CHOLESKY DECOMPOSITION

QINGWEN LIU

1. Introduction

LU decomposition is decomposing a matrix as the product of a lower triangular matrix $L$ and an upper triangular matrix $U$. For LU decomposition, if $A \in M_{n \times n}(\mathbb{R})$, then $A = LU$. In Cholesky decomposition, the lower and upper triangular matrices are transposes of each other, which means that an LU decomposition is a Cholesky decomposition if $U = L^t$, which could be represented as $A = LL^t$.

1.1. Cholesky decomposition. Cholesky decomposition, or Cholesky factorization, is used in the special case when $A$ is a square, conjugate symmetric matrix (or Hermitian) positive-definite matrix into the product of a lower triangular matrix, which makes the problem simpler. Conjugate symmetric matrix is one where the element $A_{ij}$ equals the element $A_{ji}$ conjugated, which is shown as $A_{ij} = \overline{A_{ji}}$ if $A_{ij}$ is complex.

1.2. Examples. An example of LU Factorization:

$$
\begin{pmatrix}
2 & 3 \\
2 & 7
\end{pmatrix} =
\begin{pmatrix}
1 & 0 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
2 & 3 \\
0 & 4
\end{pmatrix}
$$

An example Cholesky decomposition:

$$
\begin{pmatrix}
1 & 1 \\
1 & 2
\end{pmatrix} =
\begin{pmatrix}
1 & 0 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix} =
\begin{pmatrix}
1 & 0 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}^t.
$$

2. Definition

Let $L \in M_{n \times n}(\mathbb{R})$ be a lower triangular matrix, that is $l_{ij} = 0$ if $i < j$. Then,

$$
L =
\begin{pmatrix}
l_{11} & 0 & 0 & \ldots & 0 \\
l_{21} & l_{22} & 0 & \ldots & 0 \\
l_{31} & l_{32} & l_{33} & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
l_{n1} & l_{n2} & l_{n3} & \ldots & l_{nn}
\end{pmatrix}
$$

Matrix $U$, the upper triangular matrix, is in $\in M_{n \times n}(\mathbb{R})$, that is $u_{ij} = 0$ if $i > j$. Then,

$$
U =
\begin{pmatrix}
u_{11} & u_{12} & u_{13} & \ldots & u_{1n} \\
0 & u_{22} & u_{23} & \ldots & u_{2n} \\
0 & 0 & u_{33} & \ldots & u_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & u_{nn}
\end{pmatrix}
$$

Date: May 3, 2022.
In Cholesky decomposition, \( U = L^t \). Then,

\[
U = L^t = \begin{pmatrix}
l_{11} & l_{21} & l_{31} & \ldots & l_{n1} \\
l_{22} & l_{32} & l_{42} & \ldots & l_{n2} \\
l_{33} & l_{43} & l_{53} & \ldots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
l_{nn} & \end{pmatrix}.
\]

For each \( A \in M_{n \times n}(\mathbb{R}) \) has a unique Cholesky decomposition,

\[
\begin{pmatrix}
A_{11} & A_{12} & \ldots & A_{1n} \\
A_{21} & A_{22} & \ldots & A_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} & A_{n2} & \ldots & A_{nn}
\end{pmatrix}
= \begin{pmatrix}
l_{11} & 0 & \ldots & 0 \\
l_{21} & l_{22} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
l_{n1} & l_{n2} & \ldots & l_{nn}
\end{pmatrix}
\begin{pmatrix}
l_{11} & l_{12} & \ldots & l_{1n} \\
l_{21} & l_{22} & \ldots & l_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
l_{n1} & l_{n2} & \ldots & l_{nn}
\end{pmatrix}.
\]

3. Calculating \( L \)

Suppose \( A \) is a symmetric positive matrix whose elements are denoted by \( a_{ij} \). Let \( L \) be a lower triangular matrix whose elements are denoted by \( l_{ij} \). The Cholesky decomposition can be represented in the following form:

\[
l_{11} = \sqrt{a_{11}}
\]

\[
l_{j1} = a_{j1}/l_{11}, \quad j \in [2, n]
\]

\[
l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}, \quad i \in [2, n]
\]

\[
l_{ji} = (a_{ij} - \sum_{k=1}^{i-1} l_{jk} l_{ik})/l_{ii}, \quad i \in [2, n-1], j \in [i+1, n].
\]

3.1. Example. Solve equations \( 6x + 15y + 55z = 76, 15x + 55y + 225z = 295, 55x + 225y + 979z = 1259 \) using Cholesky decomposition. Three equations are

\[
6x + 15y + 55z = 76 \\
15x + 55y + 225z = 295 \\
55x + 225y + 979z = 1259.
\]

Converted three equations into matrix

\[
\begin{pmatrix}
6 & 15 & 55 \\
15 & 55 & 225 \\
55 & 225 & 979
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= \begin{pmatrix}
76 \\
295 \\
1259
\end{pmatrix}.
\]

