Positive definite and semidefinite matrices

A matrix $A \in \mathbb{C}^{n\times n}$ is said to be **positive semidefinite** if

$$\langle A\mathbf{x}, \mathbf{x}\rangle \geq 0 \text{ for every } \mathbf{x} \in \mathbb{C}^n; \tag{16.1}$$

it is said to be **positive definite** if

$$\langle A\mathbf{x}, \mathbf{x}\rangle > 0 \text{ for every nonzero vector } \mathbf{x} \in \mathbb{C}^n. \tag{16.2}$$

Correspondingly, $A \in \mathbb{C}^{n\times n}$ is said to be **negative semidefinite** if $-A$ is positive semidefinite, and it is said to be **negative definite** if $-A$ is positive definite.

If $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$, then the **notation**

$A \succeq B$ (resp., $A \succ B$) means that $A - B$ is positive semidefinite (resp., $A - B$ is positive definite).

Lemma 16.1. *If $A \in \mathbb{C}^{n\times n}$ and $A \succeq O$, then:*

(1) *A is automatically Hermitian.*

(2) *The eigenvalues of A are nonnegative numbers.*

Moreover,

$$A \succ O \iff A = A^H \text{ and the eigenvalues of } A \text{ are all positive} \tag{16.3}$$

$$\iff A \succeq O \text{ and } \det A > 0. \tag{16.4}$$

Proof. If $A \succeq O$, then

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

for every $\mathbf{x} \in \mathbb{C}^n$. Therefore, by a straightforward calculation,

$$\begin{aligned} 4\langle A\mathbf{x}, \mathbf{y}\rangle &= \sum_{k=1}^{4} i^k \langle A(\mathbf{x} + i^k\mathbf{y}), (\mathbf{x} + i^k\mathbf{y})\rangle \\ &= \sum_{k=1}^{4} i^k \langle (\mathbf{x} + i^k\mathbf{y}), A(\mathbf{x} + i^k\mathbf{y})\rangle = 4\langle \mathbf{x}, A\mathbf{y}\rangle\,; \end{aligned}$$

i.e., $\langle A\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A\mathbf{y}\rangle$ for every choice of $\mathbf{x}$, $\mathbf{y} \in \mathbb{C}^n$. Therefore, (1) holds.

Next, let $\mathbf{x}$ be an eigenvector of A corresponding to the eigenvalue λ. Then

$$\lambda\langle \mathbf{x}, \mathbf{x}\rangle = \langle A\mathbf{x}, \mathbf{x}\rangle \geq 0.$$

Therefore $A \succeq O \implies \lambda \geq 0$ and $A \succ O \implies \lambda > 0$, since $\langle \mathbf{x}, \mathbf{x}\rangle > 0$. Thus, (2) and, in view of (1), the implication $\implies$ in (16.3) hold. The implication $\impliedby$ in (16.3) follows from the fact that $A = A^H \implies A = WDW^H$, in which W is unitary and $D = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$; the verification of (16.4) is left to the reader. □

• **Warning:** The conclusions of Lemma 16.1 are not true under the less restrictive constraint

$$\langle A\mathbf{x}, \mathbf{x}\rangle \geq 0 \text{ for every } \mathbf{x} \in \mathbb{R}^n.$$

Thus, for example, if

$$A = \begin{bmatrix} 2 & -2 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

then

$$\langle A\mathbf{x}, \mathbf{x}\rangle = (x_1 - x_2)^2 + x_1^2 + x_2^2 > 0$$

for every nonzero vector $\mathbf{x} \in \mathbb{R}^n$. However, A is clearly not Hermitian. The next lemma serves to clarify this example.

Lemma 16.2. *If $A \in \mathbb{R}^{n\times n}$, then*

$$\langle A\mathbf{u}, \mathbf{u}\rangle \geq 0 \textit{ for every } \mathbf{u} \in \mathbb{C}^n \iff \begin{cases} \langle A\mathbf{x}, \mathbf{x}\rangle \geq 0 \textit{ for every } \mathbf{x} \in \mathbb{R}^n \\ \textit{and } A = A^T. \end{cases}$$

Proof. If the conditions on the right hold and $\mathbf{u} = \mathbf{x} + i\mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$\begin{aligned} \langle A(\mathbf{x} + i\mathbf{y}), (\mathbf{x} + i\mathbf{y})\rangle &= \langle A\mathbf{x}, \mathbf{x}\rangle - i\langle A\mathbf{x}, \mathbf{y}\rangle + i\langle A\mathbf{y}, \mathbf{x}\rangle + \langle A\mathbf{y}, \mathbf{y}\rangle \\ &= \langle A\mathbf{x}, \mathbf{x}\rangle + \langle A\mathbf{y}, \mathbf{y}\rangle \geq 0\,, \end{aligned}$$

since $\langle A\mathbf{y}, \mathbf{x}\rangle = \langle \mathbf{y}, A\mathbf{x}\rangle = \langle A\mathbf{x}, \mathbf{y}\rangle$ when $A = A^T \in \mathbb{R}^{n\times n}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Thus, the conditions on the left hold. The converse implication is justified by Lemma 16.1. □

Exercise 16.1. Show that if $A \in \mathbb{C}^{n\times n}$ and $A = A^H$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, then $\lambda_1 I_n - A \succeq O$ **(even if** $\lambda_1 \leq 0$**).**

Exercise 16.2. Show that if $V \in \mathbb{C}^{n\times k}$ and $\operatorname{rank} V = k$, then

$$A \succ O \Longrightarrow V^H A V \succ O\,,$$

but the converse implication is not true if $k < n$.

Exercise 16.3. Show that if $A \in \mathbb{C}^{n\times n}$ with entries a_{ij}, $i, j = 1, \ldots, n$, and $A \succeq O$, then $|a_{ij}|^2 \leq a_{ii}a_{jj}$.

Exercise 16.4. Show that if $A \in \mathbb{C}^{n\times n}$, $n = p + q$, and

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{C}^{p\times p}$, $A_{22} \in \mathbb{C}^{q\times q}$, then

$$A \succ O \Longleftrightarrow A_{11} \succ O\,, \quad A_{21} = A_{12}^H\,, \quad \text{and} \quad A_{22} - A_{21}A_{11}^{-1}A_{12} \succ O\,.$$

Exercise 16.5. Let $U \in \mathbb{C}^{n\times n}$ be unitary and let $A \in \mathbb{C}^{n\times n}$. Show that if $A \succ O$ and $AU \succ O$, then $U = I_n$. [HINT: Consider $\langle AU\mathbf{x}, \mathbf{x}\rangle$ for eigenvectors $\mathbf{x}$ of U.]

16.1. A detour on triangular factorization

The **notation**

$$(16.5)\qquad A_{[j,k]} = \begin{bmatrix} a_{jj} & \cdots & a_{jk} \\ \vdots & \ddots & \vdots \\ a_{kj} & \cdots & a_{kk} \end{bmatrix} \quad \text{for} \quad A \in \mathbb{C}^{n\times n} \quad \text{and} \quad 1 \leq j \leq k \leq n$$

will be convenient.

The **trade secret** behind the factorization formulas that will be considered below is that if $B, L, U \in \mathbb{C}^{n\times n}$, L is lower triangular, and U is upper triangular, then

$$(16.6)\qquad \begin{array}{lll} (LB)_{[1,k]} = L_{[1,k]}\, B_{[1,k]} & \text{and} & (BU)_{[1,k]} = B_{[1,k]}\, U_{[1,k]}\,, \\ (BL)_{[k,n]} = B_{[k,n]}\, L_{[k,n]} & \text{and} & (UB)_{[k,n]} = U_{[k,n]}\, B_{[k,n]} \end{array}$$

for $k = 1, \ldots, n$.

Exercise 16.6. Let $P_k = \operatorname{diag}\{I_k, O_{(n-k)\times(n-k)}\}$. Show that

(a) $A \in \mathbb{C}^{n\times n}$ is upper triangular $\Longleftrightarrow$ $AP_k = P_k A P_k$ for $k = 1, \ldots, n$.

