

Mathematical Tripos Part IB: Lent Term 2024

Numerical Analysis – Lecture 13

13 LU factorization (cont.)

13.1 Existence and uniqueness of the LU factorization

Definition 13.1 A square matrix A is called *strictly regular* if all its leading submatrices $A_k = (a_{ij})_{i,j=1}^k$ (the matrix A itself as well) are nonsingular.

Theorem 13.2 (Existence) A matrix A admits an LU factorization $\Leftrightarrow$ it is strictly regular.

Proof. ($\Rightarrow$) This part follows from the scheme (where we partition matrices into blocks):

$$\left[\begin{array}{c|c} A_k & B_{k,n-k} \\ \hline C_{n-k,k} & D_{n-k} \end{array} \right] = A = LU = \left[\begin{array}{c|c} L_k & O_{k,n-k} \\ \hline M_{n-k,k} & L'_{n-k} \end{array} \right] \cdot \left[\begin{array}{c|c} U_k & V_{k,n-k} \\ \hline O_{n-k,k} & U'_{n-k} \end{array} \right].$$

Multiplying blocks we obtain $A_k = L_k U_k$, and since L_k, U_k are nonsingular, so is A_k .

($\Leftarrow$) We use induction. For $n = 1$ the result is trivial. Assume the result for $(n-1) \times (n-1)$ matrices. Partition the $n \times n$ matrix A as

$$A = \left[\begin{array}{c|c} A_{n-1} & \mathbf{w} \\ \hline \mathbf{v}^T & a_{nn} \end{array} \right].$$

Let us show that we can get

$$A = LU \quad \text{with} \quad L = \left[\begin{array}{c|c} L_{n-1} & \mathbf{o} \\ \hline \mathbf{x}^T & 1 \end{array} \right], \quad U = \left[\begin{array}{c|c} U_{n-1} & \mathbf{y} \\ \hline \mathbf{o}^T & u_{nn} \end{array} \right],$$

where $L_{n-1}, U_{n-1}, \mathbf{x}^T, \mathbf{y}$ and u_{nn} are to be determined. Multiplying out these block matrices we see that we want to have

$$A = \left[\begin{array}{c|c} A_{n-1} & \mathbf{w} \\ \hline \mathbf{v}^T & a_{nn} \end{array} \right] = \left[\begin{array}{c|c} L_{n-1}U_{n-1} & L_{n-1}\mathbf{y} \\ \hline \mathbf{x}^T U_{n-1} & \mathbf{x}^T \mathbf{y} + 1 \cdot u_{nn} \end{array} \right].$$

In virtue of induction assumption, L_{n-1} and U_{n-1} exist and are non-singular. Then $L_{n-1}\mathbf{y} = \mathbf{w}$, $\mathbf{x}^T U_{n-1} = \mathbf{v}^T$ can be solved to get

$$\mathbf{y} = L_{n-1}^{-1} \mathbf{w}, \quad \mathbf{x}^T = \mathbf{v}^T U_{n-1}^{-1}, \quad u_{nn} = a_{nn} - \mathbf{x}^T \mathbf{y}.$$

So, there exists a factorization $A = LU$ with $\ell_{ii} = 1$, but we need also U to be nonsingular, which it is because $\det U = \det A$ and $\det A \neq 0$ by strict regularity.

Theorem 13.3 (Uniqueness) For a strictly regular matrix A , its LU factorization is unique, i.e.

$$A = L_1 U_1 = L_2 U_2 \quad \text{implies} \quad L_1 = L_2, \quad U_1 = U_2.$$

Proof. The equality $L_1 U_1 = L_2 U_2$ implies $L_2^{-1} L_1 = U_2 U_1^{-1} (= V, \text{ say})$. The inverse of a lower triangular matrix and the product of such matrices are lower triangular matrices. The same is true for upper triangular matrices. Consequently, V is simultaneously lower and upper triangular, hence it is diagonal. Since $L_2^{-1} L_1$ has unit diagonal, we obtain $V = I$, hence the statement. $\square$

Remark 13.4 Both theorems are for the LU factorization with a *nonsingular* U . If we required only $\ell_{ii} = 1$, then *some* singular matrices A may be LU factorized, and in many ways:

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

(See Question 3 about LU factorization with arbitrary U). There is no LU factorization (whatsoever) for regular but not strictly regular matrices, e.g. for $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. On the other hand, *every* regular matrix A admits an LU factorization with pivoting, i.e. for every nonsingular matrix A there exists a permutation matrix P such that $PA = LU$.

