

Diagonalization and Cholesky

Last modified: 2026-10-27 20:53:54 (PDT)

Similarity and diagonalization

Source

This section is adapted from Zhou [1], used under the MIT License (see the license text in [Section 2 in Subspaces and Rank](#)). These notes take eigenvalues to be real numbers ([Definition 16 in Special Matrices and Decompositions](#)), so the source's complex eigenvalues are left out, and eigenvalues are characterized through null spaces rather than through the characteristic polynomial. Only the source's material on similarity, diagonalization and the basic eigenvalue properties is adapted; its characteristic polynomial and algebraic multiplicity, its trace and determinant identities, and the part of its section on symmetric matrices that builds bases of the column space and null space from eigenvectors are left out. The rest of that section is covered by [Theorem 12 in Special Matrices and Decompositions](#) and [Theorem 13 in Special Matrices and Decompositions](#). Its result that eigenvalues of orthogonal matrices have modulus 1 appears here in the real form ± 1 . Three other parts of the source are not part of this section:

- its statement that the eigenvalues of a lower triangular matrix are also its diagonal entries (Theorem [0.4](#) states the upper triangular case)
- its three further examples of matrices with geometric multiplicity smaller than algebraic multiplicity (Example [0.2](#) gives one nilpotent example)
- its code cells The source's two-vector argument that eigenvectors for distinct eigenvalues are independent is extended here to any number of eigenvectors.

Theorem 0.1 (A square matrix is invertible exactly when it has full column rank). *An $n \times n$ matrix $\mathbf{A}$ is invertible ([Definition 6 in Special Matrices and Decompositions](#)) exactly when $\text{rank}(\mathbf{A}) = n$ (full column rank, [Definition 3 in Subspaces and Rank](#)), that is, exactly when $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}_n\}$.*

i Proof

Proof. **Invertible implies rank n .** If $\mathbf{A}\tilde{x} = \tilde{0}_n$, then

$$\begin{aligned} \tilde{x} &= \mathbf{I}_n \tilde{x} && \text{(Theorem 8 in Matrices)} \\ &= (\mathbf{A}^{-1} \mathbf{A}) \tilde{x} && \text{(Definition 5 in Special Matrices and Decompositions)} \\ &= \mathbf{A}^{-1} (\mathbf{A} \tilde{x}) && \text{(Theorem 5 in Matrices)} \\ &= \mathbf{A}^{-1} \tilde{0}_n && (\mathbf{A} \tilde{x} = \tilde{0}_n) \\ &= \tilde{0}_n, && \text{(Definition 13 in Matrices)} \end{aligned}$$

so $\mathcal{N}(\mathbf{A}) = \{\tilde{0}_n\}$, the nullity is 0, and $\text{rank}(\mathbf{A}) = n$ (Theorem 7 in Column Space, Null Space and Rank-Nullity).

Rank n implies invertible. Every system $\mathbf{A}y = \tilde{b}$ has exactly one solution (Corollary 1 in Projections and Linear Systems, both parts, with $m = n$). Let $\tilde{y}_j$ solve $\mathbf{A}\tilde{y}_j = \tilde{e}_j$ (Definition 18 in Vectors), and let $\mathbf{B}$ have columns $\tilde{y}_1, \dots, \tilde{y}_n$. Column j of $\mathbf{AB}$ is $\mathbf{A}\tilde{y}_j = \tilde{e}_j$ (1), so $\mathbf{AB} = \mathbf{I}_n$. For the other order,

$$\begin{aligned} \mathbf{A}(\mathbf{BA} - \mathbf{I}_n) &= \mathbf{ABA} - \mathbf{AI}_n && \text{(Theorem 7 in Projections and Linear Systems, Theorem 5 in Matrices)} \\ &= \mathbf{I}_n \mathbf{A} - \mathbf{AI}_n && (\mathbf{AB} = \mathbf{I}_n) \\ &= \mathbf{A} - \mathbf{A} && \text{(Theorem 8 in Matrices)} \\ &= \mathbf{0}_{n \times n}, && \text{(arithmetic)} \end{aligned}$$

so every column of $\mathbf{BA} - \mathbf{I}_n$ is in $\mathcal{N}(\mathbf{A}) = \{\tilde{0}_n\}$ (1), and $\mathbf{BA} = \mathbf{I}_n$. So $\mathbf{B}$ satisfies Definition 6 in Special Matrices and Decompositions. The two conditions in the statement are equivalent by Theorem 7 in Column Space, Null Space and Rank-Nullity: the nullity is 0 exactly when $\mathcal{N}(\mathbf{A}) = \{\tilde{0}_n\}$, because a subspace of dimension 0 has the empty list as a basis, whose span is $\{\tilde{0}\}$ (Definition 9 in Subspaces and Rank, Definition 5 in Subspaces and Rank), and $\{\tilde{0}\}$ has dimension 0 (Example 16 in Subspaces and Rank). $\square$

Example 0.1 (Rank decides invertibility).

- $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$ of Example 5 in Special Matrices and Decompositions has rank 2: $c_1(2, 0) + c_2(1, 1) = (2c_1 + c_2, c_2)$ is $\tilde{0}$ only if $c_2 = 0$ and then $c_1 = 0$. So it is invertible, as that example found by exhibiting the inverse.
- $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has rank less than 2: $(1, -1)$ is a nonzero vector in its null space. So it is singular, as Example 5 in Special Matrices and Decompositions found.

Definition 0.1 (Eigenspace). Let λ be an eigenvalue of a $p \times p$ matrix $\mathbf{A}$ (Definition 16 in Special Matrices and Decompositions). The **eigenspace** [2] of λ is $\mathcal{E}_\lambda \stackrel{\text{def}}{=} \mathcal{N}(\mathbf{A} - \lambda \mathbf{I}_p)$. It is a subspace of $\mathbb{R}^p$ (Theorem 2 in Column Space, Null Space and Rank-Nullity).

Definition 0.2 (Geometric multiplicity). The dimension (Definition 9 in Subspaces and Rank) of the eigenspace $\mathcal{E}_\lambda$ (Definition 0.1) of an eigenvalue λ is the **geometric multiplicity** of λ .

Example 0.2 (Eigenspaces of two matrices).

- For $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ of Example 24 in Special Matrices and Decompositions, $\mathbf{A} - 3\mathbf{I}_2 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ sends $\underline{v}$ to $(v_2 - v_1)(1, -1)$, so $\mathcal{E}_3 = \text{span}\{(1, 1)\}$, with geometric multiplicity 1.
- For $\mathbf{J} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{J}\underline{v} = (v_2, 0)$.

$$\begin{aligned} \mathbf{J}(1, 0) &= (0, 0) \\ &= 0(1, 0), \end{aligned}$$

so 0 is an eigenvalue. If $\mathbf{J}\underline{v} = \lambda\underline{v}$ with $\lambda \neq 0$, the second entry gives $\lambda v_2 = 0$, so $v_2 = 0$, and then the first gives $\lambda v_1 = 0$, so $\underline{v} = \underline{0}$. So 0 is the only eigenvalue, and

$$\begin{aligned} \mathcal{E}_0 &= \mathcal{N}(\mathbf{J}) \\ &= \{(t, 0) : t \in \mathbb{R}\}, \end{aligned}$$

with geometric multiplicity 1.

Example 0.3 (A number that is not an eigenvalue has no eigenspace). For $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and $\lambda = 2$, $\mathbf{A} - 2\mathbf{I}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ sends $\underline{v}$ to (v_2, v_1) , which is $\underline{0}$ only for $\underline{v} = \underline{0}$, so $\mathcal{N}(\mathbf{A} - 2\mathbf{I}_2) = \{\underline{0}\}$. And 2 is not an eigenvalue (Definition 16 in Special Matrices and Decompositions): $\mathbf{A}\underline{v} = 2\underline{v}$ reads $(2v_1 + v_2, v_1 + 2v_2) = (2v_1, 2v_2)$, so $v_2 = 0$ and $v_1 = 0$. So Definition 0.1 does not apply to 2.

