Chapter 5. Matrix Transformations and Factorizations

5.9. Factorizations of Nonnegative Definite Matrices—Proofs of Theorems
Table of contents

1. Theorem 5.9.1
2. Theorem 5.9.2
3. Theorem 5.9.A
Theorem 5.9.1. Let $A$ be a symmetric nonnegative definite matrix and let $B$ be a symmetric nonnegative definite matrix such that $B^2 = A$. Then $B = VC^{1/2}V^T = VSV^T$ where $S = C^{1/2} = \text{diag}(c_1^{1/2}, c_2^{1/2}, \ldots, c_n^{1/2})$ where $c_1, c_2, \ldots, c_n$ are the eigenvalues of $A$.

Proof. By Theorem 3.8.15(2), $A = VCV^T$ where $V$ is orthogonal and $C = \text{diag}(c_1, c_2, \ldots, c_n)$. 


Theorem 5.9.1

**Theorem 5.9.1.** Let $A$ be a symmetric nonnegative definite matrix and let $B$ be a symmetric nonnegative definite matrix such that $B^2 = A$. Then $B = VC^{1/2}V^T = VSV^T$ where $S = C^{1/2} = \text{diag}(c_1^{1/2}, c_2^{1/2}, \ldots, c_n^{1/2})$ where $c_1, c_2, \ldots, c_n$ are the eigenvalues of $A$.

**Proof.** By Theorem 3.8.15(2), $A = VCV^T$ where $V$ is orthogonal and $C = \text{diag}(c_1, c_2, \ldots, c_n)$. We have

\[
(B - VC^{1/2}V^T)^2 = (B - VC^{1/2}V^T)(V - VC^{1/2}V^T)
\]

\[
= B^2 - VC^{1/2}V^TB - BVC^{1/2}V^T + (VC^{1/2}V^T)^2
\]

\[
= A - VC^{1/2}V^TB - (VC^{1/2}V^TB^T)^T + A
\]

\[
= 2A - VC^{1/2}V^TB - (VC^{1/2}V^TB) \quad (\ast)
\]

since $B$ is symmetric.
Theorem 5.9.1

Theorem 5.9.1. Let $A$ be a symmetric nonnegative definite matrix and let $B$ be a symmetric nonnegative definite matrix such that $B^2 = A$. Then $B = VC^{1/2}V^T = VSV^T$ where $S = C^{1/2} = \text{diag}(c_1^{1/2}, c_2^{1/2}, \ldots, c_n^{1/2})$ where $c_1, c_2, \ldots, c_n$ are the eigenvalues of $A$.

Proof. By Theorem 3.8.15(2), $A = VCV^T$ where $V$ is orthogonal and $C = \text{diag}(c_1, c_2, \ldots, c_n)$. We have

\[
(B - VC^{1/2}V^T)^2 = (B - VC^{1/2}V^T)(V - VC^{1/2}V^T)
\]

\[
= B^2 - VC^{1/2}V^TB - BVC^{1/2}V^T + (VC^{1/2}V^T)^2
\]

\[
= A - VC^{1/2}V^TB - (VC^{1/2}V^TB)^T + A
\]

\[
= 2A - VC^{1/2}V^TB - (VC^{1/2}V^TB) \quad (*)
\]

since $B$ is symmetric.

Since $B$ is symmetric nonnegative definite then $B = UDU^T$ for orthogonal $U$ and diagonal $D = \text{diag}(d_1, d_2, \ldots, d_n)$ where each $d_i$ is nonnegative, by Theorem 3.8.15(2).
Theorem 5.9.1

Theorem 5.9.1. Let $A$ be a symmetric nonnegative definite matrix and let $B$ be a symmetric nonnegative definite matrix such that $B^2 = A$. Then $B = V C^{1/2} V^T = V S V^T$ where $S = C^{1/2} = \text{diag}(c_1^{1/2}, c_2^{1/2}, \ldots, c_n^{1/2})$ where $c_1, c_2, \ldots, c_n$ are the eigenvalues of $A$.

