Cholesky Factorization of Matrices

Khin Myo Aye*

Abstract

Cholesky factorization is a type of matrix factorization which is used for solving system of linear equations. In this paper, we will study the factorization of real positive definite matrices by using Cholesky factorization. This type of factorization is in the form $A = LL^T$ where L is lower triangular matrix. First, the diagonalization of a matrix will be presented, second, positive definite matrix and its properties will be discussed. Finally, the Cholesky factorization of real positive definite matrices will be discussed in numerically point of view.

1. Preliminaries

In this study, some notations and some definitions about the matrices will be discussed. Throughout in this paper, the indices *i*, *j*, *k*, *n*, *r* represent the positive integers. The notations R and C represent the set of real numbers and the set of complex number, respectively.

1.1 Definitions. Let x be *n*-dimensional column vector and $A=(a_{ij})$ be $n\times n$ matrix (*n* rows and *n* columns), where $i, j = 1,2,\dots,n.$ Then Ax is *n*-dimensional column vector defined as

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{ij}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix}.$$

Let $B=(b_{ij})$ be $n\times n$ matrix. Then the **product** $C=(c_{ij})$ of A and B is $n\times n$ matrix with each column of C is a linear combination of the columns of A:

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

The transpose of the matrix $A=(a_{ij})$ is the matrix denoted by A^T which is in the form (a_{ji}); the columns of A^T is the rows of $A.$ The matrix $A \in R^{n\times n}$ is **nonsingular or invertible** if there exists a matrix $B \in \mathbb{R}^{n\times n}$ such that $BA = AB = I$, where $I \in R^{n\times n}$ is the identity matrix.

* Lecturer, Department of Mathematics, Yangon University of Education

1.2 Definitions. Let $A \in \mathbb{R}^{n\times n}$, i.e., A is $n\times n$ matrix with real entries. A is **symmetric** if it has the same entries below the diagonal as above: $a_{ij} = a_{ji}$ for all i, j. If the matrix A is symmetric, we have $A^T = A.$ The matrix A is **positive definite** if A is symmetric and $x^T Ax > 0$ for all nonzero $x \in R^n$.

1.3 Definition. A nonzero element $x \in R^n$ is called an **eigenvector** of the matrix $A \in R^{n\times n}$ if there exists a number λ such that $Ax = \lambda x$ .This number λ is an **eigenvalue** of the matrix A, and the vector x is the eigenvector associated with λ.

1.4 Theorem. Let A be a real symmetric matrix and let λ be an eigenvalue in C. Then λ is real.

Proof.

Let $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, $z_i \in \mathbb{C}$ be an eigenvector of A associated with λ .

Then $\bar{z}^T z = (\bar{z}_1 \ldots \bar{z}_n)\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \bar{z}_1 z_1 + \cdots + \bar{z}_n z_n = |z_1|^2 + \cdots + |z_n|^2 > 0.$

$Az = \lambda z$ implies that $\bar{z}^T Az = \bar{z}^T \lambda z = \lambda \bar{z}^T z$

$$(\bar{z}^T Az)^T = \lambda(\bar{z}^T z)^T$$

$z^T A^T \bar{z} = \lambda \bar{z}^T z$

$z^T A\bar{z} = \lambda \bar{z}^T z$

But $\overline{Az} = \bar{A}\bar{z} = A\bar{z}$, $\overline{Az} = \overline{\lambda z} = \bar{\lambda}\bar{z}$ and $z^T A\bar{z} = \bar{\lambda}\bar{z}^T z$ .Therefore

$\lambda \bar{z}^T z = \bar{\lambda}\bar{z}^T z$.

Since $z^T \bar{z} \neq 0,$ it follows that $\lambda = \bar{\lambda}$ so λ is real.

1.5 Theorem. All eigenvalues of positive definite matrix is positive.

Proof. Let $A \in R^{n\times n}$ be positive definite matrix and λ be an eigenvector of A, i.e., $Ax = \lambda x$ for any nonzero eigenvector $x \in \mathbb{R}^n$. Then we have $x^T Ax = x^T \lambda x = \lambda |x|^2$. Since $x^T Ax > 0,$ we have $\lambda > 0.$

1.6 Theorem. If eigenvectors $x_1, x_2, \ldots, x_k$ correspond to different eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, of the matrix A, then those eigenvectors are linearly independent.

