

13.1 Positive Semidefinite and Positive Definite Matrices

Chapter 12 sorted the operators that have an orthonormal eigenbasis. Among them the self-adjoint ones are the matrix analogue of the real numbers, since they are exactly the operators fixed by starring. Inside the real numbers sits a smaller and even more useful set, the nonnegative ones, and this section finds its matrix analogue. The answer turns out to have three faces — a sign condition on the spectrum, a sign condition on a scalar function, and a factorization — and the whole of the rest of the chapter is built on the fact that the three agree.

Throughout, $F = \mathbb{R}$ or $F = \mathbb{C}$, and F^n carries the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x}$.

13.1.1 Which self-adjoint matrices deserve to be called nonnegative?

A real number t is nonnegative in three equivalent ways: $t \geq 0$; $tx^2 \geq 0$ for every real x ; and $t = s^2$ for some real s . Each of the three has a matrix reading. “Every eigenvalue is ≥ 0 ” copies the first. “ $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$ for every $\mathbf{x}$ ” copies the second, because $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$ is the one scalar a matrix attaches to a single vector. “ $\mathbf{A} = \mathbf{B}^* \mathbf{B}$ ” copies the third, with the star doing what nothing does for numbers but everything does for matrices.

It is not obvious that the three agree, and for a general matrix they do not: they all fail for $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ in different ways. What makes them agree is self-adjointness, and the middle condition is the one to take as the definition, because it is the one that says something about every vector at once.

A matrix is positive semidefinite when its quadratic form never goes negative.

Definition 13.1.1. Positive Semidefinite and Positive Definite

Let $\mathbf{A} \in M_n(F)$. Then $\mathbf{A}$ is **positive semidefinite**, written $\mathbf{A} \succeq \mathbf{O}$, if

- (P1) $\mathbf{A}$ is **Hermitian**, that is $\mathbf{A}^* = \mathbf{A}$ (over $\mathbb{R}$: symmetric, $\mathbf{A}^\top = \mathbf{A}$); and
- (P2) $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$ for **every** $\mathbf{x} \in F^n$.

It is **positive definite**, written $\mathbf{A} \succ \mathbf{O}$, if (P1) holds and $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle > 0$ for every **nonzero** $\mathbf{x} \in F^n$.

The same words apply to an operator $T \in \mathcal{L}(V)$ on a finite-dimensional inner product space: T is positive semidefinite if $T^* = T$ and $\langle Tv, v \rangle \geq 0$ for every $v \in V$, and positive definite if the inequality is strict for $v \neq 0$.

In words: (P1) puts $\mathbf{A}$ in the class where the analogy with real numbers is available at all, and (P2) says that the number $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^* \mathbf{A} \mathbf{x}$ is never negative. The word **every** in (P2) is the whole content; a single vector proves nothing. In the definite case the vector $\mathbf{x} = \mathbf{0}$ has to be excluded, since $\langle \mathbf{A}\mathbf{0}, \mathbf{0} \rangle = 0$ for every $\mathbf{A}$ whatsoever.

Clause (P1) is also what makes clause (P2) meaningful. For a Hermitian $\mathbf{A}$, the number $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$ is **real** for every $\mathbf{x}$, over $\mathbb{C}$ as well as over $\mathbb{R}$ (Proposition 11.6.7), so comparing it with 0 is legitimate. Without (P1) the number can be a genuine complex number, and “ ≥ 0 ” would be

meaningless.

Positive definite implies positive semidefinite, since a condition that holds strictly for $\mathbf{x} \neq \mathbf{0}$ holds weakly for $\mathbf{x} = \mathbf{0}$ too. The reverse fails, and the smallest witness is $\text{diag}(1, 0)$.

Before any examples, the trap.

Warning. Over $\mathbb{R}$, clause (P1) has to be assumed; over $\mathbb{C}$ it comes free. Let $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the quarter turn of $\mathbb{R}^2$. Then $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = -x_2x_1 + x_1x_2 = 0$ for every $\mathbf{x} \in \mathbb{R}^2$, so (P2) holds in the strongest possible way, and yet $\mathbf{A}^\top = -\mathbf{A} \neq \mathbf{A}$. A real matrix can pass the quadratic-form test and be nowhere near symmetric, because over $\mathbb{R}$ the quadratic form cannot see the skew part of a matrix at all (Exercise (C3: What the real quadratic form does determine)). So a real definition that dropped (P1) would call the quarter turn positive semidefinite, and none of the theorems below would survive.

Over $\mathbb{C}$ the situation is the opposite, and it is worth recording once.

Proposition 13.1.1. Over $\mathbb{C}$, positivity implies Hermitian

Let $\mathbf{A} \in M_n(\mathbb{C})$ satisfy $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}$ for every $\mathbf{x} \in \mathbb{C}^n$ — in particular, let $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$ for every $\mathbf{x}$. Then $\mathbf{A}^* = \mathbf{A}$.

Proof. Put $\mathbf{S} = \mathbf{A} - \mathbf{A}^*$. For every $\mathbf{x}$, moving the star across (Lemma 11.4.1) and using conjugate symmetry,

$$\langle \mathbf{A}^*\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle = \overline{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle} = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle,$$

the last equality because the number is real. Hence $\langle \mathbf{S}\mathbf{x}, \mathbf{x} \rangle = 0$ for every $\mathbf{x} \in \mathbb{C}^n$, and Theorem 12.2.6 gives $\mathbf{S} = \mathbf{O}$. ■

So over $\mathbb{C}$ one may state the definition with (P2) alone; over $\mathbb{R}$ one may not. We keep (P1) in the definition for both fields, so that a single sentence covers them and so that no reader has to remember which field licenses which shortcut.

Now the examples, all checked against (P1) and (P2) in turn.

Example. Five Positive Matrices

Decide which of the following are positive semidefinite, and which are positive definite.

- (a) $\mathbf{I}_n$, and more generally $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ with every d_i real.
- (b) $\mathbf{x}\mathbf{x}^*$ for a fixed $\mathbf{x} \in F^n$.
- (c) The Gram matrix $\mathbf{G}$ of a list $(v_1, \dots, v_k)$ in an inner product space.
- (d) The zero matrix $\mathbf{O} \in M_n(F)$.

$$(e) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Solution.

(a) $\mathbf{D}^* = \mathbf{D}$ because the d_i are real, so (P1) holds. For (P2),

$$\langle \mathbf{D}\mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^* \mathbf{D} \mathbf{x} = \sum_{i=1}^n d_i |x_i|^2.$$

So $\mathbf{D} \succeq \mathbf{O}$ exactly when every $d_i \geq 0$, and $\mathbf{D} \succ \mathbf{O}$ exactly when every $d_i > 0$: if some $d_i < 0$, the vector $\mathbf{e}_i$ refutes (P2), and if some $d_i = 0$, the vector $\mathbf{e}_i$ refutes definiteness. In particular $\mathbf{I}_n \succ \mathbf{O}$, with $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$.