(b) $A \in \mathbb{C}^{n\times n}$ is lower triangular $\Longleftrightarrow$ $P_k A = P_k A P_k$ for $k = 1, \ldots, n$.

We shall say that a matrix $A \in \mathbb{C}^{n\times n}$ admits an LU (resp., UL) factorization if there exist a lower triangular matrix $L \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $U \in \mathbb{C}^{n\times n}$ such that $A = LU$ (resp., $A = UL$).

Theorem 16.3. *If* $A \in \mathbb{C}^{n\times n}$*, then:*

(1) *A admits an LU factorization with invertible triangular factors L and U* $\iff$ $\det A_{[1,k]} \neq 0$ *for* $k = 1, \ldots, n$*.*

(2) *A admits a UL factorization with invertible triangular factors L and U* $\iff$ $\det A_{[k,n]} \neq 0$ *for* $k = 1, \ldots, n$*.*

(3) *If* $\det A_{[1,k]} \neq 0$ *for* $k = 1, \ldots, n$*, then* $A = LDU$ *for exactly one lower triangular matrix L with ones on the diagonal, one upper triangular matrix U with ones on the diagonal, and one diagonal matrix D.*

(4) *If* $\det A_{[k,n]} \neq 0$ *for* $k = 1, \ldots, n$*, then* $A = UDL$ *for exactly one lower triangular matrix L with ones on the diagonal, one upper triangular matrix U with ones on the diagonal, and one diagonal matrix D.*

Proof. The proof is divided into steps.

1. *Verification of* (1)*:*

Suppose first that $A = LU$ with invertible factors L and U. Then, by the first formula in (16.6), $A_{[1,k]} = L_{[1,k]}U_{[1,k]}$. Moreover, since L and U are triangular matrices, $L_{[1,k]}$ and $U_{[1,k]}$ are also invertible for $k = 1, \ldots, n$. Thus, $A_{[1,k]}$ is invertible for $k = 1, \ldots, n$.

Suppose next that $A_{[1,k]}$ is invertible for $k = 1, \ldots, n$ and let $X \in \mathbb{C}^{n\times n}$ be the lower triangular matrix with entries x_{ij} for $i \geq j$ that are specified by the formulas

$$\begin{bmatrix} x_{k1} & \cdots & x_{kk} \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} (A_{[1,k]})^{-1} \quad \text{for } k = 1, \ldots, n\,, \tag{16.7}$$

with the understanding that $x_{11} = 1/a_{11}$. Now, let $Y = XA$. Then, by the first formula in (16.6),

$$Y_{[1,k]} = X_{[1,k]}A_{[1,k]} \quad \text{for } k = 1, \ldots, n\,,$$

and hence, in view of (16.7),

$$\begin{bmatrix} y_{k1} & \cdots & y_{kk} \end{bmatrix} = \begin{bmatrix} x_{k1} & \cdots & x_{kk} \end{bmatrix} A_{[1,k]} = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Thus, Y is upper triangular with $y_{jj} = 1$ for $j = 1, \ldots, n$. Therefore, Y and $X = YA^{-1}$ are invertible and $A = LU$ with $L = X^{-1}$ and $U = Y$.

2. *Verification of* (2)*:*

If $A = UL$ and U and L are both invertible, then, as U and L are triangular, $U_{[k,n]}$ and $L_{[k,n]}$ are both invertible for $k = 1, \ldots, n$. Thus, as $A_{[k,n]} = U_{[k,n]}L_{[k,n]}$ for $k = 1, \ldots, n$ by the fourth formula in (16.6), $A_{[k,n]}$ is also invertible for $k = 1, \ldots, n$.

Suppose next that $A_{[k,n]}$ is invertible for $k = 1, \ldots, n$ and let $X \in \mathbb{C}^{n\times n}$ be the lower triangular matrix with entries x_{ij} for $i \geq j$ that are specified by the formulas

$$\begin{bmatrix} x_{kk} \\ \vdots \\ x_{kn} \end{bmatrix} = (A_{[k,n]})^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{for } k = 1, \ldots, n\,, \tag{16.8}$$

with the understanding that $x_{nn} = 1/a_{nn}$, and let $Y = AX$. Then, by the third formula in (16.6), $Y_{[k,n]} = A_{[k,n]}X_{[k,n]}$ for $k = 1, \ldots, n$ and hence

$$\begin{bmatrix} y_{kk} \\ \vdots \\ y_{kn} \end{bmatrix} = A_{[k,n]} \begin{bmatrix} x_{kk} \\ \vdots \\ x_{kn} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, Y is upper triangular with $y_{jj} = 1$ for $j = 1, \ldots, n$. Therefore, Y and $X = A^{-1}Y$ are invertible and $A = UL$ with $L = X^{-1}$ and $U = Y$.

3. *Verification of* (3) *and* (4)*:*

To verify (3), suppose that an invertible matrix A admits a pair of factorizations $A = L_1D_1U_1 = L_2D_2U_2$ in which the diagonal entries of the triangular factors are all equal to one. Then the identity $L_2^{-1}L_1D_1 = D_2U_2U_1^{-1}$ implies that $D_1 = D_2$ and that $L_2^{-1}L_1D_1$ is both upper and lower triangular and hence is a diagonal matrix, which must be equal to D_1. Therefore, $L_1 = L_2$ and $U_1 = U_2$.

The verification of (4) is left to the reader; it is similar to the verification of (3). □

Remark 16.4. Formulas (16.7) and (16.8) serve to make the proof of Theorem 16.3 efficient, but mysterious.

To explain where the first of these two formulas comes from, we first observe that if $A = LU$ is invertible, then L and U are invertible and the diagonal matrix $\Delta = \operatorname{diag}\{u_{11}, \ldots, u_{nn}\}$ based on the diagonal entries of U is invertible. Therefore, $Y = \Delta^{-1}U$ is an upper triangular matrix with diagonal entries $y_{jj} = 1$ for $j = 1, \ldots, n$ and

$$A = LU \iff L^{-1}A = U \iff \Delta^{-1}L^{-1}A = \Delta^{-1}U \iff XA = Y\,,$$

with $X = \Delta^{-1}L^{-1}$. Thus A admits an LU factorization if and only if there exist a lower triangular matrix $X \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $Y \in \mathbb{C}^{n\times n}$ with $y_{jj} = 1$ for $j = 1, \ldots, n$ such that $XA = Y$ (because then X is invertible and $A = LY$ with $L = X^{-1}$).

It is remarkable that the awesome looking nonlinear matrix equation $XA = Y$, which is a system of n^2 equations with $(n^2+n)/2$ unknown entries x_{ij} with $1 \le j \le i \le n$ in X and $(n^2-n)/2$ unknown entries y_{ij} with $1 \le i < j \le n$ in Y, is tractable. But, in self-evident notation,

$$(16.9)\qquad \begin{aligned} Y_{[1,k]} &= (XA)_{[1,k]} = \begin{bmatrix} I_k & O \end{bmatrix} \begin{bmatrix} X_{11} & O \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_k \\ O \end{bmatrix} \\ &= X_{11}A_{11} = X_{[1,k]}A_{[1,k]} \end{aligned}$$

for $k = 1, \ldots, n$. It is now easily seen that when $\det A_{[1,k]} \ne 0$ for $k = 1, \ldots, n$, then (16.7) is just the bottom row of (16.9).

The motivation for (16.8) is similar.

Exercise 16.7. Verify item (4) in Theorem 16.3.

Exercise 16.8. Show that if $A \in \mathbb{C}^{n\times n}$ is invertible and $A^{-1} = B$, then $A_{[1,k]}$ is invertible for $k = 1, \ldots, n$ if and only if $B_{[k,n]}$ is invertible for $k = 1, \ldots, n$.

Exercise 16.9. Show that if $A \in \mathbb{C}^{n\times n}$ and $A_{[1,k]}$ is invertible for $k = 1, \ldots, n$, then formula (16.7) implies that $x_{kk} = \det A_{[1,k-1]}/\det A_{[1,k]}$ for $k = 2, \ldots, n$, whereas, if $A_{[k,n]}$ is invertible for $k = 1, \ldots, n$, then (16.8) implies that $x_{kk} = \det A_{[k+1,n]}/\det A_{[k,n]}$ for $k = 1, \ldots, n-1$.

