Linear Algebra 2

Lecture #25

Matrix definiteness

Irena Penev

May 22, 2024

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:

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

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.
  - 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:

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

    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.
  - 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:

(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.
- 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:

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.

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. $\langle \cdot, \cdot \rangle$ 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).

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 $\mathcal{B}$.

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.

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

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

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

(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;

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

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

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

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. $\langle \cdot, \cdot \rangle$ 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 $\langle \mathbf{x}, \mathbf{x} \rangle > 0$.

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

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

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

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

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

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

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

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:

(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:

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 prove some basic results about matrix definiteness!

- Let us now prove some basic results about matrix definiteness!

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.

- Let us now prove some basic results about matrix definiteness!

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. Fix a symmetric matrix $A \in \mathbb{R}^{n \times n}$. For (a), we have the following sequence of equivalent statements:

$$\begin{aligned} A \text{ is positive definite} \quad & \Longleftrightarrow \quad \mathbf{x}^T A\mathbf{x} > 0 \ \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \\ & \Longleftrightarrow \quad -\mathbf{x}^T A\mathbf{x} < 0 \ \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \\ & \Longleftrightarrow \quad \mathbf{x}^T(-A)\mathbf{x} < 0 \ \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \\ & \Longleftrightarrow \quad -A \text{ is negative definite.} \end{aligned}$$

This proves (a). The proof of (b) is very similar. □

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.

- **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.

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.

- **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.

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.

- **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

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

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);

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

Thus, A is positive definite. This proves (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);

(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 $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n \times n}$ 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) $\mathrm{trace}(A) = \lambda_1 + \dots + \lambda_n$.

- 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 $\mathrm{trace}(A)$ are both positive.

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

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

Proof.

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

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.

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

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. $\square$

- 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 $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n \times n}$ 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 important because we will rely on it to prove the correctness of the other two tests.

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.

- Proof: Later! (We first prove a couple of technical propositions).
- 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.

- Proof: Later! (We first prove a couple of technical propositions).
- 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.")

- 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:

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

Proof of Proposition 10.2.1.

- Schematically, Proposition 10.2.1 states the following:

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

Proof of Proposition 10.2.1. Set $A_{1,1} = \begin{bmatrix} \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_n^T \end{bmatrix}$.

- Schematically, Proposition 10.2.1 states the following:

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

Proof of Proposition 10.2.1. Set $A_{1,1} = \begin{bmatrix} \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_n^T \end{bmatrix}$. Since none of the elementary row operations modified the first row of A, we see that the first row of $\widetilde{A}$ is $\left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \end{array}\right]$.

- Schematically, Proposition 10.2.1 states the following:

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

Proof of Proposition 10.2.1. Set $A_{1,1} = \begin{bmatrix} \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_n^T \end{bmatrix}$. Since none of the elementary row operations modified the first row of A, we see that the first row of $\widetilde{A}$ is $\left[\begin{array}{c:c} a_{1,1} & \mathbf{a}^T \end{array}\right]$. On the other hand, for each $i \in \{2, \dots, n\}$, the i-th row of $\widetilde{A}$ is $\left[\begin{array}{c:c} 0 & \mathbf{r}_i^T - \frac{a_{i,1}}{a_{1,1}} \mathbf{a}^T \end{array}\right]$,

- Schematically, Proposition 10.2.1 states the following:

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

Proof of Proposition 10.2.1. Set $A_{1,1} = \begin{bmatrix} \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_n^T \end{bmatrix}$. Since none of the elementary row operations modified the first row of A, we see that the first row of $\widetilde{A}$ is $\left[\begin{array}{c:c} a_{1,1} & \mathbf{a}^T \end{array}\right]$. On the other hand, for each $i \in \{2, \dots, n\}$, the i-th row of $\widetilde{A}$ is $\left[\begin{array}{c:c} 0 & \mathbf{r}_i^T - \frac{a_{i,1}}{a_{1,1}}\mathbf{a}^T \end{array}\right]$, whereas the $(i-1)$-th row of the $(n-1) \times (n-1)$ matrix $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ is $\mathbf{r}_i^T - \frac{a_{i,1}}{a_{1,1}}\mathbf{a}^T$. The result is now immediate. □

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.

Proof.

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.

Proof. Assume that A is symmetric. Then

$$\left(A - \alpha \mathbf{a}\mathbf{a}^T\right)^T \;=\; A^T - \alpha(\mathbf{a}^T)^T\mathbf{a}^T \;\overset{(*)}{=}\; A - \alpha \mathbf{a}\mathbf{a}^T,$$

where in (*), we used the fact that A is symmetric and so $A^T = A$. This proves that $A - \alpha \mathbf{a}\mathbf{a}^T$ is indeed 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 10.2.3 [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.

Proof.

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.

