

13.6 Positive Matrices in Blocks

Chapter 8 cut a matrix into blocks, ran one step of elimination, and gave a name to what was left in the far corner: the Schur complement. Section 1 of this chapter met that same quantity as a single number, in the induction that proved Sylvester's criterion. This section puts the two together. For a Hermitian matrix split into blocks, positive definiteness can be decided one block at a time, and the deciding quantity is the Schur complement. The test is then spent on a question Chapter 12 left open: which matrices carry a positive definite solution of the Lyapunov equation, a certificate of stability that names no eigenvalue and computes none. What the same block decomposition says about determinants is the next section's subject.

Throughout, $F = \mathbb{R}$ or $F = \mathbb{C}$, and F^n carries the standard inner product.

13.6.1 What the blocks of a positive matrix must satisfy

A Hermitian matrix cut into blocks of sizes k and l has a shape forced on it. Suppose

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{D} & \mathbf{C} \end{pmatrix} \in M_{k+l}(F),$$

with $\mathbf{A} \in M_k(F)$, $\mathbf{B} \in M_{k \times l}(F)$, $\mathbf{D} \in M_{l \times k}(F)$, and $\mathbf{C} \in M_l(F)$. Starring a block matrix transposes the arrangement and stars each block, so $\mathbf{M}^* = \mathbf{M}$ says exactly three things: $\mathbf{A}^* = \mathbf{A}$, $\mathbf{C}^* = \mathbf{C}$, and $\mathbf{D} = \mathbf{B}^*$. A Hermitian $\mathbf{M}$ is therefore always of the form

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix}$$

with $\mathbf{A}$ and $\mathbf{C}$ Hermitian and $\mathbf{B}$ unrestricted, and we write it this way from now on.

Two necessary conditions for positivity are immediate. The vectors of F^{k+l} whose last l coordinates vanish form a copy of F^k , and on them the quadratic form of $\mathbf{M}$ is the quadratic form of $\mathbf{A}$:

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}^* \mathbf{M} \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} = \mathbf{x}^* \mathbf{A} \mathbf{x}, \quad \begin{pmatrix} \mathbf{0} \\ \mathbf{y} \end{pmatrix}^* \mathbf{M} \begin{pmatrix} \mathbf{0} \\ \mathbf{y} \end{pmatrix} = \mathbf{y}^* \mathbf{C} \mathbf{y}.$$

So $\mathbf{M} \succeq \mathbf{0}$ forces $\mathbf{A} \succeq \mathbf{0}$ and $\mathbf{C} \succeq \mathbf{0}$. This is the statement that a principal submatrix of a positive semidefinite matrix is positive semidefinite, used already in the proof of Theorem 13.1.4, here for the two index blocks $\{1, \dots, k\}$ and $\{k+1, \dots, k+l\}$.

The two conditions are not sufficient, and the smallest counterexample has $k = l = 1$:

$$\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{A} = \mathbf{O}_1 \not\succeq \mathbf{0}, \quad \mathbf{C} = (1) \succ \mathbf{0}.$$

Here $\det \mathbf{M} = -1 < 0$, so $\mathbf{M} \not\succeq \mathbf{0}$ by Theorem 13.1.4 (e). Something must couple $\mathbf{B}$ to $\mathbf{A}$ and $\mathbf{C}$, and the 2×2 case already says what. For $\mathbf{M} = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$ Hermitian, Example (The 2×2 Test)

gives $\mathbf{M} \succ \mathbf{O}$ if and only if $a > 0$ and $ac - |b|^2 > 0$. Divide the second condition by the positive number a :

$$c - \bar{b}a^{-1}b > 0.$$

Written with $\mathbf{B} = (b)$, so that $\mathbf{B}^* = (\bar{b})$, the left-hand side is the Schur complement $\mathbf{M}/\mathbf{A}$ of Definition 8.3.1. The general statement now writes itself: replace the scalars by blocks.

13.6.2 Testing one block at a time

Theorem 13.6.1. Positive Definiteness Block by Block

Let $k, l \geq 1$ and let

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix} \in M_{k+l}(F)$$

be Hermitian, with $\mathbf{A} \in M_k(F)$, $\mathbf{B} \in M_{k \times l}(F)$, and $\mathbf{C} \in M_l(F)$.

- (a) If $\mathbf{M} \succeq \mathbf{O}$, then $\mathbf{A} \succeq \mathbf{O}$ and $\mathbf{C} \succeq \mathbf{O}$. If $\mathbf{M} \succ \mathbf{O}$, then $\mathbf{A} \succ \mathbf{O}$ and $\mathbf{C} \succ \mathbf{O}$.
- (b) Suppose $\mathbf{A} \succ \mathbf{O}$. Then $\mathbf{A}$ is invertible, the Schur complement

$$\mathbf{M}/\mathbf{A} = \mathbf{C} - \mathbf{B}^*\mathbf{A}^{-1}\mathbf{B} \in M_l(F)$$

is Hermitian, and

$$\mathbf{M} \succ \mathbf{O} \iff \mathbf{M}/\mathbf{A} \succ \mathbf{O}, \quad \mathbf{M} \succeq \mathbf{O} \iff \mathbf{M}/\mathbf{A} \succeq \mathbf{O}.$$

Idea. Everything is one congruence. Chapter 8's block elimination clears $\mathbf{B}$ and $\mathbf{B}^*$ at once and leaves $\mathbf{A}$ and $\mathbf{M}/\mathbf{A}$ on the diagonal; for a Hermitian $\mathbf{M}$ the two outer factors of that elimination are adjoint to each other, because the row operation and the column operation are the same operation seen twice. So the elimination is a congruence $\mathbf{S}^*\mathbf{M}\mathbf{S}$, and Proposition 13.1.2 says that a congruence by an invertible matrix neither creates nor destroys positivity. All that remains is to read the verdict off a block **diagonal** matrix, where the quadratic form splits into two independent pieces. This is the step (*) of Section 1's proof of Sylvester's criterion, with the last coordinate replaced by a block of l coordinates.

Proof.

- (a) Let $\mathbf{x} \in F^k$ and put $\mathbf{v} = (\mathbf{x}, \mathbf{0}) \in F^{k+l}$. By Theorem 8.1.1, $\mathbf{M}\mathbf{v} = (\mathbf{A}\mathbf{x}, \mathbf{B}^*\mathbf{x})$, and hence $\mathbf{v}^*\mathbf{M}\mathbf{v} = \mathbf{x}^*\mathbf{A}\mathbf{x}$. If $\mathbf{M} \succeq \mathbf{O}$, the left side is ≥ 0 for every $\mathbf{x}$, so $\mathbf{A} \succeq \mathbf{O}$; if $\mathbf{M} \succ \mathbf{O}$, then $\mathbf{x} \neq \mathbf{0}$ forces $\mathbf{v} \neq \mathbf{0}$ and the left side is > 0 , so $\mathbf{A} \succ \mathbf{O}$. The same argument with $\mathbf{v} = (\mathbf{0}, \mathbf{y})$ gives the statements for $\mathbf{C}$.
- (b) Since $\mathbf{A} \succ \mathbf{O}$, every eigenvalue of $\mathbf{A}$ is positive (Theorem 13.1.5 (b)), so $0 \notin \text{spec}(\mathbf{A})$ and $\mathbf{A}$ is invertible (Theorem 9.2.6). Also $(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1} = \mathbf{A}^{-1}$, so

$$(\mathbf{M}/\mathbf{A})^* = \mathbf{C}^* - \mathbf{B}^*(\mathbf{A}^{-1})^*\mathbf{B} = \mathbf{C} - \mathbf{B}^*\mathbf{A}^{-1}\mathbf{B} = \mathbf{M}/\mathbf{A},$$

and $\mathbf{M}/\mathbf{A}$ is Hermitian. Put

$$\mathbf{S} = \begin{pmatrix} \mathbf{I}_k & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_l \end{pmatrix}, \quad \mathbf{S}^* = \begin{pmatrix} \mathbf{I}_k & \mathbf{O} \\ -\mathbf{B}^*\mathbf{A}^{-1} & \mathbf{I}_l \end{pmatrix},$$

the second because $\mathbf{A}^{-1}$ is Hermitian. By Theorem 8.1.5, $\mathbf{S}$ is invertible. Two applications of Theorem 8.1.1 give

$$\begin{aligned} \mathbf{MS} &= \begin{pmatrix} \mathbf{A} & \mathbf{B} - \mathbf{A}\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{B}^* & \mathbf{C} - \mathbf{B}^*\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{B}^* & \mathbf{M}/\mathbf{A} \end{pmatrix}, \\ \mathbf{S}^*(\mathbf{MS}) &= \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{B}^* - \mathbf{B}^*\mathbf{A}^{-1}\mathbf{A} & \mathbf{M}/\mathbf{A} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}/\mathbf{A} \end{pmatrix}. \end{aligned} \quad (*)$$

