Mathematical Tripos Part IB: Easter Term 2003

Numerical Analysis – Lecture 2

2 LU factorization (cont.)

2.1 Existence and uniqueness of the LU factorization

Definition 2.1 A square matrix A is called *strictly regular* if all its leading submatrices A_k (A itself as well) are nonsingular.

Theorem 2.2 (Existence) *A matrix A admits a LU factorization $\Leftrightarrow$ it is strictly regular.*

Proof. ($\Rightarrow$) This part follows from the scheme

$$\begin{bmatrix} A_k & B_{k,n-k} \\ C_{n-k,k} & D_{n-k} \end{bmatrix} = A = LU = \begin{bmatrix} L_k & O_{k,n-k} \\ M_{n-k,k} & L'_{n-k} \end{bmatrix} \cdot \begin{bmatrix} U_k & V_{k,n-k} \\ O_{n-k,k} & U'_{n-k} \end{bmatrix} \quad \Rightarrow \quad A_k = L_k U_k.$$

Since L_k, U_k are nonsingular, so is A_k.

($\Leftarrow$) The LU factorization by Gaussian elimination runs according to the following scheme

$$A^{(1)} = L_1 A, \quad A^{(2)} = L_2 A^{(1)}, \quad \ldots, \quad A^{(k-1)} = L_{k-1} A^{(k-2)}, \ldots$$

so that with $L^{(k-1)} := L_{k-1} \cdots L_2 L_1$, after $k-1$ steps we obtain

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

The partition as above gives

$$\begin{bmatrix} A_k^{(k-1)} & B_{k,n-k}^{(k-1)} \\ C_{n-k,k}^{(k-1)} & D_{n-k}^{(k-1)} \end{bmatrix} = A^{(k-1)} = L^{(k-1)} A = \begin{bmatrix} L_k^{(k-1)} & O_{k,n-k} \\ M_{n-k,k}^{(k-1)} & L_{n-k}^{\prime(k-1)} \end{bmatrix} \cdot \begin{bmatrix} A_k & B_{k,n-k} \\ C_{n-k,k} & D_{n-k} \end{bmatrix},$$

whence

$$A_k^{(k-1)} = L_k^{(k-1)} A_k.$$

Here $A_k^{(k-1)}$ is a $k \times k$ upper triangular matrix (which in fact coincides with U_k), and we can proceed further with the algorithm if $a_{kk}^{(k-1)} \neq 0$, i.e., if $\det A_k^{(k-1)} \neq 0$. But $L_k^{(k-1)}$ is lower triangular with unit diagonal and A_k is nonsingular by assumption. Hence $A_k^{(k-1)}$ is nonsingular and we can perform the next step. ∎

Theorem 2.3 (Uniqueness) *The LU factorization is unique, i.e.*

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

Proof. The equality $L_1 U_1 = L_2 U_2$ implies $L_2^{-1} L_1 = U_2 U_1^{-1} (=: V$, 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 $V = I$. ∎

Corollary 2.4 *A strictly regular matrix A has a unique factorization $A = LDU$ where both L and U have unit diagonal.*

Corollary 2.5 *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, $U = L^T$ ∎

2.2 Symmetric positive definite and diagonally dominant matrices

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

Definition 2.6 A matrix A is called (symmetric) *positive definite* (SPD-matrix) if $x^T Ax > 0$ for all $x \neq 0$ and $A = A^T$.

Theorem 2.7 *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 $x \in \mathbb{R}^n \setminus 0$. Since L is nonsingular, $y := L^T x \neq 0$. Then $x^T Ax = y^T Dy = \sum_{k=1}^n d_{kk} y_k^2 > 0$, hence A is positive definite.

Conversely, if A is (symmetric) positive definite, then it is strictly regular. (For if $A_k x = 0$ for some $k \leq n$ and some $x \in \mathbb{R}^k$, then for $x_* = (x_1, \ldots, x_k, 0, \ldots, 0) \in \mathbb{R}^n$ we obtain $x_*^T Ax_* = x^T A_k x = 0$, a contradiction to the positive definiteness.) Thus, by Corollary 2.5 it admits an LDL^T factorization. Take x such that $L^T x = e_k = [0..1..0]^T$. Then $0 < x^T Ax = e_k^T De_k = d_{kk}$. ∎

Method 2.8 One can check if a symmetric matrix is positive definite by trying to form its LDL^T factorization. (The only application of this method is during examination hours.)

Example 2.9 The matrix below is positive definite.