Proof. Suppose that $k=2$ and $\alpha_1 x_1+\alpha_2 x_2=0, \alpha_1, \alpha_2 \in \mathbb{C}$. Then

$$A(\alpha_1 x_1+\alpha_2 x_2)=\alpha_1 A x_1+\alpha_2 A x_2$$
$$=\alpha_1 \lambda_1 x_1+\alpha_2 \lambda_2 x_2=0.$$

Substracting λ_2 times the previous equation, we have $\alpha_1(\lambda_1-\lambda_2)x_1=0.$ Since $\lambda_1 \neq \lambda_2$ and $x_1 \neq 0,$ we must have $\alpha_1=0.$ Similarly $\alpha_2=0.$ Therefore x_1 and x_2 are linearly independent.

Suppose that $x_1, x_2, \ldots, x_{k-1}$ are linearly independent. To show $x_1, x_2, \ldots, x_{k-1}, x_k$ are linearly independent, suppose that

$$\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_{k-1} x_{k-1}+\alpha_k x_k=0, \qquad \alpha_i \in \mathbb{C}$$

Multiply by A, we have

$$\alpha_1 A x_1+\alpha_2 A x_2+\ldots+\alpha_k A x_k=0$$
$$\alpha_1 \lambda_1 x_1+\alpha_2 \lambda_2 x_2+\ldots+\alpha_k \lambda_k x_k=0.$$

Multiply the previous equation by λ_k, then we have

$$\alpha_1 \lambda_k x_1+\alpha_2 \lambda_k x_2+\ldots+\alpha_k \lambda_k x_k=0.$$

Subtracting this equation, we have

$$\alpha_1(\lambda_1-\lambda_k)x_1+\alpha_2(\lambda_2-\lambda_k)x_2+\ldots+\alpha_{k-1}(\lambda_{k-1}-\lambda_k)x_{k-1}=0.$$

By our hypothesis, $\alpha_1=\alpha_2=\ldots=\alpha_{k-1}=0.$ Therefore $\alpha_k x_k=0.$ Since $x_k \neq 0,$ we have $\alpha_k=o.$ Therefore $x_1, x_2, \ldots, x_k$ are linearly independent.

1.7 Theorem. (Diagonalization of a matrix) Suppose $A\hat{I}\, R^{n'\,n}$ has n linearly independent eigenvectors. If these eigenvectors are columns of a matrix S, then $S^{-1}AS$ is a diagonal matrix D such that $D=diag(\lambda_1, \ldots, \lambda_n) \in R^{n\times n}$, $\lambda_i s'$ are eigenvalue of the matrix A.

Proof. Let $x_1, x_2,$ K $, x_n$ be eigenvectors of the matrix $A\hat{I}\, R^{n'\,n}$ associated with $\lambda_1,$ K $,\lambda_n,$ where each $x_i \in R^n (i=1,2,$ K $,n).$ Put the eigenvectors in the columns of S and

$$AS=A\begin{pmatrix} | & | & & | \\ x_1 & x_2 & \mathrm{L} & x_n \\ | & | & & | \end{pmatrix}=\begin{pmatrix} | & | & & | \\ \lambda_1 x_1 & \lambda_2 x_2 & & \lambda_n x_n \\ | & | & & | \end{pmatrix}.$$

$$=\begin{pmatrix} | & | & & | \\ x_1 & x_2 & \mathrm{L} & x_n \\ | & | & & | \end{pmatrix}\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \mathrm{O} & \\ & & & \lambda_n \end{pmatrix}.$$

Therefore $AS = SD$, where S is invertible since its columns are independent. Thus we have $S^{-1}AS = D$, or $A = SDS^{-1}$.

2. Cholesky Factorization for Real Positive Definite Matrices

Positive definite matrices can be decomposed into triangular factors twice as quickly as general matrices.

2.1 Definitions. The matrices as the form:

$$L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \in R^{n\times n} \text{ and } U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & & \ddots & \\ 0 & & & u_{nn} \end{pmatrix} \in R^{n\times n} \text{ are}$$

lower triangular matrix and **upper triangular matrix**, respectively.

