# Linear Algebra 2: Lecture 21

Irena Penev

Summer 2023

In this lecture, $\cdot$ is the standard scalar product in $\mathbb{R}^n$, and $||\cdot||$ is the induced norm.

## 1 Positive (semi-)definite matrices: definition

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

- *positive semi-definite* if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;
- *positive definite* if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

Obviously, every positive definite matrix is positive semi-definite.

**Example 1.1.** *The identity matrix $I_n$ is positive definite. This is because $I_n$ is symmetric, and for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, we have that $\mathbf{x}^T I_n\mathbf{x} = \mathbf{x}^T\mathbf{x} = \mathbf{x}\cdot\mathbf{x} > 0$.*

We note that the definition of a positive (semi-)definite matrix would also make sense without the requirement that $A$ be symmetric. However, note that for any $A \in \mathbb{R}^{n\times n}$, the matrix $\frac{1}{2}(A + A^T)$ is symmetric, and for all vectors $\mathbf{x} \in \mathbb{R}^n$, we have that

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

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

## 2 Positive definite matrices and the scalar product

One reason for interest in positive definite matrices is their role in defining scalar products. This is made precise in Theorem 2.2 below. Before stating that theorem, though, let us first recall the definition of a scalar product from Lecture 11, and let us prove a simple proposition (Propoisition 2.1) about products of the form $\mathbf{x}^T A\mathbf{y}$, where $A$ is a square matrix and $\mathbf{x}, \mathbf{y}$ are vectors.

Recall from Lecture Notes 11 that a *scalar product* (also called *inner product*) in a vector space $V$ over the field $\mathbb{R}$ is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ that satisfies the following axioms:

r.1. for all $\mathbf{x} \in V$, $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds if and only if $\mathbf{x} = \mathbf{0}$;

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

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

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

As we saw in Lecture 11, these four axioms imply the following:

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

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

**Proposition 2.1.** *Let $\mathbb{F}$ be a field. Then for all matrices $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ in $\mathbb{F}^{n\times n}$, and all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that*

$$\mathbf{x}^T A\mathbf{y} \quad = \quad \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i y_j.$$

*Proof.* Fix a matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ in $\mathbb{F}^{n\times n}$, and vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ in $\mathbb{F}^n$.

First of all, we have that

$$A\mathbf{y} \quad = \quad \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad = \quad \begin{bmatrix} \sum_{j=1}^{n} a_{1,j}y_j \\ \sum_{j=1}^{n} a_{2,j}y_j \\ \vdots \\ \sum_{j=1}^{n} a_{n,j}y_j \end{bmatrix}.$$

But now

$$
\begin{aligned}
\mathbf{x}^T A\mathbf{y} &= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^n a_{1,j}y_j \\ \sum_{j=1}^n a_{2,j}y_j \\ \vdots \\ \sum_{j=1}^n a_{n,j}y_j \end{bmatrix} \\
&= \sum_{i=1}^n x_i \Big( \sum_{j=1}^n a_{i,j}y_j \Big) \\
&= \sum_{i=1}^n \sum_{j=1}^n a_{i,j}x_i y_j,
\end{aligned}
$$

which is what we needed to show. □

**Theorem 2.2.** *For any function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the following are equivalent:*

*(i) $\langle \cdot, \cdot \rangle$ is a scalar product on $\mathbb{R}^n$;*

*(ii) there exists a positive definite matrix $A \in \mathbb{R}^{n\times n}$ such that 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.* Suppose first that (i) holds. For all indices $i, j \in \{1, \dots, n\}$, we set $a_{i,j} := \langle \mathbf{e}_i, \mathbf{e}_j \rangle$. Set $A := \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$. Then for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ in $\mathbb{R}^n$, we have that:

$$
\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \langle \sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j \rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n x_i y_i \langle \mathbf{e}_i, \mathbf{e}_j \rangle && \text{because } \langle \cdot, \cdot \rangle \text{ is a scalar product in } \mathbb{R}^n \\
&= \sum_{i=1}^n \sum_{j=1}^n x_i y_i a_{i,j} && \text{because } a_{i,j} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle \\
&= \mathbf{x}^T A\mathbf{y} && \text{by Proposition 2.1.}
\end{aligned}
$$