Write $\mathbf{N} = \mathbf{A} \oplus (\mathbf{M}/\mathbf{A})$ for that block diagonal matrix. For $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2) \in F^k \times F^l$,

$$\mathbf{y}^*\mathbf{N}\mathbf{y} = \mathbf{y}_1^*\mathbf{A}\mathbf{y}_1 + \mathbf{y}_2^*(\mathbf{M}/\mathbf{A})\mathbf{y}_2. \quad (\dagger)$$

Claim. Given $\mathbf{A} \succ \mathbf{O}$: $\mathbf{N} \succ \mathbf{O}$ if and only if $\mathbf{M}/\mathbf{A} \succ \mathbf{O}$, and $\mathbf{N} \succeq \mathbf{O}$ if and only if $\mathbf{M}/\mathbf{A} \succeq \mathbf{O}$.

Proof. Taking $\mathbf{y}_1 = \mathbf{0}$ in $(\dagger)$ shows that positivity of $\mathbf{N}$, weak or strict, passes to $\mathbf{M}/\mathbf{A}$, since $\mathbf{y}_2 \neq \mathbf{0}$ forces $\mathbf{y} \neq \mathbf{0}$. Conversely, suppose $\mathbf{M}/\mathbf{A} \succeq \mathbf{O}$; then both terms of $(\dagger)$ are ≥ 0 , so $\mathbf{N} \succeq \mathbf{O}$. If moreover $\mathbf{M}/\mathbf{A} \succ \mathbf{O}$, then for $\mathbf{y} \neq \mathbf{0}$ at least one of $\mathbf{y}_1, \mathbf{y}_2$ is nonzero, so at least one term of $(\dagger)$ is > 0 while the other is ≥ 0 ; hence $\mathbf{N} \succ \mathbf{O}$. ■

Finally $\mathbf{S}$ is invertible, so Proposition 13.1.2 (b) gives $\mathbf{M} \succ \mathbf{O}$ if and only if $\mathbf{S}^*\mathbf{MS} = \mathbf{N} \succ \mathbf{O}$. For the weak version, Proposition 13.1.2 (a) applied to $\mathbf{S}$ turns $\mathbf{M} \succeq \mathbf{O}$ into $\mathbf{N} \succeq \mathbf{O}$, and applied to $\mathbf{S}^{-1}$, together with $(\mathbf{S}^{-1})^*\mathbf{N}\mathbf{S}^{-1} = \mathbf{M}$, turns $\mathbf{N} \succeq \mathbf{O}$ back into $\mathbf{M} \succeq \mathbf{O}$. Combining with the claim proves (b). ■

The identity $(*)$ is Theorem 8.3.1 (a) in disguise. That theorem factors $\mathbf{M} = \mathbf{L}\mathbf{N}\mathbf{U}$ with $\mathbf{L}$ lower and $\mathbf{U}$ upper block triangular; here $\mathbf{L}$ and $\mathbf{U}$ are $(\mathbf{S}^{-1})^*$ and $\mathbf{S}^{-1}$, so the two outer factors are adjoint to each other and the factorization is a **congruence**, not merely an equivalence. Taking determinants in $(*)$ recovers Theorem 8.3.2,

$$\det \mathbf{M} = \det \mathbf{A} \cdot \det(\mathbf{M}/\mathbf{A}),$$

which is the form the next section uses.

Chapter 8 closed its treatment of Schur complements by observing that when $\mathbf{M}$ is real symmetric the complement $\mathbf{M}/\mathbf{A}$ is symmetric again, and promising that “Chapter 13 uses Schur complements to test whether $\mathbf{M}$ is positive definite one block at a time”. [Theorem 13.6.1](#) is that test, and part (b) says exactly what the promise says: once the corner block $\mathbf{A}$ is known to be positive definite, the whole of $\mathbf{M}$ is positive definite precisely when the smaller matrix $\mathbf{M}/\mathbf{A}$ is. A question about $k + l$ coordinates has become a question about l of them.

It also finishes what Section 1 started. There, in the induction proving Sylvester's criterion, the matrix was split with $k = n - 1$ and $l = 1$; the shear $\mathbf{S}$ was the matrix above with the column $\mathbf{b}$ in place of $\mathbf{B}$; and the congruence came out as $\mathbf{S}^* \mathbf{A} \mathbf{S} = \mathbf{A}' \oplus (s)$ with the single number $s = d - \mathbf{b}^* \mathbf{A}'^{-1} \mathbf{b}$ in the corner. That section promised that Section 6 would replace the scalar by a block. The scalar has become $\mathbf{M}/\mathbf{A}$, and the argument is word for word the same one.

Example. A 4×4 Tested in 2×2 Blocks

Decide whether each of

$$\mathbf{M} = \left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 3 & 2 \\ 0 & 1 & 2 & 4 \end{array} \right), \quad \mathbf{M}' = \left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{array} \right)$$

is positive definite, using the displayed partition. For the one that is not, exhibit a vector on which the quadratic form is negative.

Solution. Both matrices are real symmetric with the same corner and border,

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Step 1: the corner. The leading principal minors of $\mathbf{A}$ are 2 and $2 - 1 = 1$, both positive, so $\mathbf{A} \succ \mathbf{O}$ by Sylvester's criterion (Theorem 13.1.5 (d)) and Theorem 13.6.1 (b) applies. Since $\det \mathbf{A} = 1$, Theorem 0.9.8 gives $\mathbf{A}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

Step 2: the coupling term. It is the same for both matrices:

$$\mathbf{B}^\top \mathbf{A}^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}, \quad \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Step 3: the complements. For $\mathbf{M}$, with $\mathbf{C} = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$,

$$\mathbf{M}/\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

whose leading principal minors are 2 and 3. So $\mathbf{M}/\mathbf{A} \succ \mathbf{O}$, hence $\mathbf{M} \succ \mathbf{O}$. As a by-product, $\det \mathbf{M} = \det \mathbf{A} \cdot \det(\mathbf{M}/\mathbf{A}) = 1 \cdot 3 = 3$.

For $\mathbf{M}'$, with $\mathbf{C}' = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$,

$$\mathbf{M}'/\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix},$$

whose determinant is -1 . So $\mathbf{M}'/\mathbf{A}$ is not even positive semidefinite, and $\mathbf{M}'$ is not positive definite.