Proof. By Theorem 3.8.15(2), $A = V C V^T$ where $V$ is orthogonal and $C = \text{diag}(c_1, c_2, \ldots, c_n)$. We have

$$ (B - V C^{1/2} V^T)^2 = (B - V C^{1/2} V^T)(V - V C^{1/2} V^T) $$
$$ = B^2 - V C^{1/2} V^T B - B V C^{1/2} V^T + (V C^{1/2} V^T)^2 $$
$$ = A - V C^{1/2} V^T B - (V C^{1/2} V^T B^T)^T + A $$
$$ = 2A - V C^{1/2} V^T B - (V C^{1/2} V^T B)^T $$

since $B$ is symmetric.

Since $B$ is symmetric nonnegative definite then $B = U D U^T$ for orthogonal $U$ and diagonal $D = \text{diag}(d_1, d_2, \ldots, d_n)$ where each $d_i$ is nonnegative, by Theorem 3.8.15(2).
Theorem 5.9.1 (continued 1)

Proof (continued). Now

\[ V^T UD^2 = V^T UD(U^T U)D(U^T U) \] since \( U \) is orthogonal

\[ = V^T (UDU^T)(UDU^T)U = V^T B^2 U \]

\[ = V^T AU = V^T (VC^{1/2} V^T)^2 U = V^T VC^{1/2} V^T VC^{1/2} V^T U \]

\[ = C^{1/2} C^{1/2} V^T U \] since \( V \) is orthogonal

\[ = CV^T U. \quad (**) \]

Let the \((i, j)\) entry of \( V^T U \) be \( z_{ij} \). Since \( D \) is diagonal, the \((i, j)\) entry of \( V^T UD^2 \) is \( z_{ij}d_j^2 \). Since \( C \) is diagonal, the \((i, j)\) entry of \( CV^T U \) is \( c_i a_{ij} \). Since \( V^T UD^2 = CV^T U \) by \((**)\), then

\[ z_{ij}d_j^2 = c_i z_{ij} \] or \[ d_j |z_{ij}| = c_i^{1/2} |z_{ij}| \]

or \[ d_j \text{sgn}(z_{ij}) |z_{ij}| = c_i^{1/2} \text{sgn}(z_{ij}) |z_{ij}|, \] and so \( d_j z_{ij} = c_i^{1/2} z_{ij} \).
Proof (continued). Now

\[ V^T UD^2 = V^T UD(U^T U)D(U^T U) \text{ since } U \text{ is orthogonal} \]
\[ = V^T (UDU^T)(UDU^T)U = V^T B^2 U \]
\[ = V^T AU = V^T (VC^{1/2}V^T)^2 U = V^T VC^{1/2}V^T VC^{1/2}V^T U \]
\[ = C^{1/2}C^{1/2}V^T U \text{ since } V \text{ is orthogonal} \]
\[ = CV^T U. \quad (***) \]

Let the \((i, j)\) entry of \(V^T U\) be \(z_{ij}\). Since \(D\) is diagonal, the \((i, j)\) entry of \(V^T UD^2\) is \(z_{ij}d_j^2\). Since \(C\) is diagonal, the \((i, j)\) entry of \(CV^T U\) is \(c_ia_{ij}\).

