Chapter 12

Triangular factorization and positive definite matrices

Half the harm that is done in this world
Is due to people who want to feel important.
They don't mean to do harm—but the harm does not interest them.
Or they do not see it, or they justify it
Because they are absorbed in the endless struggle
To think well of themselves.

T. S. Elliot, *The Cocktail Party*

This chapter is devoted primarily to positive definite and semidefinite matrices and related applications. To add perspective, however, it is convenient to begin with some general observations on the triangular factorization of matrices. In a sense this is not new, because the formula

$$EPA = U \quad \text{or, equivalently,} \quad A = P^{-1}E^{-1}U$$

that emerged from the discussion of Gaussian elimination is almost a triangular factorization. Under appropriate extra assumptions on the matrix $A \in \mathbb{C}^{n\times n}$, the formula $A = P^{-1}E^{-1}U$ holds with $P = I_n$.

• **WARNING:** We remind the reader that from now on $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle_{st}$, the standard inner product, and $\|\mathbf{u}\| = \|\mathbf{u}\|_2$ for vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$, unless indicated otherwise. Correspondingly, $\|A\| = \|A\|_{2,2}$ for matrices A.

12.1. A detour on triangular factorization

The notation

$$A_{[j,k]} = \begin{bmatrix} a_{jj} & \cdots & a_{jk} \\ \vdots & \cdots & \vdots \\ a_{kj} & \cdots & a_{kk} \end{bmatrix} \quad \text{for} \quad A \in \mathbb{C}^{n\times n} \quad \text{and} \quad 1 \le j \le k \le n \tag{12.1}$$

will be convenient.

Theorem 12.1. *A matrix $A \in \mathbb{C}^{n\times n}$ admits a factorization of the form*

$$A = LDU\,, \tag{12.2}$$

where $L \in \mathbb{C}^{n\times n}$ is a lower triangular matrix with ones on the diagonal, $U \in \mathbb{C}^{n\times n}$ is an upper triangular matrix with ones on the diagonal and $D \in \mathbb{C}^{n\times n}$ is an invertible diagonal matrix, if and only if the submatrices

$$A_{[1,k]} \quad \textit{are invertible for} \quad k = 1, \ldots, n\,. \tag{12.3}$$

Moreover, if the conditions in (12.3) *are met, then there is only one set of matrices, L, D and U, with the stated properties for which* (12.2) *holds.*

Proof. Suppose first that the condition (12.3) is in force. Then, upon expressing

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

in block form with $A_{11} \in \mathbb{C}^{p\times p}$, $A_{22} \in \mathbb{C}^{q\times q}$ and $p + q = n$, we can invoke the first Schur complement formula

$$A = \begin{bmatrix} I_p & O \\ A_{21}A_{11}^{-1} & I_q \end{bmatrix} \begin{bmatrix} A_{11} & O \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I_p & A_{11}^{-1}A_{12} \\ O & I_q \end{bmatrix}$$

repeatedly to obtain the asserted factorization formula (12.2). Thus, if $A_{11} = A_{[1,n-1]}$, then $\alpha_n = A_{22} - A_{21}A_{11}^{-1}A_{12}$ is a nonzero number and the exhibited formula states that

$$A = L_n \begin{bmatrix} A_{[1,n-1]} & O \\ O & \alpha_n \end{bmatrix} U_n\,,$$

where $L_n \in \mathbb{C}^{n\times n}$ is a lower triangular matrix with ones on the diagonal and $U_n \in \mathbb{C}^{n\times n}$ is an upper triangular matrix with ones on the diagonal. The next step is to apply the same procedure to the $(n-1)\times(n-1)$ matrix $A_{[1,n-1]}$. This yields a factorization of the form

$$A_{[1,n-1]} = \widetilde{L}_{n-1} \begin{bmatrix} A_{[1,n-2]} & O \\ O & \alpha_{n-1} \end{bmatrix} \widetilde{U}_{n-1}\,,$$

where $\widetilde{L}_{n-1} \in \mathbb{C}^{(n-1)\times(n-1)}$ is a lower triangular matrix with ones on the diagonal and $\widetilde{U}_{n-1} \in \mathbb{C}^{(n-1)\times(n-1)}$ is an upper triangular matrix with ones on the diagonal. Therefore,

$$A = L_n \begin{bmatrix} \widetilde{L}_{n-1} & O \\ O & 1 \end{bmatrix} \begin{bmatrix} A_{[1,n-2]} & O & O \\ O & \alpha_{n-1} & O \\ O & O & \alpha_n \end{bmatrix} \begin{bmatrix} \widetilde{U}_{n-1} & O \\ O & 1 \end{bmatrix} U_n \,,$$

which is one step further down the line. The final formula is obtained by iterating this procedure $n-3$ more times.

