1 LU Factorization

The following facts will be usefull in understanding this section. Suppose A is an $n \times n$ matrix, $\mathbf{R} = [r_1\ r_2\ \ldots\ r_n]$ is a row vector and $\mathbf{C} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ is column vector. We can write A as an arrangement of row vectors in the form

$$A = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_n \end{bmatrix}$$

or as an arrangement of column vectors in the form

$$A = [\mathbf{C}_1\ \mathbf{C}_2\ \ldots\ \mathbf{C}_n].$$

The product $\mathbf{R}A$ is a row vector which can be interpretted as a linear combination of the rows of A, i.e.,

$$\mathbf{R}A = r_1\mathbf{R}_1 + r_2\mathbf{R}_2 + \ldots + r_n\mathbf{R}_n$$

and the product $A\mathbf{C}$ is a column vector which may be interpretted as a linear combination of the columns of A, i.e.,

$$A\mathbf{C} = c_1\mathbf{C}_1 + c_2\mathbf{C}_2 + c_n\mathbf{C}_n.$$

For example, let $\mathbf{R} = [1\ 3\ -1\ 2]\,, \mathbf{C} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ -3 \end{bmatrix}, \ A = \begin{bmatrix} 1 & -2 & 0 & -1 \\ 2 & 4 & 5 & -2 \\ 3 & 3 & -1 & 1 \\ 2 & -1 & 3 & -4 \end{bmatrix}.$

Then

$$\begin{aligned} \mathbf{R}A &= \begin{bmatrix} 1 & -2 & 0 & -1 \end{bmatrix} + 3\begin{bmatrix} 2 & 4 & 5 & -2 \end{bmatrix} - \begin{bmatrix} 3 & 3 & -1 & 1 \end{bmatrix} + 2\begin{bmatrix} 2 & -1 & 3 & -4 \end{bmatrix} \\ &= \begin{bmatrix} 8 & 5 & 22 & -16 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} A\mathbf{C} &= \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} -2 \\ 4 \\ 3 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 5 \\ -1 \\ 3 \end{bmatrix} - 3\begin{bmatrix} -1 \\ -2 \\ 1 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} 6 \\ 14 \\ -5 \\ 21 \end{bmatrix}. \end{aligned}$$

We can similarly understand the product BA as the rows of B forming linear combinations of the rows of A and AB as the columns of B forming linear combinations of the columns of A.

Definition 1 *A nonsingular $n \times n$ matrix A is said to have an LU factorization if twe can write*

$$A = LU,$$

where L is a lower triangular matrix with ones on the main diagonal and U is an upper traingluar matix with nonzero daigonal elements.

Example

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 & 4 & 1 \\ 2 & 8 & 6 & 4 \\ 3 & 10 & 8 & 8 \\ 4 & 12 & 10 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 \\ 4 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 & 1 \\ 0 & 4 & -2 & 2 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 6 \end{bmatrix} = LU. \end{aligned}$$

Suppose now that a matrix A has an LU factorization and we want to solve the system

$$AX = B.$$

This can be done as follows:

1. Let $Y = UX$.
2. Solve the system $LY = B$ for Y.
3. Solve the system $UX = Y$ for X.

Observe that, in this case we have

$$\begin{aligned} B &= LY \\ &= LUX \\ &= AX. \end{aligned}$$

In other words, X is the desired solution.

Example To solve

$$\begin{bmatrix} 1 & 2 & 4 & 1 \\ 2 & 8 & 6 & 4 \\ 3 & 10 & 8 & 8 \\ 4 & 12 & 10 & 6 \end{bmatrix} X = \begin{bmatrix} 21 \\ 52 \\ 79 \\ 82 \end{bmatrix}$$

we begin by solving

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 \\ 4 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 21 \\ 52 \\ 79 \\ 82 \end{bmatrix}.$$

The first equation gives

$$y_1 = 21,$$

the next ones give

$$\begin{aligned} y_2 &= 52 - 2y_1 = 10, \\ y_3 &= 79 - 3y_1 - y_2 = 6, \\ y_4 &= 82 - 4y_1 - y_2 - 2y_3 = -24. \end{aligned}$$