It remains to show that the matrix $A$ is positive definite. First of all, $A$ is symmetric because for all $i, j \in \{1, \dots, n\}$, we have that

$$
a_{i,j} \stackrel{(*)}{=} \langle \mathbf{e}_i, \mathbf{e}_j \rangle \stackrel{(**)}{=} \langle \mathbf{e}_j, \mathbf{e}_i \rangle \stackrel{(*)}{=} a_{j,i},
$$

where both instances of (*) follow from the construction of $A$, and (**) follows from the fact that $\langle \cdot, \cdot \rangle$ is a scalar product in $\mathbb{R}^n$. On the other hand, for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, we have that

$$
\mathbf{x}^T A\mathbf{x} = \langle \mathbf{x}, \mathbf{x} \rangle \stackrel{(*)}{>} 0,
$$

where (*) follows from the fact that $\langle \cdot, \cdot \rangle$ is a scalar product. So, (ii) holds.

Suppose now that (ii) holds, and let the matrix $A$ be as in (ii). We must prove (i). We verify the axioms of a scalar product one by one.

r.1. For every $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, we have that $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^T A \mathbf{x} \overset{(*)}{>} 0$, where (*) follows from the fact that $A$ is positive definite. On the other hand, we have that $\langle \mathbf{0}, \mathbf{0} \rangle = \mathbf{0}^T A \mathbf{0} = 0$. This proves that for all $\mathbf{x} \in \mathbb{R}^n$, we have that $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and that equality holds if and only if $\mathbf{x} = \mathbf{0}$.

r.2. For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, we have that:

$$\begin{aligned} \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle &= (\mathbf{x} + \mathbf{y})^T A \mathbf{z} \\ &= (\mathbf{x}^T + \mathbf{y}^T) A \mathbf{z} \\ &= \mathbf{x}^T A \mathbf{z} + \mathbf{y}^T A \mathbf{z} \\ &= \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle. \end{aligned}$$

r.3. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, we have that

$$\langle \alpha \mathbf{x}, \mathbf{y} \rangle = (\alpha \mathbf{x})^T A \mathbf{y} = \alpha (\mathbf{x}^T A \mathbf{y}) = \alpha \langle \mathbf{x}, \mathbf{y} \rangle.$$

r.4. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have that

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y} \overset{(*)}{=} (\mathbf{x}^T A \mathbf{y})^T = \mathbf{y}^T A^T \mathbf{x} \overset{(**)}{=} \mathbf{y}^T A \mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle,$$

where in (*), we used the fact that $A$ is a $1 \times 1$ (and consequently, symmetric) matrix, and in (**), we used the fact that $A$ is symmetric.

This proves that (i) holds. □

**Remark:** Positive (semi-)definite matrices also play an important role in optimization, but we shall not discuss this in this course.

# 3 Basic properties of positive (semi-)definite matrices

Let us say that the main diagonal of a matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ in $\mathbb{R}^{n \times n}$ is *non-negative* (resp. *positive*) if $a_{1,1}, \dots, a_{n,n} \geq 0$ (resp. $a_{1,1}, \dots, a_{n,n} > 0$). In other words, the main diagonal of a square matrix is non-negative (resp. positive) if all the entries on the main diagonal of that matrix are non-negative (resp. positive). The following proposition gives a necessary (but not sufficient) condition for positive (semi-)definiteness.