Conversely, if A admits a factorization of the form (12.2) with the stated properties, then, upon writing the factorization in block form as

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & O \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} D_{11} & O \\ O & D_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ O & U_{22} \end{bmatrix},$$

it is readily checked that

$$A_{11} = L_{11}D_{11}U_{11}$$

or, equivalently, that

$$A_{[1,k]} = L_{[1,k]}D_{[1,k]}U_{[1,k]} \quad \text{for} \quad k = 1, \ldots, n\,.$$

Thus, $A_{[1,k]}$ is invertible for $k = 1, \ldots, n$, as needed.

To verify uniqueness, suppose that $A = L_1D_1U_1 = L_2D_2U_2$. Then the identity $L_2^{-1}L_1D_1 = D_2U_2U_1^{-1}$ implies that $L_2^{-1}L_1D_1$ is both upper and lower triangular and hence must be a diagonal matrix, which is readily seen to be equal to D_1. Therefore, $L_1 = L_2$ and by an analogous argument $U_1 = U_2$, which then forces $D_1 = D_2$. □

Theorem 12.2. *A matrix $A \in \mathbb{C}^{n\times n}$ admits a factorization of the form*

$$A = UDL\,, \tag{12.4}$$

where $L \in \mathbb{C}^{n\times n}$ is a lower triangular matrix with ones on the diagonal, $U \in \mathbb{C}^{n\times n}$ is an upper triangular matrix with ones on the diagonal and $D \in \mathbb{C}^{n\times n}$ is an invertible diagonal matrix, if and only if the blocks

$$A_{[k,n]} \quad \text{are invertible for} \quad k = 1, \ldots, n. \tag{12.5}$$

Moreover, if the conditions in (12.5) *are met, then there is only one set of matrices, L, D and U, with the stated properties for which* (12.4) *holds.*

Proof. The details are left to the reader. They are easily filled in with the proof of Theorem 12.1 as a guide. □

Exercise 12.1. Prove Theorem 12.2.

Exercise 12.2. Let $P_k = \text{diag}\,\{I_k, O_{(n-k)\times(n-k)}\}$. Show that

(a) $A \in \mathbb{C}^{n\times n}$ is upper triangular if and only if $AP_k = P_kAP_k$ for $k = 1, \dots, n$.

(b) $A \in \mathbb{C}^{n\times n}$ is lower triangular if and only if $P_kA = P_kAP_k$ for $k = 1, \dots, n$.

Exercise 12.3. Show that if $L \in \mathbb{C}^{n\times n}$ is lower triangular, $U \in \mathbb{C}^{n\times n}$ is upper triangular and $D \in \mathbb{C}^{n\times n}$ is diagonal, then

$$(LDU)_{[1,k]} = L_{[1,k]}D_{[1,k]}U_{[1,k]} \quad \text{and} \quad (UDL)_{[k,n]} = U_{[k,n]}D_{[k,n]}L_{[k,n]} \tag{12.6}$$

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

12.2. Definite and semidefinite matrices

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

$$\langle A\mathbf{x}, \mathbf{x}\rangle \geq 0 \text{ for every } \mathbf{x} \in \mathbb{C}^n; \tag{12.7}$$

it is said to be **positive definite over** $\mathbb{C}^n$ if

$$\langle A\mathbf{x}, \mathbf{x}\rangle > 0 \text{ for every nonzero vector } \mathbf{x} \in \mathbb{C}^n. \tag{12.8}$$

The **notation**
$A \succeq O$ will be used to indicate that the matrix $A \in \mathbb{C}^{n\times n}$ is positive semidefinite over $\mathbb{C}^n$. Similarly, the notation
$A \succ O$ will be used to indicate that the matrix $A \in \mathbb{C}^{n\times n}$ is positive definite over $\mathbb{C}^n$. Moreover, if $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$, then
$A \succeq B$ and $A \succ B$ means that $A - B \succeq O$ and $A - B \succ O$, respectively. Correspondingly, a matrix $A \in \mathbb{C}^{n\times n}$ is said to be **negative semidefinite** over $\mathbb{C}^n$ if $-A \succeq O$ and **negative definite** over $\mathbb{C}^n$ if $-A \succ O$.