Proof. Suppose first that A is positive definite. By Proposition 10.1.4, we have that $\alpha > 0$, and in particular, the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is defined (i.e. we are not dividing by zero). We must show that this matrix is positive definite. First of all, by Proposition 10.2.2, the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is symmetric. Now, fix any $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then (next slide):

Proof (continued).

$$\begin{aligned}
\mathbf{x}^T(A' - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{x} &= \mathbf{x}^T A'\mathbf{x} - \tfrac{1}{\alpha}(\mathbf{x}^T\mathbf{a}\mathbf{a}^T\mathbf{x}) \\
&\overset{(*)}{=} \left[\begin{array}{c|c} -\tfrac{1}{\alpha}\mathbf{a}^T\mathbf{x} & \mathbf{x}^T \end{array}\right] \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right] \underbrace{\left[\begin{array}{c} -\tfrac{1}{\alpha}\mathbf{a}^T\mathbf{x} \\ \hline \mathbf{x} \end{array}\right]}_{:=\mathbf{y}} \\
&= \mathbf{y}^T A\mathbf{y} \overset{(**)}{>} 0
\end{aligned}$$

where (**) follows from the fact that A is positive definite and $\mathbf{y} \neq \mathbf{0}$ (because $\mathbf{x} \neq \mathbf{0}$), and (*) follows from the following computation:

$$\begin{aligned}
\left[\begin{array}{c|c} -\tfrac{1}{\alpha}\mathbf{a}^T\mathbf{x} & \mathbf{x}^T \end{array}\right] \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right] \left[\begin{array}{c} -\tfrac{1}{\alpha}\mathbf{a}^T\mathbf{x} \\ \hline \mathbf{x} \end{array}\right] &= \left[\begin{array}{c|c} -\tfrac{1}{\alpha}\mathbf{a}^T\mathbf{x} & \mathbf{x}^T \end{array}\right] \left[\begin{array}{c} -\mathbf{a}^T\mathbf{x} + \mathbf{a}^T\mathbf{x} \\ \hline -\tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T\mathbf{x} + A'\mathbf{x} \end{array}\right] \\
&= \left[\begin{array}{c|c} -\tfrac{1}{\alpha}\mathbf{a}^T\mathbf{x} & \mathbf{x}^T \end{array}\right] \left[\begin{array}{c} 0 \\ \hline A'\mathbf{x} - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T\mathbf{x} \end{array}\right] \\
&= \mathbf{x}^T A'\mathbf{x} - \tfrac{1}{\alpha}\mathbf{x}^T\mathbf{a}\mathbf{a}^T\mathbf{x}.
\end{aligned}$$

This proves that $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is indeed positive definite.

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.

Proof (continued). Suppose conversely that $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. WTS A is positive definite.

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.

Proof (continued). Suppose conversely that $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. WTS A is positive definite.

By hypothesis, A is 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.

Proof (continued). Suppose conversely that $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. WTS A is positive definite.

By hypothesis, A is symmetric. Now, fix any $\mathbf{x} \in \mathbb{R}^{n+1}$; WTS

$\mathbf{x}^T A\mathbf{x} \geq 0$, and "$=$" holds iff $\mathbf{x} = \mathbf{0}$.

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.

Proof (continued). Suppose conversely that $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. WTS A is positive definite.

By hypothesis, A is symmetric. Now, fix any $\mathbf{x} \in \mathbb{R}^{n+1}$; WTS $\mathbf{x}^T A\mathbf{x} \geq 0$, and "$=$" holds iff $\mathbf{x} = \mathbf{0}$. Set $\mathbf{x} = \left[\begin{array}{c} x_0 \\ \hline \mathbf{z} \end{array}\right]$, where $x_0 \in \mathbb{R}$ and $\mathbf{z} \in \mathbb{R}^n$.

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.

Proof (continued). Suppose conversely that $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. WTS A is positive definite.

By hypothesis, A is symmetric. Now, fix any $\mathbf{x} \in \mathbb{R}^{n+1}$; WTS $\mathbf{x}^T A\mathbf{x} \geq 0$, and "$=$" holds iff $\mathbf{x} = \mathbf{0}$. Set $\mathbf{x} = \left[\begin{array}{c} x_0 \\ \hline \mathbf{z} \end{array}\right]$, where $x_0 \in \mathbb{R}$ and $\mathbf{z} \in \mathbb{R}^n$. We now compute (next slide):

Proof (continued). Reminder: $\mathbf{x} = \left[\begin{array}{c} x_0 \\ \hline \mathbf{z} \end{array}\right]$, where $x_0 \in \mathbb{R}$, $\mathbf{z} \in \mathbb{R}^n$.

$$\begin{aligned}
\mathbf{x}^T A \mathbf{x} &= \left[\begin{array}{c:c} x_0 & \mathbf{z}^T \end{array}\right] \left[\begin{array}{c:c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right] \left[\begin{array}{c} x_0 \\ \hline \mathbf{z} \end{array}\right] \\
&= \alpha x_0^2 + x_0 \mathbf{a}^T \mathbf{z} + x_0 \mathbf{z}^T \mathbf{a} + \mathbf{z}^T A' \mathbf{z} \\
&\overset{(*)}{=} \alpha x_0^2 + 2 x_0 \mathbf{a}^T \mathbf{z} + \mathbf{z}^T A' \mathbf{z} \\
&= \mathbf{z}^T (A' - \tfrac{1}{\alpha} \mathbf{a}\mathbf{a}^T) \mathbf{z} + \tfrac{1}{\alpha} \mathbf{z}^T \mathbf{a}\mathbf{a}^T \mathbf{z} + 2 x_0 \mathbf{a}^T \mathbf{z} + \alpha x_0^2 \\
&= \mathbf{z}^T (A' - \tfrac{1}{\alpha} \mathbf{a}\mathbf{a}^T) \mathbf{z} + (\tfrac{1}{\sqrt{\alpha}} \mathbf{a}^T \mathbf{z})^2 + 2 x_0 \mathbf{a}^T \mathbf{z} + (\sqrt{\alpha} x_0)^2 \\
&= \mathbf{z}^T (A' - \tfrac{1}{\alpha} \mathbf{a}\mathbf{a}^T) \mathbf{z} + \left(\tfrac{1}{\sqrt{\alpha}} \mathbf{a}^T \mathbf{z} + \sqrt{\alpha} x_0\right)^2 \overset{(**)}{\geq} 0,
\end{aligned}$$

where in (*), we used the fact that $x_0 \mathbf{z}^T \mathbf{a}$ is a 1×1 (and consequently symmetric) matrix, and so $x_0 \mathbf{z}^T \mathbf{a} = (x_0 \mathbf{z}^T \mathbf{a})^T = x_0 \mathbf{a}^T \mathbf{z}$; and where for the inequality (**), we used the fact that $\mathbf{z}^T (A' - \frac{1}{\alpha} \mathbf{a}\mathbf{a}^T) \mathbf{z} \geq 0$, since $A' - \frac{1}{\alpha} \mathbf{a}\mathbf{a}^T$ is positive definite.

