Dr A. Iserles
Easter 1997

Numerical Analysis – Lecture 2

2 Factorization of structured matrices

2.1 Symmetric positive definite matrices

Let A be an $n \times n$ symmetric positive definite matrix (i.e., $A_{k,\ell} = A_{\ell,k}$ and $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$). An analogue of LU factorization takes advantage of symmetry: we express A in the form of the product LDL^{T}, where L is $n \times n$ lower triangular, with ones on its diagonal, whereas D is a diagonal matrix (which, A being positive definite, has positive elements along its diagonal). Subject to its existence, we can write this factorization as

$$A = \left[\begin{array}{cccc} \boldsymbol{l}_1 & \boldsymbol{l}_2 & \cdots \boldsymbol{l}_n \end{array}\right] \left[\begin{array}{cccc} D_{1,1} & 0 & \cdots & 0 \\ 0 & D_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_{n,n} \end{array}\right] \left[\begin{array}{c} \boldsymbol{l}_1^{\mathrm{T}} \\ \boldsymbol{l}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{l}_n^{\mathrm{T}} \end{array}\right] = \sum_{k=1}^{n} D_{k,k} \boldsymbol{l}_k \boldsymbol{l}_k^{\mathrm{T}}.$$

(as before, $\boldsymbol{l}_k$ is the kth column of L).
The analogy with the algorithm of Section 1.2 becomes obvious by letting $U = DL$, but the present form lends itself better to exploitation of symmetry. 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 $A_k = A_{k-1} - D_{k,k} \boldsymbol{l}_k \boldsymbol{l}_k^{\mathrm{T}}$.

Example Let $A = A_0 = \left[\begin{array}{cc} 2 & 4 \\ 4 & 11 \end{array}\right]$. Hence $\boldsymbol{l}_1 = \left[\begin{array}{c} 1 \\ 2 \end{array}\right]$, $D_{1,1} = 2$ and

$$A_1 = A_0 - D_{1,1} \boldsymbol{l}_1 \boldsymbol{l}_1^{\mathrm{T}} = \left[\begin{array}{cc} 2 & 4 \\ 4 & 11 \end{array}\right] - 2 \left[\begin{array}{cc} 1 & 2 \\ 2 & 4 \end{array}\right] = \left[\begin{array}{cc} 0 & 0 \\ 0 & 3 \end{array}\right].$$

We deduce that $\boldsymbol{l}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$, $D_{2,2} = 3$ and $A = \left[\begin{array}{cc} 1 & 0 \\ 2 & 1 \end{array}\right] \left[\begin{array}{cc} 2 & 0 \\ 0 & 3 \end{array}\right] \left[\begin{array}{cc} 1 & 2 \\ 0 & 1 \end{array}\right]$.

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

Proof. Suppose that $A = LDL^{\mathrm{T}}$ and let $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Since L is nonsingular, $\boldsymbol{y} := L^{\mathrm{T}} \boldsymbol{x} \neq \mathbf{0}$. Then $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}} D \boldsymbol{y} = \sum_{k=1}^{n} D_{k,k} y_k^2 > 0$, hence A is positive definite.
Conversely, suppose that A is positive definite. We wish to demonstrate that an LDL^{T} factorization exists. We denote by $\boldsymbol{e}_k \in \mathbb{R}^n$ the kth unit (a.k.a. coordinate) vector. Hence $\boldsymbol{e}_1^{\mathrm{T}} A \boldsymbol{e}_1 = A_{1,1} > 0$ and $\boldsymbol{l}_1$ & $D_{1,1}$ are well defined. We now show that $(A_{k-1})_{k,k} > 0$ for $k = 1, 2, \ldots$. The result is true for $k = 1$ and we continue by induction (hence may assume that $A_{k-1} = A - \sum_{j=1}^{k-1} D_{j,j} \boldsymbol{l}_j \boldsymbol{l}_j^{\mathrm{T}}$ has been computed successfully).
We 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 $\boldsymbol{l}_j^{\mathrm{T}} \boldsymbol{x} = 0$ for $j = k-1, k-2, \ldots, 1$. In other words, since $0 = \boldsymbol{l}_j^{\mathrm{T}} \boldsymbol{x} = \sum_{i=1}^{n} L_{i,j} x_i = \sum_{i=j}^{k} L_{i,j} x_i$, 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 $(A_{k-1})_{k,k} = \boldsymbol{x}^{\mathrm{T}} A_{k-1} \boldsymbol{x}$. Thus, from the definition of A_{k-1} and the choice of $\boldsymbol{x}$,

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

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^{T} 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

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

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

2.2 Sparse matrices

Frequently it is required to solve *very* large systems $A\boldsymbol{x} = \boldsymbol{b}$ ($n = 10^5$ is a modest example!) 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, if possible, to derive sparse L and U. The only tool at our disposal is the freedom to exchange rows and columns to minimize *fill-in.* For example,

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

with copious fill-in. However, reordering (symmetrically) rows and columns,

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

An important example of sparse matrices are *band matrices,* whereby all the nonzero elements occur on or close to the diagonal.

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 Exercise 1. □