

Mathematical Tripos Part IB: Lent Term 2024

Numerical Analysis – Lecture 12

12 LU factorization

12.1 Introduction

Problem 12.1 The basic problem of numerical analysis is: solve

$$A\mathbf{x} = \mathbf{b}. \quad (12.1)$$

The formula

$$\det A = \sum_i (-1)^{\sigma(i)} a_{1,i_1} a_{2,i_2} \dots a_{n,i_n}$$

gives an opportunity to solve (12.1) explicitly by Cramer's rule with the small detail that it costs $n!$ multiplications. Given a computer with 10^9 flops/sec, this way of solving an $n \times n$ system will take

$$\text{a) } n = 10, \quad t = 10^{-4} \text{ sec}, \quad \text{b) } n = 20, \quad t = 17 \text{ min}, \quad \text{c) } n = 30, \quad t = 400\,000 \text{ years}.$$

Thus we have to look at methods from a more constructive and practical point of view.

Example 12.2 (Triangular matrices) An *upper triangular* matrix U has $u_{ij} = 0$ if $i > j$. Then $\det U = \prod u_{ii}$. Also, we can solve $U\mathbf{x} = \mathbf{y}$ by the so-called *back-substitution*

$$x_n = y_n / u_{nn}, \quad x_i = (y_i - \sum_{j=i+1}^n u_{ij} x_j) / u_{ii}, \quad i = n-1, \dots, 1.$$

Similarly, a *lower triangular* matrix L (s.t. $\ell_{ij} = 0$ if $i < j$) allows us to solve $L\mathbf{y} = \mathbf{b}$ directly by the *forward substitution*. Hence, if we manage to factorize A as

$$A = LU,$$

then solution of the system $A\mathbf{x} = \mathbf{b}$ can be split into two cases

$$A\mathbf{x} = L \underbrace{U\mathbf{x}}_{\mathbf{y}} = \mathbf{b} \Rightarrow \quad 1) \quad L\mathbf{y} = \mathbf{b}, \quad 2) \quad U\mathbf{x} = \mathbf{y}, \quad (12.2)$$

both being solved sufficiently easily, as we have seen.

Example 12.3 (Orthogonal matrices) An orthogonal matrix Q satisfies $Q^T Q = I$, i.e., $Q^{-1} = Q^T$, so that a solution to $Q\mathbf{x} = \mathbf{y}$ is simply $\mathbf{x} = Q^T \mathbf{y}$. If

$$A = QR$$

where Q is orthogonal and R is upper triangular, then again one can solve the system $A\mathbf{x} = \mathbf{b}$ in two simple steps

$$A\mathbf{x} = QR\mathbf{x} = \mathbf{b} \Rightarrow \quad 1) \quad Q\mathbf{y} = \mathbf{b}, \quad 2) \quad R\mathbf{x} = \mathbf{y}.$$

12.2 LU factorization

Definition 12.4 By the LU factorization of a (nonsingular) matrix A we understand a representation of A as a product

$$A = LU,$$

where L is a lower triangular matrix with unit diagonal and U is a nonsingular upper triangular matrix.

Remark 12.5 Nonsingularity of U (and A) is a natural requirement in all practical applications. Theoretically it is not necessary; see Question 1 on Examples Sheet 3.

Method 12.6 We present a method of the LU factorization that is based on the direct approach. If $A = LU$, and ℓ_k and u_k^T are the columns and the rows of the matrices L and U respectively, then

$$A = \begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & \ddots & \\ \hline & & & \\ \hline \ell_1 & \ell_2 & & \ell_n \\ \hline \end{array} \times \begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & \ddots & \\ \hline & & & \\ \hline u_1^T & u_2^T & & u_n^T \\ \hline \end{array} = \sum_{k=1}^n \ell_k u_k^T.$$

Since the leading $k - 1$ elements of ℓ_k and u_k^T are zero for $k \geq 2$, it follows that the first $k - 1$ columns and rows of the matrix $\ell_k u_k^T$ are zeros as well. Hence u_1^T , which is the 1-st row of the matrix $\ell_1 u_1^T$, coincides with the 1-st row of A , while $\ell_1 \cdot u_{11} = \ell_1 \cdot a_{11}$ coincides with the 1-st column of A . Subtracting and letting

$$A^{(1)} := A - \ell_1 u_1^T$$

we obtain similarly that u_2^T and $\ell_2 \cdot a_{22}^{(1)}$ are the 2-nd row and the 2-nd column of $A^{(1)}$ respectively, so that we proceed further with

$$A^{(2)} := A^{(1)} - \ell_2 u_2^T,$$

and so on.