Exercise 16.10. Show that the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ admits an LU factorization but does not admit a UL factorization and find matrices L, D, U such that $A = LDU$ with L lower (resp., U upper) triangular with ones on the diagonal and D diagonal.

Exercise 16.11. Let $A \in \mathbb{C}^{n\times n}$ be a Vandermonde matrix with columns $\mathbf{v}(\lambda_1), \ldots, \mathbf{v}(\lambda_n)$ based on n distinct points $\lambda_1, \ldots, \lambda_n$. Show that A admits an LU factorization, but may not admit a UL factorization.

Exercise 16.12. Show that if $A \in \mathbb{C}^{n\times n}$ and $A^2 = A$, then A is an orthogonal projection if and only if $A \succeq O$.

16.2. Characterizations of positive definite matrices

There are a number of different characterizations of positive definite matrices:

Theorem 16.5. *If $A \in \mathbb{C}^{n\times n}$, then the following statements are equivalent:*

(1) *$A \succ O$.*
(2) *$A = A^H$ and the eigenvalues, $\lambda_1, \ldots, \lambda_n$, of A are all positive; i.e., $\lambda_j > 0$ for $j = 1, \ldots, n$.*

(3) $A = A^H$ *and* $\det A_{[1,k]} > 0$ *for* $k = 1, \ldots, n$.

(4) $A = LL^H$, *where* L *is a lower triangular invertible matrix.*

(5) $A = A^H$ *and* $\det A_{[k,n]} > 0$ *for* $k = 1, \ldots, n$.

(6) $A = UU^H$, *where* U *is an upper triangular invertible matrix.*

Proof. Lemma 16.1 ensures that (1) $\Longrightarrow$ (2). Conversely, if (2) is in force, then $A = VDV^H$ with $V \in \mathbb{C}^{n\times n}$ unitary and $D \succ O$ and diagonal. Therefore,

$$\langle A\mathbf{x}, \mathbf{x}\rangle = \langle DV^H\mathbf{x}, V^H\mathbf{x}\rangle > 0$$

for every nonzero vector $\mathbf{x} \in \mathbb{C}^n$. Thus, (2) $\Longleftrightarrow$ (1).

Next, it is clear that (1) $\Longrightarrow$ (3) and hence, in view of Theorem 16.3, that A admits exactly one factorization of the form $A = L_1DU_1$, where L_1 is lower triangular with ones on the diagonal, U_1 is upper triangular with ones on the diagonal, and D is a diagonal matrix. Since $A = A^H$, $L_1DU_1 = U_1^HD^HL_1^H$. Consequently, $U_1 = L_1^H$ and $D = D^H$. Moreover, as

$$\begin{aligned} A = L_1DL_1^H \implies & A_{[1,k]} = (L_1)_{[1,k]}D_{[1,k]}(L_1^H)_{[1,k]} \\ \implies & \det A_{[1,k]} = |\det (L_1)_{[1,k]}|^2 \det D_{[1,k]} = \det D_{[1,k]} \end{aligned}$$

for $k = 1, \ldots, n$, the diagonal entries in the matrix $D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ are positive. Thus, D admits a square root

$$D^{1/2} = \mathrm{diag}\{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\}$$

and hence (4) holds with $L = L_1D^{1/2}$. Since the implication (4) $\Longrightarrow$ (1) is clear, the implications (1) $\Longrightarrow$ (3) $\Longrightarrow$ (4) $\Longrightarrow$ (1) are justified.

To complete the proof, it suffices to check that (1) $\Longrightarrow$ (5) $\Longrightarrow$ (6) $\Longrightarrow$ (1). The details are left to the reader. ☐

Exercise 16.13. Show that if $A \in \mathbb{C}^{n\times n}$, then $A \succ O$ if and only if $A = V^HV$ for some invertible matrix $V \in \mathbb{C}^{n\times n}$.

The next three exercises are formulated in terms of the matrix

(16.10)

$$Z_n = \sum_{j=1}^{n} \mathbf{e}_j\mathbf{e}_{n+1-j}^T = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad \text{where } \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{bmatrix} = I_n\,.$$

Exercise 16.14. Show that $Z_n^H = Z_n$ and $Z_n^HZ_n = I_n$, i.e., Z_n is both Hermitian and unitary.

Exercise 16.15. Show that $U \in \mathbb{C}^{n\times n}$ is an invertible upper triangular matrix if and only if Z_nUZ_n is an invertible lower triangular matrix and then use this information to verify the equivalence of (4) and (6) in Theorem 16.5.

Exercise 16.16. Show that if $p \geq 1$, $q \geq 1$, and $p+q=n$, then

$$\begin{bmatrix} O & Z_q \\ Z_p & O \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} O & Z_p \\ Z_q & O \end{bmatrix} = \begin{bmatrix} Z_qA_{22}Z_q & Z_qA_{21}Z_p \\ Z_pA_{12}Z_q & Z_pA_{11}Z_p \end{bmatrix}$$

and then use this identity to verify the equivalence of (3) and (5) in Theorem 16.5.

16.3. Square roots

Theorem 16.6. *If $A \in \mathbb{C}^{n\times n}$ and $A \succeq O$, then there is exactly one matrix $B \in \mathbb{C}^{n\times n}$ such that $B \succeq O$ and $B^2 = A$.*

Proof. If $A \in \mathbb{C}^{n\times n}$ and $A \succeq O$, then there exists a unitary matrix U and a diagonal matrix

$$D = \operatorname{diag}\{d_{11}, \ldots, d_{nn}\}$$

with nonnegative entries such that $A = UDU^H$. Therefore, upon setting

$$D^{1/2} = \operatorname{diag}\{d_{11}^{1/2}, \ldots, d_{nn}^{1/2}\},$$

it is readily checked that the matrix $B = UD^{1/2}U^H$ is again positive semidefinite and

$$B^2 = (UD^{1/2}U^H)(UD^{1/2}U^H) = UDU^H = A.$$

This completes the proof of the existence of at least one positive semidefinite square root of A.

Suppose next that there are two positive semidefinite square roots of A, B_1 and B_2. Then there exist a pair of unitary matrices U_1 and U_2 and a pair of diagonal matrices $D_1 = \operatorname{diag}\{\alpha_1, \ldots, \alpha_n\}$ with $\alpha_1 \geq \cdots \geq \alpha_n \geq 0$ and $D_2 = \operatorname{diag}\{\beta_1, \ldots, \beta_n\}$ with $\beta_1 \geq \cdots \geq \beta_n \geq 0$ such that

$$B_1 = U_1D_1U_1^H \quad \text{and} \quad B_2 = U_2D_2U_2^H.$$

Thus, as

$$U_1D_1^2U_1^H = B_1^2 = A = B_2^2 = U_2D_2^2U_2^H,$$

it follows that

$$U_2^HU_1D_1^2 = D_2^2U_2^HU_1$$

and hence, upon setting $W = U_2^HU_1$, that

$$WD_1^2 - D_2WD_1 = D_2^2W - D_2WD_1.$$

But this in turn implies that the matrix

$$X = WD_1 - D_2W$$

with entries x_{ij} for $i, j = 1, \ldots, n$ is a solution of the equation

$$XD_1 + D_2X = O$$

and hence that

$$x_{ij}\alpha_j + \beta_i x_{ij} = 0 \quad \text{for } i, j = 1, \ldots, n\,.$$

Thus, $x_{ij} = 0$ if $\alpha_j + \beta_i > 0$. On the other hand, if $\alpha_j + \beta_i = 0$, then $\alpha_j = \beta_i = 0$ and so $x_{ij} = w_{ij}\alpha_j - \beta_i w_{ij} = 0$ in this case also. Therefore, $X = O$ is the only solution of the equation $XD_1 + D_2X = O$. Consequently,

$$U_2^H U_1 D_1 - D_2 U_2^H U_1 = X = O\,;$$

i.e.,

$$B_1 = U_1 D_1 U_1^H = U_2 D_2 U_2^H = B_2\,,$$

as claimed. □

If $A \succeq O$, the symbol $A^{1/2}$ will be used to denote the unique $n \times n$ matrix $B \succeq O$ with $B^2 = A$. Correspondingly, B will be referred to as the **square root** of A. The restriction that $B \succeq O$ is essential to ensure uniqueness. Thus, for example,

$$\begin{bmatrix} O & A \\ A^{-1} & O \end{bmatrix} \begin{bmatrix} O & A \\ A^{-1} & O \end{bmatrix} = I_{2n}\,,$$

for every invertible matrix $A \in \mathbb{C}^{n\times n}$.