Step 4: the witness. The congruence $(*)$ turns a bad vector for $\mathbf{M}'/\mathbf{A}$ into a bad vector for $\mathbf{M}'$. Here $\mathbf{y}_2 = (1, 1)$ gives $\mathbf{y}_2^\top (\mathbf{M}'/\mathbf{A}) \mathbf{y}_2 = 1 - 1 - 1 + 0 = -1$, so $\mathbf{y} = (0, 0, 1, 1)$ satisfies $\mathbf{y}^\top (\mathbf{S}^\top \mathbf{M}' \mathbf{S}) \mathbf{y} = -1$, and $\mathbf{x} = \mathbf{S} \mathbf{y}$ satisfies $\mathbf{x}^\top \mathbf{M}' \mathbf{x} = -1$. Since $-\mathbf{A}^{-1} \mathbf{B} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$, we get $\mathbf{x} = (1, -3, 1, 1)$. Checking directly, $\mathbf{M}' \mathbf{x} = (0, 0, 0, -1)$ and $\mathbf{x}^\top (\mathbf{M}' \mathbf{x}) = 0 + 0 + 0 - 1 = -1$.

One 2×2 inverse and one 2×2 product replaced four leading principal minors of a 4×4 matrix. And when the answer is “no”, the congruence hands back an explicit vector, which a determinant test never does.

Quick check. Let $\mathbf{A} \succ \mathbf{O}$. What is $\mathbf{M}/\mathbf{A}$ when $\mathbf{B} = \mathbf{O}$, and what does Theorem 13.6.1 (b) say in that case?

Solution. When $\mathbf{B} = \mathbf{O}$ the Schur complement is $\mathbf{M}/\mathbf{A} = \mathbf{C} - \mathbf{O} = \mathbf{C}$, so the criterion reads: $\mathbf{M} \succ \mathbf{O}$ if and only if $\mathbf{C} \succ \mathbf{O}$. That is the right answer, since $\mathbf{M} = \mathbf{A} \oplus \mathbf{C}$ is then block diagonal and $(\dagger)$ says its quadratic form is the sum of the two separate quadratic forms. The general case is this case after the change of variable $\mathbf{S}$.

Warning. **The corner must be definite, not merely semidefinite.** If $\mathbf{A}$ is singular, then $\mathbf{A}^{-1}$ does not exist, $\mathbf{M}/\mathbf{A}$ is undefined, and Theorem 13.6.1 (b) says nothing whatsoever. This is a real obstruction, not a gap in the proof. The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ above has $\mathbf{A} \succeq \mathbf{O}$ and $\mathbf{C} \succ \mathbf{O}$ and is **not** positive semidefinite, while $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, with the same $\mathbf{A}$ and the same $\mathbf{C}$, **is**. The two differ only in $\mathbf{B}$, so no criterion that ignores $\mathbf{B}$ can separate them.

Remark. There is a criterion for the semidefinite case with a singular corner, and it has three clauses instead of one: for Hermitian $\mathbf{M}$, one has $\mathbf{M} \succeq \mathbf{O}$ exactly when $\mathbf{A} \succeq \mathbf{O}$, every column of $\mathbf{B}$ lies in the column space of $\mathbf{A}$, and $\mathbf{C} - \mathbf{B}^* \mathbf{A}^+ \mathbf{B} \succeq \mathbf{O}$. The middle clause is what the warning’s first matrix violates. The matrix $\mathbf{A}^+$ in the third clause is the Moore–Penrose pseudoinverse; it agrees with $\mathbf{A}^{-1}$ when $\mathbf{A}$ is invertible, and it is built from the singular values of $\mathbf{A}$ in Section 12 of this chapter. The statement and its proof belong there, with the tools they need. This section keeps the definite case, which is the one the rest of the chapter consumes.

13.6.3 A positive definite certificate for stability

Chapter 12 ended its section on the Sylvester equation with a specialization it could not finish. For $\mathbf{A} \in M_n(\mathbb{C})$ and a Hermitian $\mathbf{Q} \in M_n(\mathbb{C})$, the **Lyapunov equation** is $\mathbf{A} \mathbf{X} + \mathbf{X} \mathbf{A}^* = -\mathbf{Q}$, and $\mathbf{A}$ is called **stable** when every $\lambda \in \text{spec}(\mathbf{A})$ has $\text{Re } \lambda < 0$ (Definition 12.10.2). That section showed that a stable $\mathbf{A}$ makes the equation uniquely solvable whatever $\mathbf{Q}$ is, because the spectra of $\mathbf{A}$ and of $-\mathbf{A}^*$ are then disjoint and Theorem 12.10.2 applies. It then stated, without proof, that “if $\mathbf{A}$ is stable and $\mathbf{Q}$ is positive definite, the unique solution $\mathbf{X}$ of $\mathbf{A} \mathbf{X} + \mathbf{X} \mathbf{A}^* = -\mathbf{Q}$ is positive definite; and conversely, if some positive definite $\mathbf{X}$ solves the equation for some positive definite $\mathbf{Q}$, then $\mathbf{A}$ is stable”, and added that “positivity is the subject of Chapter 13, where the tools to handle $\mathbf{x}^* \mathbf{X} \mathbf{x}$ are built”.

They are built, and the one the proof below needs is the block test just proved. The difficulty is that $\mathbf{X}$ is given by an equation rather than by a formula: there is no factorization to look at and no spectrum to compute. What one can do is take a single coordinate off at a time and let a Schur complement carry the verdict — which is exactly [Theorem 13.6.1](#) (b) with $l = 1$.

Everything rests on evaluating the equation at one well-chosen vector. Eigenvectors of $\mathbf{A}$ are the natural candidates, but $\mathbf{X}$ sits between $\mathbf{A}$ and $\mathbf{A}^*$ and blocks the usual move; the repair is to feed the equation the vector $\mathbf{X}^{-1}\mathbf{v}$ rather than $\mathbf{v}$.

Lemma 13.6.2. Reading the Lyapunov Equation at an Eigenvector

Let $\mathbf{A}, \mathbf{Q}, \mathbf{X} \in M_n(\mathbb{C})$ satisfy $\mathbf{AX} + \mathbf{XA}^* = -\mathbf{Q}$, with $\mathbf{X}$ Hermitian and invertible, and let $\mathbf{Av} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. Put $\mathbf{u} = \mathbf{X}^{-1}\mathbf{v}$. Then $\mathbf{u} \neq \mathbf{0}$, the number $\mathbf{v}^*\mathbf{X}^{-1}\mathbf{v}$ is real, and

$$2(\operatorname{Re} \lambda) \mathbf{v}^*\mathbf{X}^{-1}\mathbf{v} = -\mathbf{u}^*\mathbf{Q}\mathbf{u}. \quad (13.6.1)$$

Proof. Since $\mathbf{Xu} = \mathbf{v} \neq \mathbf{0}$, we have $\mathbf{u} \neq \mathbf{0}$. The inverse of a Hermitian invertible matrix is Hermitian, because $(\mathbf{X}^{-1})^* = (\mathbf{X}^*)^{-1} = \mathbf{X}^{-1}$, so $\mathbf{v}^*\mathbf{X}^{-1}\mathbf{v} = \langle \mathbf{X}^{-1}\mathbf{v}, \mathbf{v} \rangle$ is a real number (Proposition 11.6.7 (a)). Write $s = \mathbf{v}^*\mathbf{X}^{-1}\mathbf{v} = \mathbf{v}^*\mathbf{u}$; then $\mathbf{u}^*\mathbf{v} = \overline{\mathbf{v}^*\mathbf{u}} = s$ as well, the last step because s is real.

