Mathematical Tripos Part IB: Lent 2010

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 *column pivoting* (or just *pivoting*). This is equivalent to picking a suitable equation for eliminating the first unknown in Gaussian elimination. Specifically, column pivoting means that, having obtained A_{k-1}, we exchange two rows of A_{k-1} so that the element of largest magnitude in the kth column is in the 'pivotal position' (k, k). In other words,

$$|(A_{k-1})_{k,k}| = \max\{|(A_{k-1})_{j,k}| \,:\, j = 1, 2, \ldots, n\}.$$

Of course, the same exchange is required in the portion of L that has been formed already (i.e., 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. (The exchange of rows can be regarded as the pre-multiplication of the relevant matrix by a permutation matrix.)

Column pivoting copes with zeros at the pivot position, except when the entire kth column of A_{k-1} is zero: in that case we let $\boldsymbol{l}_k$ be the kth unit vector while, as before, choose $\boldsymbol{u}_k^\top$ as the kth row of A_k). This choice preserves the condition that the matrix $\boldsymbol{l}_k\boldsymbol{u}_k^\top$ has the same kth row and column as A_{k-1}. Thus $A_k := A_{k-1} - \boldsymbol{l}_k\boldsymbol{u}_k^\top$ still has zeros in its kth row and column as required.

An important advantage of column pivoting is that $|L_{i,j}| \le 1$ for all $i, j = 1, \ldots, n$. This avoids division by zero and tends to reduce the chance of large numbers occuring during the factorization, a phenomenon that might lead to *ill conditioning* and to accumulation of *roundoff error.*

In *row pivoting* one exchanges columns of A_{k-1}, rather than rows (sic!), whereas *total pivoting* corresponds to exchange of both rows and columns, so that the modulus of the pivotal element $(A_{k-1})_{k,k}$ is maximised.

Symmetric matrices Let A be an $n \times n$ symmetric matrix (i.e., $A_{k,\ell} = A_{\ell,k}$). An analogue of LU factorization takes advantage of symmetry: we express A in the form of the product $LDL^\top$, where L is $n \times n$ lower triangular, with ones on its diagonal, whereas D is a diagonal matrix. Subject to its existence, we can write this factorization as

$$A = \begin{bmatrix} \boldsymbol{l}_1 & \boldsymbol{l}_2 & \cdots \boldsymbol{l}_n \end{bmatrix} \begin{bmatrix} D_{1,1} & 0 & \cdots & 0 \\ 0 & D_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_{n,n} \end{bmatrix} \begin{bmatrix} \boldsymbol{l}_1^\top \\ \boldsymbol{l}_2^\top \\ \vdots \\ \boldsymbol{l}_n^\top \end{bmatrix} = \sum_{k=1}^n D_{k,k}\boldsymbol{l}_k\boldsymbol{l}_k^\top$$

where, as before, $\boldsymbol{l}_k$ is the kth column of L. The analogy with the LU algorithm becomes obvious by letting $U = DL^\top$, but the present form lends itself better to exploitation of symmetry and requires roughly half the storage of conventional LU. Specifically, to compute this factorization, we let $A_0 = A$ 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 $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.

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

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.

[1]Corrections and suggestions to these notes should be emailed to `A.Iserles@damtp.cam.ac.uk`. All handouts are available on the WWW at the URL `http://www.damtp.cam.ac.uk/user/na/PartIB/Handouts.html`.

Proof. Suppose that $A = LDL^\top$ and let $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$. Since L is nonsingular, $\boldsymbol{y} := L^\top \boldsymbol{x} \neq \boldsymbol{0}$. 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 follows. The bottom $n-k$ components are zero, $x_k = 1$ and $x_1, x_2, \ldots, x_{k-1}$ are calculated in a reverse order, each x_j being chosen so that $\boldsymbol{l}_j^\top \boldsymbol{x} = 0$ for $j = k-1, k-2, \ldots, 1$. In other words, since $0 = \boldsymbol{l}_j^\top \boldsymbol{x} = \sum_{i=1}^n L_{i,j} x_i = \sum_{i=j}^k L_{i,j} x_i$, we let $x_j = -\sum_{i=j+1}^k L_{i,j} x_i$, $j = k-1, k-2, \ldots, 1$.
Since the first $k-1$ rows & columns of A_{k-1} vanish, our choice implies that $(A_{k-1})_{k,k} = \boldsymbol{x}^\top A_{k-1} \boldsymbol{x}$. 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 only tool at our disposal at the moment is the freedom to exchange rows and columns to minimise *fill-in*.

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 Follows from the second question on Examples' Sheet 3. □

This theorem suggests that if one requires a factorization of a sparse matrix then one might try to reorder its rows and columns by a preliminary calculation so that many of the zero elements are leading zero elements in rows and columns. This will reduce the fill-in.

Example 1 The LU factorisation of

$$\begin{bmatrix} -3 & 1 & 1 & 2 & 0 \\ 1 & -3 & 0 & 0 & 1 \\ 1 & 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 & 0 & 0 \\ -\frac{1}{3} & -\frac{1}{8} & 1 & 0 & 0 \\ -\frac{2}{3} & -\frac{1}{4} & \frac{6}{19} & 1 & 0 \\ 0 & -\frac{3}{8} & \frac{1}{19} & \frac{4}{81} & 1 \end{bmatrix} \begin{bmatrix} -3 & 1 & 1 & 2 & 0 \\ 0 & -\frac{8}{3} & \frac{1}{3} & \frac{2}{3} & 1 \\ 0 & 0 & \frac{19}{8} & \frac{3}{4} & \frac{1}{8} \\ 0 & 0 & 0 & \frac{81}{19} & \frac{4}{19} \\ 0 & 0 & 0 & 0 & \frac{272}{81} \end{bmatrix},$$

has significant fill-in. However, reordering (symmetrically) rows and columns $1 \leftrightarrow 3$, $2 \leftrightarrow 4$ and $4 \leftrightarrow 5$ yields

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 \\ 1 & 2 & -3 & 0 & 1 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{2}{3} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{6}{29} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 \\ 0 & 0 & -\frac{29}{6} & 0 & 1 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & -\frac{272}{87} \end{bmatrix}.$$

Example 2 If the nonzeros of A occur only on the diagonal, in one row and in one column, then the full row and column should be placed at the bottom and on the right of A, respectively.