Linear Algebra 2

Lecture #25

Matrix definiteness

Irena Penev

May 21, 2025

Definition

A **symmetric** matrix $A \in \mathbb{R}^{n\times n}$ is said to be

- *positive definite* if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$;
- *positive semi-definite* if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;
- *negative definite* if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$;
- *negative semi-definite* if $\mathbf{x}^T A\mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;
- *indefinite* if it is neither positive semi-definite nor negative semi-definite.

- **Remark:** Obviously, any positive definite matrix is positive semi-definite, and any negative definite matrix if negative semi-definite.

- **Remark:** The definitions from the previous slide would also make sense without the requirement that A be symmetric.

- **Remark:** The definitions from the previous slide would also make sense without the requirement that A be symmetric.
  - However, for any matrix $A \in \mathbb{R}^{n \times n}$, the matrix $\frac{1}{2}(A + A^T)$ is symmetric:

$$\left(\tfrac{1}{2}(A + A^T)\right)^T \quad = \quad \tfrac{1}{2}(A^T + A) \quad = \quad \tfrac{1}{2}(A + A^T),$$

- **Remark:** The definitions from the previous slide would also make sense without the requirement that A be symmetric.
  - However, for any matrix $A \in \mathbb{R}^{n\times n}$, the matrix $\frac{1}{2}(A + A^T)$ is symmetric:

$$\left(\tfrac{1}{2}(A + A^T)\right)^T \quad = \quad \tfrac{1}{2}(A^T + A) \quad = \quad \tfrac{1}{2}(A + A^T),$$

and for all vectors $\mathbf{x} \in \mathbb{R}^n$, we have that

$$\begin{aligned} \mathbf{x}^T\left(\tfrac{1}{2}(A + A^T)\right)\mathbf{x} \quad &= \quad \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A^T\mathbf{x}) \\ &\overset{(*)}{=} \quad \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A^T\mathbf{x})^T \\ &= \quad \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) \\ &= \quad \mathbf{x}^T A\mathbf{x}, \end{aligned}$$

where (*) follows from the fact that $\mathbf{x}^T A\mathbf{x}$ is a 1×1 matrix, and is consequently symmetric.

- **Remark:** The definitions from the previous slide would also make sense without the requirement that A be symmetric.
  - However, for any matrix $A \in \mathbb{R}^{n \times n}$, the matrix $\frac{1}{2}(A + A^T)$ is symmetric:

    $$\left(\tfrac{1}{2}(A + A^T)\right)^T = \tfrac{1}{2}(A^T + A) = \tfrac{1}{2}(A + A^T),$$

    and for all vectors $\mathbf{x} \in \mathbb{R}^n$, we have that

    $$\begin{aligned} \mathbf{x}^T\left(\tfrac{1}{2}(A + A^T)\right)\mathbf{x} &= \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A^T\mathbf{x}) \\ &\overset{(*)}{=} \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A^T\mathbf{x})^T \\ &= \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) \\ &= \mathbf{x}^T A\mathbf{x}, \end{aligned}$$

    where (*) follows from the fact that $\mathbf{x}^T A\mathbf{x}$ is a 1×1 matrix, and is consequently symmetric.
  - So, instead of considering an arbitrary square matrix A, we can consider the symmetric matrix $\frac{1}{2}(A + A^T)$ instead.

- **Remark:** The definitions from the previous slide would also make sense without the requirement that A be symmetric.
  - However, for any matrix $A \in \mathbb{R}^{n \times n}$, the matrix $\frac{1}{2}(A + A^T)$ is symmetric:

$$\left(\tfrac{1}{2}(A + A^T)\right)^T \quad = \quad \tfrac{1}{2}(A^T + A) \quad = \quad \tfrac{1}{2}(A + A^T),$$

    and for all vectors $\mathbf{x} \in \mathbb{R}^n$, we have that

$$\begin{aligned} \mathbf{x}^T \left(\tfrac{1}{2}(A + A^T)\right)\mathbf{x} &= \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A^T\mathbf{x}) \\ &\overset{(*)}{=} \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A^T\mathbf{x})^T \\ &= \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) + \tfrac{1}{2}(\mathbf{x}^T A\mathbf{x}) \\ &= \mathbf{x}^T A\mathbf{x}, \end{aligned}$$

    where (*) follows from the fact that $\mathbf{x}^T A\mathbf{x}$ is a 1×1 matrix, and is consequently symmetric.
  - So, instead of considering an arbitrary square matrix A, we can consider the symmetric matrix $\frac{1}{2}(A + A^T)$ instead.
  - This is important because some tests of definiteness only work if we assume that the matrix in question is symmetric.

- **Remark:** Matrix definiteness plays an important role in optimization, though we will not cover this.

- **Remark:** Matrix definiteness plays an important role in optimization, though we will not cover this.
- Another reason for caring about positive definite matrices in particular is the following theorem.

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

(i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;

(ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

(iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

(i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;

(ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

(iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

- We start by proving Theorem 10.4.1 (plus an easy corollary).

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

(i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;

(ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

(iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

- We start by proving Theorem 10.4.1 (plus an easy corollary).
- After that, we prove a few results about matrix definiteness, and finally, we present three methods of testing whether a symmetric matrix is positive definite.

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;
3. (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

- We start by proving Theorem 10.4.1 (plus an easy corollary).
- After that, we prove a few results about matrix definiteness, and finally, we present three methods of testing whether a symmetric matrix is positive definite.
- Before proving Theorem 10.4.1, we recall a couple of definitions, plus Theorem 9.2.2 (from the previous lecture).

Definition

A *bilinear form* on a vector space V over a field $\mathbb{F}$ is a function $f : V \times V \to \mathbb{F}$ that satisfies the following four axioms:

b.1. $\forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$: $f(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = f(\mathbf{x}_1, \mathbf{y}) + f(\mathbf{x}_2, \mathbf{y})$;

b.2. $\forall \mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{F}$: $f(\alpha\mathbf{x}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$;

b.3. $\forall \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in V$: $f(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = f(\mathbf{x}, \mathbf{y}_1) + f(\mathbf{x}, \mathbf{y}_2)$;

b.4. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{F}$: $f(\mathbf{x}, \alpha\mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$.

The bilinear form f is said to be *symmetric* if it further satisfies the property that $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in V$.

Definition

A *scalar product* (also called *inner product*) in a **real** vector space V is a function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ that satisfies the following four axioms:

r.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

r.2. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;

r.3. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}$: $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$;

r.4. $\forall \mathbf{x}, \mathbf{y} \in V$: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.

r.2'. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$;

r.3'. $\forall \mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$, $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$.

Definition

A *scalar product* (also called *inner product*) in a **real** vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ that satisfies the following four axioms:

r.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

r.2. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;

r.3. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}$: $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$;

r.4. $\forall \mathbf{x}, \mathbf{y} \in V$: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.

r.2'. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$;

r.3'. $\forall \mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$, $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$.

- **Remark:** every scalar product $\langle \cdot, \cdot \rangle$ in a **real** vector space V is a symmetric bilinear form.
  - Indeed, r.2, r.3, r.2', and r.3' are precisely the axioms b.1, b.2, b.3, and b.4, respectively.
  - Moreover, by r.4, scalar products in real vector spaces are symmetric.

Theorem 9.2.2

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V.