Evaluate the quadratic form of each side of the equation at $\mathbf{u}$. Using $\mathbf{Xu} = \mathbf{v}$ and $\mathbf{u}^*\mathbf{X} = (\mathbf{Xu})^* = \mathbf{v}^*$, which is where the Hermitian hypothesis is spent,

$$\mathbf{u}^*\mathbf{AXu} = \mathbf{u}^*\mathbf{Av} = \lambda \mathbf{u}^*\mathbf{v} = \lambda s, \quad \mathbf{u}^*\mathbf{XA}^*\mathbf{u} = \mathbf{v}^*\mathbf{A}^*\mathbf{u} = (\mathbf{Av})^*\mathbf{u} = \bar{\lambda} \mathbf{v}^*\mathbf{u} = \bar{\lambda} s.$$

Adding the two gives $\mathbf{u}^*(\mathbf{AX} + \mathbf{XA}^*)\mathbf{u} = (\lambda + \bar{\lambda})s = 2(\operatorname{Re} \lambda)s$, while the right-hand side of the equation gives $\mathbf{u}^*(-\mathbf{Q})\mathbf{u} = -\mathbf{u}^*\mathbf{Q}\mathbf{u}$. This proves the identity. ■

The identity has three ingredients, and it fixes the sign of any one of them once the signs of the other two are known. Both halves of the theorem are that observation.

Theorem 13.6.3. Lyapunov's Theorem

Let $\mathbf{A} \in M_n(\mathbb{C})$ and let $\mathbf{Q} \in M_n(\mathbb{C})$ with $\mathbf{Q} \succ \mathbf{O}$.

- (a) If $\mathbf{A}$ is stable, then $\mathbf{AX} + \mathbf{XA}^* = -\mathbf{Q}$ has exactly one solution $\mathbf{X} \in M_n(\mathbb{C})$, and that solution is positive definite.
- (b) Conversely, if some $\mathbf{X} \succ \mathbf{O}$ satisfies $\mathbf{AX} + \mathbf{XA}^* = -\mathbf{Q}$, then $\mathbf{A}$ is stable.

Idea. Both halves are [Eq. 13.6.1](#) read in different directions: it ties the sign of $\operatorname{Re} \lambda$ to the sign of $\mathbf{v}^*\mathbf{X}^{-1}\mathbf{v}$, the third factor $-\mathbf{u}^*\mathbf{Q}\mathbf{u}$ being negative throughout. Knowing $\mathbf{X} \succ \mathbf{O}$ gives (b) in one line. For (a) the same identity controls the quadratic form of $\mathbf{X}^{-1}$ only along eigenvectors of $\mathbf{A}$, and the work is to get from there to every vector. Peel one eigenvector off at a time: choose an orthonormal basis whose **last** vector is a unit eigenvector of $\mathbf{A}$, so that $\mathbf{A}$ becomes block lower triangular with a 1×1 corner. The leading $(n-1) \times (n-1)$ block of the transformed equation is a Lyapunov equation again, for a stable matrix, so induction

makes the leading block of $\mathbf{X}$ positive definite. One Schur complement is then all that stands between us and the verdict — and by Theorem 8.3.4 that number is the reciprocal of the last diagonal entry of $\mathbf{X}^{-1}$, which the identity has just shown to be positive.

Proof. Existence and uniqueness. Suppose $\mathbf{A}$ is stable. The eigenvalues of $-\mathbf{A}^*$ are the numbers $-\bar{\mu}$ with $\mu \in \text{spec}(\mathbf{A})$, as Chapter 12 computes from $\det(\mathbf{M}^*) = \overline{\det \mathbf{M}}$ immediately after Definition 12.10.2. If some $\lambda \in \text{spec}(\mathbf{A})$ were also an eigenvalue of $-\mathbf{A}^*$, then $\lambda = -\bar{\mu}$ for some $\mu \in \text{spec}(\mathbf{A})$, so $\lambda + \bar{\mu} = 0$; but $\text{Re}(\lambda + \bar{\mu}) = \text{Re } \lambda + \text{Re } \mu < 0$. So $\text{spec}(\mathbf{A}) \cap \text{spec}(-\mathbf{A}^*) = \emptyset$, and Theorem 12.10.2, (a) $\Rightarrow$ (c), applied to the pair $(\mathbf{A}, -\mathbf{A}^*)$, says that $\mathbf{A}\mathbf{X} - \mathbf{X}(-\mathbf{A}^*) = -\mathbf{Q}$ has exactly one solution. That is the Lyapunov equation.

(b). Let $\mathbf{X} \succ \mathbf{O}$ satisfy the equation. Every eigenvalue of $\mathbf{X}$ is positive (Theorem 13.1.5 (b)), so $\mathbf{X}$ is invertible (Theorem 9.2.6), and $\mathbf{X}^{-1} \succ \mathbf{O}$: writing $\mathbf{X} = \mathbf{B}^*\mathbf{B}$ with $\mathbf{B}$ invertible (Theorem 13.1.5 (c)) and putting $\mathbf{C} = (\mathbf{B}^*)^{-1}$, which is invertible with $\mathbf{C}^* = \mathbf{B}^{-1}$, we get $\mathbf{C}^*\mathbf{C} = \mathbf{B}^{-1}(\mathbf{B}^*)^{-1} = (\mathbf{B}^*\mathbf{B})^{-1} = \mathbf{X}^{-1}$, and Theorem 13.1.5 (c) $\Rightarrow$ (a) applies to $\mathbf{X}^{-1}$.

Let $\lambda \in \text{spec}(\mathbf{A})$ and let $\mathbf{v} \neq \mathbf{0}$ be an eigenvector for it. With $\mathbf{u} = \mathbf{X}^{-1}\mathbf{v}$, Lemma 13.6.2 gives $\mathbf{u} \neq \mathbf{0}$ and

$$2(\text{Re } \lambda) \mathbf{v}^* \mathbf{X}^{-1} \mathbf{v} = -\mathbf{u}^* \mathbf{Q} \mathbf{u} < 0,$$

the inequality because $\mathbf{Q} \succ \mathbf{O}$ and $\mathbf{u} \neq \mathbf{0}$. Also $\mathbf{v}^* \mathbf{X}^{-1} \mathbf{v} > 0$, since $\mathbf{X}^{-1} \succ \mathbf{O}$ and $\mathbf{v} \neq \mathbf{0}$. Dividing, $\text{Re } \lambda < 0$. As λ was an arbitrary eigenvalue, $\mathbf{A}$ is stable.

(a). Let $\mathbf{A}$ be stable, $\mathbf{Q} \succ \mathbf{O}$, and let $\mathbf{X}$ be the unique solution.

Step 1: $\mathbf{X}$ is Hermitian. Starring the equation gives $(\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^*)^* = \mathbf{A}\mathbf{X}^* + \mathbf{X}^*\mathbf{A}^*$, while $(-\mathbf{Q})^* = -\mathbf{Q}$ because $\mathbf{Q}$ is Hermitian. So $\mathbf{X}^*$ solves the same equation, and uniqueness forces $\mathbf{X}^* = \mathbf{X}$.

Step 2: $\mathbf{X}$ is invertible. Let $\mathbf{v} \in \mathbb{C}^n$ with $\mathbf{X}\mathbf{v} = \mathbf{0}$. Then $\mathbf{v}^* \mathbf{X} = (\mathbf{X}\mathbf{v})^* = \mathbf{0}^*$ by Step 1, so both terms of

$$\mathbf{v}^*(\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^*)\mathbf{v} = \mathbf{v}^* \mathbf{A}(\mathbf{X}\mathbf{v}) + (\mathbf{v}^* \mathbf{X})\mathbf{A}^* \mathbf{v}$$

vanish. The left-hand side is $-\mathbf{v}^* \mathbf{Q} \mathbf{v}$, so $\mathbf{v}^* \mathbf{Q} \mathbf{v} = 0$, and $\mathbf{Q} \succ \mathbf{O}$ forces $\mathbf{v} = \mathbf{0}$. So $\text{nul}(\mathbf{X}) = \{\mathbf{0}\}$ and $\mathbf{X}$ is invertible (Theorem 3.4.4).

Step 3: induction on n . We prove by induction on $n \geq 1$ that for every stable $\mathbf{A} \in M_n(\mathbb{C})$ and every $\mathbf{Q} \succ \mathbf{O}$ the unique solution is positive definite.