Matrices L and U are **nonsingular** if and only if $\det(L) = \prod_{k=1}^{n} l_{kk} \neq 0$ and $\det(U) = \prod_{k=1}^{n} u_{kk} \neq 0$, respectively.

For every matrix $A = (a_{ij}) \in R^{n\times n}$, there are **principal submatrices** such as

$$\begin{pmatrix} a_{kk} & \cdots & a_{k,k+l} \\ \vdots & & \\ a_{k+l,k} & & a_{k+l,k+l} \end{pmatrix} \in R^{(l+1)\times(l+1)}.$$

Cholesky factorization is the decomposition of positive definite matrix such as $A = LL^T$.

2.2 Lemma. All principal submatrices of a positive definite matrix are also positive definite.

Proof. Let $A = (a_{ij}) \in R^{n\times n}$ be a positive definite matrix. Let $B \in R^{(l+1)\times(l+1)}$ be a principal submatrix of the matrix A represented by

$$B = \begin{pmatrix} a_{kk} & \cdots & a_{k,k+l} \\ \vdots & & \\ a_{k+l,K} & \cdots & a_{k+l,k+l} \end{pmatrix}.$$

Then B is symmetric. Let $0 \neq x = (x_i)_{i=k}^{k+l} \in R^{l+1}$. For $z = (z_i) \in R^n$ with

$$z_i = \begin{cases} x_i, & k \leq i \leq k+l \\ 0, & otherwise, \end{cases}$$

it follows that $z \neq 0$ and $x^T Bx = \sum_{i,j=k}^{k+l} a_{ij}x_i x_j = \sum_{i,j=1}^{n} a_{ij}z_i z_j = z^T Az > 0.$

2.3 Lemma. Let $A = (a_{ij}) \in R^{n \times n}$ be positive definite matrix. Then $\det(A) > 0$.

Proof. For the symmetric matrix $A \in R^{n \times n}$, there exists a factorization of the form $A = SDS^{-1}$, with $S \in R^{n \times n}$ is invertible, $D = diag(\lambda_1, \dots, \lambda_n) \in R^{n \times n}$. Since A is positive definite, it follows that, the eigenvalues $\lambda_1, \dots, \lambda_n$ of A are all real and positive. For $0 \neq x^{(k)} \in R^n$ with $Ax^{(k)} = \lambda x^{(k)}$, we have

$$0 < \left(x^{(k)}\right)^T A x^{(k)} = \left(\left(x^{(k)}\right)^T x^{(k)}\right)\lambda_k, \quad k = 1,2,\dots,n.$$

Therefore $\lambda_k > 0$ for $k = 1,2,\dots,n$, since $\left(x^{(k)}\right)^T x^{(k)} > 0$.

Since we have the property that $(\det S)(\det S^{-1}) = \det SS^{-1} = \det I = 1$,

$$\begin{aligned}\det(A) &= \det(S)\ \det(D)\ \det(S^{-1}) \\ &= \det(D) \\ &= \prod_{k=1}^{n} \lambda_k > 0.\end{aligned}$$

2.4 Theorem. Let $A \in R^{n \times n}$ be positive definite. Then there is unique lower triangular matrix $L = (l_{kj}) \in R^{n \times n}$ with $l_{kk} > 0$ for all k such that

$$A = LL^T. \qquad (2.1)$$

This is Cholesky factorization of the positive definite matrix A.

Proof. Let $A \in R^{n \times n}$ be positive definite matrix. By the mathematical induction on n, for $n = 1$,

$$A = (\alpha) \in R^{1 \times 1} \text{ with scalar } \alpha > 0$$

Thus $\alpha = l.l$, $l = \sqrt{\alpha}$ and $A = LL^T$ is true for $n = 1$ with $L = (\alpha) \in R^{1 \times 1}$. Assume

$$A_{n-1} = \begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & & a_{n-1,n-1} \end{pmatrix} = L_{n-1}L_{n-1}^T,$$

where $A_{n-1} \in R^{(n-1) \times (n-1)}$, and $L_{n-1} \in R^{(n-1) \times (n-1)}$ is lower triangular matrix. Consider the positive definite matrix

$$A = \left(\begin{array}{c|c} A_{n-1} & b \\ \hline b^T & a_{nn} \end{array}\right) \in R^{n \times n}$$

with vector $b \in R^{n-1}$. By Lemma 2.2, A_{n-1} is positive definite. By induction hypothesis, there is unique lower triangular matrix