Theorem 0.2 (Eigenvalues are where $\mathbf{A} - \lambda\mathbf{I}$ is singular). Let $\mathbf{A}$ be a $p \times p$ matrix and λ a real number. Then λ is an eigenvalue of $\mathbf{A}$ ([Definition 16 in Special Matrices and Decompositions](#)) exactly when $\mathbf{A} - \lambda\mathbf{I}_p$ is singular ([Definition 6 in Special Matrices and Decompositions](#)). In particular, $\mathbf{A}$ is singular exactly when 0 is one of its eigenvalues; and each eigenspace $\mathcal{E}_\lambda$ ([Definition 0.1](#)) is a subspace of $\mathbb{R}^p$ consisting of $\mathbf{0}$ and the eigenvectors for λ .

i Proof

Proof. For any $\tilde{v} \in \mathbb{R}^p$,

$$\begin{aligned} (\mathbf{A} - \lambda\mathbf{I}_p) \tilde{v} &= \mathbf{A}\tilde{v} - (\lambda\mathbf{I}_p) \tilde{v} \quad (\text{Theorem 7 in Projections and Linear Systems, the difference law}) \\ &= \mathbf{A}\tilde{v} - \lambda(\mathbf{I}_p \tilde{v}) \quad (\text{Theorem 7 in Projections and Linear Systems, the scalar factor}) \\ &= \mathbf{A}\tilde{v} - \lambda\tilde{v}, \quad (\text{Theorem 8 in Matrices}) \end{aligned}$$

so $\mathbf{A}\tilde{v} = \lambda\tilde{v}$ exactly when $\tilde{v} \in \mathcal{N}(\mathbf{A} - \lambda\mathbf{I}_p)$; the nonzero such $\tilde{v}$ are the eigenvectors for λ . So λ is an eigenvalue exactly when that null space contains a nonzero vector, which by [Theorem 0.1](#) is exactly when $\mathbf{A} - \lambda\mathbf{I}_p$ is singular. With $\lambda = 0$, $\mathbf{A} - 0\mathbf{I}_p = \mathbf{A}$. An eigenspace is a null space, so it is a subspace ([Theorem 2 in Column Space, Null Space and Rank-Nullity](#)). $\square$

Example 0.4 (Eigenvalues and singular matrices). For $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{A} - 3\mathbf{I}_2 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ is singular (its columns are negatives of each other, so its rank is 1), matching the eigenvalue 3 of [Example 24 in Special Matrices and Decompositions](#). $\mathbf{A}$ itself has rank 2, since $c_1(2, 1) + c_2(1, 2) = (2c_1 + c_2, c_1 + 2c_2)$ is $\mathbf{0}$ only if

$$\begin{aligned} c_1 &= c_2 \\ &= 0 \end{aligned}$$

(subtract twice the second entry from the first: $-3c_2 = 0$, and then the second entry gives $c_1 = 0$); so it is invertible ([Theorem 0.1](#)), and 0 is not an eigenvalue. The singular matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has eigenvalue 0, with eigenvector $(1, -1)$.

Theorem 0.3 (Eigenvalues of shifts and powers). Let $\tilde{v}$ be an eigenvector of a $p \times p$ matrix $\mathbf{A}$ for the eigenvalue λ ([Definition 16 in Special Matrices and Decompositions](#)). Then

1. $\tilde{v}$ is an eigenvector of $\mathbf{A} + s\mathbf{I}_p$ for $\lambda + s$, for every number s ;

2. v is an eigenvector of $\mathbf{A}^k$ ([Definition 2 in Special Matrices and Decompositions](#)) for $\tilde{\lambda}^k$, for every positive integer k .

i Proof

Proof. **Part 1.**

$$\begin{aligned}
 (\mathbf{A} + s\mathbf{I}_p)\tilde{v} &= \mathbf{A}\tilde{v} + (s\mathbf{I}_p)\tilde{v} && \text{(Theorem 6 in Matrices)} \\
 &= \mathbf{A}\tilde{v} + s(\mathbf{I}_p\tilde{v}) && \text{(Theorem 7 in Projections and Linear Systems)} \\
 &= \lambda\tilde{v} + s(\mathbf{I}_p\tilde{v}) && \text{(eigenvector)} \\
 &= \lambda\tilde{v} + s\tilde{v} && \text{(Theorem 8 in Matrices)} \\
 &= (\lambda + s)\tilde{v}. && \text{(add entrywise)}
 \end{aligned}$$

Part 2, by induction on k . For $k = 1$ it is the assumption. If $\mathbf{A}^{k-1}\tilde{v} = \lambda^{k-1}\tilde{v}$, then

$$\begin{aligned}
 \mathbf{A}^k\tilde{v} &= \mathbf{A}(\mathbf{A}^{k-1}\tilde{v}) && \text{(Definition 2 in Special Matrices and Decompositions, Theorem 5 in Matrices)} \\
 &= \mathbf{A}(\lambda^{k-1}\tilde{v}) && \text{(induction hypothesis)} \\
 &= \lambda^{k-1}(\mathbf{A}\tilde{v}) && \text{(Theorem 7 in Projections and Linear Systems)} \\
 &= \lambda^{k-1}(\lambda\tilde{v}) && \text{(eigenvector)} \\
 &= \lambda^k\tilde{v}. && \text{(multiply the numbers)}
 \end{aligned}$$

In both parts $\tilde{v} \neq 0$, so $\tilde{v}$ is an eigenvector. □

Example 0.5 (Shifting and squaring). For $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ with eigenvector $(1, 1)$ for 3 ([Example 24 in Special Matrices and Decompositions](#)):

- $\mathbf{A} - 2\mathbf{I}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ sends $(1, 1)$ to $(1, 1) = (3 - 2)(1, 1)$;
-

$$\begin{aligned}
 \mathbf{A}^2 &= \begin{bmatrix} 2 \cdot 2 + 1 \cdot 1 & 2 \cdot 1 + 1 \cdot 2 \\ 1 \cdot 2 + 2 \cdot 1 & 1 \cdot 1 + 2 \cdot 2 \end{bmatrix} \\
 &= \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}
 \end{aligned}$$

(1) sends $(1, 1)$ to $(9, 9) = 3^2(1, 1)$.

Theorem 0.4 (The eigenvalues of an upper triangular matrix are its diagonal entries). *The eigenvalues (Definition 16 in Special Matrices and Decompositions) of an upper triangular $p \times p$ matrix $\mathbf{U}$ (Definition 7 in Projections and Linear Systems) are exactly its diagonal entries $u_{11}, \dots, u_{pp}$.*

Proof

Proof. $\mathbf{U} - \lambda \mathbf{I}_p$ is upper triangular, with diagonal entries $u_{ii} - \lambda$ (Definition 12 in Matrices, Definition 5 in Matrices).

If λ is not a diagonal entry, every $u_{ii} - \lambda \neq 0$, so $(\mathbf{U} - \lambda \mathbf{I}_p) \tilde{v} = \tilde{0}$ has exactly one solution, $\tilde{v} = \tilde{0}$ (Theorem 17 in Projections and Linear Systems), and λ is not an eigenvalue (Theorem 0.2).

If λ is a diagonal entry, let k be the smallest index with $u_{kk} = \lambda$, and write $\mathbf{M} = \mathbf{U} - \lambda \mathbf{I}_p$, so $m_{kk} = 0$ and $m_{ii} \neq 0$ for $i < k$. Look for $\tilde{v}$ with $v_k = 1$ and $v_j = 0$ for $j > k$. Entry i of $\mathbf{M}\tilde{v}$ is

$$\begin{aligned} \sum_{j=1}^p m_{ij} v_j &= \sum_{j=i}^p m_{ij} v_j \quad (m_{ij} = 0 \text{ for } j < i) \\ &= \sum_{j=i}^k m_{ij} v_j, \quad (v_j = 0 \text{ for } j > k) \end{aligned}$$

where the last sum is empty, and so 0, when $i > k$. So:

- for $i > k$ entry i is 0;
- for $i = k$ it is $m_{kk} v_k = 0$;
- for $i < k$ it is 0 exactly when $v_i = -\frac{1}{m_{ii}} \sum_{j=i+1}^k m_{ij} v_j$, and these equations fix $v_{k-1}, \dots, v_1$ in turn, as in Theorem 17 in Projections and Linear Systems.

So $\mathbf{M}\tilde{v} = \tilde{0}$ with $v_k = 1$, hence $\tilde{v} \neq \tilde{0}$, and λ is an eigenvalue. □

Example 0.6 (Eigenvalues read off the diagonal). $\mathbf{U} = \begin{bmatrix} 2 & 1 & -1 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & -1 \end{bmatrix}$ of Example 24 in Projections and Linear Systems has eigenvalues 2, $\frac{1}{2}$ and -1 . For $\lambda = \frac{1}{2}$ (so $k = 2$), take $v_2 = 1, v_3 = 0$; row 1 of $(\mathbf{U} - \frac{1}{2} \mathbf{I}_3) \tilde{v}$ is $\frac{3}{2} v_1 + 1 = 0$, so $v_1 = -\frac{2}{3}$. Check:

$$\begin{aligned} \mathbf{U} \left(-\frac{2}{3}, 1, 0\right) &= \left(-\frac{4}{3} + 1, \frac{1}{2}, 0\right) \\ &= \frac{1}{2} \left(-\frac{2}{3}, 1, 0\right). \end{aligned}$$

The matrix $\mathbf{J}$ of Example 0.2 is upper triangular with both diagonal entries 0, and 0 is its only eigenvalue.

Theorem 0.5 (Eigenvalues of idempotent and orthogonal matrices).

