Mathematical Tripos Part IB: Lent 2020

Numerical Analysis – Lecture 13[1]

Pivoting Naive LU factorization fails when, for example, $A_{1,1} = 0$. The remedy is to exchange rows of A, a technique called *pivoting*. Specifically, at the k'th step of the algorithm we look for another row $p \geq k$ such that the entry $(A_{k-1})_{p,k}$ is nonzero. We permute rows p and k and proceed. The algorithm with pivoting can thus be written as follows:

- Let $A_0 = A$.
- For $k = 1, \ldots, n$: find $p \geq k$ such that $(A_{k-1})_{p,k} \neq 0$. Let P_k be the permutation matrix[2] that swaps positions k and p. Let $\boldsymbol{u}_k^\top$ be the k'th row of $P_k A_{k-1}$ and $\boldsymbol{l}_k$ be $\frac{1}{(P_k A_{k-1})_{k,k}} \times$ (k'th column of $P_k A_{k-1}$). Set $A_k = P_k A_{k-1} - \boldsymbol{l}_k \boldsymbol{u}_k^T$.

If we unroll the algorithm we have $A_1 = P_1 A_0 - \boldsymbol{l}_1 \boldsymbol{u}_1^T$, $A_2 = P_2 P_1 A - P_2 \boldsymbol{l}_1 \boldsymbol{u}_1^\top - \boldsymbol{l}_2 \boldsymbol{u}_2^T$, etc. and at the end, since $A_n = 0$ (and P_n the identity matrix):

$$P_{n-1} \cdots P_1 A = \tilde{\boldsymbol{l}}_{\boldsymbol{1}} \boldsymbol{u}_1^\top + \cdots + \tilde{\boldsymbol{l}}_{\boldsymbol{n}} \boldsymbol{u}_n^\top \tag{5.2}$$

where $\tilde{\boldsymbol{l}}_{\boldsymbol{k}} = P_{n-1} \ldots P_{k+1} \boldsymbol{l}_k$. Note that the first $k-1$ components of $\tilde{\boldsymbol{l}}_{\boldsymbol{k}}$ are zero since this is the case for $\boldsymbol{l}_k$ and since the permutations $P_{k+1}, \ldots, P_{n-1}$ only permute components of index $\geq k+1$. Therefore, Equation (5.2) can be rewritten as:

$$PA = \tilde{L} U$$

where $P = P_{n-1} \ldots P_1$ is a permutation matrix, and $\tilde{L} = [\tilde{\boldsymbol{l}}_{\boldsymbol{1}} \ \ \ldots \ \ \tilde{\boldsymbol{l}}_{\boldsymbol{n}}]$ is unit lower triangular, and U is upper triangular.

There is one situation where the algorithm above can still fail: this if for some k, *all* the entries in the k'th column of A_{k-1} are zero. In this case one can choose $\boldsymbol{l}_k$ to be the vector with a 1 at position k and zero elsewhere, and choose $\boldsymbol{u}_k^\top$ to be the k'th row of A_{k-1}, and $P_k = I$ (identity matrix). With this choice, the first k rows and columns of $A_k = A_{k-1} - \boldsymbol{l}_k \boldsymbol{u}_k^\top$ become zero as desired (this is not the only choice of $P_k, \boldsymbol{l}_k, \boldsymbol{u}_k$ that works in this case; other choices are possible).

We have thus shown that for any matrix A (even singular) one can find a permutation matrix P such that PA has an LU factorization.

Pivoting is not only important to find an element that is nonzero, but also for the overall numerical stability of the algorithm. A common choice of pivot p is to take $p \geq k$ such that $|(A_{k-1})_{p,k}|$ is maximum. This ensures in particular that the entries of $\boldsymbol{l}_k$ are all bounded above by 1 in magnitude.