(b) $(\mathbf{x}\mathbf{x}^*)^* = \mathbf{x}^{**}\mathbf{x}^* = \mathbf{x}\mathbf{x}^*$, so (P1) holds. For (P2), using $\langle \mathbf{u}, \mathbf{w} \rangle = \mathbf{w}^* \mathbf{u}$ twice,

$$\mathbf{y}^*(\mathbf{x}\mathbf{x}^*)\mathbf{y} = (\mathbf{y}^*\mathbf{x})(\mathbf{x}^*\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle \overline{\langle \mathbf{x}, \mathbf{y} \rangle} = |\langle \mathbf{x}, \mathbf{y} \rangle|^2 \geq 0.$$

So $\mathbf{x}\mathbf{x}^* \succeq \mathbf{O}$. It is never positive definite for $n \geq 2$, since any $\mathbf{y} \neq \mathbf{0}$ orthogonal to $\mathbf{x}$ makes the value 0; and for $\mathbf{x} = \mathbf{0}$ it is the zero matrix.

(c) $\mathbf{G}^* = \mathbf{G}$ by Theorem 11.2.5 (a), and $\mathbf{x}^* \mathbf{G} \mathbf{x} = \|\sum_j x_j v_j\|^2 \geq 0$ by part (b) of the same theorem. So $\mathbf{G} \succeq \mathbf{O}$ always, and by (c) of that theorem $\mathbf{G} \succ \mathbf{O}$ exactly when the list is linearly independent.

(d) $\mathbf{O}^* = \mathbf{O}$ and $\langle \mathbf{O}\mathbf{x}, \mathbf{x} \rangle = 0 \geq 0$, so $\mathbf{O} \succeq \mathbf{O}$. This is the degenerate case, and it matters: it is the bottom of the order introduced in Section 5, and it shows that (P2) can hold with equality everywhere.

(e) Symmetric, so (P1) holds; but $\langle \mathbf{A}\mathbf{e}_2, \mathbf{e}_2 \rangle = -1 < 0$. Not positive semidefinite. This is (a) with $d_2 = -1$.

Example (e) is the nonexample by minimal change: keep (P1), change one diagonal entry, and (P2) fails at one standard basis vector. The quarter turn of the warning is the other kind of nonexample, keeping (P2) and losing (P1).

Two remarks on the name and the convention. “Definite” is the older word, from the theory of quadratic forms: the form $q(\mathbf{x}) = \mathbf{x}^* \mathbf{A} \mathbf{x}$ has a *definite* sign. “Semi” weakens the sign to allow zeros. Some authors write “positive” for what we call positive semidefinite and “strictly positive” for positive definite; the words differ but the two classes are the same two classes everywhere.

Warning. A positive matrix is not a matrix with positive entries. The all-ones matrix $\mathbf{J} \in M_2(\mathbb{R})$ has every entry positive and is positive semidefinite, but $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ has every entry positive and is **not**: at $\mathbf{x} = (1, -1)$ its form is $1 - 2 - 2 + 1 = -2$. Conversely $\text{diag}(1, 1)$ has a zero entry and is positive definite. Entrywise positivity is a different subject, the one Chapter 9 met through stochastic matrices, and it shares only a word with this one.

13.1.2 Positivity survives a change of variable

Before the equivalences, one small result that will be used in almost every proof of this chapter. The quadratic form $\mathbf{x} \mapsto \mathbf{x}^* \mathbf{A} \mathbf{x}$ is attached to $\mathbf{A}$; substituting $\mathbf{x} = \mathbf{S} \mathbf{y}$ turns it into the quadratic form of $\mathbf{S}^* \mathbf{A} \mathbf{S}$. The operation $\mathbf{A} \mapsto \mathbf{S}^* \mathbf{A} \mathbf{S}$ is called **congruence**, and Chapter 14 studies it for its own sake. Here we need only that it preserves the sign.

Proposition 13.1.2. Congruence preserves positivity

Let $\mathbf{A} \in M_n(F)$ be Hermitian and $\mathbf{S} \in M_{n \times m}(F)$. Then $\mathbf{S}^* \mathbf{A} \mathbf{S} \in M_m(F)$ is Hermitian, and:

- (a) if $\mathbf{A} \succeq \mathbf{O}$, then $\mathbf{S}^* \mathbf{A} \mathbf{S} \succeq \mathbf{O}$;
- (b) if $\mathbf{A} \succ \mathbf{O}$ and $\text{nul}(\mathbf{S}) = \{\mathbf{0}\}$, then $\mathbf{S}^* \mathbf{A} \mathbf{S} \succ \mathbf{O}$. In particular, for $m = n$ and $\mathbf{S}$ invertible, $\mathbf{A} \succ \mathbf{O}$ if and only if $\mathbf{S}^* \mathbf{A} \mathbf{S} \succ \mathbf{O}$.

Proof. $(\mathbf{S}^* \mathbf{A} \mathbf{S})^* = \mathbf{S}^* \mathbf{A}^* \mathbf{S}^{**} = \mathbf{S}^* \mathbf{A} \mathbf{S}$, so the matrix is Hermitian. For every $\mathbf{y} \in F^m$,

$$\mathbf{y}^* (\mathbf{S}^* \mathbf{A} \mathbf{S}) \mathbf{y} = (\mathbf{S} \mathbf{y})^* \mathbf{A} (\mathbf{S} \mathbf{y}) = \langle \mathbf{A} (\mathbf{S} \mathbf{y}), \mathbf{S} \mathbf{y} \rangle. \quad (*)$$

- (a) If $\mathbf{A} \succeq \mathbf{O}$, the right-hand side of $(*)$ is ≥ 0 for every $\mathbf{y}$.
- (b) If $\mathbf{A} \succ \mathbf{O}$ and $\text{nul}(\mathbf{S}) = \{\mathbf{0}\}$, then $\mathbf{y} \neq \mathbf{0}$ forces $\mathbf{S} \mathbf{y} \neq \mathbf{0}$, so the right-hand side of $(*)$ is > 0 . For the last claim, take $\mathbf{S}$ invertible: one direction is what we just proved, and the other is the same statement applied to $\mathbf{S}^{-1}$ and the matrix $\mathbf{S}^* \mathbf{A} \mathbf{S}$, since $(\mathbf{S}^{-1})^* (\mathbf{S}^* \mathbf{A} \mathbf{S}) \mathbf{S}^{-1} = \mathbf{A}$.