Lemma 12.3. *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.*

(3) *$A \succ O \iff$ the eigenvalues of A are all positive $\iff$* $\det A > 0$.

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 $\lambda \geq 0$, since $\langle \mathbf{x}, \mathbf{x}\rangle > 0$. This justifies assertion (2); the proof of (3) is left to the reader. □

WARNING: The conclusions of Lemma 12.3 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

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

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.

Exercise 12.4. Let $A \in \mathbb{C}^{n\times n}$. Show that if $A \succeq O$, then

$$A \succ O \iff \text{all the eigenvalues of } A \text{ are positive} \iff \det A > 0\,.$$

Exercise 12.5. Show that if $V \in \mathbb{C}^{n\times n}$ is invertible, then

$$A \succ O \iff V^H A V \succ O\,.$$

Exercise 12.6. Show that if $V \in \mathbb{C}^{n\times k}$ and $\operatorname{rank} V = k$, then

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

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

Exercise 12.7. Show that if the $n \times n$ matrix $A = [a_{ij}]$, $i, j = 1, \ldots, n$, is positive semidefinite over $\mathbb{C}^n$, then $|a_{ij}|^2 \leq a_{ii}a_{jj}$.

Exercise 12.8. 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

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

Exercise 12.9. Show that if $A \in \mathbb{C}^{p\times q}$, then

$$\|A\| \leq 1 \iff I_q - A^H A \succeq O \iff I_p - AA^H \succeq O\,.$$

[HINT: Use the singular value decomposition of A.]

Exercise 12.10. Show that if $A \in \mathbb{C}^{n\times n}$ and $A = A^H$, then

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

Exercise 12.11. Show that if $A \in \mathbb{C}^{n\times n}$ and $A \succeq O$, then

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

Exercise 12.12. 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 $\langle A\mathbf{x}, \mathbf{x}\rangle$ for eigenvectors $\mathbf{x}$ of U.]

12.3. Characterizations of positive definite matrices

A basic question of interest is to check when an $n \times n$ matrix $A = [a_{ij}]$, $i, j = 1, \ldots, n$ is positive definite over $\mathbb{C}^n$. The next theorem supplies a number of equivalent characterizations.

Theorem 12.4. *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 of* A *are all positive; i.e.* $\lambda_j > 0$ *for* $j = 1, \ldots, n$.
(3) $A = V^H V$ *for some* $n \times n$ *invertible matrix* V.
(4) $A = A^H$ *and* $\det A_{[1,k]} > 0 \quad \text{for} \quad k = 1, \ldots, n$.
(5) $A = LL^H$, *where* L *is a lower triangular invertible matrix.*
(6) $A = A^H$ *and* $\det A_{[k,n]} > 0 \quad \text{for} \quad k = 1, \ldots, n$.
(7) $A = UU^H$, *where* U *is an upper triangular invertible matrix.*

Proof. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ denote an orthonormal set of eigenvectors corresponding to $\lambda_1, \ldots, \lambda_n$. Then, since $\langle \mathbf{u}_j, \mathbf{u}_j\rangle = 1$, the formula

$$\lambda_j = \lambda_j \langle \mathbf{u}_j, \mathbf{u}_j\rangle = \langle A\mathbf{u}_j, \mathbf{u}_j\rangle\,, \quad \text{for} \quad j = 1, \ldots, n\,,$$

clearly displays the fact that (1)$\Longrightarrow$(2). Next, if (2) is in force, then

$$D = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$$

admits a square root

$$D^{1/2} = \text{diag}\{\sqrt{\lambda}_1, \ldots, \sqrt{\lambda}_n\}$$

and hence the diagonalization formula

$$A = UDU^H \quad \text{with} \quad U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}$$

can be rewritten as

$$A = V^H V \quad \text{with} \quad V = D^{1/2} U^H \quad \text{invertible}\,.$$

Thus, (2) $\Longrightarrow$ (3), and, upon setting

$$\Pi_k = \begin{bmatrix} I_k \\ O_{(n-k)\times k} \end{bmatrix} \quad \text{and} \quad V_1 = V\Pi_k \,,$$

it is readily seen that

$$\begin{aligned} A_{[1,k]} &= \Pi_k^H A \Pi_k \\ &= \Pi_k^H V^H V \Pi_k \\ &= V_1^H V_1 \,. \end{aligned}$$

