

1 Positive Definite Systems & Cholesky Decomposition (1.4)

In the previous section, we learned about Gaussian Elimination with Pivoting (i.e., $\mathbf{Ax} = \mathbf{b} \iff PA = LU$), which can be used to solve *any* system with invertible matrix A . However, for some matrix A with special properties, we can use a *better* way to solve it, mainly for computational efficiency.

One such class of matrices are **positive definite** ones. In particular, if A is $n \times n$, real, and satisfies the conditions

1. $A = A^T$ (i.e., symmetric), and

2. $\mathbf{x}^T A \mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \langle x^T, Ax \rangle > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$,

then A is said to be **positive definite**.

1.1 Properties of Positive Definite Matrix

There are some useful properties of positive definite matrices.

Lemma 1.1

A positive definite matrix is invertible.

Theorem 1.1

Let M be an invertible $n \times n$ matrix. Then, $A = M^T M$ is positive definite.

Theorem 1.2: Cholesky Decomposition Theorem

Let A be positive definite. Then, there exists an upper-triangular matrix R such that

$$A = R^T R.$$

This is known as the Cholesky decomposition, and R is called the **Cholesky factor** of A . In addition, this decomposition is unique if we choose $r_{ii} > 0$ for $i = 1, 2, \dots, n$.

(Example: Cholesky Algorithm.) Suppose

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix},$$

so that

$$R^T = \begin{bmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & r_{32} & r_{33} \end{bmatrix}.$$

Then,

$$\begin{aligned} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= \begin{bmatrix} r_{11} & 0 & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \\ &= \begin{bmatrix} r_{11}^2 & r_{11}r_{12} & r_{11}r_{13} \\ r_{11}r_{12} & r_{12}^2 + r_{22}^2 & r_{12}r_{13} + r_{22}r_{23} \\ r_{11}r_{13} & r_{13}r_{12} + r_{23}r_{22} & r_{13}^2 + r_{23}^2 + r_{33}^2 \end{bmatrix} \end{aligned}$$

One thing to notice is that this matrix is symmetric. So, when finding the values of each element, we only need to find the values of specific elements (one side). We can find the value of each element in R . Starting with the first row, we have

- $r_{11}^2 = a_{11} \Rightarrow \boxed{r_{11} = \sqrt{a_{11}}}$.
- $r_{11}r_{12} = a_{12} \Rightarrow \boxed{r_{12} = \frac{a_{12}}{r_{11}}}$.
- $r_{11}r_{13} = a_{13} \Rightarrow \boxed{r_{13} = \frac{a_{13}}{r_{11}}}$.

For the second row, we have

- $r_{12}^2 + r_{22}^2 = a_{22} \Rightarrow \boxed{r_{22} = \sqrt{a_{22} - r_{12}^2}}$.
- $r_{12}r_{13} + r_{22}r_{23} = a_{23} \Rightarrow \boxed{r_{23} = \frac{a_{23} - r_{12}r_{13}}{r_{22}}}$.

For the third row, we have

- $r_{13}^2 + r_{23}^2 + r_{33}^2 = a_{33} \Rightarrow \boxed{r_{33} = \sqrt{a_{33} - r_{13}^2 - r_{23}^2}}$.

The above example is the Cholesky Algorithm for a 3×3 matrix. For the general case, suppose we have

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}}_A = \underbrace{\begin{bmatrix} r_{11} & 0 & 0 & \dots & 0 \\ r_{21} & r_{12} & 0 & \dots & 0 \\ r_{31} & r_{23} & r_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{1n} & r_{2n} & r_{3n} & \dots & r_{nn} \end{bmatrix}}_{R^T} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & \dots & r_{1n} \\ 0 & r_{22} & r_{23} & \dots & r_{2n} \\ 0 & 0 & r_{33} & \dots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & r_{nn} \end{bmatrix}}_R.$$

Here,

$$a_{ij} = \begin{cases} \sum_{k=1}^i r_{ki}r_{kj} & i < j \\ \sum_{k=1}^i r_{ki}^2 & i = j \\ \sum_{k=1}^j r_{ki}r_{kj} & i > j \end{cases}.$$

From this, we're able to solve r_{ii} :

$$r_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}.$$

After solving for r_{ii} , we can find r_{ij} :

$$r_{ij} = \frac{\left(a_{ij} - \sum_{k=1}^{i-1} r_{ki}r_{kj}\right)}{r_{ii}}$$

for $j = i + 1, \dots, n$, noting that for $j < i$, $r_{ij} = 0$.

In general, Cholesky is doing half of the LU decomposition work: $\mathcal{O}(n^3)$.