$L_{n-1} = (l_{kj}) \in R^{(n-1)\times(n-1)}$ with $l_{kk} > 0$ for $k = 1, \dots, n-1$

and $A_{n-1} = L_{n-1} L_{n-1}^T$. Let

$$L = \begin{pmatrix} L_{n-1} & 0 \\ c^T & \alpha \end{pmatrix} \in R^{n \times n}$$

be required lower triangular matrix with $c \in R^{n-1}$ and scalar α such that

$$A = \begin{pmatrix} A_{n-1} & b \\ b^T & a_{nn} \end{pmatrix} = \begin{pmatrix} L_{n-1} & 0 \\ c^T & \alpha \end{pmatrix} \begin{pmatrix} L_{n-1}^T & c \\ 0 & \alpha \end{pmatrix} \tag{2.2}$$

holds if and only if

$$\begin{aligned} L_{n-1} c &= b \\ c^T c + \alpha^2 &= a_{nn} \end{aligned} \tag{2.3}$$

where $c \in R^{n-1}$ is unique solution of $c = L_{n-1}^{-1} b$ since $L_{n-1} \in R^{(n-1)\times(n-1)}$ is unique lower triangular matrix with $l_{kk} > 0 \quad (k = 1, \dots, n-1)$ is nonsingular. In addition to $\alpha \in ℝ$ is a solution of (2.3) which satisfies (2.2). From (2.2) we have

$$\det(A) = \det\left(\begin{pmatrix} L_{n-1} & 0 \\ c^T & \alpha \end{pmatrix}\right) \det\left(\begin{pmatrix} L_{n-1}^T & c \\ 0 & \alpha \end{pmatrix}\right)$$

$= \det(L_{n-1})^2 \alpha^2$.

Since $\det(A) > 0$ and L_{n-1} is nonsingular, it follows that $\alpha^2 > 0$; this implies that for (2.3) $\alpha > 0$ can be uniquely chosen. Thus we have the factorization $A = LL^T$.

The Cholesky factorization for the positive definite matrix $A \in R^{n \times n}$ as in equation (2.1) is the system of equations as $n(n+1)/2$ equations for the $n(n+1)/2$ unknowns $l_{ik}\ (i \geq k)$ such that

$$a_{ik} = \sum_{j=1}^{k} l_{ij} l_{kj}, \qquad 1 \leq k \leq i \leq n. \tag{2.4}$$

This system can be solved by the following algorithm.

2.5 Algorithm

Column-wise calculation of the entries defining the lower triangular matrix $L \in R^{n \times n}$ using equation (2.1) gives the following algorithm.

$$
\begin{aligned}
&\textit{for} \quad k = 1,\dots,n \\
&\quad l_{kk} = \left(a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2 \right)^{\frac{1}{2}}; \\
&\quad \textit{for} \quad i = k+1,\dots,n \\
&\quad\quad l_{kk} = \left(a_{ik} - \sum_{j=1}^{k-1} l_{ij} l_{kj} \right) \Big/ l_{kk} \quad ; \quad \textit{end} \\
&\textit{end}.
\end{aligned}
$$

Figure 2.1 LL^T factorization

2.6 Theorem. Every positive definite matrix has unique Cholesky factorization.

Proof. According to 2.5 Algorithm, The factorization exists. The uniqueness is also established by 2.5 Algorithm. The value $\alpha = \sqrt{a_{11}}$ is determined by the form of the $A = LL^T$ factorization, and once α is determined, the first row of L_1^T is obtained. Similarly, we can find other quantities *ls*' are determined at each step. The factorization is unique.

Now we can apply the Cholesky factorization solving the linear system $Ax = b$.