- (a) For every matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, the function $f : V \times V \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V$$

  is a bilinear form on V, and moreover, all the following hold:
  - (a.1) $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$,
  - (a.2) $f\left(\sum_{i=1}^{n} c_i \mathbf{b}_i, \sum_{j=1}^{n} d_j \mathbf{b}_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} c_i d_j$ for all $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$,
  - (a.3) f is symmetric iff A is symmetric.
- (b) For every bilinear form f on V, there exists a unique matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, called the *matrix of the bilinear form f w.r.t. the basis $\mathcal{B}$*, that satisfies the property that

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

  Moreover, the entries of the matrix A are given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$.

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

(i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;

(ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

(iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof.

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;
3. (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof. It is enough to prove the following sequence of implications: "(i) $\implies$ (ii) $\implies$ (iii) $\implies$ (i)."

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle\cdot,\cdot\rangle$ be a bilinear form on V. Then the following are equivalent:

(i) $\langle\cdot,\cdot\rangle$ is a scalar product in V;

(ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle\cdot,\cdot\rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

(iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle\cdot,\cdot\rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof. It is enough to prove the following sequence of implications: "(i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i)." The implication "(ii) $\Longrightarrow$ (iii)" is obvious, and so in fact, we just need to prove the implications "(i) $\Longrightarrow$ (ii)" and "(iii) $\Longrightarrow$ (i)."

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). We first assume (i) and prove (ii).

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). We first assume (i) and prove (ii).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V, and let B be the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). We first assume (i) and prove (ii).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V, and let B be the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Since (i) holds, the bilinear form $\langle \cdot, \cdot \rangle$ is symmetric, and so by Theorem 9.2.2(a), the matrix B is also symmetric.

Theorem 10.4.1

1. $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). We first assume (i) and prove (ii).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V, and let B be the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Since (i) holds, the bilinear form $\langle \cdot, \cdot \rangle$ is symmetric, and so by Theorem 9.2.2(a), the matrix B is also symmetric.

Now, fix any non-zero vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{R}^n$; WTS $\mathbf{x}^T B\mathbf{x} > 0$.

Theorem 10.4.1

1. $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). We first assume (i) and prove (ii).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V, and let B be the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Since (i) holds, the bilinear form $\langle \cdot, \cdot \rangle$ is symmetric, and so by Theorem 9.2.2(a), the matrix B is also symmetric.

Now, fix any non-zero vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{R}^n$; WTS $\mathbf{x}^T B \mathbf{x} > 0$. Set $\mathbf{v} := x_1 \mathbf{b}_1 + \dots + x_n \mathbf{b}_n$, so that $\begin{bmatrix} \mathbf{v} \end{bmatrix}_{\mathcal{B}} = \mathbf{x}$.

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). We first assume (i) and prove (ii).

Fix any basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V, and let B be the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Since (i) holds, the bilinear form $\langle \cdot, \cdot \rangle$ is symmetric, and so by Theorem 9.2.2(a), the matrix B is also symmetric.

Now, fix any non-zero vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{R}^n$; WTS $\mathbf{x}^T B \mathbf{x} > 0$. Set $\mathbf{v} := x_1 \mathbf{b}_1 + \dots + x_n \mathbf{b}_n$, so that $\begin{bmatrix} \mathbf{v} \end{bmatrix}_{\mathcal{B}} = \mathbf{x}$. Since $\mathbf{x} \neq \mathbf{0}$, and since $\begin{bmatrix} \cdot \end{bmatrix}_{\mathcal{B}}$ is an isomorphism, we have that $\mathbf{v} \neq \mathbf{0}$.

Theorem 10.4.1

1. $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). Reminder: $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \mathbf{x}$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}^T B \mathbf{x} > 0$.

Theorem 10.4.1

1. $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). Reminder: $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\left[\ \mathbf{v}\ \right]_{\mathcal{B}} = \mathbf{x}$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}^T B \mathbf{x} > 0$.

Then

$$\mathbf{x}^T B \mathbf{x} \;=\; \left[\ \mathbf{v}\ \right]_{\mathcal{B}}^T B \left[\ \mathbf{v}\ \right]_{\mathcal{B}} \;\stackrel{(*)}{=}\; \langle \mathbf{v}, \mathbf{v} \rangle \;\stackrel{(**)}{>}\; 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$, and (**) follows from (i), and more precisely, from the axiom r.1.

Theorem 10.4.1

1. $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;

Proof (continued). Reminder: $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\left[\ \mathbf{v} \ \right]_{\mathcal{B}} = \mathbf{x}$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}^T B \mathbf{x} > 0$.

Then

$$\mathbf{x}^T B \mathbf{x} \quad = \quad \left[\ \mathbf{v} \ \right]_{\mathcal{B}}^T B \left[\ \mathbf{v} \ \right]_{\mathcal{B}} \quad \stackrel{(*)}{=} \quad \langle \mathbf{v}, \mathbf{v} \rangle \quad \stackrel{(**)}{>} \quad 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$, and (**) follows from (i), and more precisely, from the axiom r.1. This proves that B is positive definite, and (ii) follows.

Theorem 10.4.1

- (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
- (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). We now assume (iii) and prove (i).

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). We now assume (iii) and prove (i).

First of all, since $\langle \cdot, \cdot \rangle$ is a bilinear form, it satisfies axioms r.2 and r.3 from the definition of a scalar product;

Theorem 10.4.1

(i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;

(iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). We now assume (iii) and prove (i).

First of all, since $\langle \cdot, \cdot \rangle$ is a bilinear form, it satisfies axioms r.2 and r.3 from the definition of a scalar product; it remains to show that it satisfies axioms r.1 and r.4.

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). We now assume (iii) and prove (i).

First of all, since $\langle \cdot, \cdot \rangle$ is a bilinear form, it satisfies axioms r.2 and r.3 from the definition of a scalar product; it remains to show that it satisfies axioms r.1 and r.4.

Using (iii), we fix a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Theorem 10.4.1

- (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
- (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). We now assume (iii) and prove (i).

First of all, since $\langle \cdot, \cdot \rangle$ is a bilinear form, it satisfies axioms r.2 and r.3 from the definition of a scalar product; it remains to show that it satisfies axioms r.1 and r.4.

Using (iii), we fix a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite. Since B is positive definite, it is in particular symmetric, and so by Theorem 9.2.2(a), the bilinear form $\langle \cdot, \cdot \rangle$ is also symmetric, i.e. r.4 holds.

It remains to show that r.1 holds (next slide).

Theorem 10.4.1

- (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
- (ii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). First, we have that

$$\langle \mathbf{0}, \mathbf{0} \rangle \stackrel{(*)}{=} \left[\ \mathbf{0}\ \right]_{\mathcal{B}}^T B \left[\ \mathbf{0}\ \right]_{\mathcal{B}} = \mathbf{0}^T B \mathbf{0} = 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). First, we have that

$$\langle \mathbf{0}, \mathbf{0} \rangle \overset{(*)}{=} \left[\ \mathbf{0}\ \right]_{\mathcal{B}}^{T} B \left[\ \mathbf{0}\ \right]_{\mathcal{B}} = \mathbf{0}^{T} B \mathbf{0} = 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Now, fix any vector $\mathbf{x} \in V \setminus \{\mathbf{0}\}$. WTS $\langle \mathbf{x}, \mathbf{x} \rangle > 0$.