Algorithm 12.7 Notice that we can use the matrix A itself to store the elements of L and U , in the subdiagonal and in the upper triangular portions of A , respectively. So, the previous method allows computing of the LU factorization as follows.

```

for k=1 to n-1 do
  for i=k+1 to n do
    a(i,k) := a(i,k)/a(i,i);    #-- l(i,k) := a(i,k)/a(i,i)
    for m=k+1 to n do          #-- u(k,m) := a(k,m)
      a(i,m) := a(i,m) - a(i,k)*a(k,m)
      #-- a(i,m) := a(i,m) - l(i,k)*u(k,m)
    end
  end
end
end

```

Example 12.8

$$\begin{bmatrix} 2 & -1 & -3 & 3 \\ 4 & 0 & -3 & 1 \\ 6 & 1 & -1 & 6 \\ -2 & -5 & 4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -1 & -3 & 3 \\ 2 & 2 & 3 & -5 \\ 3 & 4 & 8 & -3 \\ -1 & -6 & 1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -1 & -3 & 3 \\ 2 & 2 & 3 & -5 \\ 3 & 2 & 2 & 7 \\ -1 & -3 & 10 & -11 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -1 & -3 & 3 \\ 2 & 2 & 3 & -5 \\ 3 & 2 & 2 & 7 \\ -1 & -3 & 5 & -46 \end{bmatrix}$$

Application 12.9 One can use the LU factorization in the following problems.

1. Solution of linear systems (see (12.2)).
2. Calculation of the determinant: $\det A = \det L \det U = \prod u_{ii}$.
3. Calculation of the inverse of A : $A^{-1} = U^{-1}L^{-1}$.

The cost. It is only multiplications and divisions that matter. At the k -th step of the LU algorithm we perform $(n - k)$ divisions to determine the components of ℓ_k and $(n - k)^2$ multiplications for finding $\ell_k u_k^T$. Hence

$$N_{LU} = \sum_{k=1}^{n-1} [(n - k)^2 + (n - k)] = \frac{1}{3}n^3 + \mathcal{O}(n^2).$$

Solving $Ly = b$ by forward substitution we use k multiplications/divisions to determine y_k , and similarly for back substitutions, thus

$$N_F = N_B \sim \frac{1}{2}n^2.$$

Remark 12.10 At the k -th step of the LU-algorithm the multipliers l_{ik} are chosen so that the outcome of the operation $A^{(k)} = A^{(k-1)} - \ell_k \mathbf{u}_k^T$ is that the k th column of $A^{(k)}$ is zero. This is equivalent to Gaussian elimination of the unknown x_k from the system $A\mathbf{x} = \mathbf{b}$. The only yet important difference between LU and Gaussian elimination is that we do not consider the right hand side $\mathbf{b}$ until the factorization is complete. This is useful e.g. when there are many right hand sides, in particular if not all the $\mathbf{b}$'s are known at the outset (as, say, in the multigrid method). In Gaussian elimination the solution for each new $\mathbf{b}$ would require $\mathcal{O}(n^3)$ computational operations, whereas with LU factorization $\mathcal{O}(n^3)$ operations are required for the initial factorization, but then the solution for each new $\mathbf{b}$ only requires $\mathcal{O}(n^2)$ operations (for back- and forward-substitutions).

12.3 Pivoting

The LU algorithm fails if $a_{kk}^{(k-1)} = 0$. A remedy is to exchange rows of $A^{(k-1)}$ by picking a suitable *pivotal equation*, i.e., the equation which is used to eliminate one unknown from certain of the other equation.

This technique is called *column pivoting* and it means that, having obtained $A^{(k-1)}$, we exchange two rows of $A^{(k-1)}$ so that the element of largest magnitude in the k -th column is in the *pivotal position* (k, k) . In other words,

$$|a_{k,k}^{(k-1)}| = \max_{j=k \dots n} |a_{j,k}^{(k-1)}|.$$

The exchange of rows k and m can be regarded as the pre-multiplication of the relevant matrix $A^{(k-1)}$ by a permutation matrix P_{km} . Since

$$P_{km}A^{(k-1)} = P_{km}A - \sum_{i=1}^{k-1} (P_{km}\ell_i)\mathbf{u}_i^T,$$

the same exchange is required in the portion of L that has been formed already (i.e., in the first $k-1$ columns). Also, we need to record the permutation of rows to solve for the right-hand side and/or to compute the determinant.

Example 12.11

$$\begin{bmatrix} 2 & 1 & 1 \\ \bullet 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \xrightarrow{\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}} \begin{bmatrix} 4 & 1 & 0 \\ 2 & 1 & 1 \\ -2 & 2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 1 \\ -\frac{1}{2} & \bullet \frac{5}{2} & 1 \end{bmatrix} \xrightarrow{\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}} \begin{bmatrix} 4 & 1 & 0 \\ -\frac{1}{2} & \frac{5}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 1 & 0 \\ -\frac{1}{2} & \frac{5}{2} & 1 \\ \frac{1}{2} & \frac{1}{5} & \frac{4}{5} \end{bmatrix},$$

$$PA = LU \Rightarrow A = P^T LU, \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ \frac{1}{2} & \frac{1}{5} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 4 & 1 & 0 \\ & \frac{5}{2} & 1 \\ & & \frac{4}{5} \end{bmatrix}.$$

An important advantage of column pivoting is that every element of L has magnitude at most one. This avoids not just division by zero (what is not the case in the previous example) but also tends to reduce the chance of very large numbers occurring during the factorization, a phenomenon that might lead to *ill conditioning* and to accumulation of *round-off error*.