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

Corollary 2.10 (Cholesky factorization) *A positive definite matrix A admits the Cholesky factorization*

$$A = \widetilde{L}\widetilde{L}^T$$

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

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

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

Definition 2.11 A matrix A is called *strictly diagonally dominant* (by rows) if

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

Theorem 2.12 *If A is strictly diagonally dominant (by rows), then it is srictly regular, hence the LU factorization exists.*

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

$$|(Ax)_k| = |a_{kk}x_k + \textstyle\sum_{j \neq k} a_{kj}x_j| \geq |a_{kk}|\,|x_k| - \underbrace{\textstyle\sum_{j \neq k} |a_{kj}|}_{< |a_{kk}|} \underbrace{|x_j|}_{\leq |x_k|} > 0.$$

Hence $Ax \neq 0$ for any 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. ∎

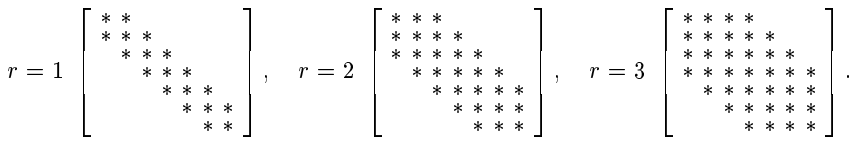
2.3 Sparse and band matrices

Definition 2.13 A matrix $A \in \mathbb{R}^n$ is called a *sparse* matrix if nearly all elements of A are zero. Most useful examples are banded matrices and block banded matrices

Definition 2.14 A matrix A is called a *band matrix* if there exists an integer $r < n$ such that

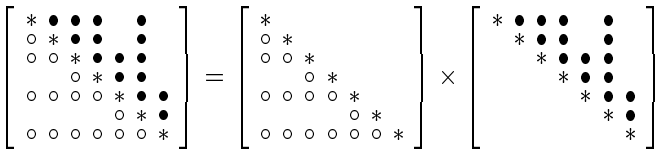
$$a_{ij} = 0 \quad \text{for all} \quad |i - j| > r.$$

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



It is frequently required to solve *very* large systems $Ax = b$ with sparse matrices ($n = 10^5$ 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 Ux, say, is comparable with that of Ax). To this end the following theorem is useful.

Theorem 2.15 *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 and*
2) *all leading zeros in the columns of A above the diagonal are inherited by U.*



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

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

Similarly for the leading zeros in the j-th column. Since $l_{kk} = 1$, it follows that

$$\begin{aligned}
0 = a_{1,j} &= l_{11}u_{1,j} && \Rightarrow \quad u_{1,j} = 0,\\
0 = a_{2,j} &= l_{21}u_{1,j} + l_{22}u_{2,j} && \Rightarrow \quad u_{2,j} = 0,\\
0 = a_{3,j} &= l_{31}u_{1,j} + l_{32}u_{2,j} + l_{33}u_{3,j} && \Rightarrow \quad u_{3,j} = 0, \quad \text{and so on.}
\end{aligned}$$

Corollary 2.16 *If A is a band matrix and $A = LU$, then $l_{ij} = u_{ij} = 0$ for all $|i - j| > r$, i.e. L and U are the band matrices with the same band width as A.*

The cost (Exercise 2.3). 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.
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 2.17 Theorem 2.15 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 2.18 The LU factorization of

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

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

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

2.4 Exercises

Exercises with a $*$-mark are *not* subject to supervision! They are just for fun!

2.1. Let A be a real $n \times n$ matrix that has the factorization $A = LU$, where L is lower triangular with ones on its diagonal and U is upper triangular. Prove that, for every integer $k \in \{1, 2, \ldots, n\}$, the first k rows of U span the same space as the first k rows of A. Prove also that the first k columns of A are in the k-dimensional subspace that is spanned by the first k columns of L. Hence deduce that no LU factorization of the given form exists if we have $\operatorname{rank} H_k < \operatorname{rank} B_k$, where H_k is the leading $k \times k$ submatrix of A and where B_k is the $n \times k$ matrix whose columns are the first k columns of A.

2.2. Calculate the Cholesky factorization of the matrix

$$\begin{bmatrix} 1 & 1 & & & & \\ 1 & 2 & 1 & & & \\ & 1 & 3 & 1 & & \\ & & 1 & 4 & 1 & \\ & & & 1 & 5 & 1 \\ & & & & 1 & \lambda \end{bmatrix}.$$

Deduce from the factorization the value of λ that makes the matrix singular. Also find this value of λ by seeking the vector in the null-space of the matrix whose first component is one.

2.3. Let A be an $n \times n$ nonsingular band matrix that satisfies the condition $a_{ij} = 0$ if $|i - j| > r$, where r is small, and let Gaussian elimination *with column pivoting* be used to solve $Ax = b$. Identify all the coefficients of the intermediate equations that can become nonzero. Hence deduce that the total number of additions and multiplications of the complete calculation can be bounded by a constant multiple of nr^2.

2.4.* Prove the following statement:

Theorem [Gerschgorin] *Let A be an $n \times n$ matrix (with real or complex coefficients). Then all of its eigenvalues are contained in the union $\cup_{i=1}^{n} \Gamma_i$, where Γ_i are the circles in the complex plane*

$$\Gamma_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}, \quad r_i := \sum_{j=1, j \neq i}^{n} |a_{ij}|, \qquad i = 1, \ldots, n.$$

[*Hint.* Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, let w be the corresponding eigenvector, and let w_k be its largest absolute value component. Consider the k-th equation of the relation $Aw = \lambda w$, and apply arguments similar to those used in the proof of Theorem 2.12.]