Theorem 10.4.1

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). First, we have that

$$\langle \mathbf{0}, \mathbf{0} \rangle \overset{(*)}{=} \left[\ \mathbf{0}\ \right]_{\mathcal{B}}^{T} B \left[\ \mathbf{0}\ \right]_{\mathcal{B}} = \mathbf{0}^T B \mathbf{0} = 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Now, fix any vector $\mathbf{x} \in V \setminus \{\mathbf{0}\}$. WTS $\langle \mathbf{x}, \mathbf{x} \rangle > 0$. Since $\left[\ \cdot\ \right]_{\mathcal{B}}$ is an isomorphism, we see that $\left[\ \mathbf{x}\ \right]_{\mathcal{B}} \neq \mathbf{0}$.

Theorem 10.4.1

1. $\langle\cdot,\cdot\rangle$ is a scalar product in V;
2. there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle\cdot,\cdot\rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). First, we have that

$$\langle \mathbf{0},\mathbf{0}\rangle \overset{(*)}{=} \left[\ \mathbf{0}\ \right]_{\mathcal{B}}^T B \left[\ \mathbf{0}\ \right]_{\mathcal{B}} = \mathbf{0}^T B\mathbf{0} = 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle\cdot,\cdot\rangle$ w.r.t. the basis $\mathcal{B}$.

Now, fix any vector $\mathbf{x} \in V \setminus \{\mathbf{0}\}$. WTS $\langle \mathbf{x},\mathbf{x}\rangle > 0$. Since $\left[\ \cdot\ \right]_{\mathcal{B}}$ is an isomorphism, we see that $\left[\ \mathbf{x}\ \right]_{\mathcal{B}} \neq \mathbf{0}$. We then have that

$$\langle \mathbf{x},\mathbf{x}\rangle \overset{(*)}{=} \left[\ \mathbf{x}\ \right]_{\mathcal{B}} B \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \overset{(**)}{>} 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle\cdot,\cdot\rangle$, and (**) follows from the fact that B is positive definite and $\left[\ \mathbf{x}\ \right]_{\mathcal{B}} \neq \mathbf{0}$.

Theorem 10.4.1

1. $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Proof (continued). First, we have that

$$\langle \mathbf{0}, \mathbf{0} \rangle \overset{(*)}{=} \left[\ \mathbf{0}\ \right]_{\mathcal{B}}^{T} B \left[\ \mathbf{0}\ \right]_{\mathcal{B}} = \mathbf{0}^T B \mathbf{0} = 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$.

Now, fix any vector $\mathbf{x} \in V \setminus \{\mathbf{0}\}$. WTS $\langle \mathbf{x}, \mathbf{x} \rangle > 0$. Since $\left[\ \cdot\ \right]_{\mathcal{B}}$ is an isomorphism, we see that $\left[\ \mathbf{x}\ \right]_{\mathcal{B}} \neq \mathbf{0}$. We then have that

$$\langle \mathbf{x}, \mathbf{x} \rangle \overset{(*)}{=} \left[\ \mathbf{x}\ \right]_{\mathcal{B}} B \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \overset{(**)}{>} 0,$$

where (*) follows from the fact that B is the matrix of the bilinear form $\langle \cdot, \cdot \rangle$, and (**) follows from the fact that B is positive definite and $\left[\ \mathbf{x}\ \right]_{\mathcal{B}} \neq \mathbf{0}$. Thus, r.1 holds. This proves (i). □

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;
3. (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Theorem 10.4.1

Let V be a non-trivial, finite-dimensional real vector space, and let $\langle \cdot, \cdot \rangle$ be a bilinear form on V. Then the following are equivalent:

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product in V;
2. (ii) for all bases $\mathcal{B}$ of V, the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite;
3. (iii) there exists a basis $\mathcal{B}$ of V s.t. the matrix B of the bilinear form $\langle \cdot, \cdot \rangle$ w.r.t. the basis $\mathcal{B}$ is positive definite.

Corollary 10.4.2

For any function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the following are equivalent:

1. (i) $\langle \cdot, \cdot \rangle$ is a scalar product on $\mathbb{R}^n$;
2. (ii) there exists a positive definite matrix $A \in \mathbb{R}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$.

- Proof: Lecture Notes (easily follows from Theorem 10.4.1).

- Let us now state some basic results about matrix definiteness.
  - We give some proofs, while omitting others.

- Let us now state some basic results about matrix definiteness.
  - We give some proofs, while omitting others.

Proposition 10.1.1

For every symmetric matrix $A \in \mathbb{R}^{n \times n}$, both the following hold:

(a) A is positive definite iff $-A$ is negative definite;

(b) A is positive semi-definite iff $-A$ is negative semi-definite.

- Proof: Lecture Notes (easy).

- Let us now state some basic results about matrix definiteness.
  - We give some proofs, while omitting others.

Proposition 10.1.1

For every symmetric matrix $A \in \mathbb{R}^{n \times n}$, both the following hold:

(a) A is positive definite iff $-A$ is negative definite;

(b) A is positive semi-definite iff $-A$ is negative semi-definite.

- Proof: Lecture Notes (easy).
- **Remark:** In view of Proposition 10.1.1, results for positive (semi-)definite matrices can easily be translated into corresponding results for negative (semi-)definite matrices.

- Let us now state some basic results about matrix definiteness.
    - We give some proofs, while omitting others.

Proposition 10.1.1

For every symmetric matrix $A \in \mathbb{R}^{n \times n}$, both the following hold:

(a) A is positive definite iff $-A$ is negative definite;

(b) A is positive semi-definite iff $-A$ is negative semi-definite.

- Proof: Lecture Notes (easy).
- **Remark:** In view of Proposition 10.1.1, results for positive (semi-)definite matrices can easily be translated into corresponding results for negative (semi-)definite matrices.
    - So, it makes sense to focus on positive (semi-)definite matrices.

- Let us now state some basic results about matrix definiteness.
    - We give some proofs, while omitting others.

Proposition 10.1.1

For every symmetric matrix $A \in \mathbb{R}^{n \times n}$, both the following hold:

(a) A is positive definite iff $-A$ is negative definite;

(b) A is positive semi-definite iff $-A$ is negative semi-definite.

- Proof: Lecture Notes (easy).
- **Remark:** In view of Proposition 10.1.1, results for positive (semi-)definite matrices can easily be translated into corresponding results for negative (semi-)definite matrices.
    - So, it makes sense to focus on positive (semi-)definite matrices.
    - In what follows, we will mostly (but not exclusively) focus on positive definite matrices, which are somewhat easier to deal with than the more general positive semi-definite ones.

- Reminder:

Corollary 8.7.4

Every symmetric matrix in $\mathbb{R}^{n\times n}$ has n real eigenvalues (with algebraic multiplicities taken into account). In other words, for every symmetric matrix $A \in \mathbb{R}^{n\times n}$, the sum of algebraic multiplicities of its distinct (real) eigenvalues is n.

- Reminder:

Corollary 8.7.4

Every symmetric matrix in $\mathbb{R}^{n\times n}$ has n real eigenvalues (with algebraic multiplicities taken into account). In other words, for every symmetric matrix $A \in \mathbb{R}^{n\times n}$, the sum of algebraic multiplicities of its distinct (real) eigenvalues is n.

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n\times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

(b) A is positive semi-definite iff $n_+ + n_0 = n$ (i.e. all eigenvalues of A are non-negative);

(c) A is negative definite iff $n_- = n$ (i.e. all eigenvalues of A are negative);

(d) A is negative semi-definite iff $n_- + n_0 = n$ (i.e. all eigenvalues of A are non-positive);

(e) A is indefinite iff n_+ and n_- are both non-zero (i.e. A has at least one positive and at least one negative eigenvalue).

