

Exercise 8.2.25 Show that every 2×2 orthogonal matrix has the form $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ or $\begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}$ for some angle θ . [Hint: If $a^2 + b^2 = 1$, then $a = \cos \theta$ and $b = \sin \theta$ for some angle θ .]

Exercise 8.2.26 Use Theorem 8.2.5 to show that every symmetric matrix is orthogonally diagonalizable.

8.3 Positive Definite Matrices

All the eigenvalues of any symmetric matrix are real; this section is about the case in which the eigenvalues are positive. These matrices, which arise whenever optimization (maximum and minimum) problems are encountered, have countless applications throughout science and engineering. They also arise in statistics (for example, in factor analysis used in the social sciences) and in geometry (see Section 8.9). We will encounter them again in Chapter 10 when describing all inner products in \mathbb{R}^n .

Definition 8.5 Positive Definite Matrices

A square matrix is called **positive definite** if it is symmetric and all its eigenvalues λ are positive, that is $\lambda > 0$.

Because these matrices are symmetric, the principal axes theorem plays a central role in the theory.

Theorem 8.3.1

If A is positive definite, then it is invertible and $\det A > 0$.

Proof. If A is $n \times n$ and the eigenvalues are $\lambda_1, \lambda_2, \dots, \lambda_n$, then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n > 0$ by the principal axes theorem (or the corollary to Theorem 8.2.5). \square

If \mathbf{x} is a column in \mathbb{R}^n and A is any real $n \times n$ matrix, we view the 1×1 matrix $\mathbf{x}^T A \mathbf{x}$ as a real number. With this convention, we have the following characterization of positive definite matrices.

Theorem 8.3.2

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

Proof. A is symmetric so, by the principal axes theorem, let $P^T A P = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ where $P^{-1} = P^T$ and the λ_i are the eigenvalues of A . Given a column \mathbf{x} in \mathbb{R}^n , write $\mathbf{y} = P^T \mathbf{x} = [y_1 \ y_2 \ \dots \ y_n]^T$. Then

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T (P D P^T) \mathbf{x} = \mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \quad (8.3)$$

If A is positive definite and $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{x}^T A \mathbf{x} > 0$ by (8.3) because some $y_j \neq 0$ and every $\lambda_i > 0$. Conversely, if $\mathbf{x}^T A \mathbf{x} > 0$ whenever $\mathbf{x} \neq \mathbf{0}$, let $\mathbf{x} = P \mathbf{e}_j \neq \mathbf{0}$ where \mathbf{e}_j is column j of I_n . Then $\mathbf{y} = \mathbf{e}_j$, so (8.3) reads $\lambda_j = \mathbf{x}^T A \mathbf{x} > 0$. \square

Note that Theorem 8.3.2 shows that the positive definite matrices are exactly the symmetric matrices A for which the quadratic form $q = \mathbf{x}^T A \mathbf{x}$ takes only positive values.

Example 8.3.1

If U is any invertible $n \times n$ matrix, show that $A = U^T U$ is positive definite.

Solution. If \mathbf{x} is in \mathbb{R}^n and $\mathbf{x} \neq \mathbf{0}$, then

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T (U^T U) \mathbf{x} = (U \mathbf{x})^T (U \mathbf{x}) = \|U \mathbf{x}\|^2 > 0$$

because $U \mathbf{x} \neq \mathbf{0}$ (U is invertible). Hence Theorem 8.3.2 applies.

It is remarkable that the converse to Example 8.3.1 is also true. In fact every positive definite matrix A can be factored as $A = U^T U$ where U is an upper triangular matrix with positive elements on the main diagonal. However, before verifying this, we introduce another concept that is central to any discussion of positive definite matrices.

If A is any $n \times n$ matrix, let ${}^{(r)}A$ denote the $r \times r$ submatrix in the upper left corner of A ; that is, ${}^{(r)}A$ is the matrix obtained from A by deleting the last $n - r$ rows and columns. The matrices ${}^{(1)}A$, ${}^{(2)}A$, ${}^{(3)}A$, ..., ${}^{(n)}A = A$ are called the **principal submatrices** of A .