For $n = 1$, write $\mathbf{A} = (\alpha)$, $\mathbf{Q} = (q)$, and $\mathbf{X} = (x)$, with $\text{Re } \alpha < 0$, with $q > 0$, and with x real by Step 1. The equation reads $2(\text{Re } \alpha)x = -q$, so $x = q/(-2 \text{Re } \alpha) > 0$ and $\mathbf{X} \succ \mathbf{O}$.

Let $n \geq 2$ and assume the statement for $n - 1$. By Theorem 9.2.4 there are $\tau \in \text{spec}(\mathbf{A})$ and $\mathbf{w} \in \mathbb{C}^n$ with $\mathbf{A}\mathbf{w} = \tau\mathbf{w}$ and $\|\mathbf{w}\| = 1$, after dividing an eigenvector by its norm; and $\text{Re } \tau < 0$. Extend $\mathbf{w}$ to an orthonormal basis $(\mathbf{w}, \mathbf{w}_2, \dots, \mathbf{w}_n)$ of $\mathbb{C}^n$ (Corollary 11.2.3) and let $\mathbf{W}$ be the matrix with columns $\mathbf{w}_2, \dots, \mathbf{w}_n, \mathbf{w}$ in that order. Reordering an orthonormal list leaves it orthonormal, so the columns of $\mathbf{W}$ are an orthonormal basis and $\mathbf{W}$ is unitary (Theorem 11.7.3 (f), Definition 11.7.2). Its last column is $\mathbf{w}$, so $\mathbf{W}\mathbf{e}_n = \mathbf{w}$ and $\mathbf{W}^* \mathbf{w} = \mathbf{W}^{-1} \mathbf{w} = \mathbf{e}_n$.

Put $\hat{\mathbf{A}} = \mathbf{W}^* \mathbf{A} \mathbf{W}$, $\hat{\mathbf{X}} = \mathbf{W}^* \mathbf{X} \mathbf{W}$, and $\hat{\mathbf{Q}} = \mathbf{W}^* \mathbf{Q} \mathbf{W}$. Inserting $\mathbf{W}\mathbf{W}^* = \mathbf{I}$ between the factors of each product,

$$\hat{\mathbf{A}}\hat{\mathbf{X}} + \hat{\mathbf{X}}\hat{\mathbf{A}}^* = \mathbf{W}^*(\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^*)\mathbf{W} = -\hat{\mathbf{Q}},$$

and $\hat{\mathbf{Q}} \succ \mathbf{O}$ by Proposition 13.1.2 (b), while $\hat{\mathbf{X}}$ is Hermitian and invertible because $\mathbf{X}$ is and $\mathbf{W}$ is unitary. The last column of $\hat{\mathbf{A}}$ is $\mathbf{W}^* \mathbf{A} \mathbf{W} \mathbf{e}_n = \mathbf{W}^* \mathbf{A} \mathbf{w} = \tau \mathbf{W}^* \mathbf{w} = \tau \mathbf{e}_n$, so in blocks of sizes $n-1$ and 1,

$$\hat{\mathbf{A}} = \begin{pmatrix} \mathbf{A}' & \mathbf{0} \\ \mathbf{g}^* & \tau \end{pmatrix}, \quad \hat{\mathbf{X}} = \begin{pmatrix} \mathbf{X}' & \mathbf{b} \\ \mathbf{b}^* & d \end{pmatrix}, \quad \hat{\mathbf{Q}} = \begin{pmatrix} \mathbf{Q}' & * \\ * & * \end{pmatrix}.$$

Claim. $\mathbf{X}' \succ \mathbf{O}$.

Proof. By Theorem 8.1.1 the leading $(n-1) \times (n-1)$ block of $\hat{\mathbf{A}} \hat{\mathbf{X}}$ is $\mathbf{A}' \mathbf{X}' + \mathbf{0} \mathbf{b}^* = \mathbf{A}' \mathbf{X}'$. Since $\hat{\mathbf{X}}$ is Hermitian, $\hat{\mathbf{X}} \hat{\mathbf{A}}^* = (\hat{\mathbf{A}} \hat{\mathbf{X}})^*$, whose leading block is $(\mathbf{A}' \mathbf{X}')^* = \mathbf{X}' \mathbf{A}'^*$. Reading off the leading block of the transformed equation,

$$\mathbf{A}' \mathbf{X}' + \mathbf{X}' \mathbf{A}'^* = -\mathbf{Q}'.$$

Here $\mathbf{Q}' \succ \mathbf{O}$ by Theorem 13.6.1 (a), applied to $\hat{\mathbf{Q}} \succ \mathbf{O}$ with this partition.

And $\mathbf{A}'$ is stable. Let $\lambda \in \text{spec}(\mathbf{A}')$, so that $\det(\lambda \mathbf{I}_{n-1} - \mathbf{A}') = 0$ by Theorem 9.2.1 (d). The matrix

$$\lambda \mathbf{I}_n - \hat{\mathbf{A}} = \begin{pmatrix} \lambda \mathbf{I}_{n-1} - \mathbf{A}' & \mathbf{0} \\ -\mathbf{g}^* & \lambda - \tau \end{pmatrix}$$

is block triangular with a zero block in the upper right, so Theorem 7.8.4 gives $\det(\lambda \mathbf{I}_n - \hat{\mathbf{A}}) = \det(\lambda \mathbf{I}_{n-1} - \mathbf{A}') \cdot (\lambda - \tau) = 0$, and $\lambda \in \text{spec}(\hat{\mathbf{A}})$ by Theorem 9.2.1 again. Finally $\hat{\mathbf{A}} = \mathbf{W}^{-1} \mathbf{A} \mathbf{W}$, so $p_{\hat{\mathbf{A}}} = p_{\mathbf{A}}$ by Theorem 7.10.3, and two matrices with the same characteristic polynomial have the same eigenvalues (Theorem 9.2.1 (e)). Hence $\lambda \in \text{spec}(\mathbf{A})$ and $\text{Re } \lambda < 0$.

So $\mathbf{A}' \in M_{n-1}(\mathbb{C})$ is stable and $\mathbf{Q}' \succ \mathbf{O}$. By the existence and uniqueness step, applied to the pair $(\mathbf{A}', \mathbf{Q}')$, the equation $\mathbf{A}' \mathbf{Y} + \mathbf{Y} \mathbf{A}'^* = -\mathbf{Q}'$ has exactly one solution; $\mathbf{X}'$ is a solution, hence is that one. The inductive hypothesis gives $\mathbf{X}' \succ \mathbf{O}$. ■

Since $\mathbf{X}' \succ \mathbf{O}$ it is invertible, so the Schur complement $\hat{\mathbf{X}}/\mathbf{X}' = d - \mathbf{b}^* \mathbf{X}'^{-1} \mathbf{b}$ is defined, and it is a 1×1 Hermitian matrix by Theorem 13.6.1 (b), that is a real number s . The matrix $\hat{\mathbf{X}}$ is invertible, so Theorem 8.3.4 (a) says that $s \neq 0$ and that the lower-right entry of $\hat{\mathbf{X}}^{-1}$ is its inverse:

$$\mathbf{e}_n^* \hat{\mathbf{X}}^{-1} \mathbf{e}_n = \frac{1}{s}.$$

On the other hand $\hat{\mathbf{A}} \mathbf{e}_n = \tau \mathbf{e}_n$ with $\mathbf{e}_n \neq \mathbf{0}$ and $\text{Re } \tau < 0$, so Lemma 13.6.2, applied to the triple $(\hat{\mathbf{A}}, \hat{\mathbf{Q}}, \hat{\mathbf{X}})$ and this eigenvector, gives

$$2(\text{Re } \tau) \mathbf{e}_n^* \hat{\mathbf{X}}^{-1} \mathbf{e}_n = -\mathbf{u}^* \hat{\mathbf{Q}} \mathbf{u} < 0, \quad \mathbf{u} = \hat{\mathbf{X}}^{-1} \mathbf{e}_n \neq \mathbf{0},$$

the inequality because $\hat{\mathbf{Q}} \succ \mathbf{O}$. Therefore $\mathbf{e}_n^* \hat{\mathbf{X}}^{-1} \mathbf{e}_n > 0$, that is $1/s > 0$, that is $s > 0$ and $\hat{\mathbf{X}}/\mathbf{X}' = (s) \succ \mathbf{O}$. Now Theorem 13.6.1 (b), with the corner block $\mathbf{X}' \succ \mathbf{O}$, gives $\hat{\mathbf{X}} \succ \mathbf{O}$. Finally $\mathbf{X} = \mathbf{W} \hat{\mathbf{X}} \mathbf{W}^* = (\mathbf{W}^*)^* \hat{\mathbf{X}} \mathbf{W}^*$ is a congruence of $\hat{\mathbf{X}}$ by the invertible matrix $\mathbf{W}^*$, so $\mathbf{X} \succ \mathbf{O}$ by Proposition 13.1.2 (b). This completes the induction and proves (a). ■

