & 4 & 4] = \frac{1}{2} [4 & 4 & 4 & 4] + [0 & 2 & 2 & 2] \end{cases},$$

so if  $P_2$  is the permutation matrix that interchanges the second and fourth rows, then

$$P_2 A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 2 & 4 & 4 & 4 \\ 4 & 5 & 5 & 6 \\ 1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ \frac{1}{2} & 1 & & \\ 1 & & 1 & \\ \frac{1}{4} & & & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & \\ 1 & 1 & 2 & \\ 0 & 1 & 1 & \end{bmatrix},$$

from which elimination can proceed as in the LU factorization. Namely,

$$P_2 A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 2 & 4 & 4 & 4 \\ 4 & 5 & 5 & 6 \\ 1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ \frac{1}{2} & 1 & & \\ 1 & \frac{1}{2} & 1 & \\ \frac{1}{4} & 0 & & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 & 4 & 4 \\ & 2 & 2 & 2 \\ & & 0 & 1 \\ & & 1 & 1 \end{bmatrix}.$$

Once again, we see that a row interchange is required. We have

$$\begin{cases} [4 & 5 & 5 & 6] = 1[4 & 4 & 4 & 4] + \frac{1}{2}[0 & 2 & 2 & 2] + [0 & 0 & 0 & 1] \\ [1 & 1 & 2 & 2] = \frac{1}{4}[4 & 4 & 4 & 4] + 0[0 & 2 & 2 & 2] + [0 & 0 & 1 & 1] \end{cases},$$

so if  $P_3$  is the permutation matrix that interchanges the third and fourth rows, then

$$P_3 P_2 A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 2 & 4 & 4 & 4 \\ 1 & 1 & 2 & 2 \\ 4 & 5 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ \frac{1}{2} & 1 & & \\ \frac{1}{4} & 0 & 1 & \\ 1 & \frac{1}{2} & & 1 \end{bmatrix} \begin{bmatrix} 4 & 4 & 4 & 4 \\ & 2 & 2 & 2 \\ & & 1 & 1 \\ & & 0 & 1 \end{bmatrix}.$$

The product of all the permutation matrices used for row interchanges becomes the permutation matrix  $P$ ; if no interchange is performed in the  $k^{\text{th}}$  step, we can regard this as multiplication by  $P_k = I$  at that step. For instance, after the final elimination step in our example (which is trivial since the (4, 3)-entry is already zero),

$$\underbrace{P_3 P_2 P_1 A}_P = \underbrace{\begin{bmatrix} 4 & 4 & 4 & 4 \\ 2 & 4 & 4 & 4 \\ 1 & 1 & 2 & 2 \\ 4 & 5 & 5 & 6 \end{bmatrix}}_{PA} = \underbrace{\begin{bmatrix} 1 & & & \\ \frac{1}{2} & 1 & & \\ \frac{1}{4} & 0 & 1 & \\ 1 & \frac{1}{2} & 0 & 1 \end{bmatrix}}_L \underbrace{\begin{bmatrix} 4 & 4 & 4 & 4 \\ & 2 & 2 & 2 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}}_U.$$

In general, given a matrix  $A \in \mathbb{F}^{n \times n}$ , either there exists a permutation matrix  $P_1 \in \mathbb{F}^{n \times n}$  such that  $P_1 A$  is of the form

$$P_1 A = \begin{bmatrix} \alpha & b^\top \\ c & D \end{bmatrix}$$

for some *nonzero*  $\alpha \in \mathbb{F}$ , or the first column of  $A$  is zero, in which case  $P_1 A$  is of the form

$$P_1 A = \begin{bmatrix} \alpha & b^\top \\ 0 & D \end{bmatrix}$$

for any permutation matrix  $P_1 \in \mathbb{F}^{n \times n}$ . Let  $\hat{c} = \frac{c}{\alpha}$  in the former case and  $\hat{c} = 0$  in the latter. Then

$$P_1 A = \begin{bmatrix} 1 & \\ \hat{c} & I \end{bmatrix} \begin{bmatrix} \alpha & b^\top \\ A' \end{bmatrix},$$

where  $A' = D - \hat{c}b^\top$ , and if  $P' A' = L' U'$  is a PLU factorization of  $A'$ , then

$$\underbrace{\begin{bmatrix} 1 & \\ & P' \end{bmatrix}}_P P_1 A = \underbrace{\begin{bmatrix} 1 & \\ P' \hat{c} & L' \end{bmatrix}}_L \underbrace{\begin{bmatrix} \alpha & b^\top \\ U' \end{bmatrix}}_U$$

is a PLU factorization of  $A$ .

## Solving linear systems

By the same reasoning as above, row-reducing  $b$  in the manner that  $A$  was reduced results in a vector  $y$  satisfying  $Pb = Ly$ . Therefore the solutions of  $Ax = b$  will be those of  $Ux = y$ . Once a PLU factorization of  $A$  has been computed, it is possible to solve  $Ax = b$  for any given  $b \in \mathbb{F}^n$  by solving  $Ly = Pb$  for  $y$  and then  $Ux = y$  for  $x$ , reusing the computed factors  $L$  and  $U$  and the permutation represented by  $P$ .

---

1. In practice, implementations of PLU factorization typically perform a row interchange that maximizes the absolute value of the pivot, regardless of whether it is needed to prevent division by zero. With this convention, all multipliers are at most 1 in absolute value. 