Since \(V^T UD^2 = CV^T U\) by (***) , then \(z_{ij}d_j^2 = c_iz_{ij}\) or \(d_j^2z_{ij}^2 = c_iz_{ij}^2\) or \(d_j|z_{ij}| = c_i^{1/2}|z_{ij}|\) or \(d_jsgn(z_{ij})|z_{ij}| = c_i^{1/2}sgn(z_{ij})|z_{ij}|\), and so \(d_jz_{ij} = c_i^{1/2}z_{ij}\).
Theorem 5.9.1 (continued 2)

Proof (continued). Now the \((i, j)\) entry of \(V^T UD\) is \(d_i z_{ij}\) and the \((i, j)\) entry of \(C^{1/2} V^T U\) is \(c_i^{1/2} z_{ij}\) and hence \(V^T UD = C^{1/2} V^T U\). We therefore have

\[
VC^{1/2}V^TB = VC^{1/2}V^T(UDU^T) \quad \text{since} \ B = UDU^T \\
= VC^{1/2}(V^T UD)U^T \\
= VC^{1/2}(C^{1/2} V^T U)U^T \quad \text{since} \ V^T UD = C^{1/2} V^T U \\
= VCV^TUU^T = VCV^T \quad \text{since} \ U \text{ is orthogonal} \\
= A \quad \text{since} \ A = VCV^T.
\]
Theorem 5.9.1 (continued 2)

**Proof (continued).** Now the \((i,j)\) entry of \(V^T UD\) is \(d_iz_{ij}\) and the \((i,j)\) entry of \(C^{1/2}V^TU\) is \(c_i^{1/2}z_{ij}\) and hence \(V^T UD = C^{1/2}V^TU\). We therefore have

\[
VC^{1/2}V^TB = VC^{1/2}V^T(UDU^T) \quad \text{since } B = UDU^T
\]

\[
= VC^{1/2}(V^T UD)U^T
\]

\[
= VC^{1/2}(C^{1/2}V^TU)U^T \quad \text{since } V^T UD = C^{1/2}V^TU
\]

\[
= VCV^TUU^T = VCV^T \quad \text{since } U \text{ is orthogonal}
\]

\[
= A \quad \text{since } A = VCV^T.
\]

From (*) we have

\[
(B - VC^{1/2}V^T)^2 = 2A = VC^{1/2}V^TB - (VC^{1/2}V^TB)^T
\]

\[
= 2A - A - A^T \quad \text{since } VC^{1/2}V^TB = A
\]

\[
= 2A - 2A \quad \text{since } A \text{ is symmetric}
\]

\[
= 0.
\]
Theorem 5.9.1 (continued 2)

Proof (continued). Now the \((i, j)\) entry of \(V^T UD\) is \(d_iz_{ij}\) and the \((i, j)\) entry of \(C^{1/2}V^T U\) is \(c_i^{1/2}z_{ij}\) and hence \(V^T UD = C^{1/2}V^T U\). We therefore have

\[
VC^{1/2}V^TB = VC^{1/2}V^T(UDU^T) \quad \text{since} \quad B = UDU^T
\]
\[
= VC^{1/2}(V^TU)U^T
\]
\[
= VC^{1/2}(C^{1/2}V^TU)U^T \quad \text{since} \quad V^TU = C^{1/2}V^TU
\]
\[
= VCV^TUU^T = VCV^T \quad \text{since} \quad U \text{ is orthogonal}
\]
\[
= A \quad \text{since} \quad A = VCV^T.
\]

From (*) we have

\[
(B - VC^{1/2}V^T)^2 = 2A = VC^{1/2}V^TB - (VC^{1/2}V^TB)^T
\]
\[
= 2A - A - A^T \quad \text{since} \quad VC^{1/2}V^TB = A
\]
\[
= 2A - 2A \quad \text{since} \quad A \text{ is symmetric}
\]
\[
= 0.
\]
Theorem 5.9.1. Let $A$ be a symmetric nonnegative definite matrix and let $B$ be a symmetric nonnegative definite matrix such that $B^2 = A$. Then $B = VC^{1/2}V^T = VSV^T$ where $S = C^{1/2} = \text{diag}(c_1^{1/2}, c_2^{1/2}, \ldots, c_n^{1/2})$ where $c_1, c_2, \ldots, c_n$ are the eigenvalues of $A$.