Corollary 13.5 A strictly regular matrix A has a unique factorization $A = LDU$ where both L and U have unit diagonal.

Corollary 13.6 A strictly regular symmetric matrix A admits the unique representation $A = LDL^T$.

Proof. By strict regularity, $A = LDU$, and by symmetry,

$$LDU = A = A^T = U^T D L^T.$$

Since the LDU factorization is unique, $U = L^T$ □

13.2 Symmetric positive definite and diagonally dominant matrices

Here we consider two important classes of matrices which are strictly regular hence possess an LU factorization.

Definition 13.7 A matrix A is called (symmetric) *positive definite* (SPD-matrix) if $\mathbf{x}^T A \mathbf{x} > 0$ for all $\mathbf{x} \neq 0$ and $A = A^T$.

Theorem 13.8 Let A be a real $n \times n$ symmetric matrix. It is positive definite if and only if it has the LDL^T factorization in which the diagonal elements of D are all positive.

Proof. Suppose that $A = LDL^T$ with $D > 0$, and let $\mathbf{x} \in \mathbb{R}^n \setminus 0$. Since L is nonsingular, $\mathbf{y} = L^T \mathbf{x} \neq 0$. Then $\mathbf{x}^T A \mathbf{x} = \mathbf{y}^T D \mathbf{y} = \sum_{k=1}^n d_{kk} y_k^2 > 0$, hence A is positive definite.

Conversely, if A is (symmetric) positive definite, then it is strictly regular. Indeed, if $A_k \mathbf{x} = 0$ for some $k \leq n$ and some $\mathbf{x} \in \mathbb{R}^k$, then for $\mathbf{x}_* = (x_1, \dots, x_k, 0, \dots, 0) \in \mathbb{R}^n$ we obtain $\mathbf{x}_*^T A \mathbf{x}_* = \mathbf{x}^T A_k \mathbf{x} = 0$, a contradiction to the positive definiteness of A . Thus, by Corollary 13.6, it admits an LDL^T factorization. Take $\mathbf{x}$ such that $L^T \mathbf{x} = \mathbf{e}_k = [0 \dots 1 \dots 0]^T$. Then $0 < \mathbf{x}^T A \mathbf{x} = \mathbf{e}_k^T D \mathbf{e}_k = d_{kk}$. □

Method 13.9 One can check if a symmetric matrix is positive definite by trying to form its LDL^T factorization.

Example 13.10 The matrix below is positive definite.

$$\begin{bmatrix} 2 & 6 & -2 \\ 6 & 21 & 0 \\ -2 & 0 & 16 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 6 & -2 \\ -3 & 3 & 6 \\ -1 & 6 & 14 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 6 & -2 \\ -3 & 3 & 6 \\ -1 & 2 & 2 \end{bmatrix} \Rightarrow L = \begin{bmatrix} 1 & & \\ 3 & 1 & \\ -1 & 2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & & \\ & 3 & \\ & & 2 \end{bmatrix}.$$

The only application of this method is during exam. Normally, we know whether the matrix in question is positive definite by some a priori reasons. Say, a Gram matrix with $a_{ij} = [\mathbf{x}^{(i)}]^T \mathbf{x}^{(j)}$, where vectors $\mathbf{x}^{(i)}$ are linearly independent, is positive definite.

Corollary 13.11 (Cholesky factorization) A positive definite matrix A admits the Cholesky factorization

$$A = \tilde{L} \tilde{L}^T,$$

where $\tilde{L}$ is a lower triangular matrix.

Proof. Since $A = LDL^T$ with a positive diagonal matrix D , we can write

$$A = LDL^T = (LD^{1/2})(D^{1/2}L^T) =: \tilde{L}\tilde{L}^T.$$

Definition 13.12 A matrix A is called *strictly diagonally dominant* (by rows) if

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad \text{for all } i, \quad \text{e.g.} \quad A = \begin{bmatrix} 6 & -2 & 1 \\ 1 & 4 & 2 \\ 2 & 1 & -5 \end{bmatrix}.$$

Theorem 13.13 If A is strictly diagonally dominant (by rows), then it is strictly regular, hence the LU factorization exists.