But this implies that

$$\begin{aligned} \langle \Pi_k^H A \Pi_k \mathbf{x}, \mathbf{x} \rangle &= \langle V_1^H V_1 \mathbf{x}, \mathbf{x} \rangle \\ &= \langle V_1 \mathbf{x}, V_1 \mathbf{x} \rangle \\ &> 0 \end{aligned}$$

for every nonzero vector $\mathbf{x} \in \mathbb{C}^k$, since V_1 has k linearly independent columns. Therefore, (3) implies (4). However, in view of Theorem 12.1, (4) implies that $A = L_1 D U_1$, where $L_1 \in \mathbb{C}^{n\times n}$ is a lower triangular matrix with ones on the diagonal, $U_1 \in \mathbb{C}^{n\times n}$ is an upper triangular matrix with ones on the diagonal and $D \in \mathbb{C}^{n\times n}$ is an invertible diagonal matrix. Thus, as $A = A^H$ in the present setting, it follows that

$$(U_1^H)^{-1} L_1 = D^H L_1^H U_1^{-1} D^{-1}$$

and therefore, since the left-hand side of the last identity is lower triangular and the right-hand side is upper triangular, the matrix $(U_1^H)^{-1} L_1$ must be a diagonal matrix. Moreover, since both U_1 and L_1 have ones on their diagonals, it follows that $(U_1^H)^{-1} L_1 = I_n$, i.e., $U_1^H = L_1$. Consequently,

$$A_{[1,k]} = \Pi_k^H A \Pi_k = \Pi_k^H U_1^H D U_1 \Pi_k = (\Pi_k^H U_1^H \Pi_k)(\Pi_k^H D \Pi_k)(\Pi_k^H U_1 \Pi_k)$$

and

$$\det A_{[1,k]} = \det\{(L_1)_{[1,k]}\} \det\{D_{[1,k]}\} \det\{(U_1)_{[1,k]}\} = d_{11} \cdots d_{kk}$$

for $k = 1, \ldots, n$. Therefore, D is positive definite over $\mathbb{C}^n$ as is $A = L_1 D L_1^H$. The formula advertised in (5) is obtained by setting $L = L_1 D^{1/2}$. It is also clear that (5) implies (1). Next, the matrix identity

$$\begin{bmatrix} O & I_{n-k} \\ I_k & O \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} O & I_k \\ I_{n-k} & O \end{bmatrix} = \begin{bmatrix} A_{22} & A_{21} \\ A_{12} & A_{11} \end{bmatrix}$$

clearly displays the fact that (4) holds if and only if (6) holds. Moreover, since (7) implies (1), it remains only to show that (6) implies (7) in order to complete the proof. This is left to the reader as an exercise.

Exercise 12.13. Verify the implication (6)$\Longrightarrow$(7) in Theorem 12.4.

Exercise 12.14. Let $A \in \mathbb{C}^{n\times n}$ and let $D_A = \mathrm{diag}\{a_{11}\ldots, a_{nn}\}$ denote the $n\times n$ diagonal matrix with diagonal entries equal to the diagonal entries of A. Show that D_A is multiplicative on upper triangular matrices in the sense that if A and B are both $n\times n$ upper triangular matrices, then $D_{AB} = D_A D_B$ and thus, if A is invertible, $D_{A^{-1}} = (D_A)^{-1}$.