Proof.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

(b) A is positive semi-definite iff $n_+ + n_0 = n$ (i.e. all eigenvalues of A are non-negative);

(c) A is negative definite iff $n_- = n$ (i.e. all eigenvalues of A are negative);

(d) A is negative semi-definite iff $n_- + n_0 = n$ (i.e. all eigenvalues of A are non-positive);

(e) A is indefinite iff n_+ and n_- are both non-zero (i.e. A has at least one positive and at least one negative eigenvalue).

Proof. Obviously, (b) and (d) together imply (e).

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);
(b) A is positive semi-definite iff $n_+ + n_0 = n$ (i.e. all eigenvalues of A are non-negative);
(c) A is negative definite iff $n_- = n$ (i.e. all eigenvalues of A are negative);
(d) A is negative semi-definite iff $n_- + n_0 = n$ (i.e. all eigenvalues of A are non-positive);
(e) A is indefinite iff n_+ and n_- are both non-zero (i.e. A has at least one positive and at least one negative eigenvalue).

Proof. Obviously, (b) and (d) together imply (e). So, we just need to prove (a)-(d).

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

(b) A is positive semi-definite iff $n_+ + n_0 = n$ (i.e. all eigenvalues of A are non-negative);

(c) A is negative definite iff $n_- = n$ (i.e. all eigenvalues of A are negative);

(d) A is negative semi-definite iff $n_- + n_0 = n$ (i.e. all eigenvalues of A are non-positive);

(e) A is indefinite iff n_+ and n_- are both non-zero (i.e. A has at least one positive and at least one negative eigenvalue).

Proof. Obviously, (b) and (d) together imply (e). So, we just need to prove (a)-(d). Here, we prove (a). The proofs of (b)-(d) are similar.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose first that A is positive definite.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose first that A is positive definite. Fix an eigenvalue λ of A, and let $\mathbf{x}$ be an associated eigenvector of A;

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose first that A is positive definite. Fix an eigenvalue λ of A, and let $\mathbf{x}$ be an associated eigenvector of A; after possibly normalizing the eigenvector $\mathbf{x}$ (i.e. replacing $\mathbf{x}$ by $\frac{\mathbf{x}}{||\mathbf{x}||}$), we may assume that $||\mathbf{x}|| = 1$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

- (a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose first that A is positive definite. Fix an eigenvalue λ of A, and let $\mathbf{x}$ be an associated eigenvector of A; after possibly normalizing the eigenvector $\mathbf{x}$ (i.e. replacing $\mathbf{x}$ by $\frac{\mathbf{x}}{||\mathbf{x}||}$), we may assume that $||\mathbf{x}|| = 1$. Then

$$0 \overset{(*)}{<} \mathbf{x}^T A\mathbf{x} \overset{(**)}{=} \mathbf{x}^T(\lambda\mathbf{x}) = \lambda(\mathbf{x}^T\mathbf{x}) = \lambda(\mathbf{x} \cdot \mathbf{x}) = \lambda||\mathbf{x}||^2 \overset{(***)}{=} \lambda,$$

where (*) follows from the fact that A is positive definite and $\mathbf{x} \neq \mathbf{0}$, (**) follows from the fact that $\mathbf{x}$ is an eigenvector of A associated with the eigenvalue λ, and (***) follows from the fact that $||\mathbf{x}|| = 1$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose first that A is positive definite. Fix an eigenvalue λ of A, and let $\mathbf{x}$ be an associated eigenvector of A; after possibly normalizing the eigenvector $\mathbf{x}$ (i.e. replacing $\mathbf{x}$ by $\frac{\mathbf{x}}{||\mathbf{x}||}$), we may assume that $||\mathbf{x}|| = 1$. Then

$$0 \overset{(*)}{<} \mathbf{x}^T A\mathbf{x} \overset{(**)}{=} \mathbf{x}^T(\lambda\mathbf{x}) = \lambda(\mathbf{x}^T\mathbf{x}) = \lambda(\mathbf{x} \cdot \mathbf{x}) = \lambda||\mathbf{x}||^2 \overset{(***)}{=} \lambda,$$

where (*) follows from the fact that A is positive definite and $\mathbf{x} \neq \mathbf{0}$, (**) follows from the fact that $\mathbf{x}$ is an eigenvector of A associated with the eigenvalue λ, and (***) follows from the fact that $||\mathbf{x}|| = 1$. We have now shown that $\lambda > 0$; since λ was an arbitrarily chosen eigenvalue of A, we deduce that all eigenvalues of A are positive, i.e. $n_+ = n$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Since A is symmetric, the spectral theorem for symmetric matrices guarantees that $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ associated with A.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Since A is symmetric, the spectral theorem for symmetric matrices guarantees that $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ associated with A. For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{x}_i$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Since A is symmetric, the spectral theorem for symmetric matrices guarantees that $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ associated with A. For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{x}_i$. Set $\lambda_0 := \min\{\lambda_1, \dots, \lambda_n\}$;

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Since A is symmetric, the spectral theorem for symmetric matrices guarantees that $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ associated with A. For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{x}_i$. Set $\lambda_0 := \min\{\lambda_1, \dots, \lambda_n\}$; since all eigenvalues of A are positive, we see that $\lambda_0 > 0$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Since A is symmetric, the spectral theorem for symmetric matrices guarantees that $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ associated with A. For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{x}_i$. Set $\lambda_0 := \min\{\lambda_1, \dots, \lambda_n\}$; since all eigenvalues of A are positive, we see that $\lambda_0 > 0$. Now, fix any $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}^T A\mathbf{x} > 0$. Since $\mathcal{B}$ is a basis of $\mathbb{R}^n$, we know that there exist scalars $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ s.t. $\mathbf{x} = \alpha_1\mathbf{x}_1 + \dots + \alpha_n\mathbf{x}_n$.

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

Proof (continued). Suppose conversely that $n_+ = n$, i.e. that all eigenvalues of A are positive. WTS A is positive definite.

Since A is symmetric, the spectral theorem for symmetric matrices guarantees that $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ associated with A. For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{x}_i$. Set $\lambda_0 := \min\{\lambda_1, \dots, \lambda_n\}$; since all eigenvalues of A are positive, we see that $\lambda_0 > 0$. Now, fix any $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}^T A \mathbf{x} > 0$. Since $\mathcal{B}$ is a basis of $\mathbb{R}^n$, we know that there exist scalars $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ s.t. $\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_n \mathbf{x}_n$. Since $\mathbf{x} \neq \mathbf{0}$, at least one of $\alpha_1, \dots, \alpha_n$ is non-zero. We now compute (next two slides):

Proof (continued). Reminder: WTS $\mathbf{x}^T A\mathbf{x} > 0$.

$$\begin{aligned}
\mathbf{x}^T A\mathbf{x} &= \Big(\sum_{i=1}^n \alpha_i \mathbf{x}_i\Big)^T A\Big(\sum_{j=1}^n \alpha_j \mathbf{x}_j\Big) \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mathbf{x}_i^T A \mathbf{x}_j \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mathbf{x}_i^T (\lambda_j \mathbf{x}_j) && \text{because each } \mathbf{x}_j \text{ is an eigenvector of } A \text{ associated with the eigenvalue } \lambda_j \\
&= \sum_{i=1}^n \sum_{j=1}^n \lambda_j \alpha_i \alpha_j (\mathbf{x}_i^T \mathbf{x}_j) \\
&= \sum_{i=1}^n \sum_{j=1}^n \lambda_j \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j) \\
&= \sum_{i=1}^n \lambda_i \alpha_i^2 (\mathbf{x}_i \cdot \mathbf{x}_i) && \text{because } \mathbf{x}_1, \dots, \mathbf{x}_n \text{ are pairwise orthogonal (by the orthonormality of } \mathcal{B}) \\
&= \sum_{i=1}^n \lambda_i \alpha_i^2 ||\mathbf{x}_i||^2 \\
&= \sum_{i=1}^n \lambda_i \alpha_i^2 && \text{because } \mathbf{x}_1, \dots, \mathbf{x}_n \text{ are unit vectors (by the orthonormality of } \mathcal{B})
\end{aligned}$$