1. Every eigenvalue of an idempotent matrix ([Definition 7 in Special Matrices and Decompositions](#)) is 0 or 1.
2. Every eigenvalue of an orthogonal matrix ([Definition 10 in Special Matrices and Decompositions](#)) is 1 or -1 .

i Proof

Proof. Let $\tilde{v} \neq 0$ with $\mathbf{A}\tilde{v} = \lambda\tilde{v}$.

Part 1. If $\tilde{\mathbf{A}}^2 = \mathbf{A}$, then

$$\begin{aligned}\lambda\tilde{v} &= \mathbf{A}\tilde{v} && \text{(eigenvector)} \\ &= \tilde{\mathbf{A}}^2\tilde{v} && \text{(idempotent)} \\ &= \lambda^2\tilde{v}, && \text{(Theorem 0.3, part 2)}\end{aligned}$$

and subtracting $\lambda\tilde{v}$ from both sides gives $(\lambda^2 - \lambda)\tilde{v} = 0$. Some entry of $\tilde{v}$ is nonzero, so

$$\begin{aligned}\lambda^2 - \lambda &= \lambda(\lambda - 1) \\ &= 0,\end{aligned}$$

and a product of two numbers is 0 only when one of them is: $\lambda = 0$ or $\lambda = 1$.

Part 2. If $\mathbf{A}$ is orthogonal, then

$$\begin{aligned}\|\tilde{v}\| &= \|\mathbf{A}\tilde{v}\| && \text{(Theorem 6 in Special Matrices and Decompositions)} \\ &= \|\lambda\tilde{v}\| && \text{(eigenvector)} \\ &= |\lambda| \|\tilde{v}\|, && \text{(Theorem 1 in Inner Products and Orthogonality, part 2)}\end{aligned}$$

and $\|\tilde{v}\| > 0$ ([Theorem 1 in Inner Products and Orthogonality](#), part 1), so dividing by $\|\tilde{v}\|$ gives $|\lambda| = 1$; since λ is real ([Definition 16 in Special Matrices and Decompositions](#)), $\lambda = 1$ or $\lambda = -1$. $\square$

Example 0.7 (Projections and reflections).

- The idempotent matrix $\begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$ of [Example 3 in Projections and Linear Systems](#) sends $(1, 1)$ to $(1, 1)$ and $(1, -1)$ to $(0, 0)$: eigenvalues 1 and 0.
- The matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which swaps the two entries, is orthogonal: its columns $(0, 1)$ and $(1, 0)$ are orthonormal ([Remark 6 in Special Matrices and Decompositions](#)). It sends $(1, 1)$ to $(1, 1)$ and $(1, -1)$ to $(-1, 1)$: eigenvalues 1 and -1 .

- The rotation $\mathbf{Q}$ of [Example 13 in Special Matrices and Decompositions](#) has no real eigenvalue. By part 2 the only candidates are ± 1 . $\mathbf{Q}\tilde{v} = \tilde{v}$ reads $-0.4v_1 - 0.8v_2 = 0$ and $0.8v_1 - 0.4v_2 = 0$, so $v_1 = -2v_2$ and

$$\begin{aligned} v_2 &= 2v_1 \\ &= -4v_2, \end{aligned}$$

forcing $\tilde{v} = 0$; $\mathbf{Q}\tilde{v} = -\tilde{v}$ reads $1.6v_1 - 0.8v_2 = 0$ and $0.8v_1 + 1.6v_2 = 0$, so $v_2 = 2v_1$ and $0.8v_1 + 3.2v_1 = 0$, forcing $\tilde{v} = 0$ again.

Definition 0.3 (Similar matrices). Two $p \times p$ matrices $\mathbf{A}$ and $\mathbf{B}$ are **similar** [3] if $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ for some invertible $p \times p$ matrix $\mathbf{P}$ ([Definition 6 in Special Matrices and Decompositions](#)).

i Remark (Similarity read both ways)

Remark 0.1 (Similarity read both ways). If $\mathbf{P}$ is invertible, then $\mathbf{P}^{-1}$ is invertible with inverse $\mathbf{P}$:

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{P} &= \mathbf{P}\mathbf{P}^{-1} \\ &= \mathbf{I}_p \end{aligned}$$

(2) is the condition of [Definition 6 in Special Matrices and Decompositions](#) for $\mathbf{P}^{-1}$ with $\mathbf{P}$ as the other factor, and that factor is unique ([Remark 3 in Special Matrices and Decompositions](#)). If $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, then

$$\begin{aligned} \mathbf{P}\mathbf{B}\mathbf{P}^{-1} &= \mathbf{P}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\mathbf{P}^{-1} && \text{(substitute } \mathbf{B}) \\ &= (\mathbf{P}\mathbf{P}^{-1})\mathbf{A}(\mathbf{P}\mathbf{P}^{-1}) && \text{(Theorem 5 in Matrices)} \\ &= \mathbf{I}_p\mathbf{A}\mathbf{I}_p && \text{(Definition 5 in Special Matrices and Decompositions)} \\ &= \mathbf{A}, && \text{(Theorem 8 in Matrices)} \end{aligned}$$

and the same steps with $\mathbf{P}$ and $\mathbf{P}^{-1}$ exchanged turn $\mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$ back into $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. So $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ exactly when $\mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$, and similarity goes both ways: $\mathbf{A}$ is similar to $\mathbf{B}$ through $\mathbf{P}^{-1}$, whose inverse is $\mathbf{P}$.

Example 0.8 (A matrix similar to a diagonal one). Let $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Then

$$\begin{aligned}
\mathbf{P}^2 &= \begin{bmatrix} 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot (-1) \\ 1 \cdot 1 + (-1) \cdot 1 & 1 \cdot 1 + (-1)(-1) \end{bmatrix} && \text{(Definition 7 in Matrices)} \\
&= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} && \text{(arithmetic)} \\
&= 2\mathbf{I}_2, && \text{(Definition 6 in Matrices, Definition 12 in Matrices)}
\end{aligned}$$

so

$$\begin{aligned}
\mathbf{P} \left(\frac{1}{2} \mathbf{P} \right) &= \frac{1}{2} \mathbf{P}^2 \\
&= \mathbf{I}_2
\end{aligned}$$

and likewise $\left(\frac{1}{2} \mathbf{P} \right) \mathbf{P} = \mathbf{I}_2$ (Theorem 7 in Projections and Linear Systems): $\mathbf{P}^{-1} = \frac{1}{2} \mathbf{P}$ (2). Then

$$\begin{aligned}
\mathbf{P}^{-1} \mathbf{A} \mathbf{P} &= \left(\frac{1}{2} \mathbf{P} \right) \mathbf{A} \mathbf{P} && \text{(substitute } \mathbf{P}^{-1} \text{)} \\
&= \frac{1}{2} (\mathbf{P} \mathbf{A} \mathbf{P}) && \text{(Theorem 7 in Projections and Linear Systems)} \\
&= \frac{1}{2} \mathbf{P} (\mathbf{A} \mathbf{P}) && \text{(Theorem 5 in Matrices)} \\
&= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 3 & -1 \end{bmatrix} && \text{(Definition 7 in Matrices, for } \mathbf{A} \mathbf{P} \text{)} \\
&= \frac{1}{2} \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix} && \text{(Definition 7 in Matrices)} \\
&= \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, && \text{(Definition 6 in Matrices)}
\end{aligned}$$

so $\mathbf{A}$ is similar to $\text{diag}(3, 1)$. At the other extreme, the identity matrix is similar only to itself:

$$\begin{aligned}
\mathbf{P}^{-1} \mathbf{I}_p \mathbf{P} &= \mathbf{P}^{-1} \mathbf{P} \\
&= \mathbf{I}_p
\end{aligned}$$

for every invertible $\mathbf{P}$ (Theorem 8 in Matrices, 2).

Theorem 0.6 (Similar matrices have the same eigenvalues). *If $\mathbf{B} = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}$ (Definition 0.3), then $\tilde{v}$ is an eigenvector of $\mathbf{A}$ for λ exactly when $\mathbf{P}^{-1} \tilde{v}$ is an eigenvector of $\mathbf{B}$ for λ . So $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues.*

i Proof

Proof. From A to B. Let $\mathbf{A}v = \lambda v$ with $v \neq 0$, and $w = \mathbf{P}^{-1}v$. Then

$$\begin{aligned}
 \mathbf{B}w &= (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})(\mathbf{P}^{-1}v) && \text{(substitute } \mathbf{B} \text{ and } w) \\
 &= \mathbf{P}^{-1}\mathbf{A}(\mathbf{P}\mathbf{P}^{-1})v && \text{(Theorem 5 in Matrices)} \\
 &= \mathbf{P}^{-1}\mathbf{A}\mathbf{I}_p v && \text{(Definition 5 in Special Matrices and Decompositions)} \\
 &= \mathbf{P}^{-1}\mathbf{A}v && \text{(Theorem 8 in Matrices)} \\
 &= \mathbf{P}^{-1}(\lambda v) && \text{(Theorem 5 in Matrices)} \\
 &= \mathbf{P}^{-1}(\lambda v) && \text{(eigenvector)} \\
 &= \lambda w. && \text{(Theorem 7 in Projections and Linear Systems)}
 \end{aligned}$$

And $w \neq 0$: if $w = 0$, then

$$\begin{aligned}
 v &= \mathbf{I}_p w \\
 &= \mathbf{P}(\mathbf{P}^{-1}w) \\
 &= \mathbf{P}0 \\
 &= 0
 \end{aligned}$$

(Theorem 8 in Matrices, 2, Theorem 5 in Matrices, Definition 13 in Matrices), which is false.