Part (b) is what makes the theorem useful, and part (a) is what makes part (b) usable. Reading (b) alone: to prove that every eigenvalue of $\mathbf{A}$ has negative real part, it is enough to exhibit **one** positive definite $\mathbf{X}$ with $\mathbf{A} \mathbf{X} + \mathbf{X} \mathbf{A}^*$ negative definite, and no eigenvalue need ever be computed.

Such a certificate is a linear object — Chapter 12 §10 turned the equation into an ordinary $n^2 \times n^2$ linear system — whereas eigenvalues are roots of a polynomial. Reading (a) with it: the certificate always exists when the claim is true, so the method never fails for lack of a witness. Pick any $\mathbf{Q} \succ \mathbf{O}$, solve, and test the answer with Sylvester's criterion (Theorem 13.1.5 (d)).

Example. A Certificate for a Shear

Take $\mathbf{Q} = \mathbf{I}_2$ and solve $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^\top = -\mathbf{I}_2$ for

$$\mathbf{A} = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix}, \quad \text{and then for} \quad \mathbf{A}_1 = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}.$$

In each case decide from the solution whether the matrix is stable.

Solution. Both matrices are real, so we look for a real symmetric $\mathbf{X} = \begin{pmatrix} x & y \\ y & z \end{pmatrix}$; the remark below says nothing is lost by doing so. For a real $\mathbf{A}$ we have $\mathbf{A}^* = \mathbf{A}^\top$, and $\mathbf{X}\mathbf{A}^\top = (\mathbf{A}\mathbf{X})^\top$ because $\mathbf{X}$ is symmetric, so the equation says that $\mathbf{A}\mathbf{X} + (\mathbf{A}\mathbf{X})^\top = -\mathbf{I}_2$.

For $\mathbf{A}$, a direct computation gives $\mathbf{A}\mathbf{X} = \begin{pmatrix} -x+2y & -y+2z \\ -y & -z \end{pmatrix}$, so

$$\mathbf{A}\mathbf{X} + (\mathbf{A}\mathbf{X})^\top = \begin{pmatrix} -2x+4y & -2y+2z \\ -2y+2z & -2z \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The corner entry gives $z = \frac{1}{2}$, the off-diagonal entry gives $y = z = \frac{1}{2}$, and the first entry gives $-2x + 2 = -1$, so $x = \frac{3}{2}$. Hence

$$\mathbf{X} = \frac{1}{2} \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix},$$

and its leading principal minors are $\frac{3}{2}$ and $\frac{3}{4} - \frac{1}{4} = \frac{1}{2}$, both positive, so $\mathbf{X} \succ \mathbf{O}$ by Sylvester's criterion. By Theorem 13.6.3 (b), $\mathbf{A}$ is stable — as it is, its only eigenvalue being -1 twice.

For $\mathbf{A}_1$, the same computation with $\mathbf{A}_1\mathbf{X} = \begin{pmatrix} x+2y & y+2z \\ 2y & 2z \end{pmatrix}$ gives

$$\mathbf{A}_1\mathbf{X} + (\mathbf{A}_1\mathbf{X})^\top = \begin{pmatrix} 2x+4y & 3y+2z \\ 3y+2z & 4z \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

so $z = -\frac{1}{4}$, then $y = -\frac{2}{3}z = \frac{1}{6}$, then $2x = -1 - \frac{2}{3}$ and $x = -\frac{5}{6}$. The equation is still uniquely solvable, since the eigenvalues 1 and 2 of $\mathbf{A}_1$ admit no relation $\lambda + \bar{\mu} = 0$, but the solution

$$\mathbf{X}_1 = \begin{pmatrix} -5/6 & 1/6 \\ 1/6 & -1/4 \end{pmatrix}$$

has $\mathbf{e}_1^\top \mathbf{X}_1 \mathbf{e}_1 = -5/6 < 0$ and is not positive definite. So Theorem 13.6.3 (a) is not contradicted, and indeed $\mathbf{A}_1$ is not stable, its eigenvalues being 1 and 2. In fact $-\mathbf{X}_1 \succ \mathbf{O}$, its leading principal minors being $5/6$ and $13/72$: replacing $\mathbf{A}_1$ by $-\mathbf{A}_1$, which is stable, replaces $\mathbf{X}_1$ by $-\mathbf{X}_1$, and Theorem 13.6.3 (a) applies to that pair.

Remark. Over $\mathbb{R}$ nothing changes. If $\mathbf{A}$ and $\mathbf{Q}$ are real with $\mathbf{Q} \succ \mathbf{O}$, then conjugating the equation entrywise turns the solution $\mathbf{X}$ into another solution $\overline{\mathbf{X}}$, because $\overline{\mathbf{A}} = \mathbf{A}$ and $\overline{\mathbf{A}^*} = \mathbf{A}^\top = \mathbf{A}^*$ and $\overline{\mathbf{Q}} = \mathbf{Q}$; by uniqueness $\overline{\mathbf{X}} = \mathbf{X}$, so $\mathbf{X}$ is a real symmetric matrix, positive definite in the sense of Definition 13.1.1 over $\mathbb{R}$ as well, since its quadratic form on $\mathbb{R}^n$ is the restriction of its quadratic form on $\mathbb{C}^n$.

Quick check. Why does Theorem 13.6.3 say nothing about the quarter turn $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$? Show that for this $\mathbf{A}$ the equation $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = -\mathbf{I}_2$ has no solution whatsoever.

Solution. The eigenvalues of $\mathbf{A}$ are $\pm i$, which have real part 0, so $\mathbf{A}$ is not stable and neither part of the theorem applies. Nothing weaker survives either: $\mathbf{A}$ is real with $\mathbf{A}^* = \mathbf{A}^\top = -\mathbf{A}$, so for **every** $\mathbf{X} \in M_2(\mathbb{C})$,

$$\operatorname{tr}(\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^*) = \operatorname{tr}(\mathbf{A}\mathbf{X}) + \operatorname{tr}(\mathbf{A}^*\mathbf{X}) = \operatorname{tr}((\mathbf{A} + \mathbf{A}^*)\mathbf{X}) = 0$$

by Theorem 0.9.6, while $\operatorname{tr}(-\mathbf{I}_2) = -2$. So no $\mathbf{X}$ solves the equation. This is Theorem 12.10.2 seen from the other side: $\lambda = i$ and $\mu = i$ give $\lambda + \overline{\mu} = 0$, the Sylvester operator is singular, and $-\mathbf{I}_2$ happens to lie outside its image.