**Proposition 3.1.**

*(a) The main diagonal of any positive semi-definite is non-negative.*

*(b) The main diagonal of any positive definite is positive.*

*Proof.* Fix a matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ in $\mathbb{R}^{n \times n}$. Note that for all indices $i \in \{1, \dots, n\}$, we have that $\mathbf{e}_i^T A \mathbf{e}_i = a_{i,i}$.[1] The result now follows from the definition of positive (semi-)definiteness. □

**Theorem 3.2.**

*(a) If $A, B \in \mathbb{R}^{n \times n}$ are both positive definite, then $A + B$ is positive definite.*

*(b) If $A \in \mathbb{R}^{n \times n}$ is positive definite and $\alpha > 0$, then $\alpha A$ is positive definite.*

*(c) If $A \in \mathbb{R}^{n \times n}$ is positive definite, then $A$ is invertible and its inverse $A^{-1}$ is positive definite.*

*Proof.* (a) and (b) are trivial. Let us prove (c). Fix a positive definite matrix $A \in \mathbb{R}^{n \times n}$. We first prove that $A$ is invertible. By Corollary 5.1 of Lecture Notes 4, it suffices to show that $A\mathbf{x} = \mathbf{0}$ only has the trivial solution. So, fix a solution $\mathbf{x}_0 \in \mathbb{R}^n$ of this equation, so that $A\mathbf{x}_0 = \mathbf{0}$. But then $\mathbf{x}_0^T A \mathbf{x}_0 = 0$; since $A$ is positive definite, it follows that $\mathbf{x}_0 = \mathbf{0}$. This proves that $A$ is invertible.

It remains to show that $A^{-1}$ is positive definite. Since $A$ is positive definite, it is symmetric; consequently, $A^{-1}$ is also symmetric.[2] Now, fix any $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Since $\mathbf{x} \neq \mathbf{0}$ and $A^{-1}$ is invertible, we see that $A^{-1}\mathbf{x} \neq 0$. But now we have the following:

$$\begin{aligned}
\mathbf{x}^T A^{-1} \mathbf{x} &= \mathbf{x}^T A^{-1} A A^{-1} \mathbf{x} \\
&= ((A^{-1})^T \mathbf{x})^T A (A^{-1}\mathbf{x}) \\
&= (A^{-1}\mathbf{x})^T A (A^{-1}\mathbf{x}) && \text{because } A^{-1} \text{ is symmetric} \\
&> 0 && \text{because } A \text{ is positive definite and } A^{-1}\mathbf{x} \neq 0.
\end{aligned}$$

So, $A^{-1}$ is positive definite. □

**Theorem 3.3.** *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Then the following are equivalent:*

[1] This can be seen directly, but it also follows from Proposition 2.1.

[2] Indeed, since $A$ is invertible and symmetric, we have that

- $(A^{-1})^T A = (A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n$;
- $A(A^{-1})^T = A^T (A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$.

So, $(A^{-1})^T$ is an inverse of $A$, and consequently, $A^{-1} = (A^{-1})^T$. Thus, $A^{-1}$ is symmetric.

*(i) A is positive definite;*

*(ii) all eigenvalues of A are positive;*

*(iii) there exists an invertible matrix $U \in \mathbb{R}^{n\times n}$ such that $A = U^TU$.*

*Proof.* We will prove "(i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i)."

"(i) $\Longrightarrow$ (ii)": Assume that (i) holds. Fix any eigenvalue $\lambda$ of $A$, and let $\mathbf{x}$ be an associated eigenvector. After possibly normalizing $\mathbf{x}$ (i.e. replacing $\mathbf{x}$ by $\frac{\mathbf{x}}{||\mathbf{x}||}$), we may assume that $||\mathbf{x}|| = 1$. Since $A$ is positive definite, we have that $\mathbf{x}^TA\mathbf{x} > 0$. But note that

$$\mathbf{x}^TA\mathbf{x} \quad = \quad \mathbf{x}^T(\lambda\mathbf{x}) \quad = \quad \lambda(\mathbf{x}^T\mathbf{x}) \quad = \quad \lambda(\mathbf{x}\cdot\mathbf{x}) \quad = \quad \lambda||\mathbf{x}||^2 \quad = \quad \lambda.$$

So, $\lambda = \mathbf{x}^TA\mathbf{x} > 0$. Thus, (ii) holds.