From B to A. By Remark 0.1,

$$\begin{aligned}
 \mathbf{A} &= \mathbf{P}\mathbf{B}\mathbf{P}^{-1} \\
 &= (\mathbf{P}^{-1})^{-1}\mathbf{B}\mathbf{P}^{-1},
 \end{aligned}$$

so the first part, with $\mathbf{P}^{-1}$ in place of $\mathbf{P}$ and the roles of $\mathbf{A}$ and $\mathbf{B}$ exchanged, shows that if $w = \mathbf{P}^{-1}v$ is an eigenvector of $\mathbf{B}$ for λ , then $\mathbf{P}w$ is an eigenvector of $\mathbf{A}$ for λ ; and

$$\begin{aligned}
 \mathbf{P}w &= \mathbf{P}\mathbf{P}^{-1}v \\
 &= v
 \end{aligned}$$

(Theorem 5 in Matrices, 2, Theorem 8 in Matrices). □

Example 0.9 (Eigenvectors carried across a similarity). In Example 0.8, $\mathbf{B} = \text{diag}(3, 1)$ has eigenvectors $(1, 0)$ for 3 and $(0, 1)$ for 1: $\text{diag}(3, 1)(1, 0) = (3, 0)$ and $\text{diag}(3, 1)(0, 1) = (0, 1)$. $\mathbf{P}(1, 0) = (1, 1)$ and $\mathbf{P}(0, 1) = (1, -1)$ are the eigenvectors of $\mathbf{A}$ for 3 and 1 found in Example 24 in Special Matrices and Decompositions.

Theorem 0.7 (Eigenvectors for distinct eigenvalues are linearly independent). *If $v_1, \dots, v_k$ are eigenvectors of a $p \times p$ matrix $\mathbf{A}$ for eigenvalues $\lambda_1, \dots, \lambda_k$ that are all different, then $v_1, \dots, v_k$ are linearly independent (Definition 1 in Subspaces and Rank).*

i Proof

Proof. By induction on k . For $k = 1$, $v_1 \neq 0$ is linearly independent on its own. Suppose the result holds for $k - 1$ eigenvectors, and $\sum_{i=1}^k c_i v_i = 0$. Then

$$\begin{aligned} 0 &= (\mathbf{A} - \lambda_k \mathbf{I}_p) \sum_{i=1}^k c_i v_i \quad (\text{multiply the supposed equation by } \mathbf{A} - \lambda_k \mathbf{I}_p) \\ &= \sum_{i=1}^k c_i (\mathbf{A} - \lambda_k \mathbf{I}_p) v_i \quad (\text{Theorem 4 in Subspaces and Rank}) \\ &= \sum_{i=1}^k c_i (\lambda_i - \lambda_k) v_i \quad (\text{Theorem 0.3, part 1, with } s = -\lambda_k) \\ &= \sum_{i=1}^{k-1} c_i (\lambda_i - \lambda_k) v_i. \quad (\text{the } i = k \text{ term is } 0) \end{aligned}$$

By the induction hypothesis $v_1, \dots, v_{k-1}$ are linearly independent, so every $c_i (\lambda_i - \lambda_k) = 0$ for $i < k$, and since $\lambda_i \neq \lambda_k$, every $c_i = 0$ for $i < k$. The supposed equation then reads $c_k v_k = 0$ with $v_k \neq 0$, so $c_k = 0$ too. $\square$

Example 0.10 (Independent eigenvectors). The eigenvectors $(1, 1)$ for 3 and $(1, -1)$ for 1 of Example 24 in Special Matrices and Decompositions are linearly independent, as Theorem 0.7 says they must be: $c_1(1, 1) + c_2(1, -1) = (c_1 + c_2, c_1 - c_2)$ is 0 only if

$$\begin{aligned} c_1 &= c_2 \\ &= 0. \end{aligned}$$

Two eigenvectors for the same eigenvalue need not be independent: $(1, 1)$ and $(2, 2)$ are both eigenvectors for 3.

Definition 0.4 (Diagonalizable matrix). A $p \times p$ matrix $\mathbf{A}$ is **diagonalizable** [4] if $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ for some invertible $p \times p$ matrix $\mathbf{X}$ and diagonal $p \times p$ matrix $\mathbf{\Lambda}$ (Definition 4 in Special Matrices and Decompositions); by Remark 0.1, that is the same as $\mathbf{\Lambda} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$, so a diagonalizable matrix is one similar to a diagonal matrix (Definition 0.3).

Example 0.11 (Diagonalizable and not).

- $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is diagonalizable: with $\mathbf{P}$ as in Example 0.8, $\text{diag}(3, 1) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.
- $\mathbf{J} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not. Suppose $\mathbf{J} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$, so $\mathbf{\Lambda} = \mathbf{X}^{-1}\mathbf{J}\mathbf{X}$ (Remark 0.1). The eigenvalues of $\mathbf{\Lambda}$ are its diagonal entries (Theorem 0.4, since a diagonal matrix is upper triangular) and are the eigenvalues of $\mathbf{J}$ (Theorem 0.6), of which 0 is the only one (Example 0.2). So $\mathbf{\Lambda} = \mathbf{0}_{2 \times 2}$, and then

$$\begin{aligned}\mathbf{J} &= \mathbf{X}\mathbf{0}_{2 \times 2}\mathbf{X}^{-1} \\ &= \mathbf{0}_{2 \times 2}\end{aligned}$$

(1), which it is not.

Theorem 0.8 (Diagonalizable means a basis of eigenvectors). *A $p \times p$ matrix $\mathbf{A}$ is diagonalizable (Definition 0.4) exactly when it has p linearly independent eigenvectors. Then $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ with the eigenvectors as the columns of $\mathbf{X}$ and their eigenvalues, in the same order, on the diagonal of $\mathbf{\Lambda}$.*

Proof

Proof. Let $\mathbf{X}$ have columns $\tilde{x}_1, \dots, \tilde{x}_p$ and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$. Column j of $\mathbf{A}\mathbf{X}$ is $\mathbf{A}\tilde{x}_j$, and column j of $\mathbf{X}\mathbf{\Lambda}$ is $\mathbf{X}$ times column j of $\mathbf{\Lambda}$ (1), which is

$$\begin{aligned}\mathbf{X}(\lambda_j \tilde{e}_j) &= \lambda_j (\mathbf{X}\tilde{e}_j) \\ &= \lambda_j \tilde{x}_j\end{aligned}$$

(Theorem 7 in Projections and Linear Systems, Theorem 3 in Subspaces and Rank), so

$$\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{\Lambda} \iff \mathbf{A}\tilde{x}_j = \lambda_j \tilde{x}_j \text{ for every } j. \quad (1)$$

Also, $\mathbf{X}$ has rank p exactly when its p columns are linearly independent (Definition 2 in Subspaces and Rank), which by Theorem 0.1 is exactly when $\mathbf{X}$ is invertible.