Proof (continued). Reminder: $\mathbf{x} = \left[\begin{array}{c} x_0 \\ \hline \mathbf{z} \end{array}\right]$, where $x_0 \in \mathbb{R}$, $\mathbf{z} \in \mathbb{R}^n$.

$$\begin{aligned}
\mathbf{x}^T A\mathbf{x} &= \left[\begin{array}{c|c} x_0 & \mathbf{z}^T \end{array}\right] \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right] \left[\begin{array}{c} x_0 \\ \hline \mathbf{z} \end{array}\right] \\
&= \alpha x_0^2 + x_0 \mathbf{a}^T \mathbf{z} + x_0 \mathbf{z}^T \mathbf{a} + \mathbf{z}^T A' \mathbf{z} \\
&\overset{(*)}{=} \alpha x_0^2 + 2x_0 \mathbf{a}^T \mathbf{z} + \mathbf{z}^T A' \mathbf{z} \\
&= \mathbf{z}^T (A' - \tfrac{1}{\alpha} \mathbf{a}\mathbf{a}^T)\mathbf{z} + \tfrac{1}{\alpha} \mathbf{z}^T \mathbf{a}\mathbf{a}^T \mathbf{z} + 2x_0 \mathbf{a}^T \mathbf{z} + \alpha x_0^2 \\
&= \mathbf{z}^T (A' - \tfrac{1}{\alpha} \mathbf{a}\mathbf{a}^T)\mathbf{z} + (\tfrac{1}{\sqrt{\alpha}} \mathbf{a}^T \mathbf{z})^2 + 2x_0 \mathbf{a}^T \mathbf{z} + (\sqrt{\alpha} x_0)^2 \\
&= \mathbf{z}^T (A' - \tfrac{1}{\alpha} \mathbf{a}\mathbf{a}^T)\mathbf{z} + \left(\tfrac{1}{\sqrt{\alpha}} \mathbf{a}^T \mathbf{z} + \sqrt{\alpha} x_0\right)^2 \overset{(**)}{\geq} 0,
\end{aligned}$$

where in (*), we used the fact that $x_0 \mathbf{z}^T \mathbf{a}$ is a 1×1 (and consequently symmetric) matrix, and so $x_0 \mathbf{z}^T \mathbf{a} = (x_0 \mathbf{z}^T \mathbf{a})^T = x_0 \mathbf{a}^T \mathbf{z}$; and where for the inequality (**), we used the fact that $\mathbf{z}^T (A' - \frac{1}{\alpha} \mathbf{a}\mathbf{a}^T)\mathbf{z} \geq 0$, since $A' - \frac{1}{\alpha} \mathbf{a}\mathbf{a}^T$ is positive definite. It remains to show that the inequality (**) is an equality iff $\mathbf{x} = \mathbf{0}$.

Proof (continued). Reminder: $\mathbf{x} = \begin{bmatrix} x_0 \\ - \\ \mathbf{z} \end{bmatrix}$, where $x_0 \in \mathbb{R}$, $\mathbf{z} \in \mathbb{R}^n$;

$\mathbf{z}^T(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} \geq 0$;

$$\mathbf{x}^T A \mathbf{x} = \mathbf{z}^T(A' - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} + \left(\tfrac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{z} + \sqrt{\alpha}x_0\right)^2 \overset{(**)}{\geq} 0;$$

WTS the inequality (**) is an equality iff $\mathbf{x} = \mathbf{0}$.

Proof (continued). Reminder: $\mathbf{x} = \begin{bmatrix} x_0 \\ \hdashline \mathbf{z} \end{bmatrix}$, where $x_0 \in \mathbb{R}$, $\mathbf{z} \in \mathbb{R}^n$; $\mathbf{z}^T(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} \geq 0$;

$$\mathbf{x}^T A\mathbf{x} = \mathbf{z}^T(A' - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} + \left(\tfrac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{z} + \sqrt{\alpha}x_0\right)^2 \overset{(**)}{\geq} 0;$$

WTS the inequality (**) is an equality iff $\mathbf{x} = \mathbf{0}$.

If $\mathbf{x} = \mathbf{0}$, then $x_0 = 0$ and $\mathbf{z} = \mathbf{0}$, and it is obvious that the inequality (**) is an equality.

Proof (continued). Reminder: $\mathbf{x} = \begin{bmatrix} x_0 \\ \hline \mathbf{z} \end{bmatrix}$, where $x_0 \in \mathbb{R}$, $\mathbf{z} \in \mathbb{R}^n$;
$\mathbf{z}^T(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} \geq 0$;

$$\mathbf{x}^T A \mathbf{x} \quad = \quad \mathbf{z}^T(A' - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} + \Big(\tfrac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{z} + \sqrt{\alpha}x_0\Big)^2 \quad \stackrel{(**)}{\geq} \quad 0;$$

WTS the inequality (**) is an equality iff $\mathbf{x} = \mathbf{0}$.