Remark 12.5. The proof that a matrix A that is positive definite over $\mathbb{C}^n$ admits a factorization of the form $A = LL^H$ for some lower triangular invertible matrix L can also be based on the general factorization formula $EPA = U$ that was established as a byproduct of Gaussian elimination. The proof may be split into two parts. The first part is to check that, since $A \succ O$, there always exists a lower triangular matrix E with ones on the diagonal such that

$$EA = U$$

is in upper echelon form and hence upper triangular. Once this is verified, the second part is easy: The identity

$$UE^H = EAE^H = (EAE^H)^H = EU^H$$

implies that $D = UE^H$ is a positive definite matrix that is both lower triangular and upper triangular. Therefore,

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

is a diagonal matrix with $d_{jj} > 0$ for $j = 1, \ldots, n$. Thus, D has a positive square root:

$$D = F^2 ,$$

where

$$F = \mathrm{diag}\{(d_{11})^{1/2}, \ldots, (d_{nn})^{1/2}\} ,$$

and consequently

$$A = (E^{-1}F)(E^{-1}F)^H .$$

This is a representation of the desired form, since $L = E^{-1}F$ is lower triangular.

Notice that d_{jj} is the j'th pivot of U and $E^{-1} = (D^{-1}U)^H$.

Exercise 12.15. Show that if $A \in \mathbb{C}^{3\times 3}$ and $A \succ O$, then there exists a lower triangular matrix E with ones on the diagonal such that EA is upper triangular.

Exercise 12.16. Show that if $A \in \mathbb{C}^{3\times 3}$ and $A \succ O$, then there exists an upper triangular matrix F with ones on the diagonal such that FA is lower triangular. [HINT: This is very much like Gaussian elimination in spirit, except that now you work from the bottom row up instead of from the top row down.]

Exercise 12.17. Let $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$, where $A_1 \in \mathbb{C}^{n\times s}$, $A_2 \in \mathbb{C}^{n\times t}$ and $s+t=r$. Show that if $\operatorname{rank} A = r$, then the matrices $A^H A$, $A_1^H A_1$, $A_2^H A_2$ and $A_2^H A_2 - A_2^H A_1 (A_1^H A_1)^{-1} A_1^H A_2$ are all positive definite (over complex spaces of appropriate sizes).

Exercise 12.18. Show that if $x \in \mathbb{R}$, then the matrix $\begin{bmatrix} 3 & 2 & x \\ 2 & 2 & 1 \\ x & 1 & 1 \end{bmatrix}$ will be positive definite over $\mathbb{C}^3$ if and only if $(x-1)^2 < 1/2$.

12.4. An application of factorization

Lemma 12.6. *Let $A \in \mathbb{C}^{n\times n}$ and suppose that $A \succ O$ and that*

$$A = \begin{bmatrix} a_{11} & \mathbf{c}^H \\ \mathbf{c} & D \end{bmatrix} \quad \text{and} \quad A^{-1} = \begin{bmatrix} b_{11} & \mathbf{d}^H \\ \mathbf{d} & E \end{bmatrix},$$

where $\mathbf{c}, \mathbf{d} \in \mathbb{C}^{n-1}$ and $D, E \in \mathbb{C}^{(n-1)\times(n-1)}$. Then

(12.9)

$$\min_{x_2,\ldots,x_n} \langle A(\mathbf{e}_1 - \sum_{j=2}^n x_j \mathbf{e}_j),\, \mathbf{e}_1 - \sum_{j=2}^n x_j \mathbf{e}_j \rangle = \{a_{11} - \mathbf{c}^H D^{-1} \mathbf{c}\}^{1/2} = b_{11}^{-1/2}.$$

Proof. In view of Theorem 12.4, $A = LL^H$, where $L \in \mathbb{C}^{n\times n}$ is an invertible lower triangular matrix. Therefore,

$$\langle A(\mathbf{e}_1 - \sum_{j=2}^n x_j \mathbf{e}_j),\, \mathbf{e}_1 - \sum_{j=2}^n x_j \mathbf{e}_j \rangle = \|L^H(\mathbf{e}_1 - \sum_{j=2}^n x_j \mathbf{e}_j)\|^2 .$$

Let $\mathbf{v}_j = L^H \mathbf{e}_j$ for $j = 1, \ldots, n$,

$$V = \begin{bmatrix} \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \quad \text{and} \quad \mathcal{V} = \operatorname{span}\{\mathbf{v}_2, \ldots, \mathbf{v}_n\}.$$