Proof (continued). Now $B$ and $VC^{1/2}V^T$ are both symmetric, so $B - VC^{1/2}V^T$ is symmetric. In a symmetric matrix $S$, $S = SS^T$ and the the $(i,j)$ entries of $S^2$ are the inner product of the $i$th row of $S$ with the $i$th column of $S^T$; that is, the $(i,j)$ entry of $S^2$ is $\|S_i\|_F^2$ (the Frobenius norm or Euclidean matrix norm) where $s_i$ is the $i$th column of $S$. So the only way $S^2 = 0$ for a symmetric matrix is when $S = 0$. Therefore we have $B = VC^{1/2}V^T$ and this is the unique square root of $A$. \hfill \blacksquare
Theorem 5.9.2. If $A$ is a symmetric positive definite matrix, then $A$ has a Cholesky factorization.

Proof. We give an inductive proof. If $A$ is $1 \times 1$, say $A = [a_{11}]$, then $a_{11} > 0$ since $A$ is positive definite and so we take $T = [\sqrt{a_{11}}]$. Then $T^T T = [\sqrt{a_{11}}][\sqrt{a_{11}}] = [a_{11}] = A$, and so $A$ has a Cholesky factorization.

Now suppose all $n \times n$ symmetric positive definite matrices have Cholesky decompositions. Consider $(n + 1) \times (n + 1)$ matrix $A$. 
Theorem 5.9.2. If $A$ is a symmetric positive definite matrix, then $A$ has a Cholesky factorization.

Proof. We give an inductive proof. If $A$ is $1 \times 1$, say $A = [a_{11}]$, then $a_{11} > 0$ since $A$ is positive definite and so we take $T = [\sqrt{a_{11}}]$. Then $T^T T = [\sqrt{a_{11}}][\sqrt{a_{11}}] = [a_{11}] = A$, and so $A$ has a Cholesky factorization.

Now suppose all $n \times n$ symmetric positive definite matrices have Cholesky decompositions. Consider $(n + 1) \times (n + 1)$ matrix $A$. Partition $A$ as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{11} = [a_{11}]$. Consider the Schur complement of $A_{11}$ in $A$, $Z = A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$. By Exercise 5.9.A, $n \times n$ matrix $Z$ is symmetric and positive definite.
Theorem 5.9.2. If $A$ is a symmetric positive definite matrix, then $A$ has a Cholesky factorization.

Proof. We give an inductive proof. If $A$ is $1 \times 1$, say $A = [a_{11}]$, then $a_{11} > 0$ since $A$ is positive definite and so we take $T = [\sqrt{a_{11}}]$. Then $T^T T = [\sqrt{a_{11}}][\sqrt{a_{11}}] = [a_{11}] = A$, and so $A$ has a Cholesky factorization.

Now suppose all $n \times n$ symmetric positive definite matrices have Cholesky decompositions. Consider $(n + 1) \times (n + 1)$ matrix $A$. Partition $A$ as

\[
A = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix} = \begin{bmatrix}
a_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\]

where $A_{11} = [a_{11}]$. Consider the Schur complement of $A_{11}$ in $A$, $Z = A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$. By Exercise 5.9.A, $n \times n$ matrix $Z$ is symmetric and positive definite. So by the induction hypothesis, $Z$ has a Cholesky factorization, say $Z = T_Z^T T_Z$ where $T_Z$ is an $n \times n$ upper triangular matrix with positive diagonal entries.
Theorem 5.9.2. If $A$ is a symmetric positive definite matrix, then $A$ has a Cholesky factorization.

Proof. We give an inductive proof. If $A$ is $1 \times 1$, say $A = [a_{11}]$, then $a_{11} > 0$ since $A$ is positive definite and so we take $T = [\sqrt{a_{11}}]$. Then $T^T T = [\sqrt{a_{11}}][\sqrt{a_{11}}] = [a_{11}] = A$, and so $A$ has a Cholesky factorization.

