Schur Complements and the Cholesky Factorization

Last modified: 2026-11-24 09:19:45 (PST)

Gram matrices, Schur complements and Cholesky

> **Source**
>
> This section is adapted from Zhou [1], used under the MIT License (see the license text in Section in Subspaces and Rank). The source's eigenvalue and quadratic-form characterizations of definiteness are covered in these notes already (Theorem 2 in Definite Matrices). These parts of the source are left out:
>
> - its covariance test, which needs probability
> - its material on minimization
> - its positive semidefinite version of the Schur complement test
> - its pivot test
> - its Hadamard-product "Schur lemma"
> - its notes on the LU factorization and computational cost
> - its statement that a nonnegative combination of positive semidefinite matrices is positive semidefinite (Theorem 0.2 states the positive definite version)
> - its section on the ellipse $ax^2+2bxy+cy^2=1$, with its worked example of a positive definite and an indefinite matrix
> - its numerical example of the Cholesky factorization
>
> The source leaves the Schur complement test unproved; it is proved here from a block expansion of the quadratic form, and the Cholesky factorization is derived from it.

Theorem 0.1 (Definite matrices are Gram matrices)**.** *Let* $\mathbf{A}$ *be a* $p \times p$ *matrix.*

1. $\mathbf{A}$ *is positive semidefinite (Definition 1 in Definite Matrices) exactly when* $\mathbf{A} =$

$\mathbf{B}^\top\mathbf{B}$ *for some matrix* $\mathbf{B}$ *with* p *columns.*

2. $\mathbf{A}$ *is positive definite (Definition 2 in Definite Matrices) exactly when* $\mathbf{A} = \mathbf{B}^\top\mathbf{B}$ *for some matrix* $\mathbf{B}$ *with* p *linearly independent columns.*

Proof

Proof. **If $\mathbf{A} = \mathbf{B}^\top\mathbf{B}$.** $\mathbf{A}$ is symmetric, since

$$\begin{aligned}(\mathbf{B}^\top\mathbf{B})^\top &= \mathbf{B}^\top(\mathbf{B}^\top)^\top \\ &= \mathbf{B}^\top\mathbf{B}\end{aligned}$$

(Theorem 6 in Matrix Multiplication, Definition 6 in Matrices). For any $\underset{\sim}{x}$,

$$\begin{aligned}\underset{\sim}{x}^\top\mathbf{A}\underset{\sim}{x} &= \underset{\sim}{x}^\top\mathbf{B}^\top\mathbf{B}\underset{\sim}{x} && \text{(substitute } \mathbf{A}\text{)} \\ &= (\mathbf{B}\underset{\sim}{x})^\top(\mathbf{B}\underset{\sim}{x}) && \text{(Theorem 6 in Matrix Multiplication, Theorem 1 in Matrix Multiplication)} \\ &= \|\mathbf{B}\underset{\sim}{x}\|^2 && \text{(Equation 5 in Vectors, squared)} \\ &\geq 0. && \text{(Theorem 1 in Inner Products and the Cauchy-Schwarz Inequality)}\end{aligned}$$

If the columns of $\mathbf{B}$ are linearly independent, then $\mathcal{N}(\mathbf{B}) = \{\underset{\sim}{0}\}$ ($\mathbf{B}\underset{\sim}{x}$ is the combination of the columns with coefficients x_j, Theorem 3 in Subspaces and Rank, and only the zero combination is $\underset{\sim}{0}$, Definition 1 in Subspaces and Rank), so for $\underset{\sim}{x} \neq \underset{\sim}{0}$, $\mathbf{B}\underset{\sim}{x} \neq \underset{\sim}{0}$ and $\|\mathbf{B}\underset{\sim}{x}\|^2 > 0$ (Theorem 1 in Inner Products and the Cauchy-Schwarz Inequality, part 1).
Only if. Let $\mathbf{A}$ be positive semidefinite. It is symmetric, so $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$ (Theorem 1 in Eigenvalues and Matrix Decompositions), and every $\lambda_i \geq 0$ (Theorem 2 in Definite Matrices). Let $\mathbf{\Lambda}^{1/2} \stackrel{\text{def}}{=} \text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_p})$, which is symmetric, with $\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2} = \mathbf{\Lambda}$ (the product of diagonal matrices multiplies their diagonals, as in the proof of Theorem 9 in Similarity and Diagonalization), and let $\mathbf{B} \stackrel{\text{def}}{=} \mathbf{\Lambda}^{1/2}\mathbf{Q}^\top$. Then

$$\begin{aligned}\mathbf{B}^\top\mathbf{B} &= (\mathbf{Q}^\top)^\top(\mathbf{\Lambda}^{1/2})^\top\mathbf{\Lambda}^{1/2}\mathbf{Q}^\top && \text{(Theorem 6 in Matrix Multiplication)} \\ &= \mathbf{Q}\,\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\,\mathbf{Q}^\top && \text{(Definition 6 in Matrices; } \mathbf{\Lambda}^{1/2} \text{ is symmetric)} \\ &= \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top && (\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2} = \mathbf{\Lambda}) \\ &= \mathbf{A}. && \text{(Theorem 1 in Eigenvalues and Matrix Decompositions)}\end{aligned}$$

If $\mathbf{A}$ is positive definite, it is positive semidefinite (Remark 1 in Definite Matrices), so the same $\mathbf{B}$ gives $\mathbf{A} = \mathbf{B}^\top\mathbf{B}$; and every $\lambda_i > 0$ (Theorem 2 in Definite Matrices), so $\mathbf{\Lambda}^{1/2}$ is diagonal with nonzero diagonal, and invertible (its inverse is $\text{diag}(1/\sqrt{\lambda_1}, \ldots, 1/\sqrt{\lambda_p})$); $\mathbf{Q}^\top$ is invertible with inverse $\mathbf{Q}$ (Remark 6 in Special Matrices, Quadratic Forms and Trace). So $\mathbf{B}$ is invertible (Theorem 1 in Special Matrices, Quadratic Forms and Trace), and its p columns are linearly independent (Theorem 1 in Similarity and Diagonalization, Definition 2 in Subspaces and Rank). □

Example 0.1 (Gram factorizations)**.**

- With $\mathbf{R} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ (1×2), $\mathbf{R}^\top \mathbf{R} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, the positive semidefinite matrix of Example 1 in Definite Matrices; the two columns of $\mathbf{R}$, 1 and 1, are dependent, so part 2 does not apply to $\mathbf{R}$; $\mathbf{R}^\top \mathbf{R}$ is in fact not positive definite (Example 4 in Definite Matrices).
- With $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, whose columns are independent (Example 2 in Subspaces and Rank), $\mathbf{X}^\top \mathbf{X} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$ (Example 1 in Design Matrices and Projections) is positive definite.
- $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ of Example 4 in Definite Matrices is symmetric but not positive semidefinite, so by part 1 it is not $\mathbf{B}^\top \mathbf{B}$ for any $\mathbf{B}$.

