NLA 2021-2022
Special linear systems

Martin Sombra

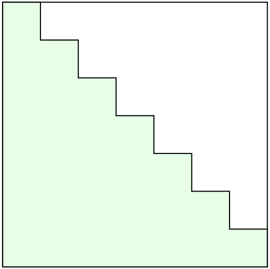
6 October 2021

Symmetric matrices

An $n \times n$ matrix A is *symmetric* if

$$A^T = A$$

A symmetric matrix needs *half the space* to store its entries



We should be able to solve a symmetric problem

$$A\,x = b$$

with $\approx \frac{n^3}{3}$ flops instead of $\approx \frac{2\,n^3}{3}$

Pivoting destroys symmetry

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} = \begin{bmatrix} c & e & f \\ b & d & e \\ a & b & c \end{bmatrix}$$

$\rightsquigarrow$ for a symmetric A, we just aim to compute the *LU factorization*

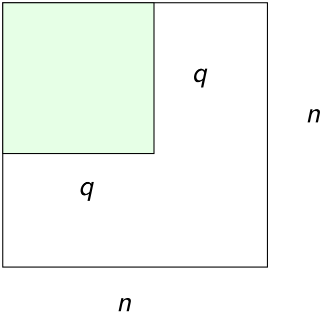
Matlab notation

For $p \leqslant q$ and $r \leqslant s$ set

$$A(p:q, r:s) = \left[a_{i,j}\right]_{p \leqslant i \leqslant q, r \leqslant j \leqslant s} \in \mathbb{R}^{(q-p+1)\times(s-r+1)}$$

For instance, the *leading* $q \times q$*-principal submatrix* of A is

$A(1:q, 1:q)$



When does the LU factorization exist?

Let A be an arbitrary $n \times n$ matrix (not necessarily symmetric)

The following are equivalent (TFAE):

1. there are unique L unit lower triangular and U upper triangular such that $A = L\,U$
2. all leading principal submatrices of A are nonsingular

The LDLT factorization

When A is symmetric and has an LU factorization, the factors are connected:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & d - \frac{bc}{a} \end{bmatrix} \begin{bmatrix} 1 & \frac{c}{a} \\ 0 & 1 \end{bmatrix}$$

In general symmetric case, set $d_i = u_{i,i}$ and write

$$U = D\,M$$

with $D = \mathrm{diag}(d_1, \ldots, d_n)$ and M *unit* upper triangular. Then

$$M = L^T$$

and so the LU factorization can be rewritten as

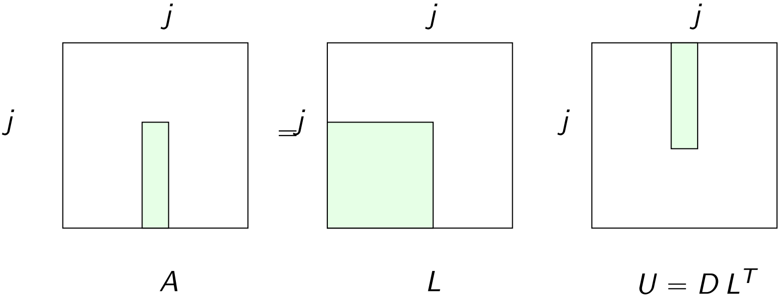
$$A = L\,D\,L^T$$

Computing the LDLT factorization

For $j = 1, \ldots, n$ set $A(j:n, j) \leftarrow L(j:n, 1:j)\, v(1:j)$ with

$$v = \begin{bmatrix} d_1\, l_{j,1} \\ \vdots \\ d_{j-1}\, l_{j,j-1} \\ d_j \end{bmatrix}$$

as in the figure



Computing the LDLT factorization (cont.)