■

13.1.3 The equivalent descriptions

The characterization needs one determinant identity, which is worth stating on its own because it also describes every coefficient of a characteristic polynomial. Recall from Definition 7.7.1 that $\mathbf{A}_{I,I}$ denotes the submatrix of $\mathbf{A}$ on the rows and columns indexed by I ; its determinant is a **principal minor** of $\mathbf{A}$, and $\mathbf{A}_{I,I}$ is a **principal submatrix**.

Lemma 13.1.3. Shifting by $t\mathbf{I}$, minor by minor

Let $\mathbf{A} \in M_n(F)$ and $t \in F$. Then

$$\det(\mathbf{A} + t\mathbf{I}_n) = \sum_{k=0}^n E_k(\mathbf{A}) t^{n-k},$$

where $E_0(\mathbf{A}) = 1$ and, for $1 \leq k \leq n$, $E_k(\mathbf{A}) = \sum_{|I|=k} \det \mathbf{A}_{I,I}$ is the sum of all $k \times k$ principal minors of $\mathbf{A}$.

Idea. Column j of $\mathbf{A} + t\mathbf{I}$ is $\mathbf{a}_j + t\mathbf{e}_j$, a sum of two columns, and the determinant is linear in each column separately. Expanding all n columns turns one determinant into 2^n of them, one for each choice of which columns contribute their $t\mathbf{e}_j$. A determinant with $\mathbf{e}_j$ in column

j collapses under cofactor expansion to the determinant with row j and column j deleted, and doing that for every chosen column leaves exactly a principal submatrix.

Proof. Write $\mathbf{a}_1, \dots, \mathbf{a}_n$ for the columns of $\mathbf{A}$, so that column j of $\mathbf{A} + t\mathbf{I}$ is $\mathbf{a}_j + t\mathbf{e}_j$. The determinant is linear in each column with the others fixed (Theorem 7.4.3). Applying that in column 1, then in column 2, and so on through column n ,

$$\det(\mathbf{A} + t\mathbf{I}) = \sum_{S \subseteq \{1, \dots, n\}} t^{|S|} \det \mathbf{M}_S,$$

where $\mathbf{M}_S$ is the matrix whose j -th column is $\mathbf{e}_j$ for $j \in S$ and $\mathbf{a}_j$ for $j \notin S$.

Fix S and let $j \in S$. Expanding $\det \mathbf{M}_S$ along column j (Theorem 7.6.1 (a)) leaves one term, since that column is $\mathbf{e}_j$: the entry is 1 in row j and 0 elsewhere, and its cofactor is $(-1)^{j+j} M_{jj} = M_{jj}$. So $\det \mathbf{M}_S$ equals the determinant of $\mathbf{M}_S$ with row j and column j deleted, which is a matrix of the same shape on the index set $\{1, \dots, n\} \setminus \{j\}$. Repeating this for each $j \in S$ deletes every row and column indexed by S and leaves

$$\det \mathbf{M}_S = \det \mathbf{A}_{S^c, S^c}, \quad S^c = \{1, \dots, n\} \setminus S.$$

Collecting the subsets by $k = |S^c|$, so that $|S| = n - k$, gives the stated formula. ■

Remark. Running the same expansion over $F[x]$ with the indeterminate $-x$ in place of t , and multiplying by $(-1)^n$, turns this into $p_{\mathbf{A}}(x) = \sum_k (-1)^k E_k(\mathbf{A}) x^{n-k}$: the coefficients of the characteristic polynomial are, up to sign, the sums of principal minors. The cases $k = 1$ and $k = n$ recover $\text{tr } \mathbf{A}$ and $\det \mathbf{A}$ from Theorem 7.10.2.

Here is the main theorem of the section.

Theorem 13.1.4. Characterizations of Positive Semidefiniteness

Let $\mathbf{A} \in M_n(F)$ be Hermitian. The following are equivalent.

- (a) $\mathbf{A} \succeq \mathbf{O}$.
- (b) Every eigenvalue of $\mathbf{A}$ is a nonnegative real number.
- (c) $\mathbf{A} = \mathbf{B}^* \mathbf{B}$ for some $\mathbf{B} \in M_n(F)$.
- (d) $\mathbf{A} = \mathbf{C}^2$ for exactly one $\mathbf{C} \in M_n(F)$ with $\mathbf{C} \succeq \mathbf{O}$.
- (e) Every principal minor of $\mathbf{A}$ is ≥ 0 .

Idea. The spectral theorem makes (a) and (b) two readings of one diagonal matrix, since an orthonormal change of basis does not change the sign of a quadratic form. From (b) the square root of Chapter 12 produces $\mathbf{C}$, and $\mathbf{C}$ is its own $\mathbf{B}$. The step (c) $\Rightarrow$ (a) is the identity $\langle \mathbf{B}^* \mathbf{B} \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{B} \mathbf{x}\|^2$, one line of moving $\mathbf{B}^*$ across. For (e), a principal submatrix is what the form does to vectors supported on a few coordinates, which gives (a) $\Rightarrow$ (e) at once; the return trip uses Lemma 13.1.3 to show that $\det(\mathbf{A} + t\mathbf{I}) > 0$ for every $t > 0$, so that no $-t$

can be an eigenvalue.

Proof. (a) $\Leftrightarrow$ (b). ($\Rightarrow$) Let $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. By Theorem 12.2.1, $\lambda \in \mathbb{R}$, and

$$0 \leq \langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle = \langle \lambda\mathbf{v}, \mathbf{v} \rangle = \lambda \|\mathbf{v}\|^2,$$

so $\lambda \geq 0$, as $\|\mathbf{v}\|^2 > 0$. ($\Leftarrow$) By Corollary 12.4.3 (over $\mathbb{C}$) or Corollary 12.5.2 (over $\mathbb{R}$), $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$ the eigenvalues. Put $\mathbf{y} = \mathbf{U}^*\mathbf{x}$. Then

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^* \mathbf{U} \mathbf{D} \mathbf{U}^* \mathbf{x} = \mathbf{y}^* \mathbf{D} \mathbf{y} = \sum_i \lambda_i |y_i|^2 \geq 0,$$

every term being a product of two nonnegative reals.

(b) $\Rightarrow$ (d). A Hermitian matrix has an orthonormal basis of eigenvectors, so the operator $T_{\mathbf{A}}$ on F^n satisfies the hypothesis of Theorem 12.8.4, whose spectral condition $\text{spec}(T_{\mathbf{A}}) \subseteq [0, \infty)$ is exactly (b). That theorem gives a unique $\mathbf{C}$ with an orthonormal eigenbasis, nonnegative spectrum, and $\mathbf{C}^2 = \mathbf{A}$, and it is self-adjoint; by the equivalence just proved, applied to $\mathbf{C}$, this says $\mathbf{C} \succeq \mathbf{O}$. Conversely any $\mathbf{C} \succeq \mathbf{O}$ with $\mathbf{C}^2 = \mathbf{A}$ is Hermitian with nonnegative spectrum, hence is that one. So there is exactly one such $\mathbf{C}$.