"(ii) $\Longrightarrow$ (iii)": Assume that (ii) holds. Since $A$ is symmetric, it is orthogonally diagonalizable (by Theorem 2.5 of Lecture Notes 20). Let $D = D(\lambda_1, \dots, \lambda_n)$ be a diagonal and $Q$ an orthogonal matrix, both in $\mathbb{R}^{n\times n}$, such that $D = Q^TAQ$, and consequently, $A = QDQ^T$. Then $\lambda_1, \dots, \lambda_n$ are all eigenvalues of $A$, and so by (ii), $\lambda_1, \dots, \lambda_n > 0$. Now, set $\widetilde{D} := D(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$; clearly, $\widetilde{D}^2 = D$. Next, set $U := \widetilde{D}Q$. Since $\widetilde{D}$ and $Q$ are both invertible,[3] so is $U$. But now

$$\begin{aligned} U^TU &= (\widetilde{D}Q)^T(\widetilde{D}Q) \\ &= Q^T\widetilde{D}^T\widetilde{D}Q \\ &= Q^T\widetilde{D}^2Q \\ &= Q^TDQ \\ &= A. \end{aligned}$$

So, (iii) holds.

"(iii) $\Longrightarrow$ (i)": Assume (iii), and fix an invertible matrix $U$ such that

[3]For $\widetilde{D}$, we observe that $\det(\widetilde{D}) = \sqrt{\lambda_1 \dots \lambda_n} > 0$, and so $\widetilde{D}$ is invertible (by Theorem 5.1 of Lecture Notes 15). On the other hand, since $Q$ is orthogonal, it is invertible (by Theorem 2.1 of Lecture Notes 14).

$A = U^T U$. Fix any vector $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then

$$\begin{aligned} \mathbf{x}^T A \mathbf{x} &= \mathbf{x}^T U^T U \mathbf{x} \\ &= (U\mathbf{x})^T (U\mathbf{x}) \\ &= (U\mathbf{x}) \cdot (U\mathbf{x}) \\ &= ||U\mathbf{x}||^2 \\ &\stackrel{(*)}{>} 0, \end{aligned}$$

where (*) follows from the fact that $U\mathbf{x} \neq \mathbf{0}$ (because $U$ is invertible and $\mathbf{x} \neq \mathbf{0}$). So, (i) holds. □

For positive semi-definite matrices, we have the following analog of Theorem 3.3.

**Theorem 3.4.** *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Then the following are equivalent:*

*(i) $A$ is positive semi-definite;*

*(ii) all eigenvalues of $A$ are non-negative;*

*(iii) there exists a matrix $U \in \mathbb{R}^{n \times n}$ such that $A = U^T U$.*

*Proof.* Analogous to the proof of Theorem 3.3. □

**Proposition 3.5.** *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix.*

*(a) If $A$ is positive semi-definite, then det($A$) and trace($A$) are both non-negative.*

*(b) If $A$ is positive definite, then det($A$) and trace($A$) are both positive.*

*Proof.* Since $A$ is symmetric, Corollary 2.4 of Lecture Notes 20 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 2.11 of Lecture Notes 18, we have that $\det(A) = \lambda_1 \dots \lambda_n$ and $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$. By Theorem 3.4, all eigenvalues of a positive semi-definite matrix are non-negative, and it follows that (a) holds. Similarly, by Theorem 3.3, all eigenvalues of a positive definite matrix are positive, and it follows that (b) holds. □

# 4 Methods of testing for positive definiteness

**Theorem 4.1** (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 if and only if $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.*

*Proof.* Suppose first that $A$ is positive definite. By Proposition 3.1(b), we have that $\alpha > 0$, and in particular, the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is defined. We must show that this matrix is positive definite. Let us first check that it is symmetric. Since $A$ is symmetric, so is $A'$. But now

$$\left(A' - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T\right)^T \quad = \quad A'^T - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T \quad = \quad A' - \tfrac{1}{\alpha}\mathbf{a}\mathbf{a}^T,$$

and so $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is indeed symmetric. Now, fix any $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then