Hence the equations

$$d_j = a_{j,j} - \sum_{k=1}^{j-1} d_k\, \ell_{j,k}^2$$

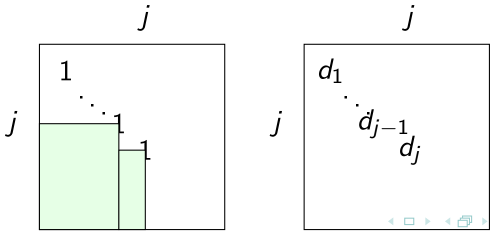
and

$$L(j+1:n,j) = \frac{1}{d_j}\big(A(j+1:n,j) - L(j+1:n,1:j-1)\, v(1:j-1)\big)$$

gives the jth diagonal entry and the jth column

$$d_j \quad \text{and} \quad L(j+1:n,j)$$

from the previous diagonal entries and the $(j-1)$th column of L



The LDLT algorithm

for $j = 1, \ldots, n$

  for $i = 1, \ldots, j-1$

    $v_i \leftarrow \ell_{j,i}\, d_i$

  end

  $d_j \leftarrow a_{j,j} - L(j, 1:j-1)\, v(1:j-1)$

$$L(j{+}1:n, j) \leftarrow \frac{1}{d_j}\big(A(j{+}1:n, j) {-} L(j{+}1, n:1:j{-}1)\, v(1:j{-}1)\big)$$

The modified LDLT algorithm

for $j = 1, \ldots, n$

for $i = 1, \ldots, j-1$

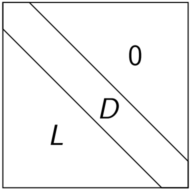
$v_i \leftarrow a_{j,i}\, a_{i,i}$

end

$a_{j,j} \leftarrow a_{j,j} - A(j, 1:j-1)\, v(1:j-1)$

$$A(j+1:n, j) \leftarrow \frac{1}{a_{i,i}} \big(A(j+1:n, j) - A(j+1, n:1:j-1)\, v(1:j-1)\big)$$

$\leadsto$ over-writing scheme



Example

Let

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 17 \end{bmatrix}$$

For $j = 1$:

$$L = \begin{bmatrix} 1 & & \\ -1 & & \\ 2 & & \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & & \\ & & \\ & & \end{bmatrix}$$

For $j = 2$:

$$[v] = -1, \quad L = \begin{bmatrix} & & \\ & & \\ & 1 & \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} & & \\ & 4 & \\ & & \end{bmatrix}$$

coming from the operations

$$v_1 = (-1) \cdot 1 = -1, \quad d_2 = 5 - (-1) \cdot (-1) = 4, \quad \ell_{3,2} = \frac{1}{4}(2 - 2 \cdot (-1))$$

Example (cont.)

For $j = 3$:

$$v = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} & & \\ & & \\ & & 9 \end{bmatrix}$$

coming from

$$v_1 = 2 \cdot 1 = 1, \quad v_2 = 1 \cdot 4 = 4, \quad d_3 = 17 - (2 \cdot 2 + 1 \cdot 4) = 9$$

Hence

$$\begin{bmatrix} 1 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 17 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

In the machine:

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 17 \end{bmatrix} \rightsquigarrow L \,\&\, D = \begin{bmatrix} 1 & & \\ -1 & 4 & \\ 2 & 1 & 9 \end{bmatrix}$$

Numerical (un)stability

The LU factorization of a symmetric matrix can be numerically unstable:

$$A = \begin{bmatrix} \eta & 1 \\ 1 & 1 \end{bmatrix}$$

with $0 < \eta < \varepsilon$ (machine epsilon)

Symmetric positive definite systems

A symmetric A is *positive definite (SPD)* if for all $x \in \mathbb{R}^n \backslash \{0\}$

$$x^T A x > 0$$

Important fact from LA:

$$A \in \mathbb{R}^{n\times n} \text{ symmetric} \iff A = Q^T \Lambda Q$$

with Q orthogonal and Λ diagonal: A is diagonalizable over the reals through an orthogonal similarity, and

$$A \text{ is SPD} \iff \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \text{ with } \lambda_i > 0$$

The Cholesky factorization

An SPD $n \times n$ matrix is nonsingular, and moreover all its leading principal submatrices are nonsingular

$\rightsquigarrow$ there is unit lower triangular L and a diagonal D such that

$$A = L\,D\,L^T \tag{1}$$

We have that $d_i > 0$ for all i and so we can write (**??**) as

$$A = G\,G^T$$

with $G = L \cdot \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})$ (*Cholesky factorization*)