Symmetric matrices Let A be an $n \times n$ symmetric matrix (i.e., $A_{k,\ell} = A_{\ell,k}$). An analogue of LU factorization that takes advantage of symmetry consists in expressing A in the form of the product $LDL^\top$, where L is $n \times n$ lower triangular, with ones on its diagonal and D is a diagonal matrix. This is a special case of an LU factorization with $U = DL^\top$. If we let $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_n$ be the columns of L then this factorization takes the form $A = \sum_{k=1}^n D_{k,k} \boldsymbol{l}_k \boldsymbol{l}_k^\top$. To compute this factorization, we can use an algorithm very similar to the one for the computation of LU factorization (without pivoting): Set $A_0 = A$ and for $k = 1, 2, \ldots, n$ let $\boldsymbol{l}_k$ be the multiple of the kth column of A_{k-1} such that $L_{k,k} = 1$. Set $D_{k,k} = (A_{k-1})_{k,k}$ and form $A_k = A_{k-1} - D_{k,k} \boldsymbol{l}_k \boldsymbol{l}_k^\top$.

Example Let $A = A_0 = \begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix}$. Hence $\boldsymbol{l}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $D_{1,1} = 2$ and

$$A_1 = A_0 - D_{1,1} \boldsymbol{l}_1 \boldsymbol{l}_1^\top = \begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix} - 2 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix}.$$

We deduce that $\boldsymbol{l}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $D_{2,2} = 3$ and $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.

[1] Corrections and suggestions to these notes should be emailed to `h.fawzi@damtp.cam.ac.uk`.

[2] A permutation matrix is a matrix with exactly one 1 in each row and in each column; the remaining entries being 0. For example $P = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ is a permutation matrix and PA exchanges the two rows of A.

Symmetric positive definite matrices Recall: A is positive definite if $\boldsymbol{x}^\top A\boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$.

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

Proof. Suppose that $A = LDL^\top$ and let $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Since L is nonsingular (it is lower triangular and all diagonal elements are equal to 1), $\boldsymbol{y} := L^\top \boldsymbol{x} \neq \mathbf{0}$. Then $\boldsymbol{x}^\top A\boldsymbol{x} = \boldsymbol{y}^\top D\boldsymbol{y} = \sum_{k=1}^{n} D_{k,k} y_k^2 > 0$, hence A is positive definite.
Conversely, suppose that A is positive definite. We wish to demonstrate that an $LDL^\top$ factorization exists. We denote by $\boldsymbol{e}_k \in \mathbb{R}^n$ the kth unit vector. Hence $\boldsymbol{e}_1^\top A\boldsymbol{e}_1 = A_{1,1} > 0$ and $\boldsymbol{l}_1$ & $D_{1,1}$ are well defined. We now show that $(A_{k-1})_{k,k} > 0$ for $k = 1, 2, \ldots$. This is true for $k = 1$ and we continue by induction, assuming that $A_{k-1} = A - \sum_{j=1}^{k-1} D_{j,j}\boldsymbol{l}_j\boldsymbol{l}_j^\top$ has been computed successfully.

Define $\boldsymbol{x} \in \mathbb{R}^n$ as the solution of the following system of equations: $\boldsymbol{l}_j^\top \boldsymbol{x} = 0$, $j = 1, \ldots, k-1$, $x_k = 1$ and $x_j = 0$ for $j = k+1, \ldots, n$. This is a system of n linear equations in the unknown $\boldsymbol{x} \in \mathbb{R}^n$. The matrix of this system of equations is upper triangular with ones on the diagonal hence it is invertible and our system has a unique solution. Now observe that since the first $k-1$ rows & columns of A_{k-1} vanish, and since $x_k = 1$ and the components $k+1, \ldots, n$ of $\boldsymbol{x}$ vanish we have $(A_{k-1})_{k,k} = \boldsymbol{x}^\top A_{k-1}\boldsymbol{x}$. Thus, from the definition of A_{k-1} and the choice of $\boldsymbol{x}$,