(d) $\Rightarrow$ (c). Take $\mathbf{B} = \mathbf{C}$; then $\mathbf{B}^* \mathbf{B} = \mathbf{C}^* \mathbf{C} = \mathbf{C}^2 = \mathbf{A}$.

(c) $\Rightarrow$ (a). Conjugate-transposing twice, $(\mathbf{B}^* \mathbf{B})^* = \mathbf{B}^* (\mathbf{B}^*)^* = \mathbf{B}^* \mathbf{B}$, which is (P1). And for every $\mathbf{x}$,

$$\langle \mathbf{B}^* \mathbf{B} \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^* \mathbf{B}^* \mathbf{B} \mathbf{x} = (\mathbf{B} \mathbf{x})^* (\mathbf{B} \mathbf{x}) = \|\mathbf{B} \mathbf{x}\|^2 \geq 0,$$

which is (P2). Both facts are recorded in operator form in Exercise (C1: The operator S^*S) (a), applied to the map $\mathbf{x} \mapsto \mathbf{B}\mathbf{x}$.

(a) $\Rightarrow$ (e). Let $I = \{i_1 < \dots < i_k\}$. Given $\mathbf{y} \in F^k$, let $\mathbf{x} \in F^n$ have $x_{i_p} = y_p$ and all other entries 0. Then $\mathbf{x}^* \mathbf{A} \mathbf{x} = \mathbf{y}^* \mathbf{A}_{I,I} \mathbf{y}$, because the terms $\overline{x_i} a_{ij} x_j$ with $i \notin I$ or $j \notin I$ vanish. So $\mathbf{A}_{I,I}$ is Hermitian and satisfies (P2), that is $\mathbf{A}_{I,I} \succeq \mathbf{O}$. By the equivalence (a) $\Leftrightarrow$ (b) applied to $\mathbf{A}_{I,I}$, all its eigenvalues are ≥ 0 . Diagonalizing it as $\mathbf{A}_{I,I} = \mathbf{U} \mathbf{D} \mathbf{U}^*$ as above, with $\mathbf{U}^* = \mathbf{U}^{-1}$, and taking determinants (Theorem 7.5.4, Corollary 7.5.6),

$$\det \mathbf{A}_{I,I} = \det \mathbf{D} = \lambda_1 \cdots \lambda_k \geq 0.$$

(e) $\Rightarrow$ (b). Let $t > 0$ be real. Every $E_k(\mathbf{A}) \geq 0$ by hypothesis and $E_0(\mathbf{A}) = 1$, so Lemma 13.1.3 gives

$$\det(\mathbf{A} + t\mathbf{I}) = \sum_{k=0}^n E_k(\mathbf{A}) t^{n-k} \geq t^n > 0.$$

In particular $\mathbf{A} + t\mathbf{I}$ is invertible (Theorem 7.5.3), so $-t$ is not an eigenvalue of $\mathbf{A}$ (Theorem 9.2.6 applied to $\mathbf{A} + t\mathbf{I}$). As $t > 0$ was arbitrary and every eigenvalue of $\mathbf{A}$ is real (Theorem 12.2.1), no eigenvalue is negative. This closes the cycle and proves the theorem. ■

Item (c) is the one to keep in mind: **every positive semidefinite matrix is a Gram matrix**, since $\mathbf{B}^* \mathbf{B}$ is the Gram matrix of the columns of $\mathbf{B}$ (Definition 11.2.3). Section 3 turns that remark into a theorem. Item (d) names the square root, which Section 2 develops. Item (b) is

how one *thinks*; item (e) is, for small matrices, how one *computes*.

13.1.4 The definite case, and Sylvester's criterion

The definite versions of (a), (b) and (c) are proved by putting “strict” in front of the same arguments. The new item is a test that can be run by hand on a numerical matrix without finding a single eigenvalue.

Theorem 13.1.5. Characterizations of Positive Definiteness, with Sylvester's Criterion

Let $\mathbf{A} \in M_n(F)$ be Hermitian, and for $1 \leq k \leq n$ let $\mathbf{A}_k$ be its leading principal submatrix, the top-left $k \times k$ corner. The following are equivalent.

- (a) $\mathbf{A} \succ \mathbf{O}$.
- (b) Every eigenvalue of $\mathbf{A}$ is a positive real number.
- (c) $\mathbf{A} = \mathbf{B}^* \mathbf{B}$ for some **invertible** $\mathbf{B} \in M_n(F)$.
- (d) (**Sylvester's criterion**) $\det \mathbf{A}_k > 0$ for every $k = 1, \dots, n$.

Idea. Steps (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) repeat the previous proof with strict inequalities. For (a) $\Rightarrow$ (d), a leading principal submatrix inherits definiteness, and a determinant is a product of eigenvalues. The content is (d) $\Rightarrow$ (a), and it is an induction on n that peels off the last row and column. The inductive hypothesis makes the corner $\mathbf{A}_{n-1}$ definite, hence invertible, and one shear — the congruence by $\mathbf{S} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{n-1}^{-1} \mathbf{b} \\ \mathbf{0}^\top & 1 \end{pmatrix}$, which is exactly the block elimination of Chapter 8 — clears the border and turns $\mathbf{A}$ into $\mathbf{A}_{n-1} \oplus (s)$ with one scalar s left over. Determinants identify the sign of s , and [Proposition 13.1.2](#) carries the verdict back.

Proof. (a) $\Leftrightarrow$ (b). As in [Theorem 13.1.4](#), with $\geq$ replaced by $>$ throughout: for ($\Rightarrow$), $0 < \lambda \|\mathbf{v}\|^2$ forces $\lambda > 0$; for ($\Leftarrow$), $\sum_i \lambda_i |y_i|^2 > 0$ whenever $\mathbf{y} = \mathbf{U}^* \mathbf{x} \neq \mathbf{0}$, which happens exactly when $\mathbf{x} \neq \mathbf{0}$ because $\mathbf{U}^*$ is invertible.

(b) $\Rightarrow$ (c). Let $\mathbf{C} \succeq \mathbf{O}$ with $\mathbf{C}^2 = \mathbf{A}$ ([Theorem 13.1.4](#) (d)). Since every eigenvalue of $\mathbf{A}$ is positive, $0 \notin \text{spec}(\mathbf{A})$ and $\mathbf{A}$ is invertible ([Theorem 9.2.6](#)). If $\mathbf{C}\mathbf{x} = \mathbf{0}$ then $\mathbf{A}\mathbf{x} = \mathbf{C}^2\mathbf{x} = \mathbf{0}$, so $\mathbf{x} = \mathbf{0}$; hence $\mathbf{C}$ is invertible too ([Theorem 3.4.4](#)). Take $\mathbf{B} = \mathbf{C}$.