If $\mathbf{x} = \mathbf{0}$, then $x_0 = 0$ and $\mathbf{z} = \mathbf{0}$, and it is obvious that the inequality (**) is an equality.

Suppose now that the inequality (**) is an equality. Then $\mathbf{z}^T(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{z} = 0$ and $\frac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{z} + \sqrt{\alpha}x_0 = 0$. The former implies that $\mathbf{z} = \mathbf{0}$ (since $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite). But now since $\frac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{z} + \sqrt{\alpha}x_0 = 0$, we deduce that $x_0 = 0$. So, $\mathbf{x} = \begin{bmatrix} x_0 \\ \hline \mathbf{z} \end{bmatrix} = \mathbf{0}$.
This proves that A is positive definite. □

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.

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.

- There are a couple of numerical examples 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.

- There are a couple of numerical examples in the Lecture Notes.
- However, Theorem 10.2.3 is not the most convenient for computational purposes.

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.

- There are a couple of numerical examples in the Lecture Notes.
- However, Theorem 10.2.3 is not the most convenient for computational purposes.
- Instead, we will use Theorem 10.2.3 to prove the correctness of two more convenient tests: the Gaussian elimination test of positive definiteness (Theorem 10.2.6) and Sylvester's criterion of positive definiteness (Theorem 10.2.9).

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 $j \in \{1, \dots, n\}$, 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 $i \in \{j+1, \dots, n\}$, 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 $j+1$.

- **Remark:** The algorithm performs a modified version of the "forward" part of the row reduction algorithm.

- **Remark:** The algorithm performs a modified version of the "forward" part of the row reduction algorithm.
  - It only performs elementary row operations of the form "$R_i \to R_i + \alpha R_j$," where $i > j$ (i.e. row i is below row j), and where α is chosen so that the i, j-th entry of the matrix becomes zero; moreover, these operations (which add scalar multiples of row j to the rows below it) are performed only in Step j.
    - Essentially, we use the j, j-th entry of the matrix A_j to "clean up" the j-th column below the main diagonal, i.e. to turn all entries of the j-th column below the main diagonal into zeros. Note that at the start of Step j, the leftmost $j - 1$ many columns have already been processed, so that they have all zeros below the main diagonal.

- **Remark:** The algorithm performs a modified version of the "forward" part of the row reduction algorithm.
  - It only performs elementary row operations of the form "$R_i \to R_i + \alpha R_j$," where $i > j$ (i.e. row i is below row j), and where α is chosen so that the i, j-th entry of the matrix becomes zero; moreover, these operations (which add scalar multiples of row j to the rows below it) are performed only in Step j.
    - Essentially, we use the j, j-th entry of the matrix A_j to "clean up" the j-th column below the main diagonal, i.e. to turn all entries of the j-th column below the main diagonal into zeros. Note that at the start of Step j, the leftmost $j - 1$ many columns have already been processed, so that they have all zeros below the main diagonal.
  - We keep modifying our matrix until we either obtain a zero or a negative number on the main diagonal (in this case, our input matrix is not positive definite), or until we transform our matrix into an upper triangular matrix with a positive main diagonal (in this case, our input matrix is positive definite).

- **Remark:** The algorithm performs a modified version of the "forward" part of the row reduction algorithm.
  - It only performs elementary row operations of the form "$R_i \to R_i + \alpha R_j$," where $i > j$ (i.e. row i is below row j), and where α is chosen so that the i, j-th entry of the matrix becomes zero; moreover, these operations (which add scalar multiples of row j to the rows below it) are performed only in Step j.
    - Essentially, we use the j, j-th entry of the matrix A_j to "clean up" the j-th column below the main diagonal, i.e. to turn all entries of the j-th column below the main diagonal into zeros. Note that at the start of Step j, the leftmost $j - 1$ many columns have already been processed, so that they have all zeros below the main diagonal.
  - We keep modifying our matrix until we either obtain a zero or a negative number on the main diagonal (in this case, our input matrix is not positive definite), or until we transform our matrix into an upper triangular matrix with a positive main diagonal (in this case, our input matrix is positive definite).
- Before proving the theorem, we take a look at a couple of examples.

Example 10.2.7

Using Theorem 10.2.6, determine whether the matrix

$$A := \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}$$

is positive definite.

Solution.

Example 10.2.7

Using Theorem 10.2.6, determine whether the matrix

$$A \quad := \quad \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}$$

is positive definite.

Solution. The matrix A is symmetric, and so Theorem 10.2.6 applies.

Example 10.2.7

Using Theorem 10.2.6, determine whether the matrix

$$A \quad := \quad \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}$$

is positive definite.