Exercise 16.17. Show that if $A \in \mathbb{C}^{n\times n}$, then

$$\begin{bmatrix} A & A \\ A & A \end{bmatrix} \succeq O \iff A \succeq O\,.$$

Exercise 16.18. Show that if $A \in \mathbb{C}^{n\times n}$, then

$$\begin{bmatrix} A^2 & A \\ A & I_n \end{bmatrix} \succeq O \iff A = A^H\,.$$

Exercise 16.19. Show that if $A, G \in \mathbb{C}^{n\times n}$, $G \succ O$, and $GA = A^H G$, then $\sigma(A) \subset \mathbb{R}$ and $A = A^*$ with respect to an appropriately defined inner product.

Exercise 16.20. Show that if A, $B \in \mathbb{C}^{n\times n}$ and $A \succeq B \succ O$, then $B^{-1} \succeq A^{-1} \succ O$. [HINT: $A - B \succ O \Longrightarrow A^{-1/2} B A^{-1/2} \prec I_n$.]

Exercise 16.21. Show that if $A, B \in \mathbb{C}^{n\times n}$ and if $A \succeq O$ and $B \succeq O$, then trace $AB \geq 0$ (even if $AB \not\succeq O$).

16.4. Polar forms and partial isometries

A matrix $A \in \mathbb{C}^{p\times q}$ is said to be a **partial isometry** if $A^HA\mathbf{x} = \mathbf{x}$ for every vector $\mathbf{x} \in \mathbb{C}^q$ that is orthogonal to $\mathcal{N}_A$. Since $\mathbb{C}^q = \mathcal{N}_A \oplus \mathcal{R}_{A^H}$, $A \in \mathbb{C}^{p\times q}$ is a partial isometry if and only if $A^HAA^H\mathbf{y} = A^H\mathbf{y}$ for every $\mathbf{y} \in \mathbb{C}^p$. Thus:

(1) $A \in \mathbb{C}^{p\times q}$ is an isometry if $A^HA = I_q$.

(2) $A \in \mathbb{C}^{p\times q}$ is a partial isometry if $A^HAA^H = A^H$.

Exercise 16.22. Show that if $A \in \mathbb{C}^{p\times q}$ is a partial isometry, then it is an isometry if and only if $\operatorname{rank} A = q$.

Theorem 16.7. *If $A \in \mathbb{C}^{p\times q}$, then there exists exactly one partial isometry $B \in \mathbb{C}^{p\times q}$ and one positive semidefinite matrix $P \in \mathbb{C}^{q\times q}$ such that $A = BP$ and $\mathcal{N}_B = \mathcal{N}_P$. In this factorization, P is* **the** *positive semidefinite square root of A^HA.*

Proof. If B and P meet the stated conditions, then $\mathbb{C}^q = \mathcal{N}_B \oplus \mathcal{R}_{B^H} = \mathcal{N}_P \oplus \mathcal{R}_{P^H}$ and $P = P^H$. Therefore, $\mathcal{R}_{B^H} = \mathcal{R}_P$ and hence

$$B^HBP = P \quad \text{and} \quad A^HA = PB^HBP = P^2\,.$$

Thus, P is the one and only positive semidefinite square root of A^HA. If $C \in \mathbb{C}^{p\times q}$ is a partial isometry such that $A = CP$ and $\mathcal{N}_C = \mathcal{N}_P$, then

$$CP\mathbf{x} = A\mathbf{x} = BP\mathbf{x} \text{ for every } \mathbf{x} \in \mathbb{C}^q \quad \text{and} \quad C\mathbf{y} = B\mathbf{y} \text{ for every } \mathbf{y} \in \mathcal{N}_P\,.$$

Therefore, $C = B$. □

Corollary 16.8. *If $A \in \mathbb{C}^{p\times q}$ and* $\operatorname{rank} A = r \geq 1$ *and A admits a pair of singular value decompositions $A = V_1S_1U_1^H = Y_1S_1X_1^H$, with isometric factors $V_1, Y_1 \in \mathbb{C}^{p\times r}$ and $U_1, X_1 \in \mathbb{C}^{q\times r}$, then $V_1U_1^H = Y_1X_1^H$.*

Proof. If $A \in \mathbb{C}^{p\times q}$ and $A = V_1S_1U_1^H$ with isometric factors $V_1 \in \mathbb{C}^{p\times r}$ and $U_1 \in \mathbb{C}^{q\times r}$ and $S_1 = \operatorname{diag}\{s_1, \ldots, s_r\} \succ O$, then

$$A = BP \quad \text{with } B = V_1U_1^H \in \mathbb{C}^{p\times q} \text{ and } P = U_1S_1U_1^H \in \mathbb{C}^{q\times q}\,. \tag{16.11}$$

The asserted uniqueness follows from Theorem 16.7, since $B^HBB^H = B^H$, $P \succeq O$, and $\mathcal{N}_B = \mathcal{N}_P$. □

Exercise 16.23. Show that the factors V_1 and U_1 in the factorization $A = V_1S_1U_1^H$ are not unique. [HINT: Diagonal matrices commute.]

The factorization BP in (16.11) is called the **right polar form** of A.

Exercise 16.24. Show that if $A \in \mathbb{C}^{p\times q}$ and $\operatorname{rank} A = r \geq 1$, then A admits exactly one **left polar form** $A = QC$ in which $Q \succeq O$ is a square root of AA^H and C is a partial isometry with $\mathcal{R}_C = \mathcal{R}_Q$.

Exercise 16.25. Show that if $P \in \mathbb{C}^{n\times n}$ is a positive semidefinite matrix and $Y_1, Y_2 \in \mathbb{C}^{n\times n}$ are such that $Y_1P = Y_2P$ and $\mathcal{N}_{Y_1} = \mathcal{N}_{Y_2} = \mathcal{N}_P$, then $Y_1 = Y_2$.

Theorem 16.9. *If* $A \in \mathbb{C}^{p\times q}$ *and* $\operatorname{rank} A = r \geq 1$, *then*

$$A^H A = I_q \iff \|A\mathbf{x}\| = \|\mathbf{x}\| \quad \textit{for every } \mathbf{x} \in \mathbb{C}^q \tag{16.12}$$

and

$$A^H AA^H = A^H \iff \|AA^H\mathbf{y}\| = \|A^H\mathbf{y}\| \quad \textit{for every } \mathbf{y} \in \mathbb{C}^p\,. \tag{16.13}$$

Proof. Since (16.12) is a special case of (16.13), it suffices to deal with the latter.

Suppose first that $\|AA^H\mathbf{y}\| = \|A^H\mathbf{y}\|$ for every $\mathbf{y} \in \mathbb{C}^p$ and let $\mathbf{x} \in \mathbb{C}^q$. Then $\mathbf{x} = \mathbf{u} + A^H\mathbf{y}$ for some choice of $\mathbf{u} \in \mathcal{N}_A$ and $\mathbf{y} \in \mathbb{C}^p$. Thus, as $\langle (I_q - A^HA)A^H\mathbf{y}, A^H\mathbf{y}\rangle = 0$, it is readily checked that

$$\langle (I_q - A^HA)\mathbf{x}, \mathbf{x}\rangle = \langle (I_q - A^HA)(\mathbf{u} + A^H\mathbf{y}), \mathbf{u} + A^H\mathbf{y}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle \geq 0\,,$$

i.e., $(I_q - A^HA) \succeq O$. Consequently,

$$\begin{aligned}\langle (I_q - A^HA)A^H\mathbf{y}, A^H\mathbf{y}\rangle = 0 &\implies \|(I_q - A^HA)^{1/2}A^H\mathbf{y}\| = 0\\ &\implies (I_q - A^HA)A^H\mathbf{y} = \mathbf{0}.\end{aligned}$$

Since these implications are valid for every $\mathbf{y} \in \mathbb{C}^p$, $(I_q - A^HA)A^H = O$.