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

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

Suppose now that $\alpha > 0$ and $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. We must show that $A$ is positive definite. Fix any $\mathbf{x} \in \mathbb{R}^{n+1}$, and set $\mathbf{x} = \left[\begin{array}{c} x_0 \\ \hline \mathbf{y} \end{array}\right]$, where $x_0 \in \mathbb{R}$ and $\mathbf{y} \in \mathbb{R}^n$. We now compute:

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

where in (*), we used the fact that $x_0\mathbf{y}^T\mathbf{a}$ is a $1 \times 1$ (and consequently symmetric) matrix, and so $x_0\mathbf{y}^T\mathbf{a} = (x_0\mathbf{y}^T\mathbf{a})^T = x_0\mathbf{a}^T\mathbf{y}$; and where for the inequality (**), we used the fact that $\mathbf{y}^T(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{y} \geq 0$, since $A$ is positive definite. It remains to show that the inequality (**) is an equality if and only if $\mathbf{x} = \mathbf{0}$. If $\mathbf{x} = \mathbf{0}$, then $x_0 = 0$ and $\mathbf{y} = \mathbf{0}$, and it is obvious that the inequality (**) is an equality. Suppose now that the inequality (**) is an equality. Then $\mathbf{y}^T(A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T)\mathbf{y} = 0$ and $\frac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{y} + \sqrt{\alpha}x_0 = 0$. The former implies that $\mathbf{y} = \mathbf{0}$ (since $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite). But now since $\frac{1}{\sqrt{\alpha}}\mathbf{a}^T\mathbf{y} + \sqrt{\alpha}x_0 = 0$, we deduce that $x_0 = 0$. So, $\mathbf{x} = \left[\begin{array}{c} x_0 \\ \hdashline \mathbf{y} \end{array}\right] = \mathbf{0}$. This proves that $A$ is positive definite. □

Theorem 4.1 allows us to reduce the problem of checking whether a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is positive definite to the problem of checking whether a $1 \times 1$ matrix (obtained in $n-1$ steps, via the reduction from Theorem 4.1) is positive definite. Obviously, a matrix in $\mathbb{R}^{1\times 1}$ is positive definite if and only if its unique entry is positive.[4]

**Example 4.2.** *Using Theorem 4.1, determine whether the matrix*

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

*is positive definite.*

[4] Indeed, suppose we are given a matrix $A = \left[\, a \,\right]$ in $\mathbb{R}^{n\times n}$. Then for all $x \in \mathbb{R}^1 = \mathbb{R}$, we have that $x^T ax = ax^2$, which has the same sign as $a$.

*Solution.* We apply Theorem 4.1 twice. First, set $\alpha_2 := 4$, $\mathbf{a}_2 := \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, and $A_2' := \begin{bmatrix} 10 & 1 \\ 1 & 6 \end{bmatrix}$, so that $A = \left[\begin{array}{c|c} \alpha_2 & \mathbf{a}_2^T \\ \hline \mathbf{a}_2 & A_2' \end{array}\right]$. We have that $\alpha_2 > 0$, and so by Theorem 4.1, $A$ is positive definite if and only if $A_2 := A_2' - \frac{1}{\alpha_2}\mathbf{a}_2\mathbf{a}_2^T$ is positive definite. We compute

$$\begin{aligned} A_2 &= A_2' - \tfrac{1}{\alpha_2}\mathbf{a}_2\mathbf{a}_2^T \\ &= \begin{bmatrix} 10 & 1 \\ 1 & 6 \end{bmatrix} - \tfrac{1}{4}\begin{bmatrix} -2 \\ 4 \end{bmatrix}\begin{bmatrix} -2 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 9 & 3 \\ 3 & 2 \end{bmatrix}. \end{aligned}$$