Solution. The matrix A is symmetric, and so Theorem 10.2.6 applies. We perform the modified version of the "forward" part of the row reduction algorithm described in Theorem 10.2.6, as follows (the dotted lines isolate the submatrix in the lower right corner that is still being processed):

Solution (continued).

$$A = \underbrace{\begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}}_{=:A_1} \qquad \text{Step 0}$$

$$\overset{R_2 \to R_2 + \frac{1}{2} R_1}{\underset{\sim}{R_3 \to R_3 - R_1}} \quad \underbrace{\left[\begin{array}{c|cc} 4 & -2 & 4 \\ \hline 0 & 9 & 3 \\ 0 & 3 & 2 \end{array}\right]}_{=:A_2} \qquad \text{Step 1}$$

$$\underset{\sim}{R_3 \to R_3 - \frac{1}{3} R_2} \quad \underbrace{\left[\begin{array}{cc|c} 4 & -2 & 4 \\ 0 & 9 & 3 \\ \hline 0 & 0 & 1 \end{array}\right]}_{=:A_3} \qquad \text{Step 2.}$$

We have now obtained an upper triangular matrix with a positive main diagonal. So, by Theorem 10.2.6, A is positive definite. (This answer is returned by Step 3 of the algorithm from Theorem 10.2.6, at which point the algorithm terminates.) □

Example 10.2.7

Using Theorem 10.2.6, determine whether the matrix

$$A \quad := \quad \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}$$

is positive definite.

- **Remark:** Normally, we do not actually number our steps, and we do not name the matrices A_i; here, we did it for the sake of extra clarity. The horizontal and vertical dotted lines are also optional, but they are useful for visually keeping track of the submatrix being processed, and so it is not a bad idea to include them.

Example 10.2.8

Using Theorem 10.2.6, determine whether the matrix

$$A := \begin{bmatrix} 2 & -2 & 2 & 0 \\ -2 & 3 & 0 & 1 \\ 2 & 0 & 6 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

is positive definite.

Solution.

Example 10.2.8

Using Theorem 10.2.6, determine whether the matrix

$$A := \begin{bmatrix} 2 & -2 & 2 & 0 \\ -2 & 3 & 0 & 1 \\ 2 & 0 & 6 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

is positive definite.

Solution. The matrix A is symmetric, and so Theorem 10.2.6 applies. We perform the modified version of the "forward" part of the row reduction algorithm described in Theorem 10.2.6, as follows (next slide):

Solution (continued).

$$A = \begin{bmatrix} 2 & -2 & 2 & 0 \\ -2 & 3 & 0 & 1 \\ 2 & 0 & 6 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

$$\underset{\sim}{\substack{R_2 \to R_2 + R_1 \\ R_3 \to R_3 - R_1}} \begin{bmatrix} 2 & -2 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 4 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

$$\underset{\sim}{\substack{R_3 \to R_3 - 2R_2 \\ R_4 \to R_4 - R_2}} \begin{bmatrix} 2 & -2 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix}.$$

We have now obtained a zero on the main diagonal of our matrix, and so by Theorem 10.2.6, the matrix A is **not** positive definite. □

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 $j+1$.

Proof. We proceed by induction on n.

Proof. We proceed by induction on n.

For $n = 1$, we fix a matrix $A = \left[\; a_{1,1} \; \right]$ in $\mathbb{R}^{1\times 1}$ (obviously, A is symmetric), and we observe that our algorithm sets $A_1 := A$, and then if $a_{1,1} > 0$, returns the answer that A is positive definite, and if $a_{1,1} \leq 0$, returns the answer that A is not positive definite. Obviously, this is correct.

Proof. We proceed by induction on n.

For $n = 1$, we fix a matrix $A = \left[\begin{array}{c} a_{1,1} \end{array}\right]$ in $\mathbb{R}^{1\times 1}$ (obviously, A is symmetric), and we observe that our algorithm sets $A_1 := A$, and then if $a_{1,1} > 0$, returns the answer that A is positive definite, and if $a_{1,1} \leq 0$, returns the answer that A is not positive definite. Obviously, this is correct.

Now, fix an integer $n \geq 2$, and assume inductively that the algorithm is correct when applied to symmetric matrices in $\mathbb{R}^{(n-1)\times(n-1)}$.

Proof. We proceed by induction on n.

For $n = 1$, we fix a matrix $A = \begin{bmatrix} a_{1,1} \end{bmatrix}$ in $\mathbb{R}^{1\times 1}$ (obviously, A is symmetric), and we observe that our algorithm sets $A_1 := A$, and then if $a_{1,1} > 0$, returns the answer that A is positive definite, and if $a_{1,1} \leq 0$, returns the answer that A is not positive definite. Obviously, this is correct.

Now, fix an integer $n \geq 2$, and assume inductively that the algorithm is correct when applied to symmetric matrices in $\mathbb{R}^{(n-1)\times(n-1)}$. Fix a symmetric matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ in $\mathbb{R}^{n\times n}$.

Proof. We proceed by induction on n.

For $n = 1$, we fix a matrix $A = \begin{bmatrix} a_{1,1} \end{bmatrix}$ in $\mathbb{R}^{1\times 1}$ (obviously, A is symmetric), and we observe that our algorithm sets $A_1 := A$, and then if $a_{1,1} > 0$, returns the answer that A is positive definite, and if $a_{1,1} \leq 0$, returns the answer that A is not positive definite. Obviously, this is correct.

Now, fix an integer $n \geq 2$, and assume inductively that the algorithm is correct when applied to symmetric matrices in $\mathbb{R}^{(n-1)\times(n-1)}$. Fix a symmetric matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ in $\mathbb{R}^{n\times n}$.

Set $A_1 := A$, as per Step 0 of the algorithm.

Proof. We proceed by induction on n.

For $n = 1$, we fix a matrix $A = \begin{bmatrix} a_{1,1} \end{bmatrix}$ in $\mathbb{R}^{1\times 1}$ (obviously, A is symmetric), and we observe that our algorithm sets $A_1 := A$, and then if $a_{1,1} > 0$, returns the answer that A is positive definite, and if $a_{1,1} \leq 0$, returns the answer that A is not positive definite. Obviously, this is correct.