The converse implication is easy and is left to the reader. $\square$

Lemma 16.10. *If* $P \in \mathbb{C}^{n\times n}$ *is a positive semidefinite matrix and* $B \in \mathbb{C}^{n\times n}$ *is a partial isometry with* $\mathcal{N}_B = \mathcal{N}_P$, *then*

$$BP = PB^H \implies B = B^H\,, \tag{16.14}$$

(i.e., $BP = (BP)^H \implies B = B^H$*).*

Proof. Under the given assumptions,

$$BPB^HBPB^H = BP^2B^H = PB^HBP = P^2 \implies BPB^H = P\,,$$

since the positive semidefinite matrix P^2 has exactly one positive semidefinite square root. But this in turn implies that

$$B^HP = B^HBPB^H = PB^H = BP\,.$$

Therefore, to complete the proof it suffices to show that $\mathcal{N}_{B^H} = \mathcal{N}_P$. But

$$B^H\mathbf{a} = \mathbf{0} \implies BPB^H\mathbf{a} = \mathbf{0} \implies P\mathbf{a} = \mathbf{0}\,,$$

i.e., $\mathcal{N}_{B^H} \subseteq \mathcal{N}_P = \mathcal{N}_B$. Thus, as $\dim \mathcal{N}_B = n - \operatorname{rank} B = n - \operatorname{rank} B^H = \dim \mathcal{N}_{B^H}$, we see that $\mathcal{N}_{B^H} = \mathcal{N}_B = \mathcal{N}_P$. $\square$

16.5. Some useful formulas

It is useful to **keep in mind** that if $A = A^H$, then a number of formulas that were established earlier assume a more symmetric form:

Theorem 16.11. *If* $A \in \mathbb{C}^{n\times n}$ *and* $\operatorname{rank} A = r$, $r \geq 1$, *then:*

(1) $A^HA = AA^H \implies \|A\| = \max\{|\langle A\mathbf{x}, \mathbf{x}\rangle| : \mathbf{x} \in \mathbb{C}^n$ *and* $\|\mathbf{x}\| = 1\}$.

(2) $A \succeq O \implies \|A\| = \max\{\langle A\mathbf{x}, \mathbf{x}\rangle : \mathbf{x} \in \mathbb{C}^n$ *and* $\|\mathbf{x}\| = 1\}$.

(3) $A \succeq O \implies$ *in the singular value decomposition* $A = V_1S_1U_1^H$ *in* (15.2), *the two* $n \times r$ *isometric matrices coincide:* $V_1 = U_1$.

(4) $A \succ O \implies \varphi(\mathbf{x}) = (\langle A\mathbf{x}, \mathbf{x}\rangle)^{1/2}$ *is a norm on* $\mathbb{C}^n$.

(5) $A \succ O \implies \varphi(\mathbf{x}, \mathbf{y}) = \langle A\mathbf{x}, \mathbf{y}\rangle$ *is an inner product on* $\mathbb{C}^n$.

Proof. We shall verify (3) and leave the justification of the rest to the reader. Since $A \succeq O \implies A = A^H$, the corresponding singular value decompositions must coincide, i.e., $V_1S_1U_1^H = U_1S_1V_1^H$. Thus, in view of Corollary 16.8, $V_1U_1^H = U_1V_1^H$ and hence $V_1 = U_1V_1^HU_1 = U_1K$ with $K = V_1^HU_1$. Consequently,

$$I_r = V_1^HV_1 = K^HU_1^HU_1K = K^HK, \quad \text{i.e., } K \text{ is unitary.}$$

Moreover, $KS_1 \succ O$, since $A = U_1KS_1U_1^H \succeq O$. Therefore, $KS_1 = S_1K^H$ and hence

$$(KS_1)^2 = S_1K^HKS_1 = S_1^2 .$$

Thus, $KS_1 = S_1$, since they are both positive definite square roots of S_1^2 . Therefore, $K = I_r$ and $V_1 = U_1$. □

Exercise 16.26. Verify items (1), (2), (4), and (5) in Theorem 16.11.

We remark that if A, $B \in \mathbb{C}^{n\times n}$, $A \succ B \succ O$, and $0 < t < 1$, then

$$A^t - B^t = \frac{\sin \pi t}{\pi} \int_0^\infty x^t(xI_n + A)^{-1}(A - B)(xI_n + B)^{-1}dx \,. \tag{16.15}$$

Exercise 16.27. Use formula (16.15) to show that if A, $B \in \mathbb{C}^{n\times n}$, then

$$A \succ B \succ O \implies A^t \succ B^t \quad \text{for} \quad 0 < t < 1 \,. \tag{16.16}$$

Exercise 16.28. Let $A = \begin{bmatrix} x & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Show that if $2 < x < 1 + \sqrt{2}$, then $A \succeq B \succeq O$, but $A^2 - B^2$ has one positive eigenvalue and one negative eigenvalue, i.e., $A \succeq B \succeq O$ does not imply that $A^2 \succeq B^2$.

16.6. Supplementary notes

This chapter is partially adapted from Chapter 12 in [**30**], which contains information on Toeplitz matrices, block Toeplitz matrices, and polynomial identities. Sections 16.4 and 16.5 are new, but (16.15) is discussed in [**30**].

Chapter 37

Matrix equations

In this chapter we shall analyze the existence and uniqueness of solutions to a number of matrix equations that occur frequently in applications. The **notation**

$$\mathbb{C}_{\mathrm{R}} = \{\lambda \in \mathbb{C} : \lambda + \overline{\lambda} > 0\} \quad \text{and} \quad \mathbb{C}_{\mathrm{L}} = \{\lambda \in \mathbb{C} : \lambda + \overline{\lambda} < 0\} \tag{37.1}$$

for the open right and open left half-plane, respectively, will be useful.

37.1. The equation $X - AXB = C$

In this section we shall study the equation $X - AXB = C$ for appropriately sized matrices A, B, C, and X. If $A = \mathrm{diag}\{\lambda_1, \ldots, \lambda_p\}$ and $B = \mathrm{diag}\{\mu_1, \ldots, \mu_q\}$ are diagonal matrices, then it is easily seen that the equation $x_{ij} - \lambda_i x_{ij} \mu_j = c_{ij}$ for the ij entry x_{ij} of X has exactly one solution if and only if $1 - \lambda_i \mu_j \neq 0$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. This condition on the eigenvalues of A and B holds for nondiagonal matrices too but requires a little more work to justify.

Lemma 37.1. *Let $A \in \mathbb{C}^{p\times p}$, $B \in \mathbb{C}^{q\times q}$ and let $\lambda_1, \ldots, \lambda_k$ and $\beta_1, \ldots, \beta_m$ denote the distinct eigenvalues of the matrices A and B, respectively; and let T denote the linear transformation from $\mathbb{C}^{p\times q}$ into $\mathbb{C}^{p\times q}$ that is defined by the rule*

$$T: X \in \mathbb{C}^{p\times q} \mapsto X - AXB \in \mathbb{C}^{p\times q}\,. \tag{37.2}$$

Then $\mathcal{N}_T = \{O_{p\times q}\}$ if and only if

$$\lambda_i \beta_j \neq 1 \quad \textit{for} \quad i = 1, \ldots, k \quad \textit{and} \quad j = 1, \ldots, m\,. \tag{37.3}$$

Proof. Let $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ and $B^T\mathbf{v}_j = \beta_j \mathbf{v}_j$ for some pair of nonzero vectors $\mathbf{u}_i \in \mathbb{C}^p$ and $\mathbf{v}_j \in \mathbb{C}^q$ and let $X = \mathbf{u}_i\mathbf{v}_j^T$. Then the formula

$$TX = \mathbf{u}_i\mathbf{v}_j^T - A\mathbf{u}_i\mathbf{v}_j^T B = (1-\lambda_i\beta_j)\mathbf{u}_i\mathbf{v}_j^T$$

clearly implies that the condition (37.3) is necessary for $\mathcal{N}_T = \{O_{p\times q}\}$.