(c) $\Rightarrow$ (a). For $\mathbf{x} \neq \mathbf{0}$, $\mathbf{B}\mathbf{x} \neq \mathbf{0}$, so $\langle \mathbf{B}^* \mathbf{B} \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{B}\mathbf{x}\|^2 > 0$.

(a) $\Rightarrow$ (d). As in the proof of (a) $\Rightarrow$ (e) of [Theorem 13.1.4](#) with $I = \{1, \dots, k\}$, the corner $\mathbf{A}_k$ satisfies $\mathbf{y}^* \mathbf{A}_k \mathbf{y} = \mathbf{x}^* \mathbf{A} \mathbf{x} > 0$ for $\mathbf{y} \neq \mathbf{0}$, so $\mathbf{A}_k \succ \mathbf{O}$. Its eigenvalues are positive, and $\det \mathbf{A}_k$ is their product, computed as in the proof of [Theorem 13.1.4](#), so $\det \mathbf{A}_k > 0$.

(d) $\Rightarrow$ (a). Induct on n . For $n = 1$, $\mathbf{A} = (a_{11})$ with $a_{11} = \det \mathbf{A}_1 > 0$, and $\langle \mathbf{A}x, x \rangle = a_{11}|x|^2 > 0$ for $x \neq 0$.

Let $n \geq 2$, and write

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}' & \mathbf{b} \\ \mathbf{b}^* & d \end{pmatrix}, \quad \mathbf{A}' = \mathbf{A}_{n-1}, \quad \mathbf{b} \in F^{n-1}, \quad d \in \mathbb{R}.$$

The leading principal submatrices of $\mathbf{A}'$ are $\mathbf{A}_1, \dots, \mathbf{A}_{n-1}$, so the inductive hypothesis gives

$\mathbf{A}' \succ \mathbf{O}$; in particular $\mathbf{A}'$ is invertible, since $\det \mathbf{A}' = \det \mathbf{A}_{n-1} > 0$. Put

$$\mathbf{S} = \begin{pmatrix} \mathbf{I}_{n-1} & -\mathbf{A}'^{-1}\mathbf{b} \\ \mathbf{0}^\top & 1 \end{pmatrix}, \quad s = d - \mathbf{b}^* \mathbf{A}'^{-1} \mathbf{b}.$$

Block multiplication (Theorem 8.1.1), together with $(\mathbf{A}'^{-1})^* = \mathbf{A}'^{-1}$ because $\mathbf{A}'$ is Hermitian, gives

$$\mathbf{S}^* \mathbf{A} \mathbf{S} = \begin{pmatrix} \mathbf{A}' & \mathbf{0} \\ \mathbf{0}^\top & s \end{pmatrix}. \quad (*)$$

Both $\mathbf{S}$ and $\mathbf{S}^*$ are triangular with 1s on the diagonal, so $\det \mathbf{S} = \det \mathbf{S}^* = 1$ by Theorem 7.4.6. Taking determinants in $(*)$ with Theorem 7.5.4 and Theorem 7.8.4,

$$0 < \det \mathbf{A} = \det(\mathbf{S}^* \mathbf{A} \mathbf{S}) = s \det \mathbf{A}',$$

and $\det \mathbf{A}' > 0$, so $s > 0$. The right-hand side of $(*)$ is therefore positive definite: for $\mathbf{y} = (\mathbf{y}', y_n) \neq \mathbf{0}$,

$$\mathbf{y}^* \begin{pmatrix} \mathbf{A}' & \mathbf{0} \\ \mathbf{0}^\top & s \end{pmatrix} \mathbf{y} = \mathbf{y}'^* \mathbf{A}' \mathbf{y}' + s|y_n|^2 > 0,$$

since the two terms are ≥ 0 and not both zero. Finally $\mathbf{S}$ is invertible, so [Proposition 13.1.2](#) (b) turns $\mathbf{S}^* \mathbf{A} \mathbf{S} \succ \mathbf{O}$ back into $\mathbf{A} \succ \mathbf{O}$. This completes the induction and proves the theorem. ■

The number $s = d - \mathbf{b}^* \mathbf{A}'^{-1} \mathbf{b}$ in that proof is the Schur complement $\mathbf{A}/\mathbf{A}'$ of Definition 8.3.1, and $(*)$ is the block $\mathbf{LDU}$ factorization of Theorem 8.3.1 written as a congruence. Section 6 replaces the scalar s by a block and gets the general test.

Warning. For semidefiniteness, the leading principal minors are not enough. Sylvester's criterion has no “ $\geq$ ” version. Take

$$\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

Its leading principal minors are $\det \mathbf{A}_1 = 0$ and $\det \mathbf{A}_2 = 0$, both ≥ 0 , and yet $\langle \mathbf{A} \mathbf{e}_2, \mathbf{e}_2 \rangle = -1$, so $\mathbf{A}$ is not positive semidefinite. What [Theorem 13.1.4](#) (e) demands is **all** principal minors, and the one this matrix fails is $\det \mathbf{A}_{\{2\},\{2\}} = -1 < 0$. The leading corners simply never look at the second coordinate on its own. This is the single most common error about the subject: for the strict test n minors suffice, for the weak test one needs all $2^n - 1$ of them.

Quick check. Is $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ positive semidefinite? Is it positive definite? Which minors did you use?

Solution. It is positive semidefinite but not positive definite. The principal minors are 1, 4, and $\det \mathbf{A} = 4 - 4 = 0$, all ≥ 0 , so [Theorem 13.1.4](#) (e) applies. It is not definite, since $\det \mathbf{A}_2 = 0$ violates Sylvester's criterion; concretely $\mathbf{A}(2, -1) = \mathbf{0}$, so $\langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle = 0$ at the nonzero $\mathbf{x} = (2, -1)$. One can also see everything at once from $\mathbf{A} = \mathbf{b} \mathbf{b}^\top$ with $\mathbf{b} = (1, 2)$,

which is [Example \(Five Positive Matrices\)](#) (b).

13.1.5 Two by two, and where these matrices come from

The 2×2 case is worth memorizing, because it is the one that appears inside larger arguments.

Example. The 2×2 Test

Let $\mathbf{A} = \begin{pmatrix} a & b \\ \bar{b} & d \end{pmatrix}$ be Hermitian, so $a, d \in \mathbb{R}$ and $b \in F$. State and justify the tests for $\mathbf{A} \succ \mathbf{O}$ and for $\mathbf{A} \succeq \mathbf{O}$.

Solution. $\mathbf{A} \succ \mathbf{O}$ if and only if $a > 0$ and $ad - |b|^2 > 0$. This is Sylvester's criterion, [Theorem 13.1.5](#) (d): the two leading principal minors are a and $\det \mathbf{A} = ad - |b|^2$.