(continued on next slide)

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

$$
\begin{aligned}
\mathbf{x}^T A \mathbf{x} &= \sum_{i=1}^{n} \lambda_i \alpha_i^2 && \text{from the previous slide} \\
&\geq \sum_{i=1}^{n} \lambda_0 \alpha_i^2 && \text{because } \lambda_0 = \min\{\lambda_1, \dots, \lambda_n\} \text{ and } \alpha_1^2, \dots, \alpha_n^2 \geq 0 \\
&> 0 && \text{because } \lambda_0 > 0 \text{ and at least one of } \alpha_1, \dots, \alpha_n \text{ is non-zero.}
\end{aligned}
$$

Thus, A is positive definite. This proves (a). $\square$

Theorem 10.1.2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and let (n_+, n_-, n_0) be the signature of A. Then all the following hold:

(a) A is positive definite iff $n_+ = n$ (i.e. all eigenvalues of A are positive);

(b) A is positive semi-definite iff $n_+ + n_0 = n$ (i.e. all eigenvalues of A are non-negative);

(c) A is negative definite iff $n_- = n$ (i.e. all eigenvalues of A are negative);

(d) A is negative semi-definite iff $n_- + n_0 = n$ (i.e. all eigenvalues of A are non-positive);

(e) A is indefinite iff n_+ and n_- are both non-zero (i.e. A has at least one positive and at least one negative eigenvalue).

- Reminder:

Theorem 8.2.10

Let $\mathbb{F}$ be a field, let $A = \left[\ a_{i,j}\ \right]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$, and assume that $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Then

(a) $\det(A) = \lambda_1 \dots \lambda_n$;

(b) $\operatorname{trace}(A) = \lambda_1 + \dots + \lambda_n$.

- Reminder:

Theorem 8.2.10

Let $\mathbb{F}$ be a field, let $A = \left[\ a_{i,j}\ \right]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$, and assume that $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Then

(a) $\det(A) = \lambda_1 \dots \lambda_n$;

(b) $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$.

- Theorem 10.1.2 (from the previous slide) and Theorem 8.2.10 together imply the following corollary.

Corollary 10.1.3

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\text{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\text{trace}(A)$ are both non-negative.

Corollary 10.1.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\operatorname{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\operatorname{trace}(A)$ are both non-negative.

Proof.

Corollary 10.1.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\operatorname{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\operatorname{trace}(A)$ are both non-negative.

Proof. Since A is symmetric, Corollary 8.7.4 guarantees that it has n real eigenvalues (with algebraic multiplicities taken into account).

Corollary 10.1.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\operatorname{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\operatorname{trace}(A)$ are both non-negative.

Proof. Since A is symmetric, Corollary 8.7.4 guarantees that it has n real eigenvalues (with algebraic multiplicities taken into account). So, let $\{\lambda_1, \ldots, \lambda_n\}$ be the spectrum of A.

Corollary 10.1.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\mathrm{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\mathrm{trace}(A)$ are both non-negative.

Proof. Since A is symmetric, Corollary 8.7.4 guarantees that it has n real eigenvalues (with algebraic multiplicities taken into account). So, let $\{\lambda_1, \dots, \lambda_n\}$ be the spectrum of A.

By Theorem 8.2.10, we have that $\det(A) = \lambda_1 \dots \lambda_n$ and $\mathrm{trace}(A) = \lambda_1 + \cdots + \lambda_n$.

Corollary 10.1.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\mathrm{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\mathrm{trace}(A)$ are both non-negative.

Proof. Since A is symmetric, Corollary 8.7.4 guarantees that it has n real eigenvalues (with algebraic multiplicities taken into account). So, let $\{\lambda_1, \dots, \lambda_n\}$ be the spectrum of A.

By Theorem 8.2.10, we have that $\det(A) = \lambda_1 \dots \lambda_n$ and $\mathrm{trace}(A) = \lambda_1 + \dots + \lambda_n$.

By Theorem 10.1.2(a), all eigenvalues of a positive definite matrix are positive, and it follows that (a) holds.

Corollary 10.1.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

(a) If A is positive definite, then $\det(A)$ and $\text{trace}(A)$ are both positive.

(b) If A is positive semi-definite, then $\det(A)$ and $\text{trace}(A)$ are both non-negative.

Proof. Since A is symmetric, Corollary 8.7.4 guarantees that it has n real eigenvalues (with algebraic multiplicities taken into account). So, let $\{\lambda_1, \dots, \lambda_n\}$ be the spectrum of A.

By Theorem 8.2.10, we have that $\det(A) = \lambda_1 \dots \lambda_n$ and $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$.

By Theorem 10.1.2(a), all eigenvalues of a positive definite matrix are positive, and it follows that (a) holds. Similarly, by Theorem 10.1.2(b), all eigenvalues of a positive semi-definite matrix are non-negative, and it follows that (b) holds. □

- The main diagonal of a square matrix $A \in \mathbb{R}^{n \times n}$ is *positive* (resp. *non-negative*, *negative*, *non-positive*) if all entries on the main diagonal of A are positive (resp. non-negative, negative, non-positive).

- The main diagonal of a square matrix $A \in \mathbb{R}^{n \times n}$ is *positive* (resp. *non-negative*, *negative*, *non-positive*) if all entries on the main diagonal of A are positive (resp. non-negative, negative, non-positive).

Proposition 10.1.4

The main diagonal of any positive definite (resp. positive semi-definite, negative definite, negative semi-definite) matrix is positive (resp. non-negative, negative, non-positive).

Proof.

- The main diagonal of a square matrix $A \in \mathbb{R}^{n \times n}$ is *positive* (resp. *non-negative*, *negative*, *non-positive*) if all entries on the main diagonal of A are positive (resp. non-negative, negative, non-positive).

Proposition 10.1.4

The main diagonal of any positive definite (resp. positive semi-definite, negative definite, negative semi-definite) matrix is positive (resp. non-negative, negative, non-positive).

Proof. Fix a matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{R}^{n \times n}$. Then for all indices $i \in \{1, \dots, n\}$, we have that $\mathbf{e}_i^T A \mathbf{e}_i = a_{i,i}$. The result now follows from the appropriate definitions.[1] $\square$

---

[1]Let us explain this in a bit more detail. Suppose that A is positive definite. Then for each $i \in \{1, \dots, n\}$, we have that $a_{i,i} = \mathbf{e}_i^T A \mathbf{e}_i > 0$, i.e. the main diagonal of A is positive. Similar remarks apply for the cases of positive semi-definiteness, negative definiteness, and negative semi-definiteness.