Eigenvectors give a diagonalization. If $\tilde{x}_1, \dots, \tilde{x}_p$ are linearly independent eigenvectors, then $\mathbf{X}$ is invertible and Equation 1 gives $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{\Lambda}$, so

$$\begin{aligned}\mathbf{A} &= \mathbf{A}\mathbf{I}_p && \text{(Theorem 8 in Matrices)} \\ &= \mathbf{A}(\mathbf{X}\mathbf{X}^{-1}) && \text{(Definition 5 in Special Matrices and Decompositions)} \\ &= (\mathbf{A}\mathbf{X})\mathbf{X}^{-1} && \text{(Theorem 5 in Matrices)} \\ &= \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}. && (\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{\Lambda})\end{aligned}$$

A diagonalization gives eigenvectors. If $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$, then

$$\begin{aligned}
\mathbf{A}\mathbf{X} &= (\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1})\mathbf{X} \quad (\text{substitute } \mathbf{A}) \\
&= \mathbf{X}\mathbf{\Lambda}(\mathbf{X}^{-1}\mathbf{X}) \quad (\text{Theorem 5 in Matrices}) \\
&= \mathbf{X}\mathbf{\Lambda}\mathbf{I}_p \quad (\text{Definition 5 in Special Matrices and Decompositions}) \\
&= \mathbf{X}\mathbf{\Lambda}, \quad (\text{Theorem 8 in Matrices})
\end{aligned}$$

so each column satisfies $\mathbf{A}\tilde{x}_j = \lambda_j\tilde{x}_j$ (Equation 1). $\mathbf{X}$ is invertible, so its columns are linearly independent, and in particular nonzero; they are p linearly independent eigenvectors. $\square$

Example 0.12 (Building the diagonalization from eigenvectors). The eigenvectors $(1, 1)$ and $(1, -1)$ of $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ are linearly independent (Example 0.10), so with $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $\mathbf{\Lambda} = \text{diag}(3, 1)$, $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$; by Remark 0.1 this is the similarity of Example 0.8, with $\mathbf{X} = \mathbf{P}$. Because $\mathbf{A}$ is symmetric, Theorem 12 in Special Matrices and Decompositions even gives such a factorization with orthonormal eigenvectors (Definition 18 in Special Matrices and Decompositions): dividing each column of $\mathbf{X}$ by its length $\sqrt{2}$ gives one. By contrast, every eigenvector of $\mathbf{J}$ in Example 0.2 has the form $(t, 0)$, so any two of them are multiples of each other, and $\mathbf{J}$ has no two linearly independent eigenvectors: another way to see that it is not diagonalizable.

Corollary 0.1 (Distinct eigenvalues make a matrix diagonalizable). *A $p \times p$ matrix with p different eigenvalues is diagonalizable.*

Proof

Proof. Choose one eigenvector for each eigenvalue. These p eigenvectors are linearly independent (Theorem 0.7), so the matrix is diagonalizable (Theorem 0.8). $\square$

Example 0.13 (A triangular matrix with distinct diagonal entries). $\mathbf{U}$ of Example 0.6 has the three different eigenvalues $2, \frac{1}{2}$ and -1 (Theorem 0.4), so it is diagonalizable, even though it is not symmetric, so Theorem 12 in Special Matrices and Decompositions does not apply to it. The converse fails: $\mathbf{I}_2$ is diagonal, hence diagonalizable ($\mathbf{I}_2 = \mathbf{I}_2 \mathbf{I}_2 \mathbf{I}_2^{-1}$), but has only the eigenvalue 1.

Theorem 0.9 (Powers of a diagonalizable matrix). *If $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ (Definition 0.4) with $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, then for every positive integer k , $\mathbf{A}^k = \mathbf{X}\mathbf{\Lambda}^k\mathbf{X}^{-1}$ and $\mathbf{\Lambda}^k = \text{diag}(\lambda_1^k, \dots, \lambda_p^k)$.*

Proof

Proof. Powers of $\mathbf{A}$. If $\mathbf{D} = \text{diag}(d_1, \dots, d_p)$ and $\mathbf{E} = \text{diag}(e_1, \dots, e_p)$, with entries d_{il} and e_{lj} (so $d_{ii} = d_i$ and $e_{jj} = e_j$), entry (i, j) of $\mathbf{DE}$ is $\sum_l d_{il} e_{lj}$ (1), where $d_{il} = 0$ unless $l = i$ and $e_{lj} = 0$ unless $l = j$. For $i \neq j$ no term survives, so the entry is 0; for $i = j$ only $l = i$ survives, giving $d_i e_i$. So $\mathbf{DE} = \text{diag}(d_1 e_1, \dots, d_p e_p)$, and [induction](#) on k gives $\mathbf{A}^k = \text{diag}(\lambda_1^k, \dots, \lambda_p^k)$.

Powers of $\mathbf{A}$, by induction on k . $k = 1$ is the assumption. If it holds for $k - 1$,

$$\begin{aligned}
 \mathbf{A}^k &= \mathbf{A}^{k-1} \mathbf{A} && \text{(Definition 2 in Special Matrices and Decompositions, Theorem 5 in Matrices)} \\
 &= \mathbf{X} \mathbf{A}^{k-1} \mathbf{X}^{-1} \mathbf{A} && \text{(induction hypothesis)} \\
 &= \mathbf{X} \mathbf{A}^{k-1} \mathbf{X}^{-1} \mathbf{X} \mathbf{A} \mathbf{X}^{-1} && \text{(the assumption)} \\
 &= \mathbf{X} \mathbf{A}^{k-1} (\mathbf{X}^{-1} \mathbf{X}) \mathbf{A} \mathbf{X}^{-1} && \text{(Theorem 5 in Matrices)} \\
 &= \mathbf{X} \mathbf{A}^{k-1} \mathbf{I}_p \mathbf{A} \mathbf{X}^{-1} && \text{(Definition 5 in Special Matrices and Decompositions)} \\
 &= \mathbf{X} \mathbf{A}^{k-1} \mathbf{A} \mathbf{X}^{-1} && \text{(Theorem 8 in Matrices)} \\
 &= \mathbf{X} \mathbf{A}^k \mathbf{X}^{-1}. && \text{(Definition 2 in Special Matrices and Decompositions)}
 \end{aligned}$$

□

Example 0.14 (A closed form for the powers). For $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, with $\mathbf{X}$ as in [Example 0.12](#) and $\mathbf{X}^{-1} = \frac{1}{2} \mathbf{X}$ ([Example 0.8](#)),

$$\begin{aligned}
 \mathbf{A}^k &= \mathbf{X} \text{diag}(3^k, 1) \mathbf{X}^{-1} && \text{(Theorem 0.9)} \\
 &= \mathbf{X} \text{diag}(3^k, 1) \left(\frac{1}{2} \mathbf{X}\right) && \text{(substitute } \mathbf{X}^{-1}) \\
 &= \frac{1}{2} \mathbf{X} \text{diag}(3^k, 1) \mathbf{X} && \text{(Theorem 7 in Projections and Linear Systems)} \\
 &= \frac{1}{2} \begin{bmatrix} 3^k & 1 \\ 3^k & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} && \text{(Definition 7 in Matrices, for } \mathbf{X} \text{diag}(3^k, 1)) \\
 &= \frac{1}{2} \begin{bmatrix} 3^k + 1 & 3^k - 1 \\ 3^k - 1 & 3^k + 1 \end{bmatrix}. && \text{(Definition 7 in Matrices)}
 \end{aligned}$$

For $k = 2$ this gives $\frac{1}{2} \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$, which matches $\mathbf{A}^2$ in [Example 0.5](#).

Gram matrices, Schur complements and Cholesky

i Source

This section is adapted from Zhou [5], used under the MIT License (see the license text in [Section 2 in Subspaces and Rank](#)). The source’s eigenvalue and quadratic-form characterizations of definiteness are covered in these notes already ([Theorem 17 in Special Matrices and Decompositions](#)). 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 0.11](#) 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 0.10 (Definite matrices are Gram matrices). *Let $\mathbf{A}$ be a $p \times p$ matrix.*

1. $\mathbf{A}$ is positive semidefinite ([Definition 21 in Special Matrices and Decompositions](#)) exactly when $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$ for some matrix $\mathbf{B}$ with p columns.
2. $\mathbf{A}$ is positive definite ([Definition 22 in Special Matrices and Decompositions](#)) exactly when $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$ for some matrix $\mathbf{B}$ with p linearly independent columns.

i 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 17 in Matrices](#), [Definition 3 in Matrices](#)). For any x ,

$$\begin{aligned}
\tilde{x}^\top \mathbf{A} \tilde{x} &= \tilde{x}^\top \mathbf{B}^\top \mathbf{B} \tilde{x} && \text{(substitute } \mathbf{A} \text{)} \\
&= (\mathbf{B} \tilde{x})^\top (\mathbf{B} \tilde{x}) && \text{(Theorem 17 in Matrices, Theorem 5 in Matrices)} \\
&= \|\mathbf{B} \tilde{x}\|^2 && \text{(Equation 2 in Vectors, squared)} \\
&\geq 0. && \text{(Theorem 1 in Inner Products and Orthogonality)}
\end{aligned}$$

If the columns of $\mathbf{B}$ are linearly independent, then $\mathcal{N}(\mathbf{B}) = \{0\}$ ($\mathbf{B}\tilde{x}$ is the combination of the columns with coefficients x_j , [Theorem 3 in Subspaces and Rank](#), and only the zero combination is 0, [Definition 1 in Subspaces and Rank](#)), so for $\tilde{x} \neq 0$, $\mathbf{B}\tilde{x} \neq 0$ and $\|\mathbf{B}\tilde{x}\|^2 > 0$ ([Theorem 1 in Inner Products and Orthogonality](#), part 1).