The two theorems of this section answer different questions with the same computation. Theorem 13.6.1 decides a positivity question about $k + l$ coordinates by asking it of l of them; Theorem 13.6.3 turns a question about the spectrum of $\mathbf{A}$ into a positivity question about a single matrix $\mathbf{X}$, which the block test then settles one coordinate at a time. One thing the congruence $(*)$ carries has gone unspent by both: the determinant identity $\det \mathbf{M} = \det \mathbf{A} \cdot \det(\mathbf{M}/\mathbf{A})$ recorded above. The next section reads two classical inequalities off it.

13.6.4 Exercises

13.6.4.1 A. Check your understanding

Exercise. A1

- Write down the general form of a Hermitian matrix in 2×2 block form, and say what each block must satisfy.
- State Theorem 13.6.1 (b) in full, including its hypothesis.
- True or false: if $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix}$ is Hermitian with $\mathbf{A} \succeq \mathbf{O}$ and $\mathbf{C} \succeq \mathbf{O}$, then $\mathbf{M} \succeq \mathbf{O}$. Justify your answer.
- State both halves of Theorem 13.6.3, and say which of the two is the one that certifies stability.
- The induction in the proof of Theorem 13.6.3 (a) conjugates by a unitary matrix whose **last** column is an eigenvector of $\mathbf{A}$. What would go wrong if it were the first column?

Solution.

- $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix}$ with $\mathbf{A} \in M_k(F)$ and $\mathbf{C} \in M_l(F)$ Hermitian and $\mathbf{B} \in M_{k \times l}(F)$ arbitrary.

The condition $\mathbf{M}^* = \mathbf{M}$ forces the lower-left block to be $\mathbf{B}^*$ and both diagonal blocks to be Hermitian.

- (b) If $\mathbf{A} \succ \mathbf{O}$, then $\mathbf{A}$ is invertible, $\mathbf{M}/\mathbf{A} = \mathbf{C} - \mathbf{B}^*\mathbf{A}^{-1}\mathbf{B}$ is Hermitian, $\mathbf{M} \succ \mathbf{O}$ if and only if $\mathbf{M}/\mathbf{A} \succ \mathbf{O}$, and $\mathbf{M} \succeq \mathbf{O}$ if and only if $\mathbf{M}/\mathbf{A} \succeq \mathbf{O}$.
- (c) False. Take $\mathbf{A} = \mathbf{O}_1$, $\mathbf{B} = (1)$, and $\mathbf{C} = (1)$, so $\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Both diagonal blocks are positive semidefinite, but $\det \mathbf{M} = -1 < 0$, so $\mathbf{M} \not\succeq \mathbf{O}$ by Theorem 13.1.4 (e).
- (d) Let $\mathbf{Q} \succ \mathbf{O}$. If $\mathbf{A}$ is stable, then $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = -\mathbf{Q}$ has exactly one solution and that solution is positive definite; conversely, if some $\mathbf{X} \succ \mathbf{O}$ satisfies the equation, then $\mathbf{A}$ is stable. The second half is the certificate: one positive definite $\mathbf{X}$ proves that every eigenvalue has negative real part, and no eigenvalue is computed.
- (e) With the eigenvector last, $\hat{\mathbf{A}} = \mathbf{W}^*\mathbf{A}\mathbf{W}$ is block **lower** triangular with the 1×1 corner at the bottom right, so the leading $(n-1) \times (n-1)$ block of the transformed equation is again a Lyapunov equation and the inductive hypothesis applies to it; and the surviving coordinate is then handled by a Schur complement whose corner block is the leading $\mathbf{X}' \succ \mathbf{O}$, which is the shape Theorem 13.6.1 (b) requires. Putting the eigenvector first would make $\hat{\mathbf{A}}$ block upper triangular, leaving the Lyapunov sub-equation in the **trailing** block and the Schur complement with a corner the theorem says nothing about.

13.6.4.2 B. Practice

Exercise. B1: Definiteness by blocks

Each matrix below is real symmetric and is displayed with a 2,2 partition. Determine whether it is positive definite. Justify your answer, and for any that is not, give a vector $\mathbf{x}$ with $\mathbf{x}^\top \mathbf{M} \mathbf{x} < 0$.

(a)

$$\mathbf{M}_1 = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ \hline 1 & 1 & 3 & 0 \\ 1 & -1 & 0 & 3 \end{array} \right)$$

(b)

$$\mathbf{M}_2 = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ \hline 1 & 1 & 2 & 0 \\ 1 & -1 & 0 & 1 \end{array} \right)$$

Solution. In both cases $\mathbf{A} = \mathbf{I}_2 \succ \mathbf{O}$ and $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, so Theorem 13.6.1 (b) applies with $\mathbf{A}^{-1} = \mathbf{I}_2$ and

$$\mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B} = \mathbf{B}^\top \mathbf{B} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

- (a) $\mathbf{M}_1/\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \mathbf{I}_2 \succ \mathbf{O}$. Hence $\mathbf{M}_1 \succ \mathbf{O}$, and $\det \mathbf{M}_1 = \det \mathbf{A} \det(\mathbf{M}_1/\mathbf{A}) = 1$.
- (b) $\mathbf{M}_2/\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$. This is the matrix of Section 1's warning: its leading principal minors are 0 and 0, and yet it is not positive semidefinite. So $\mathbf{M}_2$ is not

positive definite, and not even positive semidefinite.

For the witness, $\mathbf{y}_2 = (0, 1)$ gives $\mathbf{y}_2^\top (\mathbf{M}_2/\mathbf{A}) \mathbf{y}_2 = -1$. With $\mathbf{S} = \begin{pmatrix} \mathbf{I}_2 & -\mathbf{B} \\ \mathbf{O} & \mathbf{I}_2 \end{pmatrix}$ and $\mathbf{y} = (0, 0, 0, 1)$, the vector $\mathbf{x} = \mathbf{S}\mathbf{y} = (-1, 1, 0, 1)$ satisfies $\mathbf{x}^\top \mathbf{M}_2 \mathbf{x} = -1$.

Exercise. B2: Reading a certificate

(a) Let

$$\mathbf{A} = \begin{pmatrix} -3 & 4 \\ -2 & 2 \end{pmatrix}, \quad \mathbf{X} = \frac{1}{4} \begin{pmatrix} 22 & 16 \\ 16 & 15 \end{pmatrix}.$$

Check that $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^\top = -\mathbf{I}_2$ and that $\mathbf{X} \succ \mathbf{O}$, and conclude that $\mathbf{A}$ is stable without solving a characteristic equation. Then compute the characteristic polynomial and confirm the verdict.

(b) Show that the hypothesis $\mathbf{Q} \succ \mathbf{O}$ in [Theorem 13.6.3](#) (b) cannot be weakened to $\mathbf{Q} \succeq \mathbf{O}$: exhibit an $\mathbf{A}$ that is not stable together with an $\mathbf{X} \succ \mathbf{O}$ and a $\mathbf{Q} \succeq \mathbf{O}$ satisfying $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = -\mathbf{Q}$.

Solution.

(a) Multiplying,

$$\mathbf{A}\mathbf{X} = \frac{1}{4} \begin{pmatrix} -66 + 64 & -48 + 60 \\ -44 + 32 & -32 + 30 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -2 & 12 \\ -12 & -2 \end{pmatrix},$$

and since $\mathbf{X}$ is symmetric, $\mathbf{X}\mathbf{A}^\top = (\mathbf{A}\mathbf{X})^\top$. The off-diagonal entries of $\mathbf{A}\mathbf{X}$ are negatives of each other, so they cancel, and

$$\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^\top = \frac{1}{4} \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix} = -\mathbf{I}_2.$$

The leading principal minors of $\mathbf{X}$ are $22/4 = 11/2 > 0$ and $\det \mathbf{X} = \frac{1}{16}(22 \cdot 15 - 16^2) = \frac{74}{16} = \frac{37}{8} > 0$, so $\mathbf{X} \succ \mathbf{O}$ by Sylvester's criterion ([Theorem 13.1.5](#) (d)). With $\mathbf{Q} = \mathbf{I}_2 \succ \mathbf{O}$, [Theorem 13.6.3](#) (b) gives that $\mathbf{A}$ is stable. Checking: $\text{tr } \mathbf{A} = -1$ and $\det \mathbf{A} = -6 + 8 = 2$, so the characteristic polynomial is $\lambda^2 + \lambda + 2$, whose roots $(-1 \pm i\sqrt{7})/2$ have real part $-\frac{1}{2} < 0$.