Now, fix an integer $n \geq 2$, and assume inductively that the algorithm is correct when applied to symmetric matrices in $\mathbb{R}^{(n-1)\times(n-1)}$. Fix a symmetric matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ in $\mathbb{R}^{n\times n}$.

Set $A_1 := A$, as per Step 0 of the algorithm. If the main diagonal of A_1 contains an entry that is zero or negative, then the algorithm terminates after Step 1, having determined that A is not positive definite (this is correct by Proposition 10.1.4).

Proof (continued). From now on, we may assume that the main diagonal of A_1 is positive, and in particular, $a_{1,1} > 0$.

Proof (continued). From now on, we may assume that the main diagonal of A_1 is positive, and in particular, $a_{1,1} > 0$. In this case, Step 1 performs the following elementary row operations on the matrix $A_1 = A$:

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

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

This transforms entries $2, \dots, n-1$ of the first column into 0. The resulting matrix is A_2. But note that our matrix $A_1 = A$ is of the form

$$A_1 \quad = \quad A \quad = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{a} & A_{1,1} \end{array}\right],$$

where $\mathbf{a} = \left[\begin{array}{ccc} a_{2,1} & \dots & a_{n,1} \end{array}\right]^T$, and $A_{1,1}$ is the matrix obtained from A by deleting the first row and first column. So, by Proposition 10.2.2, the matrix A_2 that we obtain after the Step 1 is precisely the matrix (next slide):

Proof (continued).

$$A_2 = \left[\begin{array}{c:c} a_{1,1} & \mathbf{a}^T \\ \hdashline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T \end{array}\right].$$

Proof (continued).

$$A_2 = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T \end{array}\right].$$

By Proposition 10.2.2, $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ is symmetric, and the remainder of our algorithm only manipulates the $(n-1) \times (n-1)$ submatrix $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ of A_2 (while leaving the top row and leftmost column of A_2 unchanged).

Proof (continued).

$$A_2 = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T \end{array}\right].$$

By Proposition 10.2.2, $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ is symmetric, and the remainder of our algorithm only manipulates the $(n-1)\times(n-1)$ submatrix $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ of A_2 (while leaving the top row and leftmost column of A_2 unchanged). Moreover, by Theorem 10.2.3, A is positive definite iff $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ is positive definite.

Proof (continued).

$$A_2 = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{0} & A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T \end{array}\right].$$

By Proposition 10.2.2, $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ is symmetric, and the remainder of our algorithm only manipulates the $(n-1)\times(n-1)$ submatrix $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ of A_2 (while leaving the top row and leftmost column of A_2 unchanged). Moreover, by Theorem 10.2.3, A is positive definite iff $A_{1,1} - \frac{1}{a_{1,1}}\mathbf{a}\mathbf{a}^T$ is positive definite. The result now readily follows from the induction hypothesis. $\square$

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 $j \in \{1, \dots, n\}$, 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 $i \in \{j + 1, \dots, n\}$, 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 $j + 1$.

- 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 $A^{(n)} = A$.

- Reminder:

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.

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.

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. We first assume (i) and prove (ii). In view of Corollary 10.1.3(a), it suffices to show that the matrices $A^{(1)}, \dots, A^{(n)}$ are all positive definite.

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. We first assume (i) and prove (ii). In view of Corollary 10.1.3(a), it suffices to show that the matrices $A^{(1)}, \dots, A^{(n)}$ are all positive definite. Obviously, these n matrices are all symmetric (because A is symmetric).

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. We first assume (i) and prove (ii). In view of Corollary 10.1.3(a), it suffices to show that the matrices $A^{(1)}, \dots, A^{(n)}$ are all positive definite. Obviously, these n matrices are all symmetric (because A is symmetric). Now, fix an index $k \in \{1, \dots, n\}$ and a vector $\mathbf{x}_k = \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix}^T$ in $\mathbb{R}^k \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}_k^T A^{(k)} \mathbf{x}_k > 0$.

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. We first assume (i) and prove (ii). In view of Corollary 10.1.3(a), it suffices to show that the matrices $A^{(1)}, \dots, A^{(n)}$ are all positive definite. Obviously, these n matrices are all symmetric (because A is symmetric). Now, fix an index $k \in \{1, \dots, n\}$ and a vector $\mathbf{x}_k = \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix}^T$ in $\mathbb{R}^k \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}_k^T A^{(k)} \mathbf{x}_k > 0$. Set $\mathbf{x} := \left[\begin{array}{ccc:ccc} x_1 & \dots & x_k & 0 & \dots & 0 \end{array}\right]^T$ (with $n-k$ zeros to the right of the vertical dotted line, so that $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$).

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

Proof. We first assume (i) and prove (ii). In view of Corollary 10.1.3(a), it suffices to show that the matrices $A^{(1)}, \dots, A^{(n)}$ are all positive definite. Obviously, these n matrices are all symmetric (because A is symmetric). Now, fix an index $k \in \{1, \dots, n\}$ and a vector $\mathbf{x}_k = \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix}^T$ in $\mathbb{R}^k \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}_k^T A^{(k)} \mathbf{x}_k > 0$. Set $\mathbf{x} := \left[\begin{array}{ccc:ccc} x_1 & \dots & x_k & 0 & \dots & 0 \end{array}\right]^T$ (with $n-k$ zeros to the right of the vertical dotted line, so that $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$). Then

$$\mathbf{x}_k^T A^{(k)} \mathbf{x}_k = \mathbf{x}^T A \mathbf{x} \overset{(*)}{>} 0,$$

where (*) follows from the fact that A is positive definite and $\mathbf{x} \neq \mathbf{0}$ (because $\mathbf{x}_k \neq \mathbf{0}$).