$\mathbf{A} \succeq \mathbf{O}$ if and only if $a \geq 0$, $d \geq 0$, and $ad - |b|^2 \geq 0$. This is [Theorem 13.1.4](#) (e): the principal minors are a , d , and $\det \mathbf{A}$, and all three are needed. Dropping $d \geq 0$ readmits $\text{diag}(0, -1)$.

Note what the second condition says: $|b|^2 \leq ad$, so an off-diagonal entry can never be larger in modulus than the geometric mean of the two diagonal entries it sits between. For $\mathbf{A} = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}$ over $\mathbb{C}$: $a = 2 > 0$ and $ad - |b|^2 = 6 - 2 = 4 > 0$, so $\mathbf{A} \succ \mathbf{O}$.

Where do positive semidefinite matrices come from? Almost always from item (c) of [Theorem 13.1.4](#), and the standard source is data.

Example. A Covariance-style Matrix

Three quantities are measured on four occasions, and each column of

$$\mathbf{X} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 1 & 1 \\ -1 & -1 & 0 \\ 0 & -1 & -1 \end{pmatrix} \in M_{4 \times 3}(\mathbb{R})$$

has been adjusted to sum to 0. Show that $\mathbf{A} = \mathbf{X}^\top \mathbf{X}$ is positive definite, and check the verdict against Sylvester's criterion.

Solution. Multiplying out, a_{ij} is the dot product of columns i and j of $\mathbf{X}$:

$$\mathbf{A} = \begin{pmatrix} 6 & 2 & -1 \\ 2 & 4 & 2 \\ -1 & 2 & 2 \end{pmatrix}.$$

The columns of $\mathbf{X}$ are linearly independent: if $\mathbf{X}\mathbf{x} = \mathbf{0}$, rows 1 and 3 read $2x_1 + x_2 = 0$ and

$-x_1 - x_2 = 0$, which force $x_1 = x_2 = 0$, and then row 2 reads $x_3 = 0$. So $\text{nul}(\mathbf{X}) = \{\mathbf{0}\}$, and taking $\mathbf{S} = \mathbf{X}$ in [Proposition 13.1.2](#) (b) with the positive definite matrix $\mathbf{I}_4$ gives $\mathbf{A} = \mathbf{X}^\top \mathbf{I}_4 \mathbf{X} \succ \mathbf{O}$. Equivalently, $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \|\mathbf{X}\mathbf{x}\|^2 > 0$ for $\mathbf{x} \neq \mathbf{0}$.

Sylvester's criterion agrees: $\det \mathbf{A}_1 = 6$, $\det \mathbf{A}_2 = 24 - 4 = 20$, and expanding along the first row,

$$\det \mathbf{A} = 6(8 - 4) - 2(4 + 2) + (-1)(4 + 4) = 24 - 12 - 8 = 4,$$

all positive. The reason the two routes must agree is that $\mathbf{A}$ is the Gram matrix of the columns of $\mathbf{X}$, and Theorem 11.2.5 (d) already said that its determinant is positive exactly when those columns are independent.

13.1.6 The invertible matrix theorem grows

Chapter 2 collected the conditions equivalent to invertibility for an operator (Theorem 2.5.7); Chapter 3 gave the eight matrix conditions of the Invertible Matrix Theorem (Theorem 3.4.4); Chapter 7 added the determinant (Theorem 7.5.3) and Chapter 9 the eigenvalue 0 (Theorem 9.2.6). Positivity adds a pair, and they are the ones used whenever a square system is attacked through least squares.

Theorem 13.1.6. Invertible Matrix Theorem, with Positive Definiteness

Let $\mathbf{A} \in M_n(F)$ with $F = \mathbb{R}$ or $F = \mathbb{C}$. The following are equivalent, and each is equivalent to each of the conditions of Theorem 3.4.4, Theorem 7.5.3, and Theorem 9.2.6.

- (a) $\mathbf{A}$ is invertible.
- (b) $\mathbf{A}^* \mathbf{A} \succ \mathbf{O}$.
- (c) $\mathbf{A} \mathbf{A}^* \succ \mathbf{O}$.

For a general $\mathbf{A} \in M_{m \times n}(F)$, one always has $\mathbf{A}^* \mathbf{A} \succeq \mathbf{O}$, with $\mathbf{A}^* \mathbf{A} \succ \mathbf{O}$ if and only if the columns of $\mathbf{A}$ are linearly independent.

Proof. Let $\mathbf{A} \in M_{m \times n}(F)$. Then $\mathbf{A}^* \mathbf{A}$ is positive semidefinite by [Theorem 13.1.4](#) ((c) $\Rightarrow$ (a)). Whether it is definite is decided by the identity $\langle \mathbf{A}^* \mathbf{A} \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^* \mathbf{A}^* \mathbf{A} \mathbf{x} = \|\mathbf{A} \mathbf{x}\|^2$, computed in the proof of [Theorem 13.1.4](#) ((c) $\Rightarrow$ (a)) and recorded in operator form in Exercise (C1: The operator $S^* S$) (a): the left side is > 0 for every $\mathbf{x} \neq \mathbf{0}$ exactly when $\mathbf{A} \mathbf{x} \neq \mathbf{0}$ for every $\mathbf{x} \neq \mathbf{0}$, that is, exactly when $\text{nul}(\mathbf{A}) = \{\mathbf{0}\}$, that is, exactly when the columns of $\mathbf{A}$ are linearly independent. The same conclusion comes from Lemma 11.4.3, which says $\text{nul}(\mathbf{A}^* \mathbf{A}) = \text{nul}(\mathbf{A})$ and so lets one read the condition off $\mathbf{A}$ directly.

Now let $\mathbf{A}$ be square. Then “ $\text{nul}(\mathbf{A}) = \{\mathbf{0}\}$ ” is item (b) of Theorem 3.4.4, which gives (a) $\Leftrightarrow$ (b). Also $\mathbf{A}$ is invertible if and only if $\mathbf{A}^*$ is, since $(\mathbf{A}^{-1})^*$ is a two-sided inverse of $\mathbf{A}^*$ and conversely; applying (a) $\Leftrightarrow$ (b) to $\mathbf{A}^*$ therefore gives (a) $\Leftrightarrow$ (c). ■

The two new items look like a detour — why test $\mathbf{A}^* \mathbf{A}$ when $\mathbf{A}$ itself is available? Because $\mathbf{A}^* \mathbf{A}$ is Hermitian and positive, and everything in this chapter applies to it. Chapter 11 already used it once, in the normal equations of least squares; Section 9 will use it again, to build the singular value decomposition of a matrix that is not even square, and that decomposition supplies the

last item of the list: a square matrix is invertible exactly when none of its singular values is zero.

