Math 221: LINEAR ALGEBRA

Chapter 8. Orthogonality

§8-3. Positive Definite Matrices

Le Chen[1]

Emory University, 2021 Spring

(last updated on 01/25/2021)

[1] Slides are adapted from those by Karen Seyffarth from University of Calgary.

Positive Definite Matrices

Cholesky factorization – Square Root of a Matrix

Positive Definite Matrices

Cholesky factorization – Square Root of a Matrix

Positive Definite Matrices

Definition

An $n \times n$ matrix A is positive definite if it is symmetric and has positive eigenvalues, i.e., if λ is a eigenvalue of A, then $\lambda > 0$.

Theorem

If A is a positive definite matrix, then $\det(A) > 0$ and A is invertible.

Proof.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the (not necessarily distinct) eigenvalues of A. Since A is symmetric, A is orthogonally diagonalizable. In particular, $A \sim D$, where $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Similar matrices have the same determinant, so

$$\det(A) = \det(D) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Since A is positive definite, $\lambda_i > 0$ for all i, $1 \leq i \leq n$; it follows that $\det(A) > 0$, and therefore A is invertible. ∎

Theorem

A symmetric matrix A is positive definite if and only if $\vec{x}^T A\vec{x} > 0$ for all $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$.

Proof.

Since A is symmetric, there exists an orthogonal matrix P so that

$$P^T AP = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = D,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of A. Let $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$, and define $\vec{y} = P^T\vec{x}$. Then

$$\vec{x}^T A\vec{x} = \vec{x}^T(PDP^T)\vec{x} = (\vec{x}^T P)D(P^T\vec{x}) = (P^T\vec{x})^T D(P^T\vec{x}) = \vec{y}^T D\vec{y}.$$

Writing $\vec{y}^T = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}$,

$$\begin{aligned} \vec{x}^T A\vec{x} &= \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2. \end{aligned}$$

Proof. (continued)

($\Rightarrow$) Suppose A is positive definite, and $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$. Since P^T is invertible, $\vec{y} = P^T\vec{x} \neq \vec{0}$, and thus $y_j \neq 0$ for some j, implying $y_j^2 > 0$ for some j. Furthermore, since all eigenvalues of A are positive, $\lambda_i y_i^2 \geq 0$ for all i; in particular $\lambda_j y_j^2 > 0$. Therefore, $\vec{x}^T A\vec{x} > 0$.

($\Leftarrow$) Conversely, if $\vec{x}^T A\vec{x} > 0$ whenever $\vec{x} \neq \vec{0}$, choose $\vec{x} = P\vec{e}_j$, where $\vec{e}_j$ is the j^{th} column of I_n. Since P is invertible, $\vec{x} \neq \vec{0}$, and thus

$$\vec{y} = P^T\vec{x} = P^T(P\vec{e}_j) = \vec{e}_j.$$

Thus $y_j = 1$ and $y_i = 0$ when $i \neq j$, so

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2 = \lambda_j,$$

i.e., $\lambda_j = \vec{x}^T A\vec{x} > 0$. Therefore, A is positive definite. ■

Theorem (Constructing Positive Definite Matrices)

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

Proof.

Let $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$. Then

$$\begin{aligned} \vec{x}^T A \vec{x} &= \vec{x}^T (U^T U) \vec{x} \\ &= (\vec{x}^T U^T)(U\vec{x}) \\ &= (U\vec{x})^T (U\vec{x}) \\ &= ||U\vec{x}||^2. \end{aligned}$$

Since U is invertible and $\vec{x} \neq \vec{0}$, $U\vec{x} \neq \vec{0}$, and hence $||U\vec{x}||^2 > 0$, i.e., $\vec{x}^T A \vec{x} = ||U\vec{x}||^2 > 0$. Therefore, A is positive definite. ■

Definition

Let $A = [\; a_{ij} \;]$ be an $n \times n$ matrix. For $1 \le r \le n$, ${}^{(r)}A$ denotes the $r \times r$ submatrix in the upper left corner of A, i.e.,

$$ {}^{(r)}A = [\; a_{ij} \;],\; 1 \le i, j \le r. $$

${}^{(1)}A$, ${}^{(2)}A, \ldots,$ ${}^{(n)}A$ are called the principal submatrices of A.

Lemma

If A is an $n \times n$ positive definite matrix, then each principal submatrix of A is positive definite.