To prove the sufficiency of this condition, invoke the Jordan decomposition $B = UJU^{-1}$ with $U = \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}$ and $J = \text{diag}\{J_{\beta_1}, \ldots, J_{\beta_m}\}$, in which $U_j \in \mathbb{C}^{q\times k_j}$ and $J_{\beta_j} \in \mathbb{C}^{k_j\times k_j}$ is upper triangular with β_j on the diagonal for $j = 1, \ldots, m$ and k_j is equal to the algebraic multiplicity of β_j for $j = 1, \ldots, m$. Then, since $J_{\beta_j} = \beta_j I_{k_j} + N_j$ and $(N_j)^{k_j} = O$,

$$\begin{aligned} X - AXB = O &\iff X - AXUJU^{-1} = O \iff XU - AXUJ = O \\ &\iff XU_j = AXU_jJ_{\beta_j} \quad \text{for } j = 1, \ldots, m \\ &\iff (I_p - \beta_j A)XU_j = AXU_jN_j \quad \text{for } j = 1, \ldots, m\,. \end{aligned}$$

Therefore, since $(I_p - \beta_j A)$ is invertible when (37.3) is in force,

$$X - AXB = O \iff XU_j = M_jXU_jN_j \quad \text{for } j = 1, \ldots, m$$

with $M_j = (I_p - \beta_j A)^{-1}A$. But upon iterating the last displayed equality n times we obtain

$$XU_j = M_j^n XU_j N_j^n = O \quad \text{for } j = 1, \ldots, m \quad \text{if } n \geq q\,.$$

Therefore, $X = XUU^{-1} = \begin{bmatrix} XU_1 & \cdots & XU_m \end{bmatrix} U^{-1} = O$ is the only solution of the equation $X - AXB = O$ when $\lambda_i\beta_j \neq 1$. □

Theorem 37.2. *Let $A \in \mathbb{C}^{p\times p}$, $B \in \mathbb{C}^{q\times q}$ and let $\lambda_1, \ldots, \lambda_k$ and $\beta_1, \ldots, \beta_m$ denote the distinct eigenvalues of the matrices A and B, respectively. Then, for each choice of $C \in \mathbb{C}^{p\times q}$, the equation*

$$X - AXB = C \qquad (37.4)$$

has exactly one solution $X \in \mathbb{C}^{p\times q}$ if and only if (37.3) *holds.*

Proof. This is immediate from Lemma 37.1 and the principle of conservation of dimension: If T is the linear transformation that is defined by the rule (37.2), then

$$pq = \dim \mathcal{N}_T + \dim \mathcal{R}_T\,.$$

Therefore, T maps onto $\mathbb{C}^{p\times q}$ if and only if $\mathcal{N}_T = \{O_{p\times q}\}$, i.e., if and only if $\lambda_i\beta_j \neq 1$ for every choice of i and j. □

Corollary 37.3. *If $A \in \mathbb{C}^{p\times p}$, $B \in \mathbb{C}^{q\times q}$, $C \in \mathbb{C}^{p\times q}$, and $r_\sigma(A)\, r_\sigma(B) < 1$, then $X = \sum_{j=0}^{\infty} A^jCB^j$ is the only solution of equation* (37.4).

Exercise 37.1. Use Theorem 22.2, the refined fixed point theorem, to justify Corollary 37.3 a second way (that does not depend upon Theorem 37.2).

Corollary 37.4. *Let $A \in \mathbb{C}^{p\times p}$, $C \in \mathbb{C}^{p\times p}$ and let $\lambda_1, \ldots, \lambda_p$ denote the eigenvalues of the matrix A. Then the* ***Stein equation***

$$X - A^H XA = C \tag{37.5}$$

has exactly one solution $X \in \mathbb{C}^{p\times p}$ if and only if $1 - \overline{\lambda_i}\lambda_j \neq 0$ for every choice of i and j.

Exercise 37.2. Verify Corollary 37.4.

Exercise 37.3. Let $A = C_\alpha^{(2)}$ and $B = C_\beta^{(2)}$ and suppose that $\alpha\beta = 1$. Show that the equation $X - AXB = C$ has no solutions if either $c_{21} \neq 0$ or $\alpha c_{11} \neq \beta c_{22}$.

Exercise 37.4. Let $A = C_\alpha^{(2)}$ and $B = C_\beta^{(2)}$ and suppose that $\alpha\beta = 1$. Show that if $c_{21} = 0$ and $\alpha c_{11} = \beta c_{22}$, then the equation $X - AXB = C$ has infinitely many solutions.

Exercise 37.5. Find the unique solution $X \in \mathbb{C}^{p\times p}$ of equation (37.5) when $A = C_0^{(p)}$, $C = \mathbf{e}_1\mathbf{u}^H + \mathbf{u}\mathbf{e}_1^H$, and $\mathbf{u}^H = \begin{bmatrix}\overline{t_0} & \overline{t_1} & \cdots & \overline{t_{p-1}}\end{bmatrix}$.

Exercise 37.6. Let $A = C_0^{(4)}$ and let T denote the linear transformation from $\mathbb{C}^{4\times 4}$ into itself that is defined by the formula $TX = X - A^H XA$.

(a) Calculate $\dim \mathcal{N}_T$.

(b) Show that a matrix $X \in \mathbb{C}^{4\times 4}$ is a solution of the matrix equation $X - A^H XA = \begin{bmatrix} a & b & c & d \\ e & 0 & 0 & 0 \\ f & 0 & 0 & 0 \\ g & 0 & 0 & 0 \end{bmatrix}$ if and only if X is a **Toeplitz matrix** (i.e., $x_{ij} = x_{i+1,j+1}$) with $x_{1j} = c_{1j}$ and $x_{j1} = c_{j1}$ for $j = 1, \ldots, 4$.

37.2. The Sylvester equation $AX - XB = C$

The strategy for studying the equation $AX - XB = C$ is much the same as for the equation $X - AXB = C$. Again the special case in which $A = \operatorname{diag}\{\lambda_1, \ldots, \lambda_p\}$ and $B = \operatorname{diag}\{\beta_1, \ldots, \beta_q\}$ are diagonal matrices points the way: The equation $\lambda_i x_{ij} - x_{ij}\beta_j = c_{ij}$ for the ij entry x_{ij} of X has exactly one solution if and only if $\lambda_i - \beta_j \neq 0$ for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. This condition on the eigenvalues of A and B holds for nondiagonal matrices too but requires a little more work to justify.

Lemma 37.5. *Let $A \in \mathbb{C}^{p\times p}$, $B \in \mathbb{C}^{q\times q}$ and let $\lambda_1, \ldots, \lambda_k$ and $\beta_1, \ldots, \beta_m$ denote the distinct eigenvalues of the matrices A and B, respectively, and* *let T denote the linear transformation from $\mathbb{C}^{p\times q}$ into $\mathbb{C}^{p\times q}$ that is defined by the rule*

$$(37.6) \qquad T:\, X\in\mathbb{C}^{p\times q}\mapsto AX-XB\in\mathbb{C}^{p\times q}\,.$$

Then $\mathcal{N}_T=\{O_{p\times q}\}$ if and only if

$$(37.7) \qquad \lambda_i-\beta_j\neq 0 \quad \textit{for} \quad i=1,\ldots,p \quad \textit{and} \quad j=1,\ldots,q\,.$$

Proof. Let $A\mathbf{u}_i=\lambda_i\mathbf{u}_i$ and $B^T\mathbf{v}_j=\beta_j\mathbf{v}_j$ for some pair of nonzero vectors $\mathbf{u}_i\in\mathbb{C}^p$ and $\mathbf{v}_j\in\mathbb{C}^q$ and let $X=\mathbf{u}_i\mathbf{v}_j^T$. Then the formula

$$TX=A\mathbf{u}_i\mathbf{v}_j^T-\mathbf{u}_i\mathbf{v}_j^TB=(\lambda_i-\beta_j)\mathbf{u}_i\mathbf{v}_j^T$$

clearly implies that the condition (37.7) is necessary for $\mathcal{N}_T=\{O_{p\times q}\}$.