Only if. Let $\mathbf{A}$ be positive semidefinite. It is symmetric, so $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$ ([Theorem 12 in Special Matrices and Decompositions](#)), and every $\lambda_i \geq 0$ ([Theorem 17 in Special Matrices and Decompositions](#)). Let $\mathbf{\Lambda}^{1/2} \stackrel{\text{def}}{=} \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_p})$, 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 0.9](#)), and let $\mathbf{B} \stackrel{\text{def}}{=} \mathbf{\Lambda}^{1/2}\mathbf{Q}^\top$. Then

$$\begin{aligned}
\mathbf{B}^\top \mathbf{B} &= (\mathbf{Q}^\top)^\top (\mathbf{\Lambda}^{1/2})^\top \mathbf{\Lambda}^{1/2} \mathbf{Q}^\top && \text{(Theorem 17 in Matrices)} \\
&= \mathbf{Q} \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{1/2} \mathbf{Q}^\top && \text{(Definition 3 in Matrices; } \mathbf{\Lambda}^{1/2} \text{ is symmetric)} \\
&= \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^\top && (\mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{1/2} = \mathbf{\Lambda}) \\
&= \mathbf{A}. && \text{(Theorem 12 in Special Matrices and Decompositions)}
\end{aligned}$$

If $\mathbf{A}$ is positive definite, it is positive semidefinite ([Remark 19 in Special Matrices and Decompositions](#)), so the same $\mathbf{B}$ gives $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$; and every $\lambda_i > 0$ ([Theorem 17 in Special Matrices and Decompositions](#)), so $\mathbf{\Lambda}^{1/2}$ is diagonal with nonzero diagonal, and invertible (its inverse is $\text{diag}(1/\sqrt{\lambda_1}, \dots, 1/\sqrt{\lambda_p})$); $\mathbf{Q}^\top$ is invertible with inverse $\mathbf{Q}$ ([Remark 6 in Special Matrices and Decompositions](#)). So $\mathbf{B}$ is invertible ([Theorem 1 in Special Matrices and Decompositions](#)), and its p columns are linearly independent ([Theorem 0.1](#), [Definition 2 in Subspaces and Rank](#)). $\square$

Example 0.15 (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 30 in Special Matrices and Decompositions](#); 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 32 in Special Matrices and Decompositions](#)).
- 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 Projections and Linear Systems) is positive definite.

- $\mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ of Example 32 in Special Matrices and Decompositions is symmetric but not positive semidefinite, so by part 1 it is not $\mathbf{B}^\top \mathbf{B}$ for any $\mathbf{B}$.

Theorem 0.11 (Operations that preserve definiteness).

1. If $\mathbf{C}$ is a $p \times p$ positive definite matrix (Definition 22 in Special Matrices and Decompositions) 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:

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

(Theorem 17 in Matrices, twice; $\mathbf{C}$ symmetric; Definition 3 in Matrices). For $\tilde{x} \neq \tilde{0}_k$, $\mathbf{A}\tilde{x} \neq \tilde{0}_p$, since the columns of $\mathbf{A}$ are independent (Theorem 3 in Subspaces and Rank, Definition 1 in Subspaces and Rank), and

$$\begin{aligned} \tilde{x}^\top (\mathbf{A}^\top \mathbf{C} \mathbf{A}) \tilde{x} &= (\mathbf{A}\tilde{x})^\top \mathbf{C} (\mathbf{A}\tilde{x}) \quad (\text{Theorem 17 in Matrices, Theorem 5 in Matrices}) \\ &> 0. \quad (\mathbf{C} \text{ positive definite, } \mathbf{A}\tilde{x} \neq \tilde{0}) \end{aligned}$$

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 16 in Matrices, Definition 6 in Matrices, Definition 3 in Matrices). For $\tilde{x} \neq \tilde{0}$,

$$\begin{aligned} \tilde{x}^\top (\alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2) \tilde{x} &= \tilde{x}^\top (\alpha_1 \mathbf{A}_1) \tilde{x} + \tilde{x}^\top (\alpha_2 \mathbf{A}_2) \tilde{x} \quad (\text{Theorem 6 in Matrices, twice}) \\ &= \alpha_1 \tilde{x}^\top \mathbf{A}_1 \tilde{x} + \alpha_2 \tilde{x}^\top \mathbf{A}_2 \tilde{x} \quad (\text{Theorem 7 in Projections and Linear Systems}) \\ &> 0. \quad (\text{each term is positive}) \end{aligned}$$

□

Example 0.16 (Building positive definite matrices).

- With $\mathbf{C} = \mathbf{I}_3$ and $\mathbf{X}$ of Example 0.15, part 1 gives again that $\mathbf{X}^\top \mathbf{X}$ is positive definite.
- With $\alpha_1 = 2$, $\mathbf{A}_1 = \mathbf{I}_2$ (Example 32 in Special Matrices and Decompositions), $\alpha_2 = 1$ and

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

(Example 0.15), 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 $\tilde{x} \neq \tilde{0}$, so it is positive definite, in line with its eigenvalues 3 and 1 (Example 24 in Special Matrices and 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 0.12 (A positive definite matrix is invertible). *Every positive definite matrix (Definition 22 in Special Matrices and Decompositions) is invertible (Definition 6 in Special Matrices and Decompositions).*

i Proof

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

$$\begin{aligned}\tilde{x}^\top \mathbf{A} \tilde{x} &= \tilde{x}^\top (\mathbf{A} \tilde{x}) \quad (\text{Definition 11 in Special Matrices and Decompositions}) \\ &= \tilde{x}^\top \tilde{0}_p \quad (\mathbf{A} \tilde{x} = \tilde{0}_p) \\ &= 0. \quad (\text{every term of the inner product is 0})\end{aligned}$$

If $\tilde{x} \neq \tilde{0}_p$, then $\tilde{x}^\top \mathbf{A} \tilde{x} > 0$ because $\mathbf{A}$ is positive definite, which contradicts $\tilde{x}^\top \mathbf{A} \tilde{x} = 0$. So $\tilde{x} = \tilde{0}_p$, the null space $\mathcal{N}(\mathbf{A})$ is $\{\tilde{0}_p\}$, and $\mathbf{A}$ is invertible by Theorem 0.1. $\square$

Example 0.17 (Positive definite matrices are invertible; the converse fails). $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite, since

$$\begin{aligned} \tilde{x}^\top \mathbf{A} \tilde{x} &= 2x_1^2 + 2x_1x_2 + 2x_2^2 && \text{(multiply out)} \\ &= x_1^2 + x_2^2 + (x_1 + x_2)^2, && \text{(expand the square)} \end{aligned}$$

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 32 in Special Matrices and Decompositions](#) is not positive definite, yet it is invertible: $(\mathbf{D})^{-1} = \frac{1}{3} \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$.

Theorem 0.13 (Adding a positive multiple of the identity makes a matrix positive definite). *Let $\mathbf{A}$ be a $p \times p$ positive semidefinite matrix ([Definition 21 in Special Matrices and Decompositions](#)) and let $\alpha > 0$. Then $\mathbf{A} + \alpha \mathbf{I}_p$ is positive definite ([Definition 22 in Special Matrices and Decompositions](#)), so it is invertible ([Theorem 0.12](#)).*

i 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 16 in Matrices](#), [Definition 6 in Matrices](#), [Definition 3 in Matrices](#)). For $\tilde{x} \neq \mathbf{0}_p$,

$$\begin{aligned} \tilde{x}^\top (\mathbf{A} + \alpha \mathbf{I}_p) \tilde{x} &= \tilde{x}^\top \mathbf{A} \tilde{x} + \alpha \tilde{x}^\top \mathbf{I}_p \tilde{x} && \text{(\a href="#">Theorem 6 in Matrices, Theorem 7 in Projections and Linear Systems)} \\ &= \tilde{x}^\top \mathbf{A} \tilde{x} + \alpha \|\tilde{x}\|^2 && \text{(\a href="#">Theorem 8 in Matrices, Equation 2 in Vectors, squared)} \\ &> 0. && (\tilde{x}^\top \mathbf{A} \tilde{x} \geq 0, \alpha > 0, \tilde{x} \neq \mathbf{0}_p) \end{aligned}$$

□

Example 0.18 (Repairing a singular matrix). $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ of [Example 30 in Special Matrices and Decompositions](#) is positive semidefinite and singular: $\mathbf{B}(1, -1) = \mathbf{0}_2$. With $\alpha = 1$,

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

which is positive definite and invertible by [Example 0.17](#).