Then, since L is invertible, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent and hence the orthogonal projection $P_{\mathcal{V}}$ of $\mathbb{C}^n$ onto $\mathcal{V}$ is given by the formula

$$P_{\mathcal{V}} = V(V^H V)^{-1} V^H .$$

Thus the minimum of interest is equal to

$$\|\mathbf{v}_1 - P_{\mathcal{V}}\mathbf{v}_1\|^2 = \langle \mathbf{v}_1 - P_{\mathcal{V}}\mathbf{v}_1, \mathbf{v}_1 \rangle = \|\mathbf{v}_1\|^2 - \mathbf{v}_1^H V (V^H V)^{-1} V^H \mathbf{v}_1 .$$

It remains to express this number in terms of the entries in the original matrix A by taking advantage of the formulas

$$\begin{bmatrix} a_{11} & \mathbf{c}^H \\ \mathbf{c} & D \end{bmatrix} = A = LL^H = \begin{bmatrix} \mathbf{v}_1^H \\ V^H \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & V \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^H \mathbf{v}_1 & \mathbf{v}_1^H V \\ V^H \mathbf{v}_1 & V^H V \end{bmatrix}.$$

The rest is left to the reader. □

Exercise 12.19. Complete the proof of Lemma 12.6 .

Exercise 12.20. Let $A \in \mathbb{C}^{n\times n}$ and assume that $A \succ O$. Evaluate

$$\min_{x_1,\dots,x_{n-1}} \langle A(\mathbf{e}_n - \sum_{j=1}^{n-1} x_j\mathbf{e}_j),\, \mathbf{e}_n - \sum_{j=1}^{n-1} x_j\mathbf{e}_j\rangle$$

in terms of the entries in A and the entries in A^{-1}.

12.5. Positive definite Toeplitz matrices

In this section we shall sketch some applications related to factorization in the special case that the given positive definite matrix is a Toeplitz matrix.

Theorem 12.7. *Let*

$$T_n = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-n} \\ t_1 & t_0 & \cdots & t_{1-n} \\ \vdots & \ddots & \ddots & \vdots \\ t_n & \cdots & t_1 & t_0 \end{bmatrix} \succ O \quad \textit{and} \quad \Gamma_n = \begin{bmatrix} \gamma_{00}^{(n)} & \cdots & \gamma_{0n}^{(n)} \\ \vdots & & \vdots \\ \gamma_{nn}^{(n)} & \cdots & \gamma_{nn}^{(n)} \end{bmatrix} = T_n^{-1}$$

and let

$$p_n(\lambda) = \sum_{j=0}^{n} \gamma_{j0}^{(n)}\lambda^j \quad \textit{and} \quad q_n(\lambda) = \sum_{j=0}^{n} \gamma_{jn}^{(n)}\lambda^j\,.$$

Then:

(1) $\sum_{i,j=0}^{n} \lambda^i \gamma_{ij}^{(n)} \overline{\omega}^j$ *is related to the polynomials* $p_n(\lambda)$ *and* $q_n(\lambda)$ *by the formula*

$$\sum_{i,j=0}^{n} \lambda^i \gamma_{ij}^{(n)} \overline{\omega}^j = \frac{p_n(\lambda)\{\gamma_{00}^{(n)}\}^{-1}p_n(\omega)^* - \lambda\overline{\omega}q_n(\lambda)\{\gamma_{nn}^{(n)}\}^{-1}q_n(\omega)^*}{1-\lambda\overline{\omega}}\,. \tag{12.10}$$

(2) $\gamma_{ij}^{(n)} = \gamma_{n-j,n-i}^{(n)} = \overline{\gamma_{ji}^{(n)}}$ *for* $i,j = 0,\dots,n$.

(3) $q_n(\lambda) = \lambda^n \overline{p_n(1/\overline{\lambda})}$.

(4) *The polynomial* $p_n(\lambda)$ *has no roots in the closed unit disc.*

(5) *If* $S_n \succ O$ *is an* $(n+1)\times(n+1)$ *Toeplitz matrix such that* $T_n^{-1}\mathbf{e}_1 = S_n^{-1}\mathbf{e}_1$*, then* $S_n = T_n$.