Next we solve

$$\begin{bmatrix} 1 & 2 & 4 & 1 \\ 0 & 4 & -2 & 2 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 21 \\ 10 \\ 6 \\ -24 \end{bmatrix}$$

for X using the usual backword sweep. We get

$$\begin{aligned} x_4 &= -4, \\ x_3 &= (6 - 3x_4)\,/ - 2 = 3, \\ x_2 &= (10 + 2x_3 - 2x_4)\,/4 = 2, \\ x_1 &= 21 - 2x_2 - 4x_3 - x_4 = 1. \end{aligned}$$

1.1 Elementary Gaussian matrices

It turns out that in order to carry out a Gaussian elemination step on a matrix A, one can do the same operation first on the identity matrix and then multiply the resulting matrix on the right by A. For example, in order to perform the Gaussian elemination step: $r_j \leftarrow r_j + m_{ji} r_i$, one starts with the identity

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{row } i \\ \\ \leftarrow \text{row } j \\ \\ \\ \end{matrix},$$

performs the operation $r_j \leftarrow r_j + m_{ji} r_i$ to get matrix

$$E = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & m_{ji} & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{row } i \\ \\ \leftarrow \text{row } j \\ \\ \\ \end{matrix}$$

and then compute EA. To see this write A in rows as

$$A = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_i \\ \vdots \\ r_j \\ \vdots \\ r_n \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{row } i \\ \\ \leftarrow \text{row } j \\ \\ \\ \end{matrix}$$

and perforn the multiplication EA:

$$\begin{aligned} EA &= \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & m_{ji} & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_i \\ \vdots \\ r_j \\ \vdots \\ r_n \end{bmatrix} \\ &= \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_i \\ \vdots \\ r_j + m_{ji} r_i \\ \vdots \\ r_n \end{bmatrix}. \end{aligned}$$

Notice that, if $j > i$, the matrix $E = E(m_{ji})$ is lower trianglular, has ones on the main diagonal, has determinant one and has inverse $E^{-1} = E^{-1}(m_{ji}) = E(-m_{ji})$. This is becuase to undo the operation of adding $m_{ji}r_i$ to r_j ($r_j \leftarrow r_j + m_{ji}r_i$) we need to subtract $m_{ji}r_i$ from r_j ($r_j \leftarrow r_j - m_{ji}r_i$). This can also be verified by direct computation:

$$E(m_{ji})\,E(-m_{ji}) = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & m_{ji} & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & -m_{ji} & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} = I$$

One more important property of the matrix E is that when we multiply two such matrices $E(m_{ij}), E(m_{kl})$, $(i,j) \neq (k,l)$ we get a matrix $E(m_{ij}, m_{kl})$, which is the identity matix with m_{ij} in the (i,j) place and m_{kl} in the (k,l) place.

Example

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

The other operation that concerns us here is the interchanging of two rows of a matrix A. Again, this can be done by interchanging the same rows in the identity matrix and then multiplying the result by A. The identity with rows (i,j) interchanged will be called an elementary permutation matrix and denoted by $P(i,j)$. It is easy to see that

$$P(i,j)^{-1} = P(j,i) = P(i,j)^t,$$

where $(\cdot)^t$ denotes the transpose of a matrix.

Definition 2 *A matrix P is called a permutation matrix if it can be written in the form*

$$P = P_1 P_2 \cdots P_m$$

where P_l , $l = 1, 2, \ldots, m$ is an elementary permutation matrix.

Theorem 3 *If P is a permutation matrix then it is nonsingular and $P^{-1} = P^t$.*

1.2 Finding the LU factorization of a Matrix

With the help of the elementary Gaussian matrices, we can find the LU factorization of a matrix A as in the following theorem

Theorem 4 *Suppose A is a nonsingular matrix and suppose the Gaussian elemination operations taken to reduce A to upper triangular form do not involve row interchanges. Then A has an LU factorization.*

Proof. Assume the Gaussian elemination operations produce the sequence of matrices

$$A \to A^{(1)} \to A^{(2)} \to \cdots \to A^{(k)}$$

with $A^{(k)}$ upper triangular. Then

$$A^{(k)} = E_k A^{(k-1)},$$

where

$$E_k = E\left(-m_{n,n-1}\right).$$

Therefore,

$$A^{(k-1)} = E_k^{-1} A^{(k)}.$$

Similarly, since no row interchanges are needed,

$$\begin{aligned} A^{(k-2)} &= E_{k-1}^{-1} A^{(k-1)} \\ &= E_{k-1}^{-1} E_k^{-1} A^{(k)}, \\ &\vdots, \\ A &= E_1^{-1} E_2^{-1} \cdots E_k^{-1} A^{(k)}. \end{aligned}$$

Thus we have

$$A = LU,$$

with

$$L = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$

and

$$U = A^{(k)}.$$

■

Clearly, if A is singular then it does not have an LU factorization.