Corollary 0.2 ($\mathbf{X}^\top \mathbf{X}$ 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 and Decompositions](#)) exactly when $\text{rank}(\mathbf{X}) = p$ ([Definition 3 in Subspaces and Rank](#)).

i Proof

Proof. The following statements are equivalent:

$$\begin{aligned} \mathbf{X}^\top \mathbf{X} \text{ is invertible} &\iff \mathcal{N}(\mathbf{X}^\top \mathbf{X}) = \{\mathbf{0}_{\sim p}\} && \text{(Theorem 0.1)} \\ &\iff \mathcal{N}(\mathbf{X}) = \{\mathbf{0}_{\sim p}\} && \text{(Theorem 5 in Column Space, Null Space and Rank-Nullity)} \\ &\iff \mathbf{X} \text{ has full column rank.} && \text{(Theorem 3 in Column Space, Null Space and Rank-Nullity)} \end{aligned}$$

□

Example 0.19 (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 $\text{rank} \ 1 < 2$. 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 $\mathbf{X}(2, -1) = \mathbf{0}_{\sim 3}$ ([Example 5 in Column Space, Null Space and Rank-Nullity](#)). For the full-column-rank matrix of [Example 1 in Projections and Linear Systems](#), $\mathbf{X}^\top \mathbf{X}$ is invertible.

Definition 0.5 (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 $1 \leq k < m$ and $1 \leq l < n$. 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 $i \leq k$ and $j \leq l$, $\mathbf{B}$ ($k \times (n - l)$) those with $i \leq k$ and $j > l$, $\mathbf{C}$ ($(m - k) \times l$) those with $i > k$ and $j \leq l$, 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** [6].

Example 0.20 (Cutting a 3×3 matrix into blocks). With

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

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

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 0.6 (Leading block of a square matrix, top-left block). When $\mathbf{X}$ is $n \times n$ and $1 \leq k \leq n$, the $k \times k$ matrix of the entries x_{ij} with $i, j \leq k$ is the **leading** (or **top-left**) $k \times k$ **block** of $\mathbf{X}$; for $k < n$ it is the block $\mathbf{A}$ of Definition 0.5 with $l = k$, and for $k = n$ it is $\mathbf{X}$ itself.

Theorem 0.14 (Leading blocks of a positive definite matrix are positive definite). *Let $\mathbf{X}$ be a $p \times p$ positive definite matrix (Definition 22 in Special Matrices and Decompositions), and for $1 \leq k \leq p$ let $\mathbf{X}_k$ be its leading $k \times k$ block (Definition 0.5), with entries x_{ij} for $i, j \leq k$. Then $\mathbf{X}_k$ is positive definite. Also, every diagonal entry of a positive definite matrix is positive.*

i Proof

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

$$\begin{aligned}
\tilde{u}^\top \mathbf{X}_k \tilde{u} &= \sum_{i=1}^k \sum_{j=1}^k u_i x_{ij} u_j \quad (\text{Definition 13 in Matrices, Definition 8 in Vectors}) \\
&= \sum_{i=1}^k \sum_{j=1}^k y_i x_{ij} y_j \quad (y_i = u_i \text{ for } i \leq k) \\
&= \sum_{i=1}^p \sum_{j=1}^p y_i x_{ij} y_j \quad (\text{each added term has a factor } y_i = 0 \text{ or } y_j = 0) \\
&= \tilde{y}^\top \mathbf{X} \tilde{y} \quad (\text{Definition 13 in Matrices, Definition 8 in Vectors}) \\
&> 0. \quad (\mathbf{X} \text{ positive definite, } \tilde{y} \neq 0)
\end{aligned}$$

For the diagonal entry x_{ii} , take $\tilde{e}_i$ (Definition 18 in Vectors), which is nonzero:

$$\begin{aligned}
\tilde{e}_i^\top \mathbf{X} \tilde{e}_i &= \sum_{r=1}^p \sum_{s=1}^p (\tilde{e}_i)_r x_{rs} (\tilde{e}_i)_s \quad (\text{Definition 13 in Matrices, Definition 8 in Vectors}) \\
&= x_{ii} \quad (\text{only the term with } r \text{ and } s \text{ both equal to } i \text{ is nonzero}) \\
&> 0. \quad (\mathbf{X} \text{ positive definite})
\end{aligned}$$

□

Example 0.21 (A quick necessary test). $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite (Example 0.16), 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 26 in Special Matrices and Decompositions); 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 32 in Special Matrices and Decompositions has them but is not positive definite.

Theorem 0.15 (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 0.5)*

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

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 and Decompositions](#)):

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

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

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

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

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

i Proof

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

$$\begin{aligned} (\mathbf{X}v)_i &= \sum_{j=1}^{k+m} x_{ij} v_j && \text{(Definition 13 in Matrices)} \\ &= \sum_{j=1}^k x_{ij} v_j + \sum_{j=1}^m x_{i,k+j} v_{k+j} && \text{(split the sum at } j = k) \\ &= \sum_{j=1}^k a_{ij} u_j + \sum_{j=1}^m b_{ij} w_j && \text{(the blocks of } \mathbf{X} \text{ and of } v) \\ &= (\mathbf{A}u)_i + (\mathbf{B}w)_i, && \text{(Definition 13 in Matrices)} \end{aligned}$$

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

The expansion.

$$\begin{aligned} \underset{\sim}{v}^\top \mathbf{X} \underset{\sim}{v} &= \sum_{i=1}^k u_i (\mathbf{X}v)_i + \sum_{i=1}^m w_i (\mathbf{X}v)_{k+i} && \text{(Definition 8 in Vectors, split the sum at } k) \\ &= \underset{\sim}{u} \cdot (\mathbf{A}u + \mathbf{B}w) + \underset{\sim}{w} \cdot (\mathbf{B}^\top u + \mathbf{C}w) && \text{(the blocks of } \mathbf{X}v) \\ &= \underset{\sim}{u}^\top \mathbf{A} \underset{\sim}{u} + \underset{\sim}{u}^\top \mathbf{B} \underset{\sim}{w} + \underset{\sim}{w}^\top \mathbf{B}^\top u + \underset{\sim}{w}^\top \mathbf{C} \underset{\sim}{w} && \text{(Theorem 5 in Direct Sums and Orthogonal Complements)} \\ &= \underset{\sim}{u}^\top \mathbf{A} \underset{\sim}{u} + \underset{\sim}{u}^\top \mathbf{B} \underset{\sim}{w} + \underset{\sim}{u}^\top \mathbf{B} \underset{\sim}{w} + \underset{\sim}{w}^\top \mathbf{C} \underset{\sim}{w} && (w^\top \mathbf{B}^\top u = (\underset{\sim}{w}^\top \mathbf{B}^\top u)^\top, \text{ which equals } \underset{\sim}{u}^\top \mathbf{B} \underset{\sim}{w}; 1 \times 1, \text{ Definition 3}) \\ &= \underset{\sim}{u}^\top \mathbf{A} \underset{\sim}{u} + 2 \underset{\sim}{u}^\top \mathbf{B} \underset{\sim}{w} + \underset{\sim}{w}^\top \mathbf{C} \underset{\sim}{w}. && \text{(combine like terms)} \end{aligned}$$

Completing the square. Write $\underline{d} \stackrel{\text{def}}{=} \mathbf{A}^{-1}\mathbf{B}\underline{w}$, so $\underline{z} = \underline{u} + \underline{d}$ and $\mathbf{A}\underline{d} = \mathbf{B}\underline{w}$ (Theorem 5 in Matrices, 2, Theorem 8 in Matrices); $\mathbf{A}^{-1}$ is symmetric (Corollary 1 in Special Matrices and Decompositions). Then

$$\begin{aligned}
\underline{z}^\top \mathbf{A} \underline{z} &= (\underline{u} + \underline{d}) \cdot \mathbf{A} (\underline{u} + \underline{d}) && \text{(substitute } \underline{z} = \underline{u} + \underline{d} \text{)} \\
&= (\underline{u} + \underline{d}) \cdot (\mathbf{A}\underline{u} + \mathbf{A}\underline{d}) && \text{(Theorem 6 in Matrices)} \\
&= \underline{u}^\top \mathbf{A}\underline{u} + \underline{u}^\top \mathbf{A}\underline{d} + \underline{d}^\top \mathbf{A}\underline{u} + \underline{d}^\top \mathbf{A}\underline{d} && \text{(Theorem 5 in Direct Sums and Orthogonal Complements, both)} \\
&= \underline{u}^\top \mathbf{A}\underline{u} + \underline{u}^\top \mathbf{A}\underline{d} + \underline{u}^\top \mathbf{A}\underline{d} + \underline{d}^\top \mathbf{A}\underline{d} && \text{(as in the expansion, with } \mathbf{A}^\top = \mathbf{A} \text{)} \\
&= \underline{u}^\top \mathbf{A}\underline{u} + 2\underline{u}^\top \mathbf{A}\underline{d} + \underline{d}^\top \mathbf{A}\underline{d} && \text{(combine like terms)} \\
&= \underline{u}^\top \mathbf{A}\underline{u} + 2\underline{u}^\top \mathbf{B}\underline{w} + \underline{d}^\top \mathbf{B}\underline{w} && \text{(}\mathbf{A}\underline{d} = \mathbf{B}\underline{w}\text{, twice)} \\
&= \underline{u}^\top \mathbf{A}\underline{u} + 2\underline{u}^\top \mathbf{B}\underline{w} + \underline{w}^\top \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}\underline{w}, && (\underline{d}^\top = \underline{w}^\top \mathbf{B}^\top \mathbf{A}^{-1}, \text{ Theorem 17 in Matrices, Corollary 1 in Spec})
\end{aligned}$$

and

$$\begin{aligned}
\underline{w}^\top \mathbf{S}\underline{w} &= \underline{w}^\top (\mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}) \underline{w} && \text{(substitute } \mathbf{S} \text{)} \\
&= \underline{w}^\top (\mathbf{C}\underline{w} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}\underline{w}) && \text{(Theorem 7 in Projections and Linear Systems, difference rule on the right)} \\
&= \underline{w}^\top \mathbf{C}\underline{w} - \underline{w}^\top \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}\underline{w}. && \text{(Theorem 7 in Projections and Linear Systems, difference rule on the left)}
\end{aligned}$$

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

Example 0.22 (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

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

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

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

and $b = 4$:

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

Definition 0.7 (Schur complement). For a block matrix $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$ as in Theorem 0.15 with $\mathbf{A}$ invertible, the **Schur complement** [7] 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 0.15.