Proof. Since $T_n \succ O \Longrightarrow \Gamma_n \succ O$ we can invoke the Schur complement formulas to write

(12.11)

$$\begin{aligned} \Gamma_n &= \begin{bmatrix} \gamma_{00}^{(n)} & \mathbf{x}^H \\ \mathbf{x} & X \end{bmatrix} \\ &= \begin{bmatrix} 1 & O \\ \mathbf{x}\{\gamma_{00}^{(n)}\}^{-1} & I_n \end{bmatrix} \begin{bmatrix} \gamma_{00}^{(n)} & O \\ O & X - \mathbf{x}\{\gamma_{00}^{(n)}\}^{-1}\mathbf{x}^H \end{bmatrix} \begin{bmatrix} 1 & \{\gamma_{00}^{(n)}\}^{-1}\mathbf{x}^H \\ O & I_n \end{bmatrix} \end{aligned}$$

and

(12.12)

$$\begin{aligned}\Gamma_n &= \begin{bmatrix} Y & \mathbf{y} \\ \mathbf{y}^H & \gamma_{nn}^{(n)} \end{bmatrix} \\ &= \begin{bmatrix} I_n & \mathbf{y}\{\gamma_{nn}^{(n)}\}^{-1} \\ O & 1 \end{bmatrix} \begin{bmatrix} Y - \mathbf{y}\{\gamma_{nn}^{(n)}\}^{-1}\mathbf{y}^H & O \\ O & \gamma_{nn}^{(n)} \end{bmatrix} \begin{bmatrix} I_n & O \\ \{\gamma_{nn}^{(n)}\}^{-1}\mathbf{y}^H & 1 \end{bmatrix},\end{aligned}$$

where $\mathbf{x}^H = \begin{bmatrix}\gamma_{01}^{(n)} & \cdots & \gamma_{0n}^{(n)}\end{bmatrix}$, $\mathbf{y}^H = \begin{bmatrix}\gamma_{n0}^{(n)} & \cdots & \gamma_{n,n-1}^{(n)}\end{bmatrix}$, X denotes the lower right-hand $n \times n$ corner of Γ_n and Y denotes the upper left-hand $n \times n$ corner of Γ_n. Thus,

$$\begin{aligned}\begin{bmatrix}1 & \lambda & \cdots & \lambda^n\end{bmatrix} \Gamma_n \begin{bmatrix} 1 \\ \overline{\omega} \\ \vdots \\ \overline{\omega}^n \end{bmatrix} &= p_n(\lambda)\{\gamma_{00}^{(n)}\}^{-1} p_n(\omega)^* \\ &\quad + \begin{bmatrix}\lambda & \cdots & \lambda^n\end{bmatrix} \left[X - \mathbf{x}\{\gamma_{00}^{(n)}\}^{-1}\mathbf{x}^H\right] \begin{bmatrix} \overline{\omega} \\ \vdots \\ \overline{\omega}^n \end{bmatrix},\end{aligned}$$

and a second development based on the second Schur complement formula yields the identity

$$\begin{aligned}\begin{bmatrix}1 & \lambda & \cdots & \lambda^n\end{bmatrix} \Gamma_n \begin{bmatrix} 1 \\ \overline{\omega} \\ \vdots \\ \overline{\omega}^n \end{bmatrix} &= q_n(\lambda)\{\gamma_{nn}^{(n)}\}^{-1} q_n(\omega)^* \\ &\quad + \begin{bmatrix}1 & \cdots & \lambda^{n-1}\end{bmatrix} \left[Y - \mathbf{y}\{\gamma_{nn}^{(n)}\}^{-1}\mathbf{y}^H\right] \begin{bmatrix} 1 \\ \vdots \\ \overline{\omega}^{n-1} \end{bmatrix}.\end{aligned}$$

The proof of formula (12.10) is now completed by verifying that

$$X - \mathbf{x}\{\gamma_{00}^{(n)}\}^{-1}\mathbf{x}^H = \Gamma_{n-1} = Y - \mathbf{y}\{\gamma_{nn}^{(n)}\}^{-1}\mathbf{y}^H, \tag{12.13}$$

where

(12.14)

$$\Gamma_k = \begin{bmatrix} \gamma_{00}^{(k)} & \cdots & \gamma_{0k}^{(k)} \\ \vdots & & \vdots \\ \gamma_{kk}^{(k)} & \cdots & \gamma_{kk}^{(k)} \end{bmatrix} = T_k^{-1} \quad \text{and} \quad T_k = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-k} \\ \vdots & \vdots & & \vdots \\ t_k & t_{k-1} & \cdots & t_0 \end{bmatrix}$$

for $k = 0, \ldots, n$. The details are left to the reader as an exercise.