2.6 Example. Consider the linear system

$$4x_1 + 2x_2 + 14x_3 = 14$$
$$2x_1 + 17x_2 - 5x_3 = -101$$
$$14x_1 - 5x_2 + 83x_3 = 155.$$

The coefficient matrix $A = \begin{pmatrix} 4 & 2 & 14 \\ 2 & 17 & -5 \\ 14 & -5 & 83 \end{pmatrix}$ is positive definite. Then we have Cholesky factorization

$$A = \begin{pmatrix} 4 & 2 & 14 \\ 2 & 17 & -5 \\ 14 & -5 & 83 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{pmatrix}.$$

$$l_{11} = \sqrt{a_{11}} = 2$$

$$l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{2} = 1$$

$$l_{31} = \frac{a_{31}}{l_{11}} = \frac{14}{2} = 7$$

$$l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{17-1} = 4$$

$$l_{32} = \frac{1}{l_{22}}(a_{32} - l_{31}l_{21}) = \frac{1}{4}(-5 - 7\times 1) = -3$$

$$l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{83 - 7^2 - (-3)^2} = 5 .$$

First we solve the system $Ly = b$. That is

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 7 & -3 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 14 \\ -101 \\ 155 \end{pmatrix}.$$

The solution is $y = \begin{pmatrix} 7 \\ -27 \\ 5 \end{pmatrix}$.

As the second step, we have to solve $L^T x = y$, that is,

$$\begin{pmatrix} 2 & 1 & 7 \\ 0 & 4 & -3 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ -27 \\ 5 \end{pmatrix}.$$

The solution is $x = \begin{pmatrix} 3 \\ -6 \\ 1 \end{pmatrix}$.

2.8 Theorem

In the calculation of Cholesky factorization, the total number of arithmetic operations is

$$\frac{n^3}{3}\left(1 + O\left(\frac{1}{n}\right)\right).$$

Proof. By the Algorithm 2.5, the total number of operations is given by,

$$\sum_{k=1}^{n}\left((2k-1)+\sum_{i=k+1}^{n}(2k-1)\right)$$

$$=\sum_{k=1}^{n}\big((n+1-k)(2k-1)\big)$$

$$=-\sum_{k=1}^{n}(n+1-k)+2\sum_{k=1}^{n}(n+1-k)k$$

$$=-\sum_{k=1}^{n}k+2(n+1)\sum_{k=1}^{n}k-2\sum_{k=1}^{n}k^{2}$$

$$=(2n+1)\frac{n(n+1)}{2}-2\frac{n(n+1)(2n+1)}{6}$$

$$=\frac{1}{6}(2n+1)(n^{2}+n)$$

$$=\frac{1}{6}(2n^{3}+3n^{2}+n)$$

$$=\frac{2n^{3}}{6}(1+\frac{3}{2n}+\frac{1}{2n^{2}})$$

$$=\frac{2n^{3}}{6}(1+\frac{3}{2n}+\frac{1}{2n^{2}})$$

$$=\frac{n^{3}}{3}\left(1+\mathrm{O}\left(\frac{1}{n}\right)\right)$$

2.9 Example.

If $A\in\mathbb{R}^{3\times3}$ is a positive definite matrix, then the total number of operations is given by $\frac{27}{3}\left(1+\mathrm{O}\left(\frac{1}{3}\right)\right)=14$, where $\mathrm{O}\left(\frac{1}{3}\right)=\frac{5}{9}$.

Acknowledgement

First of all, we would like to express our heartfelt thanks to Dr. Pyone Pyone Aung (Pro Rector, Yangon University of Education) and Dr. Kay Thwe Hling (Pro Rector, Yangon University of Education).

We would like to offer our gratitude to Dr. Lwin Mar Htun (Professor and Head of Department, Mathematics Department, Yangon University of Education) for her continuous encouragement, invaluable guidance and timely advice, expert judgments and critical remarks throughout the study.

References

Lang.S.; 'Linear Algebra', 3rd ed., Springer, 2004.

Lloyd N. Trefethen David Bau, ' Numerical Linear Algebra' SIAM, 1997.

Plato.R.; ' Concise Numerical Mathematics', AMS, Rhode, 2004.

Strang. G.; ' Linear Algebra and its Applications', 4th ed., Thomson Brooks/Cole, 2006.