Example 0.23 (Schur complements of 2×2 matrices).

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

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

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 0.16 (Schur complement test). Let $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$ be as in Theorem 0.15. Then $\mathbf{X}$ is positive definite (Definition 22 in Special Matrices and Decompositions) exactly when $\mathbf{A}$ is positive definite and its Schur complement $\mathbf{S}$ (Definition 0.7) is positive definite.

i Proof

Proof. **If.** $\mathbf{A}$ is positive definite, so invertible (Theorem 18 in Special Matrices and Decompositions), and $\mathbf{S}$ is defined. $\mathbf{X}$ is symmetric. Let $\underline{v} = (\underline{u}, \underline{w}) \neq \underline{0}$ and $\underline{z} = \underline{u} + \mathbf{A}^{-1} \mathbf{B} \underline{w}$. By Theorem 0.15, $\underline{v}^\top \mathbf{X} \underline{v} = \underline{z}^\top \mathbf{A} \underline{z} + \underline{w}^\top \mathbf{S} \underline{w}$, and both terms are at least 0. If $\underline{w} \neq \underline{0}$, the second term is positive. If $\underline{w} = \underline{0}$, then $\underline{u} \neq \underline{0}$ and $\underline{z} = \underline{u}$, so the first term is positive. Either way $\underline{v}^\top \mathbf{X} \underline{v} > 0$.

Only if. $\mathbf{A}$ is the top-left $k \times k$ block of $\mathbf{X}$, so it is positive definite (Theorem 0.14), and invertible (Theorem 18 in Special Matrices and Decompositions). $\mathbf{S}$ is symmetric:

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

(Theorem 17 in Matrices, Definition 3 in Matrices for $(\mathbf{B}^\top)^\top = \mathbf{B}$, Corollary 1 in Special Matrices and Decompositions), 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 3 in Matrices). For $w \neq 0$, take $u = -\mathbf{A}^{-1} \mathbf{B} w$, so $z = 0$ and $v = (u, w) \neq 0$; then $w^\top \mathbf{S} w = v^\top \mathbf{X} v > 0$ (Theorem 0.15). $\square$

Example 0.24 (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 0.23), 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 32 in Special Matrices and Decompositions found directly.

Theorem 0.17 (Cholesky factorization). A $p \times p$ matrix $\mathbf{A}$ is positive definite (Definition 22 in Special Matrices and Decompositions) exactly when $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$ for some lower triangular $\mathbf{L}$ (Definition 8 in Projections and Linear Systems) with positive diagonal entries. For a positive definite $\mathbf{A}$ that $\mathbf{L}$ is unique; it is the **Cholesky factor** [8] of $\mathbf{A}$.

Proof

Proof. **If.** $\mathbf{L}^\top$ is upper triangular with positive diagonal (Definition 3 in Matrices), so $\mathbf{L}^\top x = 0$ only for $x = 0$ (Theorem 17 in Projections and Linear Systems), 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 3 in Matrices) is positive definite (Theorem 0.10, part 2).

Only if, with uniqueness, by induction on p . For $p = 1$, $\mathbf{A} = [a]$ with $a > 0$ (Theorem 0.14), 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:

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

with $\tilde{b}, \tilde{g} \in \mathbb{R}^{p-1}$, 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 3 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 $\tilde{\mathbf{L}}_{22}$ is lower triangular with positive diagonal; $\tilde{g}$ can be anything. So choosing $\mathbf{L}$ means choosing $\ell_{11} > 0$, $\tilde{g}$ and such an $\tilde{\mathbf{L}}_{22}$.

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 $\mathbf{L}$ (1, [Definition 3 in Matrices](#)). Row 1 of $\mathbf{L}$ is $(\ell_{11}, 0, \dots, 0)$, and row $i+1$ is $(g_i, (\mathbf{L}_{22})_{i1}, \dots, (\mathbf{L}_{22})_{i,p-1})$. So, for $i, j = 1, \dots, p-1$:

$$\begin{aligned} (\mathbf{L}\mathbf{L}^\top)_{11} &= \ell_{11} \ell_{11} + \sum_{r=2}^p 0 \cdot 0 \\ &= \ell_{11}^2, & (\text{rows 1 and 1}) \\ (\mathbf{L}\mathbf{L}^\top)_{i+1,1} &= g_i \ell_{11} + \sum_{r=2}^p (\mathbf{L}_{22})_{i,r-1} \cdot 0 \\ &= \ell_{11} g_i, & (\text{rows } i+1 \text{ and } 1) \\ (\mathbf{L}\mathbf{L}^\top)_{i+1,j+1} &= g_i g_j + \sum_{r=2}^p (\mathbf{L}_{22})_{i,r-1} (\mathbf{L}_{22})_{j,r-1} & (\text{rows } i+1 \text{ and } j+1) \\ &= (g\tilde{g}^\top)_{ij} + (\mathbf{L}_{22}\mathbf{L}_{22}^\top)_{ij}. & (\text{Definition 7 in Matrices, with } s = r-1) \end{aligned}$$

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 $\tilde{b}^\top$ exactly when the $(2, 1)$ block matches $\tilde{b}$, which is where the symmetry of $\mathbf{A}$ is used. So $\mathbf{L}\mathbf{L}^\top = \mathbf{A}$ says exactly

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

$a_{11} > 0$ (Theorem [0.14](#)), so the first equation with $\ell_{11} > 0$ forces $\ell_{11} = \sqrt{a_{11}}$; the second then forces $\tilde{g} = \frac{1}{\ell_{11}} \tilde{b}$; and the third becomes

$$\begin{aligned} \mathbf{L}_{22}\mathbf{L}_{22}^\top &= \mathbf{A}_{22} - \tilde{g}\tilde{g}^\top & (\text{subtract } \tilde{g}\tilde{g}^\top) \\ &= \mathbf{A}_{22} - \left(\frac{1}{\ell_{11}} \tilde{b}\right) \left(\frac{1}{\ell_{11}} \tilde{b}\right)^\top & (\text{substitute } \tilde{g}) \\ &= \mathbf{A}_{22} - \frac{1}{\ell_{11}^2} \tilde{b}\tilde{b}^\top & (\text{Definition 3 in Matrices, Theorem 7 in Projections and Linear Systems}) \\ &= \mathbf{A}_{22} - \frac{1}{a_{11}} \tilde{b}\tilde{b}^\top. & (\ell_{11}^2 = a_{11}) \end{aligned}$$

Here $[a_{11}]$ is invertible with inverse $[1/a_{11}]$, so with $\mathbf{B} = \tilde{b}^\top$ in Definition [0.7](#),

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

(Theorem 7 in Projections and Linear Systems), and the right side is the Schur complement of $[a_{11}]$ in $\mathbf{A}$. It is positive definite by Theorem 0.16 (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. $\square$

Example 0.25 (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

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

the Schur complement of Example 0.23, 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 0.24).

- [1] H. Zhou, “Eigen-decomposition.” 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/10-eig/10-eig.html>
- [2] Wikipedia contributors, “Eigenvalues and eigenvectors — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors
- [3] Wikipedia contributors, “Matrix similarity — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Matrix_similarity
- [4] Wikipedia contributors, “Diagonalizable matrix — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Diagonalizable_matrix
- [5] 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>
- [6] Wikipedia contributors, “Block matrix — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Block_matrix
- [7] Wikipedia contributors, “Schur complement — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Schur_complement
- [8] Wikipedia contributors, “Cholesky decomposition — Wikipedia, the free encyclopedia.” 2026. Available: https://en.wikipedia.org/wiki/Cholesky_decomposition