Example 8.3.2

If $A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$ then ${}^{(1)}A = [10]$, ${}^{(2)}A = \begin{bmatrix} 10 & 5 \\ 5 & 3 \end{bmatrix}$ and ${}^{(3)}A = A$.

Lemma 8.3.1

If A is positive definite, so is each principal submatrix ${}^{(r)}A$ for $r = 1, 2, \dots, n$.

Proof. Write $A = \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix}$ in block form. If $\mathbf{y} \neq \mathbf{0}$ in \mathbb{R}^r , write $\mathbf{x} = \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix}$ in \mathbb{R}^n .

Then $\mathbf{x} \neq \mathbf{0}$, so the fact that A is positive definite gives

$$0 < \mathbf{x}^T A \mathbf{x} = \begin{bmatrix} \mathbf{y}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} = \mathbf{y}^T ({}^{(r)}A) \mathbf{y}$$

This shows that ${}^{(r)}A$ is positive definite by Theorem 8.3.2.⁵ □

If A is positive definite, Lemma 8.3.1 and Theorem 8.3.1 show that $\det({}^{(r)}A) > 0$ for every r . This proves part of the following theorem which contains the converse to Example 8.3.1, and characterizes the positive definite matrices among the symmetric ones.

⁵A similar argument shows that, if B is any matrix obtained from a positive definite matrix A by deleting certain rows and deleting the *same* columns, then B is also positive definite.

Theorem 8.3.3

The following conditions are equivalent for a symmetric $n \times n$ matrix A :

1. A is positive definite.
2. $\det({}^{(r)}A) > 0$ for each $r = 1, 2, \dots, n$.
3. $A = U^T U$ where U is an upper triangular matrix with positive entries on the main diagonal.

Furthermore, the factorization in (3) is unique (called the **Cholesky factorization**⁶ of A).

Proof. First, (3) \Rightarrow (1) by Example 8.3.1, and (1) \Rightarrow (2) by Lemma 8.3.1 and Theorem 8.3.1.

(2) \Rightarrow (3). Assume (2) and proceed by induction on n . If $n = 1$, then $A = [a]$ where $a > 0$ by (2), so take $U = [\sqrt{a}]$. If $n > 1$, write $B = {}^{(n-1)}A$. Then B is symmetric and satisfies (2) so, by induction, we have $B = U^T U$ as in (3) where U is of size $(n-1) \times (n-1)$. Then, as A is symmetric, it has block form $A = \begin{bmatrix} B & \mathbf{p} \\ \mathbf{p}^T & b \end{bmatrix}$ where \mathbf{p} is a column in \mathbb{R}^{n-1} and b is in \mathbb{R} . If we write $\mathbf{x} = (U^T)^{-1} \mathbf{p}$ and $c = b - \mathbf{x}^T \mathbf{x}$, block multiplication gives

$$A = \begin{bmatrix} U^T U & \mathbf{p} \\ \mathbf{p}^T & b \end{bmatrix} = \begin{bmatrix} U^T & 0 \\ \mathbf{x}^T & 1 \end{bmatrix} \begin{bmatrix} U & \mathbf{x} \\ 0 & c \end{bmatrix}$$

as the reader can verify. Taking determinants and applying Theorem 3.1.5 gives $\det A = \det(U^T) \det U \cdot c = c(\det U)^2$. Hence $c > 0$ because $\det A > 0$ by (2), so the above factorization can be written

$$A = \begin{bmatrix} U^T & 0 \\ \mathbf{x}^T & \sqrt{c} \end{bmatrix} \begin{bmatrix} U & \mathbf{x} \\ 0 & \sqrt{c} \end{bmatrix}$$

Since U has positive diagonal entries, this proves (3).