For matrix

\[
\begin{pmatrix}
6 & 15 & 55 \\
15 & 55 & 225 \\
55 & 225 & 979
\end{pmatrix}
\]

could be reduced to eliminated matrix by using row operations, which is

\[
\begin{pmatrix}
6 & 15 & 55 \\
0 & 3\frac{1}{2} & 17\frac{1}{2} \\
0 & 0 & 1\frac{1}{2}
\end{pmatrix}.
\]
CHOLESKY DECOMPOSITION

Pivots are the first non-zero element in each row of eliminated matrix, so pivots are 6, \(\frac{35}{2}, \frac{112}{3}\). Thus, matrix is symmetric positive definite here, and Cholesky decomposition is possible. By using the formula provided above, we get

\[
\begin{align*}
l_{11} &= \sqrt{a_{11}} = \sqrt{6} = 2.4495 \\
l_{21} &= a_{21} l_{11} = 15 \times \frac{2.4495}{2} = 6.1237 \\
l_{22} &= \sqrt{a_{22} - l_{21}^2} = \sqrt{55 - (6.1237)^2} = \sqrt{55 - 37.5} = 4.1833 \\
l_{31} &= a_{31} l_{11} = \frac{55}{2.4495} = 22.4537 \\
l_{32} &= \frac{a_{32} - l_{31} l_{21}}{l_{22}} = \frac{225 - (22.4527) \times (6.1237)}{4.1833} = \frac{225 - 137.5}{4.1833} = 20.9165 \\
l_{33} &= \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{979 - (22.4537)^2 - (20.9165)^2} = \sqrt{979 - 941.6667} = 6.1101
\end{align*}
\]

So,

\[
L = \begin{pmatrix}
l_{11} & 0 & 0 \\
l_{21} & l_{22} & 0 \\
l_{31} & l_{32} & l_{33}
\end{pmatrix} = \begin{pmatrix}
2.4495 & 0 & 0 \\
6.1237 & 4.1833 & 0 \\
22.4537 & 20.9165 & 6.1101
\end{pmatrix}
\]

and

\[
L^t = \begin{pmatrix}
2.4495 & 6.1237 & 22.4537 \\
0 & 4.1833 & 20.9165 \\
0 & 0 & 6.1101
\end{pmatrix}.
\]

Thus,

\[
LL^t = \begin{pmatrix}
2.4495 & 0 & 0 \\
6.1237 & 4.1833 & 0 \\
22.4537 & 20.9165 & 6.1101
\end{pmatrix} \begin{pmatrix}
2.4495 & 0 & 0 \\
6.1237 & 4.1833 & 0 \\
22.4537 & 20.9165 & 6.1101
\end{pmatrix} = \begin{pmatrix}
6 & 15 & 55 \\
15 & 55 & 225 \\
55 & 225 & 979
\end{pmatrix}.
\]

Now, \(Ax = B\), and \(A = LL^t \Rightarrow LL^t x = B\). Let \(L^t x = y\), then

\[
Ly = B \Rightarrow \begin{pmatrix}
2.4495 & 0 & 0 \\
6.1237 & 4.1833 & 0 \\
22.4537 & 20.9165 & 6.1101
\end{pmatrix} \begin{pmatrix}
y_1 \\
y_2 \\
y_3
\end{pmatrix} = \begin{pmatrix}
76 \\
295 \\
1259
\end{pmatrix}.
\]

Then, we can get \(y_1 = 31.0269, y_2 = 25.0998, y_3 = 6.1101\). Since \(L^t x = y\), then

\[
\begin{pmatrix}
2.4495 & 6.1237 & 22.4537 \\
0 & 4.1833 & 20.9165 \\
0 & 0 & 6.1101
\end{pmatrix} \begin{pmatrix}
x \\
y \\
z
\end{pmatrix} = \begin{pmatrix}
31.0269 \\
25.0998 \\
6.1101
\end{pmatrix}
\]

and we get \(x = 1, y = 1, z = 1\). Thus, the solution by Cholesky decomposition method is \(x = 1, y = 1, z = 1\).

4. BEHIND THE MODELS

André-Louis Cholesky discovered the linear algebra method that carries his name through his work as late 19th century map maker, and it continues to be an efficient trick that fuels many machine learning models. His work was published after he died in the battle during World War I. The decomposition has applications in Least–Squares Regression and Monte-Carlo Simulation.
4.1. **Least-Squares Regression.** A linear regression takes the form $Y = X\beta$, where $Y$ is a vector of dependent variables, and $X$ is a vector of independent variables. Least-Squares Regression refers to finding the vector $\beta$ where the squared differences between the predicted and actual $Y$ values are minimized. The Cholesky decomposition is roughly twice as efficient as other methods for solving systems of linear equations.

4.2. **Monte-Carlo Simulation.** Cholesky decomposition helps us to simulate uncorrelated normal variables and transform them into correlated normal variables. Take normal random variables for example, we go from uncorrelated,

![uncorrelated data](image1)

... to correlated

![correlated data](image2)

4.3. **Conclusion.** Cholesky decomposition underlies many machine learning applications, and Least-Squares Regression and Monte-Carlo Simulation are two of applications. It is important to remember that Cholesky decomposition is an efficient method for notating system of linear equations.