13.1.7 Exercises

13.1.7.1 A. Check your understanding

Exercise. A1

- State the definition of $\mathbf{A} \succeq \mathbf{O}$ and of $\mathbf{A} \succ \mathbf{O}$, including every quantifier and the field.
- Explain why clause (P1) cannot be dropped from the real definition, with a concrete matrix.
- True or false: if $\mathbf{A}$ is Hermitian and every leading principal minor of $\mathbf{A}$ is ≥ 0 , then $\mathbf{A} \succeq \mathbf{O}$. Justify your answer.
- Write down the three conditions of the 2×2 test for $\mathbf{A} \succeq \mathbf{O}$.
- Name the item of [Theorem 13.1.4](#) that says every positive semidefinite matrix is a Gram matrix.

Solution.

- $\mathbf{A} \in M_n(F)$ with $F = \mathbb{R}$ or $\mathbb{C}$ is positive semidefinite if $\mathbf{A}^* = \mathbf{A}$ and $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$ for every $\mathbf{x} \in F^n$; positive definite if $\mathbf{A}^* = \mathbf{A}$ and $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle > 0$ for every **nonzero** $\mathbf{x} \in F^n$ ([Definition 13.1.1](#)).
- Because over $\mathbb{R}$ the quadratic form does not see the skew part. The quarter turn $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = 0$ for every $\mathbf{x} \in \mathbb{R}^2$ and is not symmetric.
- False. $\text{diag}(0, -1)$ has leading principal minors 0 and 0, and $\langle \mathbf{A}\mathbf{e}_2, \mathbf{e}_2 \rangle = -1$. All principal minors are needed ([Theorem 13.1.4](#) (e)).
- With $\mathbf{A} = \begin{pmatrix} a & b \\ \bar{b} & d \end{pmatrix}$ Hermitian: $a \geq 0$, $d \geq 0$, and $ad - |b|^2 \geq 0$.
- Item (c): $\mathbf{A} = \mathbf{B}^*\mathbf{B}$, and $\mathbf{B}^*\mathbf{B}$ is the Gram matrix of the columns of $\mathbf{B}$.

13.1.7.2 B. Practice

Exercise. B1: Determine the definiteness

For each matrix below, determine whether it is positive definite, positive semidefinite but not positive definite, or neither. Justify your answer.

- $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ over $\mathbb{R}$.
- $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ over $\mathbb{R}$.
- $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ over $\mathbb{R}$.
- The all-ones matrix $\mathbf{J} \in M_3(\mathbb{R})$.

- (e) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ over $\mathbb{R}$.
- (f) $\begin{pmatrix} 2 & 1+i \\ 1-i & 1 \end{pmatrix}$ over $\mathbb{C}$.

Solution.

- (a) Positive definite. Symmetric, with leading principal minors 2 and $4 - 1 = 3$, both positive ([Theorem 13.1.5](#) (d)).
- (b) Neither. The leading principal minors are 0 and 0, which tempts one into “semidefinite”, but the principal minor $\det \mathbf{A}_{\{2\},\{2\}} = -1$ is negative, and $\langle \mathbf{A}\mathbf{e}_2, \mathbf{e}_2 \rangle = -1 < 0$.
- (c) Positive definite. Symmetric, with leading principal minors 2, $4 - 1 = 3$, and $2 \cdot 3 - (-1)(-2) = 4$, all positive.
- (d) Positive semidefinite, not positive definite. $\mathbf{J} = \mathbf{1}\mathbf{1}^\top$, which is [Example \(Five Positive Matrices\)](#) (b) with $\mathbf{x} = \mathbf{1}$; so $\langle \mathbf{J}\mathbf{x}, \mathbf{x} \rangle = |\langle \mathbf{x}, \mathbf{1} \rangle|^2 \geq 0$. It vanishes at the nonzero $\mathbf{x} = (1, -1, 0)$, so it is not definite. (Equivalently, $\det \mathbf{J} = 0$.)
- (e) Neither: it is not symmetric, so (P1) fails and it is not in the class at all. Its quadratic form is identically 0, which shows how little the form sees over $\mathbb{R}$.
- (f) Positive semidefinite, not positive definite. Hermitian, with $a = 2 \geq 0$, $d = 1 \geq 0$, and $ad - |b|^2 = 2 - 2 = 0$. The determinant is 0, so definiteness fails; indeed $\mathbf{A}(1+i, -2) = \mathbf{0}$.

Exercise. B2: A parameter

For which $t \in \mathbb{R}$ is

$$\mathbf{A}(t) = \begin{pmatrix} 1 & t & 0 \\ t & 2 & t \\ 0 & t & 3 \end{pmatrix}$$

positive definite? Justify your answer.

Solution. $\mathbf{A}(t)$ is symmetric for every real t , so Sylvester’s criterion applies ([Theorem 13.1.5](#) (d)). The leading principal minors are

$$\det \mathbf{A}_1 = 1, \quad \det \mathbf{A}_2 = 2 - t^2, \quad \det \mathbf{A}_3 = 6 - 4t^2,$$

the last by expansion along the first row: $1 \cdot (6 - t^2) - t \cdot (3t - 0) = 6 - 4t^2$. All three are positive exactly when $t^2 < 2$ and $t^2 < 3/2$, that is when $t^2 < 3/2$. Hence $\mathbf{A}(t) \succ \mathbf{O}$ if and only if $|t| < \sqrt{6}/2$.

At $t = \sqrt{6}/2$ the matrix is positive semidefinite but singular, and for $|t| > \sqrt{6}/2$ the determinant is negative, so $\mathbf{A}(t)$ has an odd number of negative eigenvalues and is not even semidefinite.

Exercise. B3: The inverse of a positive definite matrix

Let $\mathbf{A} \in M_n(F)$ with $\mathbf{A} \succ \mathbf{O}$. Prove that $\mathbf{A}$ is invertible and that $\mathbf{A}^{-1} \succ \mathbf{O}$.

Solution. By [Theorem 13.1.5](#) (b), every eigenvalue of $\mathbf{A}$ is positive, so $0 \notin \text{spec}(\mathbf{A})$ and $\mathbf{A}$ is invertible (Theorem 9.2.6). Then $(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1} = \mathbf{A}^{-1}$, so $\mathbf{A}^{-1}$ is Hermitian. Apply [Proposition 13.1.2](#) (b) with $\mathbf{S} = \mathbf{A}^{-1}$, which is invertible:

$$\mathbf{S}^* \mathbf{A} \mathbf{S} = \mathbf{A}^{-1} \mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \succ \mathbf{O}.$$

(Alternatively: the eigenvalues of $\mathbf{A}^{-1}$ are the reciprocals $1/\lambda_i$ of those of $\mathbf{A}$, hence positive, and [Theorem 13.1.5](#) (b) applies again.)

13.1.7.3 C. Going deeper

Exercise. C1: Zero trace forces zero

Let $\mathbf{A} \in M_n(F)$ with $\mathbf{A} \succeq \mathbf{O}$.