To verify (2), let δ_{ij} denote the Kronecker delta symbol, i.e.,

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} ,$$

and write

$$\begin{aligned} \sum_{s=0}^{n} t_{i-s}\gamma_{sj}^{(n)} &= \delta_{ij} = \delta_{n-i,n-j} = \sum_{s=0}^{n} t_{n-i-s}\gamma_{s,n-j}^{(n)} \\ &= \sum_{s=0}^{n} t_{n-i-(n-s)}\gamma_{n-s,n-j}^{(n)} = \sum_{s=0}^{n} \gamma_{n-s,n-j}^{(n)} t_{s-i} \\ &= \sum_{s=0}^{n} \gamma_{js}^{(n)} t_{s-i} , \end{aligned}$$

which, upon comparing the last two sums, yields the first formula in (2); the second follows from the fact that T_n and Γ_n are Hermitian matrices.

Suppose next that $p_n(\omega) = 0$. Then formula (12.10) implies that

$$-|\omega|^2 q_n(\omega) \left\{ \gamma_{nn}^{(n)} \right\}^{-1} q_n(\omega)^* = (1 - |\omega|^2) \sum_{i,j=0}^{n} \omega^i \gamma_{ij}^{(n)} \overline{\omega}^j , \tag{12.15}$$

which is impossible if $|\omega| < 1$, because then the left-hand side of the identity (12.15) is less than or equal to zero, whereas the right-hand side is positive. Thus, $|p_n(\omega)| > 0$ if $|\omega| < 1$. Moreover, if $|\omega| = 1$, then formula (12.15) implies that $q_n(\omega) = 0$ also. Thus, formula (12.10) implies that

$$0 = \begin{bmatrix} 1 & \lambda & \cdots & \lambda^n \end{bmatrix} \Gamma_n \begin{bmatrix} 1 \\ \overline{\omega} \\ \vdots \\ \overline{\omega}^n \end{bmatrix} \tag{12.16}$$

for all $\lambda \in \mathbb{C}$, which is impossible.

Finally, in view of items (2) and (3), formula (12.10) can be rewritten as

$$\sum_{i,j=0}^{n} \lambda^i \gamma_{ij}^{(n)} \overline{\omega}^j = \frac{p_n(\lambda)\{\gamma_{00}^{(n)}\}^{-1} p_n(\omega)^* - \lambda^{n+1} p_n(1/\lambda)\{\gamma_{00}^{(n)}\}^{-1} \overline{\omega}^{n+1} p_n(1/\omega)^*}{1 - \lambda\overline{\omega}} , \tag{12.17}$$

which exhibits the fact that if $T_n \succ O$, then all the entries $\gamma_{ij}^{(n)}$ are completely determined by the first column of Γ_n, and hence serves to verify (5). □

Exercise 12.21. Verify the identity (12.13) in the setting of Theorem 12.7. [HINT: Use the Schur complement formulas to calculate Γ_n^{-1} alias T_n from the two block decompositions (12.11) and (12.12).]

Theorem 12.7 is just the tip of the iceberg; it can be generalized in many directions. Some indications are sketched in the next several exercises and the next section, all of which can be skipped without loss of continuity.

Exercise 12.22. Show that if $T_n \succ O$, then Γ_n admits the triangular factorization

$$\Gamma_n = L_n D_n L_n^H \,, \tag{12.18}$$

where

(12.19)

$$L_n = \begin{bmatrix} \gamma_{00}^{(n)} & O & \cdots & O \\ \gamma_{10}^{(n)} & \gamma_{00}^{(n-1)} & \cdots & O \\ \vdots & & \ddots & \vdots \\ \gamma_{n0}^{(n)} & \gamma_{n-1,0}^{(n-1)} & \cdots & \gamma_{00}^{(0)} \end{bmatrix}, \quad L_n^H = \begin{bmatrix} \gamma_{00}^{(n