Theorem 5 *If A is nonsingular, then there exists a permutation matrix P such that PA has an LU factorization.*

In the general case then in order to solve the system $AX = B$, we begin by multiplying both sides by P to get

$$PAX = PB.$$

since PA has an LU factorization,

$$LUX = PB.$$

Thus, in order to slve the system $AX = B$ we proceed as follows:

1. Compute the matrices P, L, U.
2. Solve the system $LY = PB$ for Y
3. Solve the system $UX = Y$ for X.

1.3 Storage Organization

We can save the LU factorization of the matrix A in the same array by replacing the zeros of the lower triangular part of U by the lower triangular elements of L. The resulting array has the form

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ l_{21} & u_{22} & u_{23} & \cdots & u_{2n} \\ l_{31} & l_{32} \vdots & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & u_{nn} \end{bmatrix}.$$

Observe that the elements l_{ij} are simply the multipliers m_{ij} used in the Gauss elemination steps. They are found and stored successfully as they are produced in place of the zeros produced during the elimination steps.

1.4 Operations count

To find the LU factorization of a matrix A, one uses the three loops:

For $i = 1 : n - 1$
  For $r = i + 1 : n$
    $m_r = a_{ri}/a_{ii}$
    For $j = r + 1 : n$
      $a_{rj} \leftarrow a_{rj} - m_r a_{ij}$
    end
  end
end

We want to count the number of multiplications and additions requried to execute these loops. The innermost loop requires one multiplication μ and on addition α per loop. Therefore, each time the innermost loop is executed, it takes $\mu + \alpha$ additions and multiplications to execute. The innermost loop is repeated $n - (r + 1) + 1 = (n - r)$ times. Each time the middle loop executes it requires one multiplication to compute m_r and $(\mu + \alpha)(n - r)$ additions and multiplication to execute its inner loop. Thus, the number of iteration for each execution of the middle loop is $(\mu + \alpha)(n - r) + \mu$. Summing up over the number of executions of the middle loop, we have the number of operations

$$
\begin{aligned}
& \sum_{r=i+1}^{n} (\mu+\alpha)(n-r)+\mu \\
= \;& (\mu+\alpha)\sum_{r=i+1}^{n}(n-r)+\mu(n-i) \\
= \;& (\mu+\alpha)\left[n((n-i))-\sum_{r=i+1}^{n} r\right]+\mu(n-i) \\
= \;& (\mu+\alpha)\left[n(n-i)-\frac{(n-i)(n+i+1)}{2}\right]+\mu(n-i) \\
= \;& (\mu+\alpha)\frac{(n-i)(n-i-1)}{2}+\mu(n-i) \\
= \;& \mu\frac{(n-i)(n-i+1)}{2}+\alpha\frac{(n-i)(n-i-1)}{2}.
\end{aligned}
$$

Then we sum up over the number of executions of the outermost loop to compute the total number N of operatons.

$$
\begin{aligned}
N &= \sum_{i=1}^{n-1}\mu\frac{(n-i)(n-i+1)}{2}+\alpha\frac{(n-i)(n-i-1)}{2} \\
&= \mu\frac{n^3-n}{6}+\alpha\frac{n(n-1)(n-2)}{6}.
\end{aligned}
$$

Thus, the LU factorization of a matrix A requires $\frac{n^3-n}{6}$ multiplications and $\frac{n(n-1)(n-2)}{6}$ additions. Both terms are $O\left(n^3\right)$. In a similar way we can show that the forward or backword sweeps for solving a linear system with the LU factorization requires $O\left(n^2\right)$ operations.

Observe that solving a linear system $AX=B$ by $B=A^{-1}X$, that is by computing the inverse of the matrix A rquires $O(n!)$ multiplications and additions. Therefore, solving a system of equations by the LU factorization is a lot less costly than computing the inverse of a matrix. When Gaussian elemination is used to find the inverse of a marix by solving repeatedly linear systems, the cost is $O\left(n^4\right)$ operations.