Cholesky Decomposition for Positive Definite Matrices

Positive Definite Matrices

Positive Definite Matrices: An $n \times n$ real matrix A is said to be (symmetric) positive definite if the following hold:

1. A is symmetric, that is, $A^T = A$;
2. $x^T Ax > 0$ for all $x \in \mathbb{R}^n$ with $x \neq 0$.

In the literature very often positive definite matrices are defined as those satisfying only 2. Such matrices need not be symmetric. Find an example!

However, in application real matrices that satisfy 2 are most often also symmetric. In accordance with the textbook, we will refer to symmetric positive definite matrices as (just) positive definite matrices.

The counterpart of transpose (^{T}) for complex matrices is complex conjugate transpose (*). If an $n \times n$ matrix is non real, that is it is a complex matrix with non zero imaginary part, then 2 with ^{T} replaced by * implies that $A^* = A$, that is, A is Hermitian. Prove this!

Therefore if A is an $n \times n$ non real matrix , then A is said to be positive definite if $x^* Ax > 0$ for all nonzero $x \in \mathbb{C}^n$.

Properties and Applications

Positive definite matrices arise in many applications involving optimization and discretization of partial differential equations. Consider systems $Ax = b$.

- When associated with electrical circuits with x representing loop currents, $x^T Ax$ is the total power drawn by the resistors.
- When associated with mass spring systems, $(x^T Ax)/2$ is the strain energy of the system.

(For examples see Section 1.2 of *D. S. Watkins, Fundamentals of Matrix Computations (second edition)*)

Important properties:

- A positive definite matrix is a nonsingular matrix.
- If A is an $n \times n$ positive definite matrix, then for any nonsingular $n \times n$ matrix X, $X^T AX$ is also positive definite.
- All principal submatrices of positive definite matrices are positive definite.

To know more about the theory and applications of positive definite matrices, check out the following link:

MIT OCW on Symmetric Positive Definite Matrices

Cholesky Decomposition for Positive Definite Matrices

Let A be a positive definite matrix. Then there exists a unique upper triangular matrix G with positive diagonal entries such that $A = G^T G$. This is called the **Cholesky Decomposition** of A and G is called the Cholesky factor of A.

André-Louis Cholesky
(1875-1918)

Cholesky Decomposition Theorem

Theorem An $n \times n$ matrix A is positive definite if and only if there exists an upper triangular matrix G with positive diagonal entries such that

$$A = G^T G.$$

Proof: Suppose $A = G^T G$ for some upper triangular matrix G with positive diagonal entries. Then

$$A^T = (G^T G)^T = G^T G = A.$$

For any $x \neq 0$, $\;x^T Ax = x^T G^T Gx = (Gx)^T(Gx) = \|Gx\|_2^2 \geq 0.$

As G is upper triangular with positive diagonal entries, it is nonsingular. Therefore as $x \neq 0$, $Gx \neq 0 \Rightarrow \|Gx\|_2 > 0$ and from above, $x^T Ax > 0$ for all $x \neq 0$.

Conversely, suppose A is an $n \times n$ positive definite matrix. Then A is nonsingular and all its leading principal submatrices are also positive definite and hence nonsingular. Moreover $x^T Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Therefore, there is a unique unit lower triangular matrix L and a matrix $D = \mathrm{diag}(d_{11}, \cdots, d_{nn})$ $d_{ii} > 0$, such that $A = LDL^T$. Let $G = D^{1/2}L^T$ where, $D^{1/2} = \mathrm{diag}(\sqrt{d_{11}}, \cdots, \sqrt{d_{nn}})$. Then G is upper triangular with positive diagonal entries such that $A = G^T G$. Hence the proof. $\square$

Algorithms for Computing Cholesky Decomposition

Inner Product Form

Suppose

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{nn} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}}_{=:A} = \begin{bmatrix} g_{11} & & & \\ g_{12} & g_{22} & & \\ \vdots & \vdots & \ddots & \\ g_{1n} & g_{2n} & \cdots & g_{nn} \end{bmatrix} \underbrace{\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ & g_{22} & \cdots & g_{2n} \\ & & \ddots & \vdots \\ & & & g_{nn} \end{bmatrix}}_{=:G}$$