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

Proof. We first assume (i) and prove (ii). In view of Corollary 10.1.3(a), it suffices to show that the matrices $A^{(1)}, \dots, A^{(n)}$ are all positive definite. Obviously, these n matrices are all symmetric (because A is symmetric). Now, fix an index $k \in \{1, \dots, n\}$ and a vector $\mathbf{x}_k = \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix}^T$ in $\mathbb{R}^k \setminus \{\mathbf{0}\}$; WTS $\mathbf{x}_k^T A^{(k)} \mathbf{x}_k > 0$. Set $\mathbf{x} := \begin{bmatrix} x_1 & \dots & x_k & \vdots & 0 & \dots & 0 \end{bmatrix}^T$ (with $n-k$ zeros to the right of the vertical dotted line, so that $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$). Then

$$\mathbf{x}_k^T A^{(k)} \mathbf{x}_k \quad = \quad \mathbf{x}^T A \mathbf{x} \quad \stackrel{(*)}{>} \quad 0,$$

where (*) follows from the fact that A is positive definite and $\mathbf{x} \neq \mathbf{0}$ (because $\mathbf{x}_k \neq \mathbf{0}$). So, $A^{(k)}$ is positive definite, and we deduce that (ii) holds.

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 (continued). We now assume (ii) and prove (i).

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 (continued). We now assume (ii) and prove (i). We may assume inductively that the statement is true for smaller matrices. More precisely, we assume that the following holds.

> ***Induction hypothesis:*** *For all $n' \in \{1, \ldots, n-1\}$, and all symmetric matrices $B \in \mathbb{R}^{n'\times n'}$, if $\det(B^{(1)}), \ldots, \det(B^{(n')}) > 0$, then B is positive definite.*

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

Proof (continued). We now assume (ii) and prove (i). We may assume inductively that the statement is true for smaller matrices. More precisely, we assume that the following holds.

> ***Induction hypothesis:*** *For all $n' \in \{1, \ldots, n-1\}$, and all symmetric matrices $B \in \mathbb{R}^{n' \times n'}$, if $\det\left(B^{(1)}\right), \ldots, \det\left(B^{(n')}\right) > 0$, then B is positive definite.*

Set $A = \left[\ a_{i,j}\ \right]_{n \times n}$.

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) $\det\big(A^{(1)}\big), \dots, \det\big(A^{(n)}\big) > 0$.

Proof (continued). We now assume (ii) and prove (i). We may assume inductively that the statement is true for smaller matrices. More precisely, we assume that the following holds.

> ***Induction hypothesis:*** *For all $n' \in \{1, \dots, n-1\}$, and all symmetric matrices $B \in \mathbb{R}^{n' \times n'}$, if $\det\big(B^{(1)}\big), \dots, \det\big(B^{(n')}\big) > 0$, then B is positive definite.*

Set $A = \left[\ a_{i,j}\ \right]_{n \times n}$. Then $a_{1,1} = \det\big(A^{(1)}\big) > 0$.

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 (continued). We now assume (ii) and prove (i). We may assume inductively that the statement is true for smaller matrices. More precisely, we assume that the following holds.

> ***Induction hypothesis:*** *For all $n' \in \{1, \ldots, n-1\}$, and all symmetric matrices $B \in \mathbb{R}^{n' \times n'}$, if $\det(B^{(1)}), \ldots, \det(B^{(n')}) > 0$, then B is positive definite.*

Set $A = \left[\ a_{i,j}\ \right]_{n \times n}$. Then $a_{1,1} = \det(A^{(1)}) > 0$. If $n = 1$, so that $A = \left[\ a_{1,1}\ \right]$, then it is clear that A is positive definite, and we are done.

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 (continued). We now assume (ii) and prove (i). We may assume inductively that the statement is true for smaller matrices. More precisely, we assume that the following holds.

> ***Induction hypothesis:*** *For all $n' \in \{1, \ldots, n-1\}$, and all symmetric matrices $B \in \mathbb{R}^{n' \times n'}$, if $\det(B^{(1)}), \ldots, \det(B^{(n')}) > 0$, then B is positive definite.*

Set $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$. Then $a_{1,1} = \det(A^{(1)}) > 0$. If $n = 1$, so that $A = \begin{bmatrix} a_{1,1} \end{bmatrix}$, then it is clear that A is positive definite, and we are done. So, from now, we assume that $n \geq 2$.

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

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

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

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

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

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

Set $B := A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T$; obviously, $B \in \mathbb{R}^{(n-1)\times(n-1)}$, and by Proposition 10.2.2, B is symmetric.

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

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

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

Set $B := A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T$; obviously, $B \in \mathbb{R}^{(n-1)\times(n-1)}$, and by Proposition 10.2.2, B is symmetric. By Proposition 10.2.1, we have that

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

and since we have already checked that $a_{1,1} > 0$, Theorem 10.2.3 guarantees that A is positive definite iff B is positive definite.

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

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

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

Set $B := A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T$; obviously, $B \in \mathbb{R}^{(n-1)\times(n-1)}$, and by Proposition 10.2.2, B is symmetric. By Proposition 10.2.1, we have that

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

and since we have already checked that $a_{1,1} > 0$, Theorem 10.2.3 guarantees that A is positive definite iff B is positive definite. Thus, it is enough to show that the symmetric matrix $B \in \mathbb{R}^{(n-1)\times(n-1)}$ is positive definite.