Proof.

Suppose A is an $n \times n$ positive definite matrix. For any integer r, $1 \leq r \leq n$, write A in block form as

$$A = \begin{bmatrix} {}^{(r)}A & B \\ C & D \end{bmatrix},$$

where B is an $r \times (n-r)$ matrix, C is an $(n-r) \times r$ matrix, and D is an $(n-r) \times (n-r)$ matrix. Let $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{bmatrix} \neq \vec{0}$ and let $\vec{x} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. Then $\vec{x} \neq \vec{0}$, and by the previous theorem, $\vec{x}^T A \vec{x} > 0$.

Proof. (continued)

But

$$\vec{x}^T A \vec{x} = \begin{bmatrix} y_1 & \cdots & y_r & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} {}^{(r)}A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{y}^T \left({}^{(r)}A \right) \vec{y},$$

and therefore $\vec{y}^T \left({}^{(r)}A \right) \vec{y} > 0$. Then ${}^{(r)}A$ is positive definite again by the previous theorem. ■

Positive Definite Matrices

Cholesky factorization – Square Root of a Matrix

Cholesky factorization – Square Root of a Matrix

$$4 = 2 \times 2^{\mathrm{T}}$$

$$\begin{bmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 6 & 1 & 0 \\ -8 & 5 & 3 \end{bmatrix} \begin{bmatrix} 2 & 6 & -8 \\ 0 & 1 & 5 \\ 0 & 0 & 3 \end{bmatrix}$$

Theorem

Let A be an $n \times n$ symmetric matrix. Then the following conditions are equivalent.

1. A is positive definite.
2. $\det({}^{(r)}A) > 0$ for $r = 1, 2, \ldots, n$.
3. $A = U^{\mathrm{T}}U$ where U is upper triangular and has positive entries on its main diagonal. Furthermore, U is unique. The expression $A = U^{\mathrm{T}}U$ is called the Cholesky factorization of A.

Algorithm for Cholesky Factorization

Let A be a positive definite matrix. The Cholesky factorization $A = U^T U$ can be obtained as follows.

1. Using only type 3 elementary row operations, with multiples of rows added to lower rows, put A in upper triangular form. Call this matrix $\widehat{U}$; then $\widehat{U}$ has positive entries on its main diagonal (this can be proved by induction on n).
2. Obtain U from $\widehat{U}$ by dividing each row of $\widehat{U}$ by the square root of the diagonal entry in that row.

Problem

Show that $A = \begin{bmatrix} 9 & -6 & 3 \\ -6 & 5 & -3 \\ 3 & -3 & 6 \end{bmatrix}$ is positive definite, and find the Cholesky factorization of A.

Solution

$$^{(1)}A = \begin{bmatrix} 9 \end{bmatrix} \quad \text{and} \quad ^{(2)}A = \begin{bmatrix} 9 & -6 \\ -6 & 5 \end{bmatrix},$$

so $\det(^{(1)}A) = 9$ and $\det(^{(2)}A) = 9$. Since $\det(A) = 36$, it follows that A is positive definite.

$$\begin{bmatrix} 9 & -6 & 3 \\ -6 & 5 & -3 \\ 3 & -3 & 6 \end{bmatrix} \to \begin{bmatrix} 9 & -6 & 3 \\ 0 & 1 & -1 \\ 0 & -1 & 5 \end{bmatrix} \to \begin{bmatrix} 9 & -6 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix}$$

Now divide the entries in each row by the square root of the diagonal entry in that row, to give

$$U = \begin{bmatrix} 3 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad U^T U = A.$$

■

Problem

Verify that

$$A = \begin{bmatrix} 12 & 4 & 3 \\ 4 & 2 & -1 \\ 3 & -1 & 7 \end{bmatrix}$$

is positive definite, and find the Cholesky factorization of A.

Solution (Final Answer)

$\det\left({}^{(1)}A\right) = 12$, $\det\left({}^{(2)}A\right) = 8$, $\det(A) = 2$; by the previous theorem, A is positive definite.

$$U = \begin{bmatrix} 2\sqrt{3} & 2\sqrt{3}/3 & \sqrt{3}/2 \\ 0 & \sqrt{6}/3 & -\sqrt{6} \\ 0 & 0 & 1/2 \end{bmatrix}$$

and $U^T U = A$.