Proposition 10.1.5

Let $A, B \in \mathbb{R}^{n\times n}$ and $\alpha \in \mathbb{R}$. Then all the following hold:

(a) if A and B are both positive definite (resp. positive semi-definite, negative definite, negative semi-definite), then $A + B$ is positive definite (resp. positive semi-definite, negative definite, negative semi-definite);

(b) if A is positive definite (resp. positive semi-definite, negative definite, negative semi-definite) and $\alpha > 0$, then αA is positive definite (resp. positive semi-definite, negative definite, negative semi-definite);

(c) if A is positive definite (resp. positive semi-definite, negative definite, negative semi-definite) and $\alpha < 0$, then αA is negative definite (resp. negative semi-definite, positive definite, positive semi-definite);

(d) if $A \in \mathbb{R}^{n\times n}$ is positive definite (respectively: negative definite), then A is invertible and its inverse A^{-1} is positive definite (respectively: negative definite).

- Parts (a)-(c) are trivial.
- The proof of (d) is in the Lecture Notes.

- We present three tests of positive definiteness:

- We present three tests of positive definiteness:
  - the recursive test of positive definiteness (see Theorem 10.2.3);

- We present three tests of positive definiteness:
  - the recursive test of positive definiteness (see Theorem 10.2.3);
  - the Gaussian elimination test of positive definiteness (see Theorem 10.2.6);

- We present three tests of positive definiteness:
    - the recursive test of positive definiteness (see Theorem 10.2.3);
    - the Gaussian elimination test of positive definiteness (see Theorem 10.2.6);
    - Sylvester's criterion of positive definiteness (see Theorem 10.2.9.

- We present three tests of positive definiteness:
    - the recursive test of positive definiteness (see Theorem 10.2.3);
    - the Gaussian elimination test of positive definiteness (see Theorem 10.2.6);
    - Sylvester's criterion of positive definiteness (see Theorem 10.2.9.
- Of these three tests, the first is arguably the least convenient for computing (at least if we are computing by hand), but it is a key ingredient in the proof of correctness of the other two tests.

- We present three tests of positive definiteness:
  - the recursive test of positive definiteness (see Theorem 10.2.3);
  - the Gaussian elimination test of positive definiteness (see Theorem 10.2.6);
  - Sylvester's criterion of positive definiteness (see Theorem 10.2.9.
- Of these three tests, the first is arguably the least convenient for computing (at least if we are computing by hand), but it is a key ingredient in the proof of correctness of the other two tests.
- The proofs of these tests can be found in the Lecture Notes.

Theorem 10.2.3 [The recursive test of positive definiteness]

Let n be a positive integer, and let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$ (with $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$) be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. Then A is positive definite iff $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.

- Note that A is an $(n+1) \times (n+1)$ matrix, whereas $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is an $n \times n$ matrix. (This is why the test is called "recursive.")

Theorem 10.2.3 [The recursive test of positive definiteness]

Let n be a positive integer, and let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$ (with $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$) be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. Then A is positive definite iff $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.

- Note that A is an $(n+1) \times (n+1)$ matrix, whereas $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is an $n \times n$ matrix. (This is why the test is called "recursive.")
- The proof of Theorem 10.2.3 is in the Lecture Notes.
  - However, let's try to gain some intuition for where the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ came from.

Theorem 10.2.3 [The recursive test of positive definiteness]

Let n be a positive integer, and let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$ (with $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$) be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. Then A is positive definite iff $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.

- Note that A is an $(n+1)\times(n+1)$ matrix, whereas $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is an $n \times n$ matrix. (This is why the test is called "recursive.")
- The proof of Theorem 10.2.3 is in the Lecture Notes.
  - However, let's try to gain some intuition for where the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ came from.
- In what follows, for a matrix $A \in \mathbb{R}^{n \times n}$ $(n \geq 2)$ and indices $i, j \in \{1, \dots, n\}$, we will denote by $A_{i,j}$ the submatrix of A obtained by deleting the i-th row and j-th column of A.

Proposition 10.2.1

Let $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ ($n \geq 2$) be a symmetric matrix in $\mathbb{R}^{n\times n}$, assume that $a_{1,1} \neq 0$, and set $\mathbf{a} := \begin{bmatrix} a_{2,1} & \dots a_{n,1} \end{bmatrix}^T$, so that

$$A = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{a} & A_{1,1} \end{array}\right].$$

Let $\widetilde{A}$ be the matrix obtained from A by (sequentially or simultaneously) performing the following elementary row operations on A:

- $R_2 \to R_2 - \frac{a_{2,1}}{a_{1,1}} R_1$;
- $R_3 \to R_3 - \frac{a_{3,1}}{a_{1,1}} R_1$;

  $\vdots$
- $R_n \to R_n - \frac{a_{n,1}}{a_{1,1}} R_1$.

Then

$$\widetilde{A} = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T \end{array}\right].$$

- Schematically, Proposition 10.2.1 states the following:

$$\underbrace{\left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{a} & A_{1,1} \end{array}\right]}_{=A} \quad \overset{\substack{R_2 \to R_2 - \frac{a_{2,1}}{a_{1,1}} R_1 \\ R_3 \to R_3 - \frac{a_{3,1}}{a_{1,1}} R_1 \\ \vdots \\ R_n \to R_n - \frac{a_{n,1}}{a_{1,1}} R_1}}{\sim} \quad \underbrace{\left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T \end{array}\right]}_{=\widetilde{A}}$$

- Schematically, Proposition 10.2.1 states the following:

$$\underbrace{\left[\begin{array}{c:c} a_{1,1} & \mathbf{a}^T \\ \hdashline \mathbf{a} & A_{1,1} \end{array}\right]}_{=A} \overset{\substack{R_2 \to R_2 - \frac{a_{2,1}}{a_{1,1}} R_1 \\ R_3 \to R_3 - \frac{a_{3,1}}{a_{1,1}} R_1 \\ \vdots \\ R_n \to R_n - \frac{a_{n,1}}{a_{1,1}} R_1}}{\sim} \underbrace{\left[\begin{array}{c:c} a_{1,1} & \mathbf{a}^T \\ \hdashline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T \end{array}\right]}_{=\widetilde{A}}$$

Proposition 10.2.2

Let $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A \in \mathbb{R}^{n \times n}$. If A is symmetric, then $A - \alpha \mathbf{a}\mathbf{a}^T$ is also symmetric.

Theorem 10.2.3 [The recursive test of positive definiteness]

Let n be a positive integer, and let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$ (with $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$) be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. Then A is positive definite iff $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.

- **Remark:** If $\alpha \neq 0$, then Proposition 10.2.2 guarantees that the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is symmetric, and Proposition 10.2.1 guarantees that

$$\left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{0} & A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T \end{array}\right]$$

  is the matrix obtained from A by (sequentially or simultaneously) performing the elementary row operations of the form "$R_i \to R_i + \beta_i R_1$" (for $i \in \{2, \dots, n\}$), with the β_i's chosen so that, with the exception of the $1,1$-th entry, the leftmost column becomes zero.

Theorem 10.2.6 [The Gaussian elim. test of positive definiteness]

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Then the following algorithm correctly determines whether A is positive definite.

- **Step 0:** Set $A_1 := A$, and go to Step 1.
- For $j \in \{1, \ldots, n\}$, and assuming the matrix A_j has already been generated, we proceed as follows.

  Step j**:**

  - If the main diagonal of A_j is **not** positive, then the algorithm returns the answer that A is **not** positive definite and terminates.
  - If the main diagonal of A_j is positive and $j = n$, then the algorithm returns the answer that A is positive definite and terminates.
  - If the main diagonal of A_j is positive and $j \leq n - 1$, then for each index $i \in \{j + 1, \ldots, n\}$, we add a suitable scalar multiple of the j-th row of A_j to the i-th row of A_j so that the i, j-th entry of the matrix becomes zero; we call the resulting matrix A_{j+1}, and we go to Step $j + 1$.