Next, set $\alpha_1 := 9$, $\mathbf{a}_1 := \begin{bmatrix} 3 \end{bmatrix}$, and $A_1' := \begin{bmatrix} 2 \end{bmatrix}$, so that $A_2 = \left[\begin{array}{c|c} \alpha_1 & \mathbf{a}_1^T \\ \hline \mathbf{a}_1 & A_1' \end{array}\right]$. We have that $\alpha_1 > 0$, and so by Theorem 4.1, $A_2$ is positive definite if and only if $A_1 := A_1' - \frac{1}{\alpha_1}\mathbf{a}_1\mathbf{a}_1^T$ is positive definite. We compute

$$\begin{aligned} A_1 &= A_1' - \tfrac{1}{\alpha_1}\mathbf{a}_1\mathbf{a}_1^T \\ &= \begin{bmatrix} 2 \end{bmatrix} - \tfrac{1}{9}\begin{bmatrix} 3 \end{bmatrix}\begin{bmatrix} 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 \end{bmatrix}. \end{aligned}$$

Since the only entry of $A_1$ is positive, we see that $A_1$ is positive definite. So, $A$ is positive definite. □

Theorem 4.1 has a corollary (Corollary 4.3 below), which essentially states that we can check for positive definiteness via a modified version of Gaussian elimination (i.e. row reduction). Corollary 4.3 is arguably more convenient to use than Theorem 4.1 itself for checking whether a symmetric matrix is positive definite.

**Corollary 4.3** (Gaussian elimination test of positive definiteness)**.** *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then $A$ is positive definite if and only if the following sequence of $n-1$ steps can be performed and it transforms the matrix $A$ into an upper triangular matrix with a positive main diagonal:*

> *For $j \in \{1, \ldots, n-1\}$:*
>
> > ***Step** $j$**:** For each $i \in \{j+1, \ldots, n\}$, add a suitable scalar multiple of the $j$-th row to the $i$-th row so that the $(i, j)$-th entry of the matrix becomes zero.*

*Proof.* We may assume inductively that the theorem is true for symmetric matrices in $\mathbb{R}^{n' \times n'}$, for all $n' \in \{1, \dots, n-1\}$.

Now, fix a symmetric matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{R}^{n \times n}$. Suppose first that $a_{1,1} \leq 0$. In this case, $A$ is not positive definite (by Proposition 3.1(b)), and our sequence of steps either cannot be performed, or it can be performed but produces a matrix whose main diagonal has at least one negative or zero entry (this is because the $(1,1)$-th entry remains unchanged throughout, and by supposition, $a_{1,1} \leq 0$).

From now on, we may assume that $a_{1,1} > 0$. If $n = 1$, then $A = \left[\ a_{1,1}\ \right]$ is positive definite, and our sequence of $n-1$ steps is empty and produces the matrix $A$ itself, which is indeed in upper triangular with a positive main diagonal.

From now on, we may assume that $n \geq 2$. In this case, Step 1 can be performed, and it consists of the following $n-1$ elementary row operations:

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

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

(This transforms entries $2, \dots, n-1$ of the first column into 0). But note that if we write our original matrix $A$ in the form $A = \left[\begin{array}{c|c} a_{1,1} & \mathbf{a}^T \\ \hline \mathbf{a} & A_{1,1} \end{array}\right]$ (where $\mathbf{a} = \left[\ a_{2,1} \ \ \dots \ \ a_{n,1}\ \right]^T$, and $A_{1,1}$ is the matrix obtained from $A$ by deleting the first row and first column), then the matrix that we obtain after the Step 1 is precisely the matrix

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

By Theorem 4.1, $A$ is positive definite if and only if $A_{1,1} - \frac{1}{a_{1,1}} \mathbf{a}\mathbf{a}^T$ is positive definite. The result now follows immediately from the induction hypothesis. $\square$

**Example 4.4.** *Using Corollary 4.3, determine whether the matrix*

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

*is positive definite.*

*Solution.* We apply the sequence of steps from Corollary 4.3. Step 1 of Corollary 4.3 can be performed: we perform elementary row operations $R_2 \to R_2 - \frac{-2}{4}R_1$ and $R_3 \to R_3 - \frac{4}{4}R_1$, and we obtain the matrix

$$\begin{bmatrix} 4 & -2 & 4 \\ 0 & 9 & 3 \\ 0 & 3 & 2 \end{bmatrix}.$$

Step 2 of Corollary 4.3 can be performed: we perform the elementary row operation $R_3 \to R_3 - \frac{3}{9}R_2$, and we obtain the matrix

$$\begin{bmatrix} 4 & -2 & 4 \\ 0 & 9 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