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

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

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

Set $B := A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T$; obviously, $B \in \mathbb{R}^{(n-1)\times(n-1)}$, and by Proposition 10.2.2, B is symmetric. By Proposition 10.2.1, we have that

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

and since we have already checked that $a_{1,1} > 0$, Theorem 10.2.3 guarantees that A is positive definite iff B is positive definite. Thus, it is enough to show that the symmetric matrix $B \in \mathbb{R}^{(n-1)\times(n-1)}$ is positive definite. By the induction hypothesis, it suffices to show that $\det(B^{(1)}), \dots, \det(B^{(n-1)}) > 0$.

Proof (continued). Reminder:

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

WTS $\det(B^{(1)}), \dots, \det(B^{(n-1)}) > 0$.

Proof (continued). Reminder:

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

WTS $\det\big(B^{(1)}\big), \ldots, \det\big(B^{(n-1)}\big) > 0$.

Fix an index $k \in \{1, \ldots, n-1\}$; WTS $\det\big(B^{(k)}\big) > 0$.

Proof (continued). Reminder:

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

WTS $\det(B^{(1)}), \dots, \det(B^{(n-1)}) > 0$.

Fix an index $k \in \{1, \dots, n-1\}$; WTS $\det(B^{(k)}) > 0$. First of all, by performing Laplace expansion along the first column of $\widetilde{A}^{(k+1)}$, we obtain $\det(\widetilde{A}^{(k+1)}) = a_{1,1}\det(B^{(k)})$;

Proof (continued). Reminder:

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

WTS $\det(B^{(1)}), \dots, \det(B^{(n-1)}) > 0$.

Fix an index $k \in \{1, \dots, n-1\}$; WTS $\det(B^{(k)}) > 0$. First of all, by performing Laplace expansion along the first column of $\widetilde{A}^{(k+1)}$, we obtain $\det(\widetilde{A}^{(k+1)}) = a_{1,1}\det(B^{(k)})$; since $a_{1,1} > 0$ (and in particular, $a_{1,1} \neq 0$), it follows that $\det(B^{(k)}) = \frac{1}{a_{1,1}}\det(\widetilde{A}^{(k+1)})$.

Proof (continued). Reminder:

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

WTS $\det(B^{(1)}), \dots, \det(B^{(n-1)}) > 0$.

Fix an index $k \in \{1, \dots, n-1\}$; WTS $\det(B^{(k)}) > 0$. First of all, by performing Laplace expansion along the first column of $\widetilde{A}^{(k+1)}$, we obtain $\det(\widetilde{A}^{(k+1)}) = a_{1,1}\det(B^{(k)})$; since $a_{1,1} > 0$ (and in particular, $a_{1,1} \neq 0$), it follows that $\det(B^{(k)}) = \frac{1}{a_{1,1}}\det(\widetilde{A}^{(k+1)})$.

On the other hand, note that $\widetilde{A}^{(k+1)}$ is obtained from $A^{(k+1)}$ by repeatedly adding a scalar multiple of one row to another; by Theorem 7.3.2(c), this type of elementary row operation does not alter the value of the determinant, and it follows that $\det(\widetilde{A}^{(k+1)}) = \det(A^{(k+1)})$.

Proof (continued). Reminder:

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

WTS $\det(B^{(1)}), \dots, \det(B^{(n-1)}) > 0$.

Fix an index $k \in \{1, \dots, n-1\}$; WTS $\det(B^{(k)}) > 0$. First of all, by performing Laplace expansion along the first column of $\widetilde{A}^{(k+1)}$, we obtain $\det(\widetilde{A}^{(k+1)}) = a_{1,1}\det(B^{(k)})$; since $a_{1,1} > 0$ (and in particular, $a_{1,1} \neq 0$), it follows that $\det(B^{(k)}) = \frac{1}{a_{1,1}}\det(\widetilde{A}^{(k+1)})$. On the other hand, note that $\widetilde{A}^{(k+1)}$ is obtained from $A^{(k+1)}$ by repeatedly adding a scalar multiple of one row to another; by Theorem 7.3.2(c), this type of elementary row operation does not alter the value of the determinant, and it follows that $\det(\widetilde{A}^{(k+1)}) = \det(A^{(k+1)})$. We now deduce that

$$\det(B^{(k)}) = \frac{1}{a_{1,1}}\det(\widetilde{A}^{(k+1)}) = \frac{1}{a_{1,1}}\det(A^{(k+1)}) \overset{(*)}{>} 0,$$

where (*) follows from the fact that $a_{1,1} > 0$ and $\det(A^{(k+1)}) > 0$. We now conclude that (ii) holds, and we are done. □

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

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 $A = \left[\begin{array}{c} a \end{array}\right]$ in $\mathbb{R}^{1 \times 1}$, and we note that $a > 0$ (because A is positive definite). We set $L := \left[\begin{array}{c} \sqrt{a} \end{array}\right]$, 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 $A = \left[\begin{array}{c} a \end{array}\right]$ in $\mathbb{R}^{1 \times 1}$, and we note that $a > 0$ (because A is positive definite). We set $L := \left[\begin{array}{c} \sqrt{a} \end{array}\right]$, 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

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

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

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

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 $A = L_1 L_1^T$; 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. Then

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

Proof (continued). We now have that

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

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 $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$.

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. $\square$

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 $i \in \{2, \ldots, n\}$.
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 $i \in \{j+1, \ldots, n\}$.

- 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.