- The proof of Theorem 10.2.6 (the Gaussian elimination test of positive definiteness) follows from Theorem 10.2.3 (the recursive test of positive definiteness) via an induction.
  - Essentially, the steps of the Gaussian elim. test keep generating ever smaller "bottom right corners" $(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)$.

Theorem 10.2.3 [The recursive test of positive definiteness]

Let n be a positive integer, and let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$ (with $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$) be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. Then A is positive definite iff $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.

$$\underbrace{\left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{a} & A_{1,1} \end{array}\right]}_{=A} \quad \overset{\substack{R_2 \to R_2 - \frac{a_{2,1}}{a_{1,1}} R_1 \\ R_3 \to R_3 - \frac{a_{3,1}}{a_{1,1}} R_1 \\ \vdots \\ R_n \to R_n - \frac{a_{n,1}}{a_{1,1}} R_1}}{\sim} \quad \underbrace{\left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T \end{array}\right]}_{\sim A}$$

- We now turn to Sylvester's criterion of positive definiteness.

- We now turn to Sylvester's criterion of positive definiteness.
- Given any $n \times n$ matrix A, and any index $k \in \{1, \ldots, n\}$, we let $A^{(k)}$ be the $k \times k$ matrix in the upper left corner of A.

- We now turn to Sylvester's criterion of positive definiteness.
- Given any $n \times n$ matrix A, and any index $k \in \{1, \ldots, n\}$, we let $A^{(k)}$ be the $k \times k$ matrix in the upper left corner of A.
- For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

then we have that

$$A^{(1)} = \begin{bmatrix} 1 \end{bmatrix}, \qquad A^{(2)} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \qquad A^{(3)} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

- We now turn to Sylvester's criterion of positive definiteness.
- Given any $n \times n$ matrix A, and any index $k \in \{1, \ldots, n\}$, we let $A^{(k)}$ be the $k \times k$ matrix in the upper left corner of A.
- For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix},$$

then we have that

$$A^{(1)} = \begin{bmatrix} 1 \end{bmatrix}, \qquad A^{(2)} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \qquad A^{(3)} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

- Clearly, for any $n \times n$ matrix A, we have that $A^{(n)} = A$.

Theorem 10.2.9 [Sylvester's criterion of positive definiteness]

For all symmetric matrices $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

1. (i) A is positive definite;
2. (ii) $\det(A^{(1)}), \ldots, \det(A^{(n)}) > 0$.

- Proof: Lecture Notes.

Theorem 10.2.9 [Sylvester's criterion of positive definiteness]

For all symmetric matrices $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

(i) A is positive definite;

(ii) $\det(A^{(1)}), \ldots, \det(A^{(n)}) > 0$.

- Proof: Lecture Notes.
- Sylvester's criterion of positive definiteness essentially follows from the recursive test of positive definiteness (Theorem 10.2.3) by induction on the size of the matrix, where we also use the following:

Theorem 10.2.9 [Sylvester's criterion of positive definiteness]

For all symmetric matrices $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

1. (i) A is positive definite;
2. (ii) $\det(A^{(1)}), \dots, \det(A^{(n)}) > 0$.

- Proof: Lecture Notes.
- Sylvester's criterion of positive definiteness essentially follows from the recursive test of positive definiteness (Theorem 10.2.3) by induction on the size of the matrix, where we also use the following:
  - the determinant of any positive definite matrix is positive (by Corollary 10.1.3(a));

Theorem 10.2.9 [Sylvester's criterion of positive definiteness]

For all symmetric matrices $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

(i) A is positive definite;

(ii) $\det(A^{(1)}), \ldots, \det(A^{(n)}) > 0$.

- Proof: Lecture Notes.
- Sylvester's criterion of positive definiteness essentially follows from the recursive test of positive definiteness (Theorem 10.2.3) by induction on the size of the matrix, where we also use the following:
  - the determinant of any positive definite matrix is positive (by Corollary 10.1.3(a));
  - adding a scalar multiple of one row of a square matrix to another ("$R_i \to R_i + \alpha R_j$") does not change the value of the determinant (by Theorem 7.3.2(c)).

Proposition 10.3.1

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

- Proof: Lecture Notes (easy).

Proposition 10.3.1

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

- Proof: Lecture Notes (easy).

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

- Proof: Next slide.

Proposition 10.3.1

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

- Proof: Lecture Notes (easy).

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

- Proof: Next slide.
- **Remark:** The main reason for interest in the Cholesky decomposition for positive definite matrices is that it allows us to solve equations of the form $A\mathbf{x} = \mathbf{b}$ (where A is positive definite) faster, as well as to compute the inverse of A faster. We omit the details.

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof.

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof. We proceed by induction on n.

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof. We proceed by induction on n.

For $n = 1$, we fix a positive definite matrix $A = \begin{bmatrix} a \end{bmatrix}$ in $\mathbb{R}^{1\times 1}$, and we note that $a > 0$ (because A is positive definite). We set $L := \begin{bmatrix} \sqrt{a} \end{bmatrix}$, and we observe that $A = LL^T$. The uniqueness of L is obvious.

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof. We proceed by induction on n.

For $n = 1$, we fix a positive definite matrix $A = \left[\; a \; \right]$ in $\mathbb{R}^{1 \times 1}$, and we note that $a > 0$ (because A is positive definite). We set $L := \left[\; \sqrt{a} \; \right]$, and we observe that $A = LL^T$. The uniqueness of L is obvious.

Now, fix a positive integer n, and assume the theorem is true for positive definite matrices in $\mathbb{R}^{n \times n}$. Fix a positive definite matrix $A \in \mathbb{R}^{(n+1) \times (n+1)}$, and set

$$A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array} \right],$$

where $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$.

Proof (continued). Reminder: $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$.

Proof (continued). Reminder: $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$.

By Theorem 10.2.3, we have that $\alpha > 0$ and that the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. By the induction hypothesis, there exists a unique lower triangular matrix $L' \in \mathbb{R}^{n\times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$. We now set

$$L \quad := \quad \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n\times n}.$$

Clearly, L is lower triangular with a positive main diagonal.

Proof (continued). Reminder: $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$.

By Theorem 10.2.3, we have that $\alpha > 0$ and that the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. By the induction hypothesis, there exists a unique lower triangular matrix $L' \in \mathbb{R}^{n\times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$. We now set

$$L \quad := \quad \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n\times n}.$$

Clearly, L is lower triangular with a positive main diagonal. Moreover, we have that

$$\begin{aligned} LL^T \quad &= \quad \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0}^T \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right] \left[\begin{array}{c|c} \sqrt{\alpha} & \frac{1}{\sqrt{\alpha}}\mathbf{a}^T \\ \hline \mathbf{0} & L'^T \end{array}\right] \\ &= \quad \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T \end{array}\right] \\ &= \quad \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right] \quad = \quad A. \end{aligned}$$

Proof (continued). We have now proven existence: $A = LL^T$.

Proof (continued). We have now proven existence: $A = LL^T$.

It remains to show that L is unique.

Proof (continued). We have now proven existence: $A = LL^T$.

It remains to show that L is unique. So far, our set-up is the following:

- $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$,
- $L = \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n\times n}$,

where L' is the unique lower triangular matrix $L' \in \mathbb{R}^{n\times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$ (equivalently: $\frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T = A'$).

Proof (continued). We have now proven existence: $A = LL^T$.