We have now obtained an upper triangular matrix with a positive main diagonal. So, by Corollary 4.3, $A$ is positive definite. □

Before stating our next theorem, we need some notation. Given any $n \times n$ matrix $A$, and any index $k \in \{1, \dots, n\}$, we let $A^{(k)}$ be the $k \times k$ matrix in the upper left corner of $A$. For example, if

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

then we have that

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

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

**Theorem 4.5** (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.* Suppose first that (i) holds. 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 $\mathbb{R}^k \setminus \{\mathbf{0}\}$. We must show that $\mathbf{x}_k^T A^{(k)} \mathbf{x}_k > 0$. Let $\mathbf{x}$ be the vector in $\mathbb{R}^n$ obtained from $\mathbf{x}_k$ by adding $n-k$ zeros to the bottom of $\mathbf{x}_k$, i.e. $\mathbf{x} := \begin{bmatrix} x_1 & \dots & x_k & 0 & \dots & 0 \end{bmatrix}^T$ (with $n-k$ zeros at the end). Then

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

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. But now Proposition 3.5(b) guarantees that $\det(A^{(k)}) > 0$, and it follows that (ii) holds.

Suppose now that (ii) holds. Since $\det(A^{(1)}) > 0$, we know that $a_{1,1} > 0$. If $n = 1$, it follows that $A = \begin{bmatrix} a_{1,1} \end{bmatrix}$ is positive definite (because its only entry is positive). So, we may assume that $n \geq 2$. We now start performing the steps described in Corollary 4.3, and we proceed until we either complete all $n-1$ steps, or until a step cannot be performed. Since $n \geq 2$ and $a_{1,1} \neq 0$, Step 1 can be performed. Now, let Step $\ell$ be the last step that we perform, and let $B = \begin{bmatrix} b_{i,j} \end{bmatrix}_{n \times n}$ be the matrix that we obtain after performing our $\ell$ steps. We only performed one type of elementary row operation, namely, that of adding a scalar multiple of one row to another. By Theorem 4.2(c) of Lecture Notes 15, this type of elementary row operation does not change the value of the determinant. So, $\det(A) = \det(B)$, and moreover, $\det(A^{(k)}) = \det(B^{(k)})$ for all $k \in \{1, \dots, n\}$.[5] Now $B^{(1)}, \dots, B^{(\ell+1)}$ are all upper triangular matrices with a positive determinant; so, for all $k \in \{1, \dots, \ell+1\}$, we have that $\det(B^{(k)}) = b_{1,1} \dots b_{k,k}$ is positive, and we deduce that $b_{1,1}, \dots, b_{\ell+1,\ell+1}$ are all positive. Thus, $\ell = n-1$, for otherwise, step $\ell+1$ could also be performed, a contradiction. But now $b_{1,1}, \dots, b_{n,n} > 0$, and so by Corollary 4.3, $A$ is positive definite. □

**Example 4.6.** *Using Sylvester's criterion of positive definiteness, determine whether the matrix*

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

*is positive definite.*

*Solution.* First, we have that

- $A^{(1)} = \begin{bmatrix} 4 \end{bmatrix}$;
- $A^{(2)} = \begin{bmatrix} 4 & -2 \\ -2 & 10 \end{bmatrix}$;
- $A^{(3)} = \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}$.

We compute $\det(A^{(1)}) = 4$, $\det(A^{(2)}) = 36$, and $\det(A^{(3)}) = 36$. All three determinants are positive, and so $A$ is positive definite. □

[5]Indeed, for each $k \in \{1, \dots, n\}$, we obtain $B^{(k)}$ from $A^{(k)}$ via the same sequence of elementary row operations that we used to obtain $B$ from $A$, except that we do not perform those operations that involve rows $R_{k+1}, \dots, R_n$; by Theorem 4.2(c) of Lecture Notes 15, $\det(A^{(k)}) = \det(B^{(k)})$.