$$(A_{k-1})_{k,k} = \boldsymbol{x}^\top A_{k-1}\boldsymbol{x} = \boldsymbol{x}^\top \left(A - \sum_{j=1}^{k-1} D_{j,j}\boldsymbol{l}_j\boldsymbol{l}_j^\top \right) \boldsymbol{x} = \boldsymbol{x}^\top A\boldsymbol{x} - \sum_{j=1}^{k-1} D_{j,j}(\boldsymbol{l}_j^\top \boldsymbol{x})^2 = \boldsymbol{x}^\top A\boldsymbol{x} > 0,$$

as required. Hence $(A_{k-1})_{k,k} > 0$, $k = 1, 2, \ldots, n$, and the factorization exists. □

Conclusion It is possible to check if a symmetric matrix is positive definite by trying to form its $LDL^\top$ factorization.

Cholesky factorization Define $D^{1/2}$ as the diagonal matrix whose (k,k) element is $D_{k,k}^{1/2}$, hence $D^{1/2}D^{1/2} = D$. Then, A being positive definite, we can write

$$A = (LD^{1/2})(D^{1/2}L^\top) = (LD^{1/2})(LD^{1/2})^\top.$$

In other words, letting $\tilde{L} := LD^{1/2}$, we obtain the *Cholesky factorization* $A = \tilde{L}\tilde{L}^\top$.

Sparse matrices It is often required to solve *very* large systems $A\boldsymbol{x} = \boldsymbol{b}$ ($n = 10^5$ is considered small in this context!) where nearly all the elements of A are zero. Such a matrix is called *sparse* and efficient solution of $A\boldsymbol{x} = \boldsymbol{b}$ should exploit sparsity. In particular, we wish the matrices L and U to inherit as much as possible of the sparsity of A and for the cost of computation to be determined by the number of nonzero entries, rather than by n. The following theorem shows that certain zeros of A are always inherited by an LU factorization.

Theorem Let $A = LU$ be an LU factorization (without pivoting) of a sparse matrix. Then all leading zeros in the rows of A to the left of the diagonal are inherited by L and all the leading zeros in the columns of A above the diagonal are inherited by U.

Proof We assume that $U_{k,k} \neq 0$ for all $k = 1, \ldots, n$ which is the same as saying that $(A_{k-1})_{k,k} \neq 0$ when running the LU factorization algorithm (without pivoting). If $A_{i,1} = 0$ this means that $L_{i,1}U_{1,1} = 0$ and so $L_{i,1} = 0$. If furthermore $A_{i,2} = 0$ we get $L_{i,1}U_{1,2} + L_{i,2}U_{2,2} = 0$ which implies $L_{i,2} = 0$ since $L_{i,1} = 0$. In general we get that if $A_{i,1} = \cdots = A_{i,j} = 0$ where $j < i$ then $L_{i,1} = \cdots = L_{i,j} = 0$. A similar reasoning applies for leading zeros in the columns of A above the diagonal. □

Banded matrices The matrix A is a *banded matrix* if there exists an integer $r < n$ such that $A_{i,j} = 0$ for $|i-j| > r$, $i, j = 1, 2, \ldots, n$. In other words, all the nonzero elements of A reside in a band of width $2r+1$ along the main diagonal. In that case, according to the previous theorem, $A = LU$ implies that $L_{i,j} = U_{i,j} = 0$ $\forall\ |i-j| > r$ and sparsity structure is inherited by the factorization.
In general, the expense of calculating an LU factorization of an $n \times n$ *dense* matrix A is $\mathcal{O}(n^3)$ operations and the expense of solving $A\boldsymbol{x} = \boldsymbol{b}$, provided that the factorization is known, is $\mathcal{O}(n^2)$. However, in the case of a banded A, we need just $\mathcal{O}(r^2 n)$ operations to factorize and $\mathcal{O}(rn)$ operations to solve a linear system. If $r \ll n$ this represents a very substantial saving!