As to the uniqueness, suppose that $A = U^T U = U_1^T U_1$ are two Cholesky factorizations. Now write $D = U U_1^{-1} = (U^T)^{-1} U_1^T$. Then D is upper triangular, because $D = U U_1^{-1}$, and lower triangular, because $D = (U^T)^{-1} U_1^T$, and so it is a diagonal matrix. Thus $U = D U_1$ and $U_1 = D U$, so it suffices to show that $D = I$. But eliminating U_1 gives $U = D^2 U$, so $D^2 = I$ because U is invertible. Since the diagonal entries of D are positive (this is true of U and U_1), it follows that $D = I$. \square

The remarkable thing is that the matrix U in the Cholesky factorization is easy to obtain from A using row operations. The key is that Step 1 of the following algorithm is *possible* for any positive definite matrix A . A proof of the algorithm is given following Example 8.3.3.

Algorithm for the Cholesky Factorization

If A is a positive definite matrix, the Cholesky factorization $A = U^T U$ can be obtained as follows:

Step 1. Carry A to an upper triangular matrix U_1 with positive diagonal entries using row operations each of which adds a multiple of a row to a lower row.

Step 2. Obtain U from U_1 by dividing each row of U_1 by the square root of the diagonal entry in that row.

⁶Andre-Louis Cholesky (1875–1918), was a French mathematician who died in World War I. His factorization was published in 1924 by a fellow officer.

Example 8.3.3

Find the Cholesky factorization of $A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$.

Solution. The matrix A is positive definite by Theorem 8.3.3 because $\det^{(1)}A = 10 > 0$, $\det^{(2)}A = 5 > 0$, and $\det^{(3)}A = \det A = 3 > 0$. Hence Step 1 of the algorithm is carried out as follows:

$$A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 10 & 5 & 2 \\ 0 & \frac{1}{2} & 1 \\ 0 & 1 & \frac{13}{5} \end{bmatrix} \rightarrow \begin{bmatrix} 10 & 5 & 2 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & \frac{3}{5} \end{bmatrix} = U_1$$

Now carry out Step 2 on U_1 to obtain $U = \begin{bmatrix} \sqrt{10} & \frac{5}{\sqrt{10}} & \frac{2}{\sqrt{10}} \\ 0 & \frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & 0 & \frac{\sqrt{3}}{\sqrt{5}} \end{bmatrix}$.

The reader can verify that $U^T U = A$.

Proof of the Cholesky Algorithm. If A is positive definite, let $A = U^T U$ be the Cholesky factorization, and let $D = \text{diag}(d_1, \dots, d_n)$ be the common diagonal of U and U^T . Then $U^T D^{-1}$ is lower triangular with ones on the diagonal (call such matrices LT-1). Hence $L = (U^T D^{-1})^{-1}$ is also LT-1, and so $I_n \rightarrow L$ by a sequence of row operations each of which adds a multiple of a row to a lower row (verify; modify columns right to left). But then $A \rightarrow LA$ by the same sequence of row operations (see the discussion preceding Theorem 2.5.1). Since $LA = [D(U^T)^{-1}][U^T U] = DU$ is upper triangular with positive entries on the diagonal, this shows that Step 1 of the algorithm is possible.

Turning to Step 2, let $A \rightarrow U_1$ as in Step 1 so that $U_1 = L_1 A$ where L_1 is LT-1. Since A is symmetric, we get

$$L_1 U_1^T = L_1 (L_1 A)^T = L_1 A^T L_1^T = L_1 A L_1^T = U_1 L_1^T \quad (8.4)$$

Let $D_1 = \text{diag}(e_1, \dots, e_n)$ denote the diagonal of U_1 . Then (8.4) gives $L_1 (U_1^T D_1^{-1}) = U_1 L_1^T D_1^{-1}$. This is both upper triangular (right side) and LT-1 (left side), and so must equal I_n . In particular, $U_1^T D_1^{-1} = L_1^{-1}$. Now let $D_2 = \text{diag}(\sqrt{e_1}, \dots, \sqrt{e_n})$, so that $D_2^2 = D_1$. If we write $U = D_2^{-1} U_1$ we have

$$U^T U = (U_1^T D_2^{-1})(D_2^{-1} U_1) = U_1^T (D_2^2)^{-1} U_1 = (U_1^T D_1^{-1}) U_1 = (L_1^{-1}) U_1 = A$$

This proves Step 2 because $U = D_2^{-1} U_1$ is formed by dividing each row of U_1 by the square root of its diagonal entry (verify). \square