To prove the sufficiency of this condition, invoke the Jordan decomposition $B=UJU^{-1}$ with $U=\begin{bmatrix}U_1 & \cdots & U_m\end{bmatrix}$ and $J=\operatorname{diag}\{J_{\beta_1},\ldots,J_{\beta_m}\}$, in which $U_j\in\mathbb{C}^{q\times k_j}$ and $J_{\beta_j}\in\mathbb{C}^{k_j\times k_j}$ for $j=1,\ldots,m$ and k_j is equal to the algebraic multiplicity of β_j for $j=1,\ldots,m$. Then, since $J_{\beta_j}=\beta_jI_{k_j}+N_j$ and $(N_j)^{k_j}=O$,

$$\begin{aligned} AX-XB=O &\iff AX-XUJU^{-1}=O \iff AXU-XUJ=O\\ &\iff AXU_j=XU_jJ_{\beta_j} \quad \text{for } j=1,\ldots,m\\ &\iff (A-\beta_jI_p)XU_j=XU_jN_j \quad \text{for } j=1,\ldots,m\,.\end{aligned}$$

Therefore, since $(A-\beta_jI_p)$ is invertible when (37.7) is in force,

$$AX-XB=O \iff XU_j=M_jXU_jN_j \quad \text{for } j=1,\ldots,m$$

with $M_j=(A-\beta_jI_p)^{-1}A$. But upon iterating the last displayed equality n times we obtain

$$XU_j=M_j^nXU_jN_j^n=O \quad \text{for } j=1,\ldots,m \quad \text{if } n\geq q\,.$$

Therefore $X=XUU^{-1}=\begin{bmatrix}XU_1 & \cdots & XU_m\end{bmatrix}U^{-1}=O$ is the only solution of the equation $AX-XB=O$ when $\lambda_i\beta_j\neq 1$. □

Theorem 37.6. *Let $A\in\mathbb{C}^{p\times p}$, $B\in\mathbb{C}^{q\times q}$ and let $\lambda_1,\ldots,\lambda_k$ and $\beta_1,\ldots,\beta_m$ denote the distinct eigenvalues of the matrices A and B, respectively. Then, for each choice of $C\in\mathbb{C}^{p\times q}$, the equation*

$$AX-XB=C$$

has exactly one solution $X\in\mathbb{C}^{p\times q}$ if and only if (37.7) *holds, i.e., if and only if $\sigma(A)\cap\sigma(B)=\emptyset$.*

Proof. This is an immediate corollary of Lemma 37.5 and the principle of conservation of dimension. □

Exercise 37.7. Let $A \in \mathbb{C}^{n\times n}$. Show that the **Lyapunov equation**

$$A^H X + XA = Q \tag{37.8}$$

has exactly one solution for each choice of $Q \in \mathbb{C}^{n\times n}$ if and only if $\sigma(A) \cap \sigma(-A^H) = \emptyset$.

Lemma 37.7. *If $A, Q \in \mathbb{C}^{n\times n}$ and if $\sigma(A) \subset \mathbb{C}_\mathrm{L}$ and $-Q \succeq O$, then the Lyapunov equation* (37.8) *has exactly one solution $X \in \mathbb{C}^{n\times n}$. Moreover, this solution is positive semidefinite.*

Proof. Since $\sigma(A) \subset \mathbb{C}_\mathrm{L}$, the matrix

$$Z = -\int_0^\infty e^{tA^H} Q e^{tA} dt$$

is well-defined and is positive semidefinite. Moreover,

$$\begin{aligned} A^H Z &= -\int_0^\infty A^H e^{tA^H} Q e^{tA} dt \\ &= -\int_0^\infty \left(\frac{d}{dt} e^{tA^H}\right) Q e^{tA} dt \\ &= -\left\{ e^{tA^H} Q e^{tA} \,\Big|_{t=0}^\infty - \int_0^\infty e^{tA^H} \frac{d}{dt}(Q e^{tA}) dt \right\} \\ &= Q + \int_0^\infty e^{tA^H} Q e^{tA} dt\, A \\ &= Q - ZA\,. \end{aligned}$$

Thus, the matrix Z is a solution of the Lyapunov equation (37.8) and hence, as the assumption $\sigma(A) \subset \mathbb{C}_\mathrm{L}$ implies that $\sigma(A) \cap \sigma(-A^H) = \emptyset$, Theorem 37.6 (as reformulated for (37.8) in Exercise 37.7) ensures that this equation has only one solution. Therefore, $X = Z$ is positive semidefinite. □

Exercise 37.8. Show that if, in addition to the assumptions that $\sigma(A) \subset \mathbb{C}_\mathrm{L}$ and $-Q \succeq O$, it is also assumed in Lemma 37.7 that $A, Q \in \mathbb{R}^{n\times n}$, then the solution X of (37.8) belongs to $\mathbb{R}^{n\times n}$.

Exercise 37.9. Let $A \in \mathbb{C}^{n\times n}$. Show that if $\sigma(A) \subset \mathbb{C}_\mathrm{R}$, the open right half-plane, then the equation $A^H X + XA = Q$ has exactly one solution for every choice of $Q \in \mathbb{C}^{n\times n}$ and that this solution can be expressed as

$$X = \int_0^\infty e^{-tA^H} Q e^{-tA} dt$$

for every choice of $Q \in \mathbb{C}^{n\times n}$. [HINT: Integrate the formula

$$A^H \int_0^\infty e^{-tA^H} Q e^{-tA} dt = -\int_0^\infty \frac{d}{dt}\left(e^{-tA^H} Q\right) e^{-tA} dt$$

by parts.]

Exercise 37.10. Show that in the setting of Exercise 37.9, the solution X can also be expressed as

$$X = -\frac{1}{2\pi}\int_{-\infty}^{\infty}(i\mu I_n + A^H)^{-1}Q(i\mu I_n - A)^{-1}d\mu\,.$$

[HINT: Write $A^H X = -\lim_{R\uparrow\infty}\frac{1}{2\pi i}\int_{-R}^{R}(A^H + i\mu I_n - i\mu I_n)\{\cdots\}d\mu$ and evaluate the integral by adding a semicircle of radius R to complete the contour keeping (35.4) in mind.]

Exercise 37.11. Let $A = \mathrm{diag}\{A_{11}, A_{22}\}$ be a block diagonal matrix in $\mathbb{C}^{n\times n}$ with $\sigma(A_{11}) \subset \mathbb{C}_R$ and $\sigma(A_{22}) \subset \mathbb{C}_L$, let $Q \in \mathbb{C}^{n\times n}$, and let $Y \in \mathbb{C}^{n\times n}$ and $Z \in \mathbb{C}^{n\times n}$ be solutions of the Lyapunov equation $A^H X + XA = Q$. Show that if Y and Z are written in block form consistent with the block decomposition of A, then $Y_{11} = Z_{11}$ and $Y_{22} = Z_{22}$.

Exercise 37.12. Let A, $Q \in \mathbb{C}^{n\times n}$. Show that if $\sigma(A)\cap i\mathbb{R} = \emptyset$ and if Y and Z are both solutions of the same Lyapunov equation $A^H X + XA = Q$ and $Y - Z \succeq O$, then $Y = Z$. [HINT: To warm up, suppose first that $A = \mathrm{diag}\{A_{11}, A_{22}\}$, where $\sigma(A_{11}) \subset \mathbb{C}_R$ and $\sigma(A_{22}) \subset \mathbb{C}_L$ and consider Exercise 37.11.]

Exercise 37.13. Let $A = \sum_{j=1}^{3}\mathbf{e}_j\mathbf{e}_{j+1}^T = C_0^{(4)}$ and let T denote the linear transformation from $\mathbb{C}^{4\times 4}$ into itself that is defined by the formula $TX = A^H X - XA$.