Equating entries on both sides for $j = 1 : n$,

$$\sum_{k=1}^{j} g_{kj}^2 = a_{jj} \;\&\; \sum_{i=1}^{j} g_{ij} g_{ik} = a_{jk};$$

$$\Rightarrow g_{jj} = \underbrace{\left(a_{jj} - \sum_{k=1}^{j-1} g_{kj}^2\right)^{1/2}}_{\text{costs } 2(j-1) \text{ flops + one square root}} \;\&\; g_{jk} = \underbrace{\left(a_{jk} - \sum_{i=1}^{j-1} g_{ij} g_{ik}\right) \Big/ g_{jj}}_{\text{costs } 2j-1 \text{ flops for each } k};\; k = j+1 : n$$

This is the *inner product formulation* for finding the Cholesky factor G one row at a time.

Flop count: $n^3/3 + O(n^2)$ flops. (Exercise!)

Outer Product Form

Let

$$b = \text{A(1,2:n)}^T, \widehat{A} = \text{A(2:n,2:n)},$$
$$g = \text{G(1,2:n)}^T, \widehat{G} = \text{G(2:n,2:n)}.$$

Then

$$\left[\begin{array}{c|c} a_{11} & b^T \\ \hline b & \widehat{A} \end{array}\right] = \left[\begin{array}{c|c} g_{11} & \\ \hline g & \widehat{G}^T \end{array}\right] \left[\begin{array}{c|c} g_{11} & g^T \\ \hline & \widehat{G} \end{array}\right] \Rightarrow \begin{cases} g_{11} = \sqrt{a_{11}} \\ g = b/g_{11} \\ \widehat{G}^T\widehat{G} = \widehat{A} - gg^T \end{cases}$$

which gives the pseudocode

1. Compute $g_{11} = \sqrt{a_{11}}$.
2. Compute $g = b/g_{11}$.
3. Compute the Cholesky factor $\widehat{G}$ of $\widehat{A} - gg^T$.

for a recursive algorithm to find the Cholesky factor of A.

This is the *outer product formulation* for finding the Cholesky factor as it involves the outer product gg^T.

Verify that this also costs $n^3/3 + O(n^2)$ flops!

Bordered Form

For j = 2:n let

$$A_{j-1} = \text{A(1:j-1, 1:j-1)};\ c = \text{A(1:j-1,j)}$$
$$G_{j-1} = \text{G(1:j-1;1:j-1)};h = \text{G(1:j-1,j)}$$

Then for j = 2:n

$$A = G^T G \Rightarrow \left[\begin{array}{c|c} A_{j-1} & c \\ \hline c^T & a_{jj} \end{array}\right] = \left[\begin{array}{c|c} G_{j-1}^T & \\ \hline h^T & g_{jj} \end{array}\right] \left[\begin{array}{c|c} G_{j-1} & h \\ \hline & g_{jj} \end{array}\right]$$
$$\Rightarrow A_{j-1} = G_{j-1}^T G_{j-1};\ c = G_{j-1}^T h;\ a_{jj} = h^T h + g_{jj}^2;$$

This gives the following pseudocode for computing G :

1. Set $G = \mathrm{zeros}(n, n)$; & $G(1, 1) = \sqrt{A(1, 1)}$;
2. for j = 2:n
   Solve G(1:j-1,1:j-1)$^T h =$ A(1:j-1,j) for h.
   Set G(1:j-1,j) $= h$;
   Set G(j,j) $= \sqrt{\text{A(j,j)} - h^T h}$;
   end

This is the *bordered* form of finding G which also costs $n^3/3 + O(n^2)$ flops

Solving a Positive Definite System of Equations

Pseudocode for solving an $n \times n$ system $Ax = b$ where A is a positive definite matrix:

1. Find the Cholesky factor G of A. (costs $n^3/3 + O(n^2)$ flops)
2. Solve $G^T y = b$ for y. (costs n^2 flops)
3. Solve $Gx = y$ for x. (costs n^2 flops)

Total cost is $n^3/3 + O(n^2)$ flops.