Exercises for 8.3

Exercise 8.3.1 Find the Cholesky decomposition of each of the following matrices.

$$\begin{array}{ll} \text{a. } \begin{bmatrix} 4 & 3 \\ 3 & 5 \end{bmatrix} & \text{b. } \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \\ \text{c. } \begin{bmatrix} 12 & 4 & 3 \\ 4 & 2 & -1 \\ 3 & -1 & 7 \end{bmatrix} & \text{d. } \begin{bmatrix} 20 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 5 \end{bmatrix} \end{array}$$

Exercise 8.3.2

- If A is positive definite, show that A^k is positive definite for all $k \geq 1$.
- Prove the converse to (a) when k is odd.
- Find a symmetric matrix A such that A^2 is positive definite but A is not.

Exercise 8.3.3 Let $A = \begin{bmatrix} 1 & a \\ a & b \end{bmatrix}$. If $a^2 < b$, show that A is positive definite and find the Cholesky factorization.

Exercise 8.3.4 If A and B are positive definite and $r > 0$, show that $A + B$ and rA are both positive definite.

Exercise 8.3.5 If A and B are positive definite, show that $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ is positive definite.

Exercise 8.3.6 If A is an $n \times n$ positive definite matrix and U is an $n \times m$ matrix of rank m , show that $U^T A U$ is positive definite.

Exercise 8.3.7 If A is positive definite, show that each diagonal entry is positive.

Exercise 8.3.8 Let A_0 be formed from A by deleting rows 2 and 4 and deleting columns 2 and 4. If A is positive definite, show that A_0 is positive definite.

Exercise 8.3.9 If A is positive definite, show that $A = CC^T$ where C has orthogonal columns.

Exercise 8.3.10 If A is positive definite, show that $A = C^2$ where C is positive definite.

Exercise 8.3.11 Let A be a positive definite matrix. If a is a real number, show that aA is positive definite if and only if $a > 0$.

Exercise 8.3.12

- Suppose an invertible matrix A can be factored in \mathbf{M}_m as $A = LDU$ where L is lower triangular with 1s on the diagonal, U is upper triangular with 1s on the diagonal, and D is diagonal with positive diagonal entries. Show that the factorization is unique: If $A = L_1 D_1 U_1$ is another such factorization, show that $L_1 = L$, $D_1 = D$, and $U_1 = U$.
- Show that a matrix A is positive definite if and only if A is symmetric and admits a factorization $A = LDU$ as in (a).

Exercise 8.3.13 Let A be positive definite and write $d_r = \det {}^{(r)}A$ for each $r = 1, 2, \dots, n$. If U is the upper triangular matrix obtained in step 1 of the algorithm, show that the diagonal elements $u_{11}, u_{22}, \dots, u_{nn}$ of U are given by $u_{11} = d_1$, $u_{jj} = d_j/d_{j-1}$ if $j > 1$. [Hint: If $LA = U$ where L is lower triangular with 1s on the diagonal, use block multiplication to show that $\det {}^{(r)}A = \det {}^{(r)}U$ for each r .]

8.4 QR-Factorization⁷

One of the main virtues of orthogonal matrices is that they can be easily inverted—the transpose is the inverse. This fact, combined with the factorization theorem in this section, provides a useful way to simplify many matrix calculations (for example, in least squares approximation).

⁷This section is not used elsewhere in the book