Now suppose all $n \times n$ symmetric positive definite matrices have Cholesky decompositions. Consider $(n+1) \times (n+1)$ matrix $A$. Partition $A$ as

$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$

where $A_{11} = [a_{11}]$. Consider the Schur complement of $A_{11}$ in $A$, $Z = A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$. By Exercise 5.9.A, $n \times n$ matrix $Z$ is symmetric and positive definite. So by the induction hypothesis, $Z$ has a Cholesky factorization, say $Z = T_Z^T T_Z$ where $T_Z$ is an $n \times n$ upper triangular matrix with positive diagonal entries.
Theorem 5.9.2 (continued)

Proof (continued). Define $T$ as $T = \begin{bmatrix} \sqrt{a_{11}} & \frac{1}{\sqrt{a_{11}}} A_{12} \\ 0 & T_Z \end{bmatrix}$. Since $T_Z$ is upper triangular with positive diagonal entries, then $T$ also has these two properties. Finally,

$$T^T T = \begin{bmatrix} \sqrt{a_{11}} & 0 \\ \frac{1}{\sqrt{a_{11}}} A_{12}^T & T_Z^T \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{a_{11}} & \frac{1}{\sqrt{a_{11}}} A_{12} \\ 0 & T_Z \end{bmatrix}$$

$$= \begin{bmatrix} 1_{11} & \frac{1}{a_{11}} A_{12}^T A_{12} + T_Z^T T_Z \\ A_{12}^T & \frac{1}{a_{11}} A_{21} A_{12} + A_{21} A_{12} + Z \end{bmatrix} = \begin{bmatrix} a_{11} & A_{12} \\ A_{21}^T & A_{21} A_{12} + Z \end{bmatrix}$$

since $A_{21}^T = A_{21}$ because $A$ is symmetric

$$= \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A.$$

So $(n + 1) \times (n + 1)$ matrix $A$ has a Cholesky factorization and so the claim holds by induction.
Theorem 5.9.2

Proof (continued). Define $T$ as $T = \begin{bmatrix} \sqrt{a_{11}} & \frac{1}{\sqrt{a_{11}}} A_{12} \\ 0 & T_Z \end{bmatrix}$. Since $T_Z$ is upper triangular with positive diagonal entries, then $T$ also has these two properties. Finally,

\[
TT = \begin{bmatrix} \sqrt{a_{11}} & 0 \\ \frac{1}{\sqrt{a_{11}}} A_{12}^T & T_Z \end{bmatrix} \begin{bmatrix} \sqrt{a_{11}} & \frac{1}{\sqrt{a_{11}}} A_{12} \\ 0 & T_Z \end{bmatrix}
\]

\[
= \begin{bmatrix} 1_{11} & A_{12} \\ A_{12}^T & \frac{1}{a_{11}} A_{12}^T A_{12} + T_Z^T T_Z \end{bmatrix} = \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & \frac{1}{a_{11}} A_{21} A_{12} + Z \end{bmatrix}
\]

since $A_{21}^T = A_{21}$ because $A$ is symmetric

\[
= \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A.
\]

So $(n+1) \times (n+1)$ matrix $A$ has a Cholesky factorization and so the claim holds by induction.
Theorem 5.9.A

**Theorem 5.9.A.** An invertible matrix $A$ has a Cholesky factorization if and only if $A$ is symmetric and positive definite.

**Proof.** If $A$ is symmetric and positive definite, then it has a Cholesky factorization by Theorem 5.9.2 (whether $A$ is invertible or not).
Theorem 5.9.A

Theorem 5.9.A. An invertible matrix $A$ has a Cholesky factorization if and only if $A$ is symmetric and positive definite.

Proof. If $A$ is symmetric and positive definite, then it has a Cholesky factorization by Theorem 5.9.2 (whether $A$ is invertible or not).