- (a) Prove that $a_{ii} \geq 0$ for every i , and that $\text{tr } \mathbf{A} \geq 0$.
- (b) Prove that if $\text{tr } \mathbf{A} = 0$ then $\mathbf{A} = \mathbf{O}$.
- (c) Deduce that if $\mathbf{A} \succeq \mathbf{O}$ and $-\mathbf{A} \succeq \mathbf{O}$, then $\mathbf{A} = \mathbf{O}$.

Solution.

- (a) $a_{ii} = \mathbf{e}_i^* \mathbf{A} \mathbf{e}_i = \langle \mathbf{A} \mathbf{e}_i, \mathbf{e}_i \rangle \geq 0$ by (P2). Summing, $\text{tr } \mathbf{A} = \sum_i a_{ii} \geq 0$.
- (b) Write $\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$ the eigenvalues (Corollary 12.4.3, Corollary 12.5.2). Then $\text{tr } \mathbf{A} = \text{tr}(\mathbf{D} \mathbf{U}^* \mathbf{U}) = \text{tr } \mathbf{D} = \sum_i \lambda_i$ by Theorem 0.9.6, and each $\lambda_i \geq 0$ by [Theorem 13.1.4](#) (b). A sum of nonnegative reals is 0 only if every term is 0, so $\mathbf{D} = \mathbf{O}$ and hence $\mathbf{A} = \mathbf{O}$.
Alternatively, without eigenvalues: by (a) each $a_{ii} \geq 0$, so $\text{tr } \mathbf{A} = 0$ forces every $a_{ii} = 0$. Writing $\mathbf{A} = \mathbf{B}^* \mathbf{B}$ ([Theorem 13.1.4](#) (c)), the i -th diagonal entry is $\|\mathbf{b}_i\|^2$ for the i -th column $\mathbf{b}_i$ of $\mathbf{B}$; so every column of $\mathbf{B}$ is $\mathbf{0}$, that is $\mathbf{B} = \mathbf{O}$ and $\mathbf{A} = \mathbf{O}$.
- (c) If both $\mathbf{A} \succeq \mathbf{O}$ and $-\mathbf{A} \succeq \mathbf{O}$, then by (a) $\text{tr } \mathbf{A} \geq 0$ and $\text{tr}(-\mathbf{A}) = -\text{tr } \mathbf{A} \geq 0$, so $\text{tr } \mathbf{A} = 0$ and (b) gives $\mathbf{A} = \mathbf{O}$.

Exercise. C2: The cone of positive matrices

Let $\mathcal{P}_n \subseteq M_n(F)$ be the set of positive semidefinite matrices.

- (a) Prove that if $\mathbf{A}, \mathbf{B} \in \mathcal{P}_n$ and $c \geq 0$ is real, then $\mathbf{A} + \mathbf{B} \in \mathcal{P}_n$ and $c\mathbf{A} \in \mathcal{P}_n$. (A set with these two properties is a **convex cone**: it contains $(1-t)\mathbf{A} + t\mathbf{B}$ for $0 \leq t \leq 1$ as well.)
- (b) Prove that $\mathcal{P}_n$ contains no line through $\mathbf{O}$: if $\mathbf{A} \in \mathcal{P}_n$ and $-\mathbf{A} \in \mathcal{P}_n$ then $\mathbf{A} = \mathbf{O}$.
- (c) Prove that $\mathcal{P}_n$ is closed under congruence: if $\mathbf{A} \in \mathcal{P}_n$ and $\mathbf{S} \in M_n(F)$, then $\mathbf{S}^* \mathbf{A} \mathbf{S} \in \mathcal{P}_n$. Give an example showing that $\mathcal{P}_n$ is **not** closed under multiplication.

Hint for (c): for the example, two positive semidefinite matrices whose product is not even Hermitian will do.

Solution.

- (a) Sums and real multiples of Hermitian matrices are Hermitian. For every $\mathbf{x}$,

$$\langle (\mathbf{A} + \mathbf{B})\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle \geq 0,$$

and $\langle c\mathbf{A}\mathbf{x}, \mathbf{x} \rangle = c\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$ since $c \geq 0$. Convexity follows by taking $c = 1 - t$ and $c = t$ and adding.

- (b) This is [Exercise \(C1: Zero trace forces zero\)](#) (c).
 (c) Immediate from [Proposition 13.1.2](#) (a). For the failure of multiplicativity, take

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

both positive semidefinite ($\mathbf{A} = \mathbf{e}_1\mathbf{e}_1^\top$ and $\mathbf{B} = \mathbf{1}\mathbf{1}^\top$). Then

$$\mathbf{AB} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

which is not symmetric, so it is not positive semidefinite — it fails (P1), never mind (P2). Section 2 shows what does survive: $\text{tr}(\mathbf{AB}) \geq 0$, and $\mathbf{AB} \succeq \mathbf{O}$ whenever $\mathbf{A}$ and $\mathbf{B}$ commute.

Exercise. C3: The largest entry sits on the diagonal

Let $\mathbf{A} \in M_n(F)$ with $\mathbf{A} \succeq \mathbf{O}$.

- (a) Prove that $|a_{ij}|^2 \leq a_{ii}a_{jj}$ for all i, j .
 (b) Deduce that if $a_{ii} = 0$ for some i , then the whole i -th row and i -th column of $\mathbf{A}$ are zero.
 (c) Deduce that $\max_{i,j} |a_{ij}| = \max_i a_{ii}$.

Hint for (a): look at a 2×2 principal submatrix.

Solution.

- (a) For $i = j$ the claim is $a_{ii}^2 \leq a_{ii}^2$. For $i \neq j$, take $I = \{i, j\}$. By the proof of [Theorem 13.1.4](#) ((a) $\Rightarrow$ (e)), $\mathbf{A}_{I,I} \succeq \mathbf{O}$, and [Example \(The \$2 \times 2\$ Test\)](#) gives $a_{ii}a_{jj} - |a_{ij}|^2 = \det \mathbf{A}_{I,I} \geq 0$.
 (b) If $a_{ii} = 0$, then (a) gives $|a_{ij}|^2 \leq 0 \cdot a_{jj} = 0$ for every j , so $a_{ij} = 0$; and $a_{ji} = \overline{a_{ij}} = 0$ since $\mathbf{A}$ is Hermitian.
 (c) Let $M = \max_i a_{ii}$, a maximum of nonnegative reals by [Exercise \(C1: Zero trace forces zero\)](#) (a). By (a), $|a_{ij}| \leq \sqrt{a_{ii}a_{jj}} \leq M$ for all i, j , so $\max_{i,j} |a_{ij}| \leq M$. The reverse inequality holds because the diagonal entries are among the entries. Hence the two maxima are equal.