Proof. Take any $\mathbf{x} = (x_1, \dots, x_n) \neq 0$ and let x_i be its largest absolute value component, i.e., $|x_i| \geq |x_j|$. Then, for the i -th component of $A\mathbf{x}$, the strict diagonal dominance gives the value

$$|(A\mathbf{x})_i| = |a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j| \geq |a_{ii}||x_i| - \underbrace{\sum_{j \neq i} |a_{ij}|}_{< |a_{ii}|} \underbrace{|x_j|}_{\leq |x_i|} > 0.$$

Hence $A\mathbf{x} \neq 0$ for any $\mathbf{x}$, i.e. A is nonsingular. The leading submatrices of a strictly diagonally dominant matrix are (even more) strictly diagonally dominant, hence the strict regularity of A . $\square$

13.3 Sparse and band matrices

Definition 13.14 A matrix $A \in \mathbb{R}^n$ is called a *sparse* matrix if nearly all elements of A are zero. Most useful examples are band matrices and block band matrices

Definition 13.15 A matrix A is called a *band matrix* if there exists an integer $r < n$ such that

$$a_{ij} = 0 \quad \text{for all } |i - j| > r.$$

In other words, all nonzero elements of A reside in a band of width $2r+1$ along the main diagonal.

$$r = 1 \quad \begin{bmatrix} * & * & & & \\ * & * & * & & \\ & * & * & * & \\ & & * & * & * \\ & & & * & * \\ & & & & * \end{bmatrix}, \quad r = 2 \quad \begin{bmatrix} * & * & * & & \\ * & * & * & * & \\ & * & * & * & * \\ & & * & * & * \\ & & & * & * \\ & & & & * \end{bmatrix}, \quad r = 3 \quad \begin{bmatrix} * & * & * & * & \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}.$$

It is frequently required to solve *very* large systems $A\mathbf{x} = \mathbf{b}$ with sparse matrices (say, $n = 10^5$ is considered small in this context). The LU factorization of such A makes sense only if L and U inherit much of the sparsity of A (so that the cost of computing $U\mathbf{x}$, say, is comparable with that of $A\mathbf{x}$). To this end the following theorem is useful.

Theorem 13.16 Let $A = LU$ be the LU factorization (without pivoting) of a sparse matrix. Then

- 1) all leading zeros in the rows of A to the left of diagonal are inherited by L ,
- 2) all leading zeros in the columns of A above the diagonal are inherited by U .

$$\begin{bmatrix} * & \bullet & \bullet & \bullet & \bullet \\ \circ & * & & & \\ \circ & \circ & * & & \\ & \circ & \circ & * & \\ \circ & \circ & \circ & \circ & * \\ \circ & \circ & \circ & \circ & \circ & * \end{bmatrix} = \begin{bmatrix} * & & & & \\ \circ & * & & & \\ \circ & \circ & * & & \\ \circ & \circ & \circ & * & \\ \circ & \circ & \circ & \circ & * \end{bmatrix} \times \begin{bmatrix} * & \bullet & \bullet & \bullet & \bullet \\ & * & & & \\ & & * & & \\ & & & * & \\ & & & & * \end{bmatrix}$$

Proof. Let $a_{i,1} = a_{i,2} = \dots = 0$ be the leading zeros in the i -th row. Then, since $u_{kk} \neq 0$, we obtain

$$\begin{aligned} 0 = a_{i,1} &= \ell_{i,1}u_{11} && \Rightarrow \ell_{i,1} = 0, \\ 0 = a_{i,2} &= \ell_{i,1}u_{12} + \ell_{i,2}u_{22} && \Rightarrow \ell_{i,2} = 0, \\ 0 = a_{i,3} &= \ell_{i,1}u_{13} + \ell_{i,2}u_{23} + \ell_{i,3}u_{33} && \Rightarrow \ell_{i,3} = 0, \quad \text{and so on.} \end{aligned}$$

Similarly for the leading zeros in the j -th column.

Corollary 13.17 If A is a band matrix and $A = LU$, then $\ell_{ij} = u_{ij} = 0$ for all $|i - j| > r$, i.e. L and U are the band matrices with the same band width as A .

The cost (Question 5). In the case of a banded A , we need just

- 1) $\mathcal{O}(nr^2)$ operations to factorize and
- 2) $\mathcal{O}(nr)$ operations to solve a linear system (after factorization).

This must be compared with $\mathcal{O}(n^3)$ and $\mathcal{O}(n^2)$ operations respectively for an $n \times n$ dense matrix. If $r \ll n$ this represents a very substantial saving!

Method 13.18 Theorem 13.16 suggests that for a factorization of a sparse but not nicely structured matrix A one might try to reorder its rows and columns by a preliminary calculation so that many of the zero elements become leading zero elements in rows and columns, thus reducing the fill-in in L and U .

Example 13.19 The LU factorization of

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

has significant fill-in. However, exchanging the first and the last rows and columns yields

$$PAP = \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ 1 & 1 & 1 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}.$$