(b) Take the quarter turn $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\mathbf{X} = \mathbf{I}_2 \succ \mathbf{O}$. Then $\mathbf{A}^* = \mathbf{A}^\top = -\mathbf{A}$, so $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{A}^* = \mathbf{A} - \mathbf{A} = \mathbf{O}$, and the equation holds with $\mathbf{Q} = \mathbf{O} \succeq \mathbf{O}$. But the eigenvalues of $\mathbf{A}$ are $\pm i$, with real part 0, so $\mathbf{A}$ is not stable. A positive definite solution on its own certifies nothing; it is the strict positivity of $\mathbf{Q}$ that makes $-\mathbf{u}^* \mathbf{Q} \mathbf{u}$ strictly negative in [Eq. 13.6.1](#), and that is what forces $\text{Re } \lambda < 0$.

13.6.4.3 C. Going deeper

Exercise. C1: A zero corner kills its border

(a) Prove that if $\mathbf{P} \succeq \mathbf{O}$ and $\mathbf{x}^* \mathbf{P} \mathbf{x} = 0$, then $\mathbf{P} \mathbf{x} = \mathbf{O}$.

(b) Deduce that if $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix} \succeq \mathbf{O}$ and $\mathbf{A} = \mathbf{O}$, then $\mathbf{B} = \mathbf{O}$.

- (c) Deduce that if $\mathbf{M} \succeq \mathbf{O}$ and $\mathbf{A}$ is singular, then $\mathbf{M}$ is singular. What does [Theorem 13.6.1](#) say about such an $\mathbf{M}$?

Hint for (a): write $\mathbf{P} = \mathbf{N}^*\mathbf{N}$.

Solution.

- (a) By Theorem 13.1.4 (c), $\mathbf{P} = \mathbf{N}^*\mathbf{N}$ for some $\mathbf{N}$. Then $0 = \mathbf{x}^*\mathbf{P}\mathbf{x} = \|\mathbf{N}\mathbf{x}\|^2$, so $\mathbf{N}\mathbf{x} = \mathbf{0}$ and $\mathbf{P}\mathbf{x} = \mathbf{N}^*(\mathbf{N}\mathbf{x}) = \mathbf{0}$.
- (b) Let $\mathbf{x} \in F^k$ and put $\mathbf{v} = (\mathbf{x}, \mathbf{0})$. As computed in the proof of [Theorem 13.6.1](#) (a), $\mathbf{v}^*\mathbf{M}\mathbf{v} = \mathbf{x}^*\mathbf{A}\mathbf{x} = 0$. By (a), $\mathbf{M}\mathbf{v} = \mathbf{0}$. But $\mathbf{M}\mathbf{v} = (\mathbf{A}\mathbf{x}, \mathbf{B}^*\mathbf{x}) = (\mathbf{0}, \mathbf{B}^*\mathbf{x})$, so $\mathbf{B}^*\mathbf{x} = \mathbf{0}$ for every $\mathbf{x} \in F^k$, that is $\mathbf{B}^* = \mathbf{O}$ and $\mathbf{B} = \mathbf{O}$.
- (c) Choose $\mathbf{x} \neq \mathbf{0}$ with $\mathbf{A}\mathbf{x} = \mathbf{0}$, which is possible since $\mathbf{A}$ is singular (Theorem 3.4.4), and put $\mathbf{v} = (\mathbf{x}, \mathbf{0}) \neq \mathbf{0}$. Then $\mathbf{v}^*\mathbf{M}\mathbf{v} = \mathbf{x}^*\mathbf{A}\mathbf{x} = 0$, so $\mathbf{M}\mathbf{v} = \mathbf{0}$ by (a), and $\mathbf{M}$ is singular. [Theorem 13.6.1](#) says nothing at all about such an $\mathbf{M}$: part (a) is a one-way implication, and part (b) needs $\mathbf{A} \succ \mathbf{O}$, which a singular $\mathbf{A}$ is not. This is the situation the warning after the theorem describes, and the criterion that does cover it waits for Section 12.

Exercise. C2: A Cauchy–Schwarz inequality for blocks

Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{pmatrix} \succeq \mathbf{O}$ with $\mathbf{A} \in M_k(F)$ and $\mathbf{C} \in M_l(F)$.

- (a) Prove that for all $\mathbf{x} \in F^k$ and $\mathbf{y} \in F^l$,

$$|\mathbf{x}^*\mathbf{B}\mathbf{y}|^2 \leq (\mathbf{x}^*\mathbf{A}\mathbf{x})(\mathbf{y}^*\mathbf{C}\mathbf{y}).$$

- (b) Deduce the Cauchy–Schwarz inequality $|\langle \mathbf{y}, \mathbf{x} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ in F^n .

Hint for (a): apply the definition of $\mathbf{M} \succeq \mathbf{O}$ to the vector $(\mathbf{x}, t\mathbf{y})$, and choose the scalar t well.

Solution.

- (a) Write $p = \mathbf{x}^*\mathbf{A}\mathbf{x} \geq 0$, $q = \mathbf{y}^*\mathbf{C}\mathbf{y} \geq 0$, and $z = \mathbf{x}^*\mathbf{B}\mathbf{y}$. For a scalar $t \in F$, block multiplication (Theorem 8.1.1) gives, for $\mathbf{v} = (\mathbf{x}, t\mathbf{y})$,

$$\mathbf{v}^*\mathbf{M}\mathbf{v} = p + \bar{t}z + tz + |t|^2q = p + 2\operatorname{Re}(tz) + |t|^2q \geq 0,$$

where the middle step uses $\bar{t}z = \overline{tz}$ and $w + \bar{w} = 2\operatorname{Re}(w)$. Now take $t = r\omega$ with $r \geq 0$ real and ω a scalar of modulus 1 chosen so that $\omega z = -|z|$ (over $\mathbb{R}$, $\omega = \pm 1$). Then

$$p - 2r|z| + r^2q \geq 0 \quad \text{for every } r \geq 0. \quad (*)$$

If $q = 0$, then $(*)$ says $p \geq 2r|z|$ for all $r \geq 0$, which forces $|z| = 0$, and the claim reads $0 \leq 0$. If $q > 0$, take $r = |z|/q \geq 0$ in $(*)$:

$$p - \frac{2|z|^2}{q} + \frac{|z|^2}{q} = p - \frac{|z|^2}{q} \geq 0,$$

so $|z|^2 \leq pq$, as claimed.

- (b) Take $k = l = n$ and $\mathbf{A} = \mathbf{B} = \mathbf{C} = \mathbf{I}_n$, so that $\mathbf{M} = \begin{pmatrix} \mathbf{I}_n & \mathbf{I}_n \\ \mathbf{I}_n & \mathbf{I}_n \end{pmatrix}$. This $\mathbf{M}$ is positive

semidefinite, since $\mathbf{M} = \mathbf{W}^* \mathbf{W}$ for $\mathbf{W} = \begin{pmatrix} \mathbf{I}_n & \mathbf{I}_n \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$ (Theorem 13.1.4 (c)). Part (a) then reads $|\mathbf{x}^* \mathbf{y}|^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$, and $\mathbf{x}^* \mathbf{y} = \langle \mathbf{y}, \mathbf{x} \rangle$ by the definition of the standard inner product. This is Theorem 11.1.5 for F^n .