It remains to show that L is unique. So far, our set-up is the following:

- $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$,
- $L = \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n\times n}$,

where L' is the unique lower triangular matrix $L' \in \mathbb{R}^{n\times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$ (equivalently: $\frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T = A'$).

Suppose that $L_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is a lower triangular matrix with a positive main diagonal and satisfying $A = L_1L_1^T$; WTS $L_1 = L$.

Proof (continued). We have now proven existence: $A = LL^T$.

It remains to show that L is unique. So far, our set-up is the following:

- $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$,
- $L = \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n\times n}$,

where L' is the unique lower triangular matrix $L' \in \mathbb{R}^{n\times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$ (equivalently: $\frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T = A'$).

Suppose that $L_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is a lower triangular matrix with a positive main diagonal and satisfying $A = L_1 L_1^T$; WTS $L_1 = L$. Set

$$L_1 = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right],$$

where β is some positive real number, $\mathbf{b}$ is some vector in $\mathbb{R}^n$, and L_1' is some lower triangular matrix in $\mathbb{R}^{n\times n}$ with a positive main diagonal.

Proof (continued). We have now proven existence: $A = LL^T$.

It remains to show that L is unique. So far, our set-up is the following:

- $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]_{(n+1)\times(n+1)}$,
- $L = \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n\times n}$,

where L' is the unique lower triangular matrix $L' \in \mathbb{R}^{n\times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$ (equivalently: $\frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T = A'$).

Suppose that $L_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is a lower triangular matrix with a positive main diagonal and satisfying $A = L_1L_1^T$; WTS $L_1 = L$. Set

$$L_1 = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right],$$

where β is some positive real number, $\mathbf{b}$ is some vector in $\mathbb{R}^n$, and L_1' is some lower triangular matrix in $\mathbb{R}^{n\times n}$ with a positive main diagonal. Then

$$A = L_1L_1^T = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right]\left[\begin{array}{c|c} \beta & \mathbf{b}^T \\ \hline \mathbf{0} & L_1'^T \end{array}\right] = \left[\begin{array}{c|c} \beta^2 & \beta\mathbf{b}^T \\ \hline \beta\mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1'L_1'^T \end{array}\right].$$

Proof (continued). We now have that

$$\left[\begin{array}{c|c} \beta^2 & \beta \mathbf{b}^T \\ \hline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

Proof (continued). We now have that

$$\left[\begin{array}{c:c} \beta^2 & \beta \mathbf{b}^T \\ \hdashline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c:c} \alpha & \mathbf{a}^T \\ \hdashline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta \mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$.

Proof (continued). We now have that

$$\left[\begin{array}{c|c} \beta^2 & \beta \mathbf{b}^T \\ \hline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta \mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$. This, together with the fact that $\beta > 0$, yields the fact that $\beta = \sqrt{\alpha}$, $\mathbf{b} = \frac{1}{\sqrt{\alpha}}\mathbf{a}$, and $L_1' L_1'^T = L' L'^T$.

Proof (continued). We now have that

$$\left[\begin{array}{c:c} \beta^2 & \beta \mathbf{b}^T \\ \hdashline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c:c} \alpha & \mathbf{a}^T \\ \hdashline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta \mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$. This, together with the fact that $\beta > 0$, yields the fact that $\beta = \sqrt{\alpha}$, $\mathbf{b} = \frac{1}{\sqrt{\alpha}}\mathbf{a}$, and $L_1' L_1'^T = L' L'^T$.

Moreover, by the uniqueness of L', we have that $L_1' = L'$.

- Indeed, L' is the unique lower triangular matrix in $\mathbb{R}^{n \times n}$ with a positive main diagonal s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L' L'^T$. Since L_1' is a lower triangular matrix in $\mathbb{R}^{n \times n}$ with a positive main diagonal s.t. $L_1' L_1'^T = L' L'^T = A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$, the uniqueness of L' guarantees that $L_1' = L'$.

Proof (continued). We now have that

$$\left[\begin{array}{c|c} \beta^2 & \beta\mathbf{b}^T \\ \hline \beta\mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1'L_1'^T \end{array}\right] = A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta\mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1'L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T$. This, together with the fact that $\beta > 0$, yields the fact that $\beta = \sqrt{\alpha}$, $\mathbf{b} = \frac{1}{\sqrt{\alpha}}\mathbf{a}$, and $L_1'L_1'^T = L'L'^T$.

Moreover, by the uniqueness of L', we have that $L_1' = L'$.

- Indeed, L' is the unique lower triangular matrix in $\mathbb{R}^{n\times n}$ with a positive main diagonal s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$. Since L_1' is a lower triangular matrix in $\mathbb{R}^{n\times n}$ with a positive main diagonal s.t. $L_1'L_1'^T = L'L'^T = A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$, the uniqueness of L' guarantees that $L_1' = L'$.

Thus,

$$L_1 = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right] = \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right] = L.$$

This proves the uniqueness of L. □

Theorem 10.3.2 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

- There is also an algorithm that, for a positive definite matrix $A = \left[\ a_{i,j}\ \right]_{n\times n}$ in $\mathbb{R}^{n\times n}$, computes the Cholesky decomposition of A, i.e. computes the (unique) lower triangular matrix $L = \left[\ \ell_{i,j}\ \right]_{n\times n}$ in $\mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.
- We construct the matrix L column by column, from left to right. Each column is constructed from top to bottom.
  - Algorithm: next slide.

1. We construct the first (i.e. leftmost) column of L as follows:
    - $\ell_{1,1} := \sqrt{a_{1,1}}$,
    - $\ell_{i,1} := \frac{a_{i,1}}{\sqrt{a_{1,1}}}$ for all $i \in \{2, \ldots, n\}$.
2. For all $j \in \{2, \ldots, n\}$, assuming we have constructed the first (i.e. leftmost) $j-1$ columns of L, we construct the j-th column of L as follows (from top to bottom):
    - $\ell_{i,j} := 0$ for all $i \in \{1, \ldots, j-1\}$,
    - $\ell_{j,j} := \sqrt{a_{j,j} - \sum_{k=1}^{j-1} \ell_{j,k}^2}$,
    - $\ell_{i,j} := \frac{1}{\ell_{j,j}}\left(a_{i,j} - \sum_{k=1}^{j-1} \ell_{i,k}\ell_{j,k}\right)$ for all $i \in \{j+1, \ldots, n\}$.

1. We construct the first (i.e. leftmost) column of L as follows:
   - $\ell_{1,1} := \sqrt{a_{1,1}}$,
   - $\ell_{i,1} := \frac{a_{i,1}}{\sqrt{a_{1,1}}}$ for all $i \in \{2, \dots, n\}$.
2. For all $j \in \{2, \dots, n\}$, assuming we have constructed the first (i.e. leftmost) $j-1$ columns of L, we construct the j-th column of L as follows (from top to bottom):
   - $\ell_{i,j} := 0$ for all $i \in \{1, \dots, j-1\}$,
   - $\ell_{j,j} := \sqrt{a_{j,j} - \sum_{k=1}^{j-1} \ell_{j,k}^2}$,
   - $\ell_{i,j} := \frac{1}{\ell_{j,j}}\left(a_{i,j} - \sum_{k=1}^{j-1} \ell_{i,k}\ell_{j,k}\right)$ for all $i \in \{j+1, \dots, n\}$.

- We omit the proof of correctness of the construction above, but it essentially follows from Theorem 10.2.3 and from the proof of Theorem 10.3.2.
- Numerical example: Lecture Notes.