Computing the Cholesky factorization

Note that

$$a_{i,i} = \sum_{k=0}^{j} g_{j,k}\, g_{i,k}$$

and so

$$g_{j,j}\, g_{i,j} = a_{i,j} - \sum_{k=1}^{j-1} g_{j,k}\, g_{i,k}$$

$\rightsquigarrow$ we can compute the j-th column of G from the previous ones

The Cholesky algorithm

for $j = 1, \dots, n$

$$g_{j,j} \leftarrow (a_{j,j} - \textstyle\sum_{k=1}^{j-1} g_{j,k}^2)^{1/2}$$

for $i = j+1, \dots, n$

$$g_{i,j} \leftarrow \frac{1}{g_{j,j}} \left(a_{i,j} - \sum_{k=1}^{j-1} g_{i,k}\, g_{j,k} \right)$$

or alternatively:

for $j = 1, \dots, n$

$$a_{j,j} \leftarrow (a_{j,j} - \textstyle\sum_{k=1}^{j-1} a_{j,k}^2)^{1/2}$$

for $i = j+1, \dots, n$

$$a_{i,j} \leftarrow \frac{1}{g_{j,j}} \left(a_{i,j} - \sum_{k=1}^{j-1} a_{i,k}\, a_{j,k} \right)$$

Can be overwritten over A and does not need the auxiliary vector v

Consider again

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 17 \end{bmatrix}$$

For $j = 1$:

$$G = \begin{bmatrix} 1 & & \\ -1 & & \\ 2 & & \end{bmatrix}$$

and for $j = 2$:

$$G = \begin{bmatrix} & & \\ & 2 & \\ & 2 & \end{bmatrix}$$

coming from the operations

$$g_{2,2} = (5 - (-1)^2)^{1/2} = 2 \quad \text{and} \quad g_{3,2} = \frac{1}{2}(2 - (-1) \cdot 2) = 2$$

The example revisited (cont.)

For $j = 3$:

$$G = \begin{bmatrix} & & \\ & & \\ & & 3 \end{bmatrix}$$

coming from $g_{3,3} = (17 - (2^2 + 2^2))^{1/2} = 3$

Hence

$$\begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

In the machine

$$A = \begin{bmatrix} 1 & & \\ -1 & 1 & \\ 2 & 1 & 1 \end{bmatrix} \rightsquigarrow G = \begin{bmatrix} 1 & & \\ -1 & 2 & \\ 2 & 2 & 3 \end{bmatrix}$$

The complexity of this algorithm is

$$\begin{aligned}
\sum_{j=1}^{n}\left(2j-1+\sum_{i=j+1}^{n}(2j-1)\right) &= \sum_{j=1}^{n}(2j-1)\,(n-j+1) \\
&= 2n\Big(\sum_{j=1}^{n} j\Big) - \sum_{j=1}^{n} j^2 + O(n^2) \\
&= 2n\left(\frac{n^2}{2}+O(n)\right) - \frac{n^3}{3} + O(n^2) \\
&= \frac{1}{3}n^3 + O(n^2)
\end{aligned}$$

$\rightsquigarrow$ half the complexity of the LU factorization

Numerical stability

Pivoting is not necessary for the Cholesky algorithm to be numerically stable: the same analysis of GEPP shows that the Cholesky solution $\widehat{x}$ satisfies $(A + \delta A)\,\widehat{x} = b$ with

$$|\delta A| \leqslant 3\,n\,\varepsilon\,|G|\,|G^T|$$

By the *Cauchy-Schwartz inequality*, for each i, j we have that

$$\begin{aligned}(|G|\,|G^T|)_{i,j} &\leqslant \sum_{k=1}^{n} |g_{i,k}|\,|g_{j,k}| \\ &\leqslant \Big(\sum_{k=1}^{n} g_{i,k}^2\Big)^{1/2}\Big(\sum_{k=1}^{n} g_{j,k}^2\Big)^{1/2} = a_{i,i}^{1/2}\,a_{j,j}^{1/2} \leqslant \max_{i,j} |a_{i,j}|\end{aligned}$$

Hence $\||G|\,|G^T|\|_\infty \leqslant n\,\|A\|_\infty$ and so

$$\|\delta A\|_\infty \leqslant 3\,n^2\,\varepsilon\,\|A\|_\infty$$

$\rightsquigarrow$ Cholesky algorithm is backward stable 

Is my A an SPD matrix?

Cholesky is the cheapest way of testing is a given symmetric $n \times n$ matrix is definite positive:

it will be the case if and only if the algorithm concludes!