12.4 Examples

Example 12.12 (LU factorization)

$$\begin{aligned}
 & \begin{array}{l} 2^{\text{nd}} - 1^{\text{st}} \cdot (-2) \\ 3^{\text{rd}} - 1^{\text{st}} \cdot 3 \\ 4^{\text{th}} - 1^{\text{st}} \cdot (-4) \\ \rightarrow \end{array} \begin{bmatrix} -3 & 2 & 3 & -1 \\ 6 & -2 & -6 & 0 \\ -9 & 4 & 10 & 3 \\ 12 & -4 & -13 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} -3 & 2 & 3 & -1 \\ -2 & 2 & 0 & -2 \\ 3 & -2 & 1 & 6 \\ -4 & 4 & -1 & -9 \end{bmatrix} \\
 & \begin{array}{l} 3^{\text{rd}} - 2^{\text{nd}} \cdot (-1) \\ 4^{\text{th}} - 2^{\text{nd}} \cdot 2 \\ \rightarrow \end{array} \begin{bmatrix} -3 & 2 & 3 & -1 \\ -2 & 2 & 0 & -2 \\ 3 & -1 & 1 & 4 \\ -4 & 2 & -1 & -5 \end{bmatrix} \xrightarrow{4^{\text{th}} - 3^{\text{rd}} \cdot (-1)} \begin{bmatrix} -3 & 2 & 3 & -1 \\ -2 & 2 & 0 & -2 \\ 3 & -1 & 1 & 4 \\ -4 & 2 & -1 & -1 \end{bmatrix}
 \end{aligned}$$

i.e.

$$A = LU, \quad L = \begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 3 & -1 & 1 & \\ -4 & 2 & -1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} -3 & 2 & 3 & -1 \\ & 2 & 0 & -2 \\ & & 1 & 4 \\ & & & -1 \end{bmatrix}$$

Example 12.13 (Forward and back substitutions) Then the solution to the system

$$Ax = b, \quad b = [3, -2, 2, 0]^T$$

proceeds in two steps

$$Ax = L \underbrace{Ux}_y = b \Rightarrow 1) \quad Ly = b, \quad 2) \quad Ux = y.$$

1) Forward substitution

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 3 & -1 & 1 & \\ -4 & 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 2 \\ 0 \end{bmatrix} \begin{matrix} \downarrow \\ \downarrow \\ \downarrow \\ \downarrow \end{matrix} \Rightarrow y = \begin{bmatrix} 3 \\ 4 \\ -3 \\ 1 \end{bmatrix},$$

2) Back substitution

$$\begin{bmatrix} -3 & 2 & 3 & -1 \\ & 2 & 0 & -2 \\ & & 1 & 4 \\ & & & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ -3 \\ 1 \end{bmatrix} \begin{matrix} \uparrow \\ \uparrow \\ \uparrow \\ \uparrow \end{matrix} \Rightarrow x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix},$$

Example 12.14 (LU factorization with pivoting)

$$\begin{aligned}
 & \begin{bmatrix} -3 & 2 & 3 & -1 \\ 6 & -2 & -6 & 0 \\ -9 & 4 & 10 & 3 \\ \bullet 12 & -4 & -13 & -5 \end{bmatrix} \xrightarrow{\begin{bmatrix} 4 \\ 2 \\ 3 \\ 1 \end{bmatrix}} \begin{bmatrix} 12 & -4 & -13 & -5 \\ 6 & -2 & -6 & 0 \\ -9 & 4 & 10 & 3 \\ -3 & 2 & 3 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 12 & -4 & -13 & -5 \\ -\frac{1}{2} & 0 & \frac{1}{2} & \frac{5}{2} \\ -\frac{3}{4} & \bullet 1 & \frac{1}{4} & -\frac{3}{4} \\ -\frac{1}{4} & 1 & -\frac{1}{4} & -\frac{9}{4} \end{bmatrix} \\
 & \xrightarrow{\begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}} \begin{bmatrix} 12 & -4 & -13 & -5 \\ -\frac{3}{4} & 1 & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{5}{2} \\ -\frac{1}{4} & 1 & -\frac{1}{4} & -\frac{9}{4} \end{bmatrix} \rightarrow \begin{bmatrix} 12 & -4 & -13 & -5 \\ -\frac{3}{4} & 1 & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{5}{2} \\ -\frac{1}{4} & 1 & -\frac{1}{4} & -\frac{9}{4} \end{bmatrix} \rightarrow \begin{bmatrix} 12 & -4 & -13 & -5 \\ -\frac{3}{4} & 1 & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{5}{2} \\ -\frac{1}{4} & 1 & -1 & 1 \end{bmatrix}
 \end{aligned}$$

i.e.

$$A = P^T LU, \quad P = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{bmatrix}, \quad L = \begin{bmatrix} 1 & & & \\ -\frac{3}{4} & 1 & & \\ \frac{1}{2} & 0 & 1 & \\ -\frac{1}{4} & 1 & 1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 12 & -4 & -13 & -5 \\ & 1 & \frac{1}{4} & -\frac{3}{4} \\ & & \frac{1}{2} & \frac{5}{2} \\ & & & 1 \end{bmatrix}$$