If $A$ is invertible and has a Cholesky factorization, then $A = T^T T$ where $T$ is an upper triangular matrix with positive diagonal entries. Then $A^T = (T^T T)^T = T^T (T^T) T = T^T T = A$ and so $A$ is symmetric. Let $x$ be a nonzero in $\mathbb{R}^n$. 

**Theorem 5.9.A.** An invertible matrix $A$ has a Cholesky factorization if and only if $A$ is symmetric and positive definite.

**Proof.** If $A$ is symmetric and positive definite, then it has a Cholesky factorization by Theorem 5.9.2 (whether $A$ is invertible or not).

If $A$ is invertible and has a Cholesky factorization, then $A = T^T T$ where $T$ is an upper triangular matrix with positive diagonal entries. Then $A^T = (T^T T)^T = T^T (T^T)^T = T^T T = A$ and so $A$ is symmetric. Let $x$ be a nonzero in $\mathbb{R}^n$. Then

$$x^T A x = x^T T^T T x = (x^T)^T T x = \langle T x, T x \rangle = \| T x \|_F \text{ (the Fobenius norm or Euclidean matrix norm of } T x).$$
Theorem 5.9.A. An invertible matrix $A$ has a Cholesky factorization if and only if $A$ is symmetric and positive definite.

Proof. If $A$ is symmetric and positive definite, then it has a Cholesky factorization by Theorem 5.9.2 (whether $A$ is invertible or not).

If $A$ is invertible and has a Cholesky factorization, then $A = T^T T$ where $T$ is an upper triangular matrix with positive diagonal entries. Then $A^T = (T^T T)^T = T^T (T^T)^T = T^T T = A$ and so $A$ is symmetric. Let $x$ be a nonzero in $\mathbb{R}^n$. Then

$$x^T A x = x^T T^T T x = (x^T T) T x = \langle T x, T x \rangle$$

$$= \|T x\|_F \text{ (the Fobenius norm or Euclidean matrix norm of } T x).$$
Theorem 5.9.A (continued)

**Theorem 5.9.A.** An invertible matrix $A$ has a Cholesky factorization if and only if $A$ is symmetric and positive definite.

**Proof (continued).** Since $A$ is hypothesized to be invertible, then $\det(A) \neq 0$ by Theorem 3.3.16 and

\[
\begin{align*}
\det(A) &= \det(T^T T) \\
&= \det(T^T) \det(T) \text{ by Theorem 3.2.4} \\
&= \det(T) \det(T) \text{ by Theorem 3.1.A} \\
&= \det(T)^2
\end{align*}
\]

and so $\det(T) \neq 0$; that is, $T$ is invertible. So for $x \neq 0$ we have $Tx \neq 0$ (since $T$ is invertible implies a unique solution to $Tx = 0$ and, of course, $0$ is that unique solution, see Note 3.5.A). Therefore $\|Tx\|_F \neq 0$ (since $\| \cdot \|_F$ is a norm) and so by $(\ast)$, $x^T Ax > 0$ and $A$ is positive definite. \qed
Theorem 5.9.A (continued)

**Theorem 5.9.A.** An invertible matrix $A$ has a Cholesky factorization if and only if $A$ is symmetric and positive definite.

**Proof (continued).** Since $A$ is hypothesized to be invertible, then $\det(A) \neq 0$ by Theorem 3.3.16 and

$$\det(A) = \det(T^T T)$$

$$= \det(T^T) \det(T) \text{ by Theorem 3.2.4}$$

$$= \det(T) \det(T) \text{ by Theorem 3.1.A}$$

$$= \det(T)^2$$

and so $\det(T) \neq 0$; that is, $T$ is invertible. So for $x \neq 0$ we have $Tx \neq 0$ (since $T$ is invertible implies a unique solution to $Tx = 0$ and, of course, $0$ is that unique solution, see Note 3.5.A). Therefore $\|Tx\|_F \neq 0$ (since $\| \cdot \|_F$ is a norm) and so by $(\ast)$, $x^T Ax > 0$ and $A$ is positive definite. \qed