(a) Calculate $\dim \mathcal{N}_T$.

(b) Show that a matrix $X \in \mathbb{C}^{4\times 4}$ with entries x_{ij} is a solution of the matrix equation $A^H X - XA = \begin{bmatrix} 0 & -a & -b & -c \\ a & 0 & 0 & 0 \\ b & 0 & 0 & 0 \\ c & 0 & 0 & 0 \end{bmatrix}$ if and only if X is a **Hankel matrix** (i.e., $x_{ij} = x_{i+1,j-1}$) with $x_{11} = a$, $x_{12} = b$, and $x_{13} = c$.

Exercise 37.14. Show that if A, B, $C \in \mathbb{C}^{n\times n}$, then

$$X(t) = e^{tA}Ce^{-tB}$$

is a solution of the differential equation

$$X'(t) = AX(t) - X(t)B$$

that meets the initial condition $X(0) = C$.

37.3. $AX = XB$

The next result complements Lemma 37.5. It deals with the case where the nullspace $\mathcal{N}_T$ of the linear transformation T introduced in (37.6) is not equal to zero.

Lemma 37.8. *Let $A, X, B \in \mathbb{C}^{n\times n}$ and suppose that $AX = XB$. Then there exists a matrix $C \in \mathbb{C}^{n\times n}$ such that $AX = X(B+C)$ and $\sigma(B+C) \subseteq \sigma(A)$.*

Proof. If X is invertible, then $\sigma(A) = \sigma(B)$; i.e., the matrix $C = O$ does the trick. Suppose therefore that X is not invertible and that $C_\beta^{(k)}$ is a $k\times k$ Jordan cell in the Jordan decomposition of $B = UJU^{-1}$ such that $\beta \notin \sigma(A)$. Then there exists a subblock $W \in \mathbb{C}^{n\times k}$ of U such that $BW = WC_\beta^{(k)}$. Therefore,

$$AXW - XWC_\beta^{(k)} = AXW - XBW = O$$

and hence, as $\beta \notin \sigma(A)$, Lemma 37.5 implies that $XW = O$. Thus, if

$$B_1 = B + W(C_\alpha^{(k)} - C_\beta^{(k)})V^H\,,$$

where $V^H \in \mathbb{C}^{k\times n}$ is the block of rows in U^{-1} corresponding to the columns in W (i.e., $UJU^{-1} = WC_\beta^{(k)}V^H + \cdots$) and $\alpha \in \sigma(A)$, then

$$XB_1 = XB = AX,$$

and the diagonal entry of the block under consideration in the Jordan decomposition of B_1 now belongs to $\sigma(A)$ and not to $\sigma(B)$. Moreover, none of the other Jordan blocks in the Jordan decomposition of B are affected by this change. The same procedure can now be applied to change the diagonal entry of any Jordan cell in the Jordan decomposition of B_1 from a point that is not in $\sigma(A)$ to a point that is in $\sigma(A)$. The proof is completed by repeating this procedure. □

Exercise 37.15. Let $A, X, B \in \mathbb{C}^{n\times n}$. Show that if $AX = XB$ and the columns of $V \in \mathbb{C}^{n\times k}$ form a basis for $\mathcal{N}_X$, then there exists a matrix $L \in \mathbb{C}^{k\times n}$ such that $\sigma(B + VL) \subseteq \sigma(A)$.

37.4. Special classes of solutions

Let $A \in \mathbb{C}^{n\times n}$ and now let:

- $\mathcal{E}_+(A)$ = the number of zeros of det $(\lambda I_n - A)$ in $\mathbb{C}_\mathrm{R}$,
- $\mathcal{E}_-(A)$ = the number of zeros of det $(\lambda I_n - A)$ in $\mathbb{C}_\mathrm{L}$,
- $\mathcal{E}_0(A)$ = the number of zeros of det $(\lambda I_n - A)$ in $i\mathbb{R}$,

counting multiplicities in all three. The triple $(\mathcal{E}_+(A), \mathcal{E}_-(A), \mathcal{E}_0(A))$ is called the **inertia** of A. Since multiplicities are counted,

$$\mathcal{E}_+(A) + \mathcal{E}_-(A) + \mathcal{E}_0(A) = n\,.$$

Theorem 37.9. *Let $A \in \mathbb{C}^{n\times n}$ and suppose that $\sigma(A) \cap i\mathbb{R} = \emptyset$. Then there exists a Hermitian matrix $G \in \mathbb{C}^{n\times n}$ such that:*

(1) $A^H G + GA \succ O$.

(2) $\mathcal{E}_+(G) = \mathcal{E}_+(A)$, $\mathcal{E}_-(G) = \mathcal{E}_-(A)$, *and* $\mathcal{E}_0(G) = \mathcal{E}_0(A) = 0$.

Proof. Suppose first that $\mathcal{E}_+(A) = p \geq 1$ and $\mathcal{E}_-(A) = q \geq 1$. Then the assumption $\sigma(A) \cap i\mathbb{R} = \emptyset$ guarantees that $p + q = n$ and hence that A admits a Jordan decomposition UJU^{-1} of the form

$$A = U \begin{bmatrix} J_1 & O \\ O & J_2 \end{bmatrix} U^{-1}$$

with $J_1 \in \mathbb{C}^{p\times p}$, $\sigma(J_1) \subset \mathbb{C}_{\mathrm{R}}$, $J_2 \in \mathbb{C}^{q\times q}$, and $\sigma(J_2) \subset \mathbb{C}_{\mathrm{L}}$.

Let $P_{11} \in \mathbb{C}^{p\times p}$ and $P_{22} \in \mathbb{C}^{q\times q}$ be positive definite matrices. Then it is readily checked, much as in the proof of Lemma 37.7, that

$$X_{11} = \int_0^\infty e^{-tJ_1^H} P_{11} e^{-tJ_1} dt$$

is a positive definite solution of the equation

$$J_1^H X_{11} + X_{11} J_1 = P_{11}$$

and that

$$X_{22} = -\int_0^\infty e^{tJ_2^H} P_{22} e^{tJ_2} dt$$

is a negative definite solution of the equation

$$J_2^H X_{22} + X_{22} J_2 = P_{22}\,.$$

(The two integrals are well-defined because $\sigma(J_1) \subset \mathbb{C}_{\mathrm{R}} \implies \sigma(J_1^H) \subset \mathbb{C}_{\mathrm{R}}$ and $\sigma(J_2) \subset \mathbb{C}_{\mathrm{L}} \implies \sigma(J_2^H) \subset \mathbb{C}_{\mathrm{L}}$.) Let

$$X = \mathrm{diag}\,\{X_{11}, X_{22}\} \quad \text{and} \quad P = \mathrm{diag}\,\{P_{11}, P_{22}\}\,.$$

Then $J^H X + XJ = P$ and hence

$$(U^H)^{-1} J^H U^H (U^H)^{-1} X U^{-1} + (U^H)^{-1} X U^{-1} U J U^{-1} = (U^H)^{-1} P U^{-1}\,.$$

Thus, the matrix $G = (U^H)^{-1} X U^{-1}$ is a solution of the equation

$$A^H G + GA = (U^H)^{-1} P U^{-1}$$

and hence, as $(U^H)^{-1} P U^{-1} \succ O$, (1) holds when $p > 0$ and $q > 0$. The cases $p = 0,\ q = n$ and $p = n,\ q = 0$ are left to the reader.

Sylvester's inertia theorem (which is discussed in Section 28.6) guarantees that $\mathcal{E}_\pm(G) = \mathcal{E}_\pm(X) = \mathcal{E}_\pm(A)$, which justifies (2). □

Exercise 37.16. Complete the proof of Theorem 37.9 by verifying the cases $p = 0$ and $p = n$.

Exercise 37.17. Find a Hermitian matrix $G \in \mathbb{C}^{2\times 2}$ that fulfills the conditions of Theorem 37.9 when $A = \text{diag}\{1+i, 1-i\}$.

37.5. Supplementary notes

This chapter is adapted from Chapter 18 and Section 20.8 of [**30**]. A number of refinements may be found in Lancaster and Tismenetsky [**55**].