Theorem 0.2 (Operations that preserve definiteness)**.**

1. *If $\mathbf{C}$ is a $p \times p$ positive definite matrix (Definition 2 in Definite Matrices) and $\mathbf{A}$ is a $p \times k$ matrix with k linearly independent columns, then the $k \times k$ matrix $\mathbf{A}^\top \mathbf{C} \mathbf{A}$ is positive definite.*
2. *If $\mathbf{A}_1$ and $\mathbf{A}_2$ are positive definite $p \times p$ matrices and $\alpha_1, \alpha_2 > 0$, then $\alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2$ is positive definite.*

Proof

Proof. **Part 1.** $\mathbf{A}^\top \mathbf{C} \mathbf{A}$ is symmetric:

$$\begin{aligned} \left(\mathbf{A}^\top \mathbf{C} \mathbf{A}\right)^\top &= \mathbf{A}^\top \mathbf{C}^\top \left(\mathbf{A}^\top\right)^\top \\ &= \mathbf{A}^\top \mathbf{C} \mathbf{A} \end{aligned}$$

(Theorem 6 in Matrix Multiplication, twice; $\mathbf{C}$ symmetric; Definition 6 in Matrices). For $\underset{\sim}{x} \neq \underset{\sim}{0}_k$, $\mathbf{A}\underset{\sim}{x} \neq \underset{\sim}{0}_p$, since the columns of $\mathbf{A}$ are independent (Theorem 3 in Subspaces and Rank, Definition 1 in Subspaces and Rank), and

$$\begin{aligned} \underset{\sim}{x}^\top \left(\mathbf{A}^\top \mathbf{C} \mathbf{A}\right) \underset{\sim}{x} &= \left(\mathbf{A}\underset{\sim}{x}\right)^\top \mathbf{C} \left(\mathbf{A}\underset{\sim}{x}\right) && \text{(Theorem 6 in Matrix Multiplication, Theorem 1 in Matrix Multiplication)} \\ &> 0. && \text{($\mathbf{C}$ positive definite, $\mathbf{A}\underset{\sim}{x} \neq \underset{\sim}{0}$)} \end{aligned}$$

Part 2. $\alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2$ is symmetric: its transpose is $\alpha_1 \mathbf{A}_1{}^\top + \alpha_2 \mathbf{A}_2{}^\top = \alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2$ (Theorem 5 in Matrix Multiplication, Definition 9 in Matrices, Definition 6 in Matrices). For $\underset{\sim}{x} \neq \underset{\sim}{0}$,

$$
\begin{aligned}
\underset{\sim}{x}^\top (\alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2)\, \underset{\sim}{x} &= \underset{\sim}{x}^\top (\alpha_1 \mathbf{A}_1)\, \underset{\sim}{x} + \underset{\sim}{x}^\top (\alpha_2 \mathbf{A}_2)\, \underset{\sim}{x} && \text{(Theorem 2 in Matrix Multiplication, twice)} \\
&= \alpha_1\, \underset{\sim}{x}^\top \mathbf{A}_1 \underset{\sim}{x} + \alpha_2\, \underset{\sim}{x}^\top \mathbf{A}_2 \underset{\sim}{x} && \text{(Theorem 1 in Generalized Inverses and Linear Systems} \\
&> 0. && \text{(each term is positive)}
\end{aligned}
$$

□

Example 0.2 (Building positive definite matrices)**.**

- With $\mathbf{C} = \mathbf{I}_3$ and $\mathbf{X}$ of Example 0.1, part 1 gives again that $\mathbf{X}^\top\mathbf{X}$ is positive definite.
- With $\alpha_1 = 2$, $\mathbf{A}_1 = \mathbf{I}_2$ (Example 4 in Definite Matrices), $\alpha_2 = 1$ and

$$
\begin{aligned}
\mathbf{A}_2 &= \mathbf{X}^\top\mathbf{X} \\
&= \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}
\end{aligned}
$$

  (Example 0.1), part 2 gives that $2\mathbf{I}_2 + \mathbf{X}^\top\mathbf{X} = \begin{bmatrix} 5 & 6 \\ 6 & 16 \end{bmatrix}$ is positive definite.
- An aside beyond part 2: $\mathbf{I}_2 + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite plus positive semidefinite; its quadratic form is $x_1^2 + x_2^2 + (x_1 + x_2)^2 > 0$ for $\underset{\sim}{x} \neq \underset{\sim}{0}$, so it is positive definite, in line with its eigenvalues 3 and 1 (Example 1 in Eigenvalues and Matrix Decompositions).
- Independence is needed in part 1: with $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, whose columns are equal, $\mathbf{A}^\top \mathbf{I}_2\, \mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$, and $(1, -1)$ gives quadratic form 0.

Theorem 0.3 (A positive definite matrix is invertible)**.** *Every positive definite matrix (Definition 2 in Definite Matrices) is invertible (Definition 6 in Special Matrices, Quadratic Forms and Trace).*

Proof

Proof. Let $\mathbf{A}$ be a $p \times p$ positive definite matrix, and let $\underset{\sim}{x}$ be a vector of length p with $\mathbf{A}\underset{\sim}{x} = \underset{\sim}{0}_p$. Then

$$
\begin{aligned}
\underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} &= \underset{\sim}{x}^\top (\mathbf{A}\underset{\sim}{x}) && \text{(Definition 13 in Special Matrices, Quadratic Forms and Trace)} \\
&= \underset{\sim}{x}^\top \underset{\sim}{0}_p && (\mathbf{A}\underset{\sim}{x} = \underset{\sim}{0}_p) \\
&= 0. && \text{(every term of the inner product is 0)}
\end{aligned}
$$

If $\underset{\sim}{x} \neq \underset{\sim}{0}_p$, then $\underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} > 0$ because $\mathbf{A}$ is positive definite, which contradicts $\underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} = 0$. So $\underset{\sim}{x} = \underset{\sim}{0}_p$, the null space $\mathcal{N}(\mathbf{A})$ is $\{\underset{\sim}{0}_p\}$, and $\mathbf{A}$ is invertible by Theorem 1 in Similarity and Diagonalization. □

Example 0.3 (Positive definite matrices are invertible; the converse fails)**.** $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite, since

$$
\begin{aligned}
\underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} &= 2x_1^2 + 2x_1 x_2 + 2x_2^2 && \text{(multiply out)} \\
&= x_1^2 + x_2^2 + (x_1 + x_2)^2, && \text{(expand the square)}
\end{aligned}
$$

which is 0 only when $x_1 = x_2 = 0$. It is invertible, with $(\mathbf{A})^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

The converse fails. $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ of Example 4 in Definite Matrices is not positive definite, yet it is invertible: $(\mathbf{D})^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$.

Theorem 0.4 (Adding a positive multiple of the identity makes a matrix positive definite)**.** *Let $\mathbf{A}$ be a $p \times p$ positive semidefinite matrix (Definition 1 in Definite Matrices) and let $\alpha > 0$. Then $\mathbf{A} + \alpha \mathbf{I}_p$ is positive definite (Definition 2 in Definite Matrices), so it is invertible (Theorem 0.3).*

Proof

Proof. $\mathbf{A} + \alpha \mathbf{I}_p$ is symmetric: its transpose is $\mathbf{A}^\top + \alpha \mathbf{I}_p = \mathbf{A} + \alpha \mathbf{I}_p$ (Theorem 5 in Matrix Multiplication, Definition 9 in Matrices, Definition 6 in Matrices). For $\underset{\sim}{x} \neq \underset{\sim}{0}_p$,

$$
\begin{aligned}
\underset{\sim}{x}^\top (\mathbf{A} + \alpha \mathbf{I}_p) \underset{\sim}{x} &= \underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} + \alpha \, \underset{\sim}{x}^\top \mathbf{I}_p \, \underset{\sim}{x} && \text{(Theorem 2 in Matrix Multiplication, Theorem 1 in Generalized Invers} \\
&= \underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} + \alpha \, \|\underset{\sim}{x}\|^2 && \text{(Theorem 4 in Matrix Multiplication, Equation 5 in Vectors, squared)} \\
&> 0. && (\underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x} \geq 0, \ \alpha > 0, \ \underset{\sim}{x} \neq \underset{\sim}{0}_p)
\end{aligned}
$$

□

Example 0.4 (Repairing a singular matrix)**.** $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ of Example 1 in Definite Matrices is positive semidefinite and singular: $\mathbf{B}(1, -1) = \underset{\sim}{0}_2$. With $\alpha = 1$,

$$\mathbf{B} + \mathbf{I}_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},$$

which is positive definite and invertible by Example 0.3.

Corollary 0.1 ($\mathbf{X}^\top\mathbf{X}$ is invertible exactly when $\mathbf{X}$ has full column rank)**.** *For an $n \times p$ matrix $\mathbf{X}$, the $p \times p$ matrix $\mathbf{X}^\top\mathbf{X}$ is invertible (Definition 6 in Special Matrices, Quadratic Forms and Trace) exactly when* $\text{rank}(\mathbf{X}) = p$ *(Definition 3 in Subspaces and Rank).*

Proof

Proof. The following statements are equivalent:

$$\begin{aligned}
\mathbf{X}^\top\mathbf{X} \text{ is invertible} &\iff \mathcal{N}(\mathbf{X}^\top\mathbf{X}) = \{\underset{\sim}{0}_p\} && \text{(Theorem 1 in Similarity and Diagonalization)} \\
&\iff \mathcal{N}(\mathbf{X}) = \{\underset{\sim}{0}_p\} && \text{(Theorem 6 in Column Space and Null Space)} \\
&\iff \mathbf{X} \text{ has full column rank.} && \text{(Theorem 3 in Column Space and Null Space)}
\end{aligned}$$

□

Example 0.5 (A design matrix with a repeated column)**.** $\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}$ of Example 2 in Subspaces and Rank has rank $1 < 2$. Its cross-product matrix is

$$\mathbf{X}^\top\mathbf{X} = \begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix},$$

which is singular: $(2, -1)$ is in its null space, because $\mathbf{X}(2, -1) = \underset{\sim}{0}_3$ (Example 5 in Column Space and Null Space). For the full-column-rank matrix of Example 1 in Design Matrices and Projections, $\mathbf{X}^\top\mathbf{X}$ is invertible.

Definition 0.1 (Block matrix)**.** Split the rows of an $m \times n$ matrix $\mathbf{X}$ into the first k and the last $m - k$, and its columns into the first l and the last $n - l$, with $1 \le k < m$ and $1 \le l < n$. This cuts $\mathbf{X}$ into four smaller matrices, its **blocks**:

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

where $\mathbf{A}$ $(k \times l)$ holds the entries x_{ij} with $i \le k$ and $j \le l$, $\mathbf{B}$ $(k \times (n-l))$ those with $i \le k$ and $j > l$, $\mathbf{C}$ $((m-k) \times l)$ those with $i > k$ and $j \le l$, and $\mathbf{D}$ $((m-k) \times (n-l))$ those with $i > k$ and $j > l$, each in its original order. A matrix written this way is a **block matrix** [2].

Example 0.6 (Cutting a 3×3 matrix into blocks)**.** With

$$\begin{aligned} k &= l \\ &= 1, \end{aligned}$$

$$\mathbf{X} = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 6 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 5 & 3 \\ 3 & 6 \end{bmatrix}.$$

For instance, entry $(1, 2)$ of $\mathbf{D}$ is

$$\begin{aligned} x_{1+1,2+1} &= x_{23} \\ &= 3. \end{aligned}$$

With

$$\begin{aligned} k &= l \\ &= 2 \end{aligned}$$

instead, the leading 2×2 block is $\begin{bmatrix} 4 & 2 \\ 2 & 5 \end{bmatrix}$.

Definition 0.2 (Leading block of a square matrix, top-left block)**.** When $\mathbf{X}$ is $n \times n$ and $1 \le k \le n$, the $k \times k$ matrix of the entries x_{ij} with $i, j \le k$ is the **leading** (or **top-left**) $k \times k$ **block** of $\mathbf{X}$; for $k < n$ it is the block $\mathbf{A}$ of Definition 0.1 with $l = k$, and for $k = n$ it is $\mathbf{X}$ itself.

Theorem 0.5 (Leading blocks of a positive definite matrix are positive definite)**.** *Let* $\mathbf{X}$ *be a* $p \times p$ *positive definite matrix (Definition 2 in Definite Matrices), and for* $1 \le k \le p$ *let* $\mathbf{X}_k$ *be its leading* $k \times k$ *block (Definition 0.1), with entries* x_{ij} *for* $i, j \le k$*. Then* $\mathbf{X}_k$ *is positive definite. Also, every diagonal entry of a positive definite matrix is positive.*

Proof

Proof. $\mathbf{X}_k$ is symmetric, because $\mathbf{X}$ is. For $\underset{\sim}{u} \in \mathbb{R}^k$ with $\underset{\sim}{u} \neq \underset{\sim}{0}$, let $\underset{\sim}{y} = (u_1, \dots, u_k, 0, \dots, 0) \in \mathbb{R}^p$, which is nonzero, with entries $y_i = u_i$ for $i \leq k$ and $y_i = 0$ for $i > k$. Then

$$\begin{aligned}
\underset{\sim}{u}^\top \mathbf{X}_k \underset{\sim}{u} &= \sum_{i=1}^{k}\sum_{j=1}^{k} u_i\, x_{ij}\, u_j && \text{(Definition 8 in Matrix Multiplication, Definition 10 in Vectors)} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k} y_i\, x_{ij}\, y_j && (y_i = u_i \text{ for } i \leq k) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{p} y_i\, x_{ij}\, y_j && \text{(each added term has a factor } y_i = 0 \text{ or } y_j = 0) \\
&= \underset{\sim}{y}^\top \mathbf{X} \underset{\sim}{y} && \text{(Definition 8 in Matrix Multiplication, Definition 10 in Vectors)} \\
&> 0. && (\mathbf{X} \text{ positive definite}, \underset{\sim}{y} \neq \underset{\sim}{0})
\end{aligned}$$

For the diagonal entry x_{ii}, take $\underset{\sim}{e}_i$ (Definition 21 in Vectors), which is nonzero:

$$\begin{aligned}
\underset{\sim}{e}_i^\top \mathbf{X} \underset{\sim}{e}_i &= \sum_{r=1}^{p}\sum_{s=1}^{p} (\underset{\sim}{e}_i)_r\, x_{rs}\, (\underset{\sim}{e}_i)_s && \text{(Definition 8 in Matrix Multiplication, Definition 10 in Vectors)} \\
&= x_{ii} && \text{(only the term with } r \text{ and } s \text{ both equal to } i \text{ is nonzero)} \\
&> 0. && (\mathbf{X} \text{ positive definite})
\end{aligned}$$

□

Example 0.7 (A quick necessary test)**.** $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite (Example 0.2), and its diagonal entries $2, 2$ and its 1×1 leading block $[2]$ are positive. The test can only rule matrices out: $-\mathbf{I}_2$ has negative diagonal entries, so it is not positive definite, even though

$$\begin{aligned}
\det(-\mathbf{I}_2) &= (-1)(-1) - 0 \\
&= 1 \\
&> 0
\end{aligned}$$

(Definition 1 in Determinants); a positive determinant alone does not make a matrix positive definite. And positive diagonal entries are not enough: $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ of Example 4 in Definite Matrices has them but is not positive definite.

Theorem 0.6 (The quadratic form of a block matrix)**.** *Let $k, m \geq 1$, let $\mathbf{A}$ be a symmetric $k \times k$ matrix, $\mathbf{B}$ a $k \times m$ matrix and $\mathbf{C}$ a symmetric $m \times m$ matrix, and let $\mathbf{X}$ be the $(k+m) \times (k+m)$ block matrix (Definition 0.1)*

$$\mathbf{X} = \begin{bmatrix} \underset{\sim}{\mathbf{A}} & \underset{\sim}{\mathbf{B}} \\ {\scriptstyle k \times k} & {\scriptstyle k \times m} \\ \underset{\sim}{\mathbf{B}}^{\top} & \underset{\sim}{\mathbf{C}} \\ {\scriptstyle m \times k} & {\scriptstyle m \times m} \end{bmatrix},$$

that is, $x_{ij} = a_{ij}$, $x_{i,k+j} = b_{ij}$, $x_{k+i,j} = b_{ji}$ and $x_{k+i,k+j} = c_{ij}$ for indices in range. $\mathbf{X}$ is symmetric (Definition 3 in Special Matrices, Quadratic Forms and Trace):

$$\begin{aligned} x_{k+i,j} &= b_{ji} \\ &= x_{j,k+i}, \end{aligned}$$

and $\mathbf{A}$ and $\mathbf{C}$ are symmetric. Split $\underset{\sim}{v} \in \mathbb{R}^{k+m}$ as $\underset{\sim}{v} = (\underset{\sim}{u}, \underset{\sim}{w})$, with $\underset{\sim}{u} \in \mathbb{R}^k$ its first k entries and $\underset{\sim}{w} \in \mathbb{R}^m$ the rest. Then

$$\underset{\sim}{v}^{\top} \mathbf{X} \underset{\sim}{v} = \underset{\sim}{u}^{\top} \mathbf{A} \underset{\sim}{u} + 2\, \underset{\sim}{u}^{\top} \mathbf{B} \underset{\sim}{w} + \underset{\sim}{w}^{\top} \mathbf{C} \underset{\sim}{w}, \tag{1}$$

and if $\mathbf{A}$ is invertible, with $\mathbf{S} \overset{def}{=} \mathbf{C} - \mathbf{B}^{\top}\mathbf{A}^{-1}\mathbf{B}$ and $\underset{\sim}{z} \overset{def}{=} \underset{\sim}{u} + \mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}$,

$$\underset{\sim}{v}^{\top} \mathbf{X} \underset{\sim}{v} = \underset{\sim}{z}^{\top} \mathbf{A} \underset{\sim}{z} + \underset{\sim}{w}^{\top} \mathbf{S} \underset{\sim}{w}.$$

Proof

Proof. **$\mathbf{X}\underset{\sim}{v}$ in blocks.** For $i \leq k$,

$$\begin{aligned} (\mathbf{X}\underset{\sim}{v})_i &= \sum_{j=1}^{k+m} x_{ij}\, v_j && \text{(Definition 8 in Matrix Multiplication)} \\ &= \sum_{j=1}^{k} x_{ij}\, v_j + \sum_{j=1}^{m} x_{i,k+j}\, v_{k+j} && \text{(split the sum at } j = k\text{)} \\ &= \sum_{j=1}^{k} a_{ij}\, u_j + \sum_{j=1}^{m} b_{ij}\, w_j && \text{(the blocks of } \mathbf{X} \text{ and of } \underset{\sim}{v}\text{)} \\ &= (\mathbf{A}\underset{\sim}{u})_i + (\mathbf{B}\underset{\sim}{w})_i, && \text{(Definition 8 in Matrix Multiplication)} \end{aligned}$$

and in the same way $(\mathbf{X}\underset{\sim}{v})_{k+i} = (\mathbf{B}^{\top}\underset{\sim}{u})_i + (\mathbf{C}\underset{\sim}{w})_i$ for $i \leq m$.
The expansion.

$$
\begin{aligned}
\underset{\sim}{v}^{\top}\mathbf{X}\underset{\sim}{v} &= \sum_{i=1}^{k} u_i\,(\mathbf{X}\underset{\sim}{v})_i + \sum_{i=1}^{m} w_i\,(\mathbf{X}\underset{\sim}{v})_{k+i} && \text{(Definition 10 in Vectors, split the sum at } k\text{)}\\
&= \underset{\sim}{u}\cdot(\mathbf{A}\underset{\sim}{u}+\mathbf{B}\underset{\sim}{w}) + \underset{\sim}{w}\cdot(\mathbf{B}^{\top}\underset{\sim}{u}+\mathbf{C}\underset{\sim}{w}) && \text{(the blocks of } \mathbf{X}\underset{\sim}{v}\text{)}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + \underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{w}^{\top}\mathbf{B}^{\top}\underset{\sim}{u} + \underset{\sim}{w}^{\top}\mathbf{C}\underset{\sim}{w} && \text{(Theorem 1 in Orthogonal Complements and the Fundamental}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + \underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{w}^{\top}\mathbf{C}\underset{\sim}{w} && (\underset{\sim}{w}^{\top}\mathbf{B}^{\top}\underset{\sim}{u} = (\underset{\sim}{w}^{\top}\mathbf{B}^{\top}\underset{\sim}{u})^{\top}\text{, which equals } \underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w};\ 1\times 1\text{, Definition 6}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + 2\,\underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{w}^{\top}\mathbf{C}\underset{\sim}{w}. && \text{(combine like terms)}
\end{aligned}
$$

Completing the square. Write $\underset{\sim}{d} \overset{\text{def}}{=} \mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}$, so $\underset{\sim}{z} = \underset{\sim}{u} + \underset{\sim}{d}$ and $\mathbf{A}\underset{\sim}{d} = \mathbf{B}\underset{\sim}{w}$ (Theorem 1 in Matrix Multiplication, Definition 5 in Special Matrices, Quadratic Forms and Trace, Theorem 4 in Matrix Multiplication); $\mathbf{A}^{-1}$ is symmetric (Corollary 1 in Special Matrices, Quadratic Forms and Trace). Then

$$
\begin{aligned}
\underset{\sim}{z}^{\top}\mathbf{A}\underset{\sim}{z} &= (\underset{\sim}{u}+\underset{\sim}{d})\cdot\mathbf{A}\,(\underset{\sim}{u}+\underset{\sim}{d}) && \text{(substitute } \underset{\sim}{z} = \underset{\sim}{u}+\underset{\sim}{d}\text{)}\\
&= (\underset{\sim}{u}+\underset{\sim}{d})\cdot(\mathbf{A}\underset{\sim}{u}+\mathbf{A}\underset{\sim}{d}) && \text{(Theorem 2 in Matrix Multiplication)}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{d} + \underset{\sim}{d}^{\top}\mathbf{A}\underset{\sim}{u} + \underset{\sim}{d}^{\top}\mathbf{A}\underset{\sim}{d} && \text{(Theorem 1 in Orthogonal Complements and the Fundamental}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{d} + \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{d} + \underset{\sim}{d}^{\top}\mathbf{A}\underset{\sim}{d} && \text{(as in the expansion, with } \mathbf{A}^{\top}=\mathbf{A}\text{)}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + 2\,\underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{d} + \underset{\sim}{d}^{\top}\mathbf{A}\underset{\sim}{d} && \text{(combine like terms)}\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + 2\,\underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{d}^{\top}\mathbf{B}\underset{\sim}{w} && (\mathbf{A}\underset{\sim}{d} = \mathbf{B}\underset{\sim}{w}\text{, twice})\\
&= \underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + 2\,\underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{w}^{\top}\,\mathbf{B}^{\top}\mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}, && (\underset{\sim}{d}^{\top} = \underset{\sim}{w}^{\top}\mathbf{B}^{\top}\mathbf{A}^{-1}\text{, Theorem 6 in Matrix Multiplication, Corolla}
\end{aligned}
$$

and

$$
\begin{aligned}
\underset{\sim}{w}^{\top}\mathbf{S}\underset{\sim}{w} &= \underset{\sim}{w}^{\top}\,(\mathbf{C}-\mathbf{B}^{\top}\mathbf{A}^{-1}\mathbf{B})\,\underset{\sim}{w} && \text{(substitute } \mathbf{S}\text{)}\\
&= \underset{\sim}{w}^{\top}\,(\mathbf{C}\underset{\sim}{w}-\mathbf{B}^{\top}\mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}) && \text{(Theorem 1 in Generalized Inverses and Linear Systems, difference rule c}\\
&= \underset{\sim}{w}^{\top}\mathbf{C}\underset{\sim}{w} - \underset{\sim}{w}^{\top}\,\mathbf{B}^{\top}\mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}. && \text{(Theorem 1 in Generalized Inverses and Linear Systems, difference rule c}
\end{aligned}
$$

Adding the two displays, the $\underset{\sim}{w}^{\top}\,\mathbf{B}^{\top}\mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}$ terms cancel, leaving $\underset{\sim}{u}^{\top}\mathbf{A}\underset{\sim}{u} + 2\,\underset{\sim}{u}^{\top}\mathbf{B}\underset{\sim}{w} + \underset{\sim}{w}^{\top}\mathbf{C}\underset{\sim}{w}$, which is $\underset{\sim}{v}^{\top}\mathbf{X}\underset{\sim}{v}$ by Equation 1. □

Example 0.8 (The 2×2 case)**.** For $\mathbf{X} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ ($k=1$ and $m=1$) and $a\neq 0$, the theorem reads

$$au^2 + 2buw + cw^2 = a\,(u + \tfrac{b}{a}\,w)^2 + \left(c - \tfrac{b^2}{a}\right)w^2,$$

with $\underset{\sim}{v} = (u, w)$: the familiar completing of the square. With

$$\begin{aligned} a &= c \\ &= 5 \end{aligned}$$

and $b = 4$:

$$\begin{aligned} 5\,(u + 0.8\,w)^2 + 1.8\,w^2 &= 5\,(u^2 + 1.6\,uw + 0.64\,w^2) + 1.8\,w^2 \\ &= 5u^2 + 8uw + 3.2\,w^2 + 1.8\,w^2 \\ &= 5u^2 + 8uw + 5w^2. \end{aligned}$$

Definition 0.3 (Schur complement)**.** For a block matrix $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$ as in Theorem 0.6 with $\mathbf{A}$ invertible, the **Schur complement** [3] of $\mathbf{A}$ in $\mathbf{X}$ is the $m \times m$ matrix $\mathbf{S} = \mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}$ defined in Theorem 0.6.

Example 0.9 (Schur complements of 2×2 matrices)**.**

- For $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ with $\mathbf{A} = [5]$:

$$\begin{aligned} \mathbf{S} &= 5 - 4 \cdot \tfrac{1}{5} \cdot 4 \\ &= \tfrac{25}{5} - \tfrac{16}{5} \\ &= \tfrac{9}{5}, \end{aligned}$$

  that is, $\mathbf{S} = [\frac{9}{5}]$.

- For $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ with $\mathbf{A} = [1]$:

$$\begin{aligned} \mathbf{S} &= [1 - 2 \cdot 1 \cdot 2] \\ &= [-3]. \end{aligned}$$

- For $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ the top-left block $[0]$ is not invertible, so there is no Schur complement of it.

Theorem 0.7 (Schur complement test)**.** *Let* $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$ *be as in Theorem 0.6. Then* $\mathbf{X}$ *is positive definite (Definition 2 in Definite Matrices) exactly when* $\mathbf{A}$ *is positive definite and its Schur complement* $\mathbf{S}$ *(Definition 0.3) is positive definite.*

Proof

Proof. **If.** $\mathbf{A}$ is positive definite, so invertible (Theorem 4 in Definite Matrices), and $\mathbf{S}$ is defined. $\mathbf{X}$ is symmetric. Let $\underset{\sim}{v} = (\underset{\sim}{u}, \underset{\sim}{w}) \neq \underset{\sim}{0}$ and $\underset{\sim}{z} = \underset{\sim}{u} + \mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}$. By Theorem 0.6, $\underset{\sim}{v}^\top \mathbf{X} \underset{\sim}{v} = \underset{\sim}{z}^\top \mathbf{A} \underset{\sim}{z} + \underset{\sim}{w}^\top \mathbf{S} \underset{\sim}{w}$, and both terms are at least 0. If $\underset{\sim}{w} \neq \underset{\sim}{0}$, the second term is positive. If $\underset{\sim}{w} = \underset{\sim}{0}$, then $\underset{\sim}{u} \neq \underset{\sim}{0}$ and $\underset{\sim}{z} = \underset{\sim}{u}$, so the first term is positive. Either way $\underset{\sim}{v}^\top \mathbf{X} \underset{\sim}{v} > 0$.
Only if. $\mathbf{A}$ is the top-left $k \times k$ block of $\mathbf{X}$, so it is positive definite (Theorem 0.5), and invertible (Theorem 4 in Definite Matrices). $\mathbf{S}$ is symmetric:

$$\begin{aligned}\left(\mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}\right)^\top &= \mathbf{B}^\top \left(\mathbf{A}^{-1}\right)^\top \mathbf{B} \\ &= \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}\end{aligned}$$

(Theorem 6 in Matrix Multiplication, Definition 6 in Matrices for $\left(\mathbf{B}^\top\right)^\top = \mathbf{B}$, Corollary 1 in Special Matrices, Quadratic Forms and Trace), and $\mathbf{C}$ is symmetric, so $\mathbf{S} = \mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}$ is symmetric (the transpose of a difference is the difference of the transposes, entry by entry, Definition 6 in Matrices). For $\underset{\sim}{w} \neq \underset{\sim}{0}$, take $\underset{\sim}{u} = -\mathbf{A}^{-1}\mathbf{B}\underset{\sim}{w}$, so $\underset{\sim}{z} = \underset{\sim}{0}$ and $\underset{\sim}{v} = (\underset{\sim}{u}, \underset{\sim}{w}) \neq \underset{\sim}{0}$; then $\underset{\sim}{w}^\top \mathbf{S} \underset{\sim}{w} = \underset{\sim}{v}^\top \mathbf{X} \underset{\sim}{v} > 0$ (Theorem 0.6). □

Example 0.10 (Testing 2×2 matrices)**.** A symmetric $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite exactly when $a > 0$ and $c - \frac{b^2}{a} > 0$, that is, $a > 0$ and $ac > b^2$ (a 1×1 matrix $[s]$ is positive definite exactly when $s > 0$, since $x^\top s x = s x^2$).

- $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$: $5 > 0$ and $\frac{9}{5} > 0$ (Example 0.9), so it is positive definite.
- $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$: $1 > 0$ but $\mathbf{S} = -3$, so it is not, as Example 4 in Definite Matrices found directly.

Theorem 0.8 (The inverse of a positive definite block matrix)**.** *Let* $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$ *be as in Theorem 0.6, with* $\mathbf{X}$ *positive definite. Then* $\mathbf{X}$ *is invertible, and with the Schur complement* $\mathbf{S}$ *of* $\mathbf{A}$*, its inverse is the following, where* $\mathbf{P}$ *is the block of the first* k *rows and columns of* $\mathbf{X}^{-1}$*:*

$$\mathbf{X}^{-1} = \begin{bmatrix} \mathbf{P} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{B}^\top \mathbf{A}^{-1} & \mathbf{S}^{-1} \end{bmatrix},$$

Proof

Proof. $\mathbf{X}$ is invertible, because it is positive definite. By the Schur complement test, $\mathbf{A}$ and $\mathbf{S}$ are positive definite as well, and so both are invertible. Let $\mathbf{W} = \begin{bmatrix} -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{I}_m \end{bmatrix}$. Each entry of $\mathbf{XW}$ adds one group of k summands to one group of m summands, so $\mathbf{XW}$ can be computed one block row at a time:

$$\begin{aligned} \mathbf{XW} &= \begin{bmatrix} -\mathbf{A}\mathbf{A}^{-1}\mathbf{B} + \mathbf{B} \\ -\mathbf{B}^\top\mathbf{A}^{-1}\mathbf{B} + \mathbf{C} \end{bmatrix} && \text{(one block row at a time)} \\ &= \begin{bmatrix} -\mathbf{I}_k\mathbf{B} + \mathbf{B} \\ \mathbf{S} \end{bmatrix} && (\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_k, \text{ and definition of } \mathbf{S}) \\ &= \begin{bmatrix} \mathbf{0} \\ \mathbf{S} \end{bmatrix} && (\mathbf{I}_k\mathbf{B} = \mathbf{B}) \end{aligned}$$

Multiplying both sides on the left by $\mathbf{X}^{-1}$ gives $\mathbf{W}$ on the left, since $\mathbf{X}^{-1}\mathbf{X} = \mathbf{I}_{k+m}$. Cut $\mathbf{X}^{-1}$ into the four pieces $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{T}$, with $\mathbf{Q}$ and $\mathbf{T}$ in the last m columns. Then the right side is $\mathbf{X}^{-1}\begin{bmatrix} \mathbf{0} \\ \mathbf{S} \end{bmatrix}$, and the last m columns of $\mathbf{X}^{-1}$ multiply $\mathbf{S}$:

$$\begin{bmatrix} -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \mathbf{QS} \\ \mathbf{TS} \end{bmatrix}.$$

So $\mathbf{TS} = \mathbf{I}_m$, and $\mathbf{QS} = -\mathbf{A}^{-1}\mathbf{B}$. Multiplying both equations on the right by $\mathbf{S}^{-1}$ gives $\mathbf{T} = \mathbf{S}^{-1}$, and $\mathbf{Q} = -\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1}$. $\mathbf{X}^{-1}$ is symmetric, by the inverse of a symmetric matrix is symmetric, so $\mathbf{R} = \mathbf{Q}^\top = -\mathbf{S}^{-1}\mathbf{B}^\top\mathbf{A}^{-1}$, using $(\mathbf{S}^{-1})^\top = \mathbf{S}^{-1}$ and $(\mathbf{A}^{-1})^\top = \mathbf{A}^{-1}$. □

Example 0.11 (The bottom-right entry of the inverse of a 2×2 matrix)**.**
Let $\mathbf{X} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ with $\mathbf{A} = [5]$, whose Schur complement is $\mathbf{S} = [\frac{9}{5}]$ (Example 0.9). It is positive definite (Example 0.10), so by Theorem 0.8:

$$\begin{aligned} \mathbf{T} &= \mathbf{S}^{-1} \\ &= [\tfrac{5}{9}]. \end{aligned}$$

Directly, $\mathbf{X}^{-1} = \frac{1}{9}\begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$, because

$$\begin{aligned} \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \cdot \tfrac{1}{9}\begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix} &= \tfrac{1}{9}\begin{bmatrix} 25-16 & -20+20 \\ 20-20 & -16+25 \end{bmatrix} \\ &= \mathbf{I}_2, \end{aligned}$$

and its bottom-right entry is $\frac{5}{9}$. The top-right entry is $-\mathbf{A}^{-1}\mathbf{B}\mathbf{S}^{-1}$:

$$-\tfrac{1}{5} \cdot 4 \cdot \tfrac{5}{9} = -\tfrac{4}{9}.$$

Theorem 0.9 (Cholesky factorization)**.** *A $p \times p$ matrix $\mathbf{A}$ is positive definite (Definition 2 in Definite Matrices) exactly when $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$ for some lower triangular $\mathbf{L}$ (Definition 2 in Solving Linear Systems and Least Squares) with positive diagonal entries. For a positive definite $\mathbf{A}$ that $\mathbf{L}$ is unique; it is the **Cholesky factor** [4] of $\mathbf{A}$.*

Proof

Proof. **If.** $\mathbf{L}^\top$ is upper triangular with positive diagonal (Definition 6 in Matrices), so $\mathbf{L}^\top \underset{\sim}{x} = \underset{\sim}{0}$ only for $\underset{\sim}{x} = \underset{\sim}{0}$ (Theorem 1 in Solving Linear Systems and Least Squares), and the columns of $\mathbf{L}^\top$ are linearly independent (Theorem 3 in Subspaces and Rank, Definition 1 in Subspaces and Rank). So $\mathbf{A} = (\mathbf{L}^\top)^\top \mathbf{L}^\top$ (Definition 6 in Matrices) is positive definite (Theorem 0.1, part 2).
Only if, with uniqueness, by induction on p. For $p = 1$, $\mathbf{A} = [a]$ with $a > 0$ (Theorem 0.5), and $[c]\,[c] = [a]$ with $c > 0$ holds exactly for $c = \sqrt{a}$. For $p > 1$, assume every $(p-1) \times (p-1)$ positive definite matrix has exactly one such factor, and split off the first row and column. $\mathbf{A}$ is symmetric, so its first row is the transpose of its first column:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \underset{\sim}{b}^\top \\ \underset{\sim}{b} & \mathbf{A}_{22} \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} \ell_{11} & \underset{\sim}{0}_{1\times(p-1)} \\ \underset{\sim}{g} & \mathbf{L}_{22} \end{bmatrix}, \qquad \mathbf{L}^\top = \begin{bmatrix} \ell_{11} & \underset{\sim}{g}^\top \\ \underset{\sim}{0}_{(p-1)\times 1} & \mathbf{L}_{22}^\top \end{bmatrix},$$

with $\underset{\sim}{b}, \underset{\sim}{g} \in \mathbb{R}^{p-1}$, where b_i and g_i are entries $(i+1, 1)$ of $\mathbf{A}$ and $\mathbf{L}$, and $(\mathbf{A}_{22})_{ij}$ and $(\mathbf{L}_{22})_{ij}$ are their entries $(i+1, j+1)$; $\mathbf{L}^\top$ has this form by Definition 6 in Matrices. A $p \times p$ matrix $\mathbf{L}$ is lower triangular with positive diagonal exactly when $\ell_{11} > 0$, its first row is otherwise zero, and $\mathbf{L}_{22}$ is lower triangular with positive diagonal; $\underset{\sim}{g}$ can be anything. So choosing $\mathbf{L}$ means choosing $\ell_{11} > 0$, g and such an $\mathbf{L}_{22}$.
Entry (i, j) of $\mathbf{L}\mathbf{L}^\top$ is $\sum_{r=1}^{p} \ell_{ir}\,\ell_{jr}$, with ℓ_{ir} entry (i, r) of $\mathbf{L}$ (Definition 1 in Matrix Multiplication, Definition 6 in Matrices). Row 1 of $\mathbf{L}$ is $(\ell_{11}, 0, \ldots, 0)$, and row $i+1$ is $(g_i, (\mathbf{L}_{22})_{i1}, \ldots, (\mathbf{L}_{22})_{i,p-1})$. So, for $i, j = 1, \ldots, p-1$:

$$
\begin{aligned}
(\mathbf{L}\mathbf{L}^\top)_{11} &= \ell_{11}\,\ell_{11} + \sum_{r=2}^{p} 0 \cdot 0 \\
&= \ell_{11}^2, && \text{(rows 1 and 1)} \\
(\mathbf{L}\mathbf{L}^\top)_{i+1,1} &= g_i\,\ell_{11} + \sum_{r=2}^{p} (\mathbf{L}_{22})_{i,r-1} \cdot 0 \\
&= \ell_{11}\,g_i, && \text{(rows } i+1 \text{ and 1)} \\
(\mathbf{L}\mathbf{L}^\top)_{i+1,j+1} &= g_i\,g_j + \sum_{r=2}^{p} (\mathbf{L}_{22})_{i,r-1}\,(\mathbf{L}_{22})_{j,r-1} && \text{(rows } i+1 \text{ and } j+1\text{)} \\
&= (\underset{\sim}{g}\underset{\sim}{g}^\top)_{ij} + (\mathbf{L}_{22}\mathbf{L}_{22}{}^\top)_{ij}. && \text{(Definition 1 in Matrix Multiplication, with } s = r-1\text{)}
\end{aligned}
$$

Entry $(1, j+1)$ is $\ell_{11}\,g_j$ by the same computation with the roles swapped, the transpose of the $(2,1)$ block; it matches $\underset{\sim}{b}^\top$ exactly when the $(2,1)$ block matches $\underset{\sim}{b}$, which is where the symmetry of $\mathbf{A}$ is used. So $\mathbf{L}\mathbf{L}^\top = \mathbf{A}$ says exactly

$$
\ell_{11}^2 = a_{11}, \qquad \ell_{11}\,\underset{\sim}{g} = \underset{\sim}{b}, \qquad \underset{\sim}{g}\underset{\sim}{g}^\top + \mathbf{L}_{22}\mathbf{L}_{22}{}^\top = \mathbf{A}_{22}.
$$

$a_{11} > 0$ (Theorem 0.5), so the first equation with $\ell_{11} > 0$ forces $\ell_{11} = \sqrt{a_{11}}$; the second then forces $\underset{\sim}{g} = \frac{1}{\ell_{11}}\,\underset{\sim}{b}$; and the third becomes

$$
\begin{aligned}
\mathbf{L}_{22}\mathbf{L}_{22}{}^\top &= \mathbf{A}_{22} - \underset{\sim}{g}\underset{\sim}{g}^\top && \text{(subtract } \underset{\sim}{g}\underset{\sim}{g}^\top\text{)} \\
&= \mathbf{A}_{22} - \Big(\tfrac{1}{\ell_{11}}\,\underset{\sim}{b}\Big)\Big(\tfrac{1}{\ell_{11}}\,\underset{\sim}{b}\Big)^\top && \text{(substitute } \underset{\sim}{g}\text{)} \\
&= \mathbf{A}_{22} - \tfrac{1}{\ell_{11}^2}\,\underset{\sim}{b}\underset{\sim}{b}^\top && \text{(Definition 6 in Matrices, Theorem 1 in Generalized Inverses and Linear} \\
&= \mathbf{A}_{22} - \tfrac{1}{a_{11}}\,\underset{\sim}{b}\underset{\sim}{b}^\top. && (\ell_{11}^2 = a_{11})
\end{aligned}
$$

Here $[a_{11}]$ is invertible with inverse $[1/a_{11}]$, so with $\mathbf{B} = \underset{\sim}{b}^\top$ in Definition 0.3,

$$
\begin{aligned}
\mathbf{B}^\top\,[a_{11}]^{-1}\mathbf{B} &= \underset{\sim}{b}\,\tfrac{1}{a_{11}}\,\underset{\sim}{b}^\top \\
&= \tfrac{1}{a_{11}}\,\underset{\sim}{b}\underset{\sim}{b}^\top
\end{aligned}
$$

(Theorem 1 in Generalized Inverses and Linear Systems), and the right side is the Schur complement of $[a_{11}]$ in $\mathbf{A}$. It is positive definite by Theorem 0.7 (with $k = 1$; $\mathbf{A}_{22}$ is symmetric because $\mathbf{A}$ is). By the induction hypothesis it has exactly one factor $\mathbf{L}_{22}$ of the required kind. So $\mathbf{L}$ exists and is unique. □

Example 0.12 (Two Cholesky factors)**.**

- $\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 5 \end{bmatrix}$:

$$\begin{aligned} \ell_{11} &= \sqrt{4} \\ &= 2, \end{aligned}$$

$$\begin{aligned} \ell_{21} &= 2/2 \\ &= 1, \end{aligned}$$

and

$$\begin{aligned} \ell_{22}^2 &= 5 - 1^2 \\ &= 4, \end{aligned}$$

so $\ell_{22} = 2$: $\mathbf{L} = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$, and

$$\begin{aligned} \mathbf{L}\mathbf{L}^\top &= \begin{bmatrix} 2 \cdot 2 & 2 \cdot 1 \\ 1 \cdot 2 & 1 \cdot 1 + 2 \cdot 2 \end{bmatrix} \\ &= \mathbf{A}. \end{aligned}$$

- $\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$: $\ell_{11} = \sqrt{5}$, $\ell_{21} = 4/\sqrt{5}$, and

$$\begin{aligned} \ell_{22}^2 &= 5 - \tfrac{16}{5} \\ &= \tfrac{9}{5}, \end{aligned}$$

the Schur complement of Example 0.9, so $\ell_{22} = \frac{3}{\sqrt{5}}$.

- For $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ the method breaks down:

$$\begin{aligned} \ell_{22}^2 &= 1 - 4 \\ &= -3 \end{aligned}$$

has no real solution, matching the failure of the Schur complement test (Example 0.10).

[1] H. Zhou, "Symmetric positive definite matrices." Lecture notes for Biostat 216, Mathematical Methods for Biostatistics, University of California, Los Angeles, 2024. Accessed: Oct. 05, 2026. [Online]. Available: https://ucla-biostat-216.github.io/2024fall/slides/11-pd/11-pd.html

[2] Wikipedia contributors, "Block matrix — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Block_matrix

[3] Wikipedia contributors, "Schur complement — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Schur_complement

[4] Wikipedia contributors, "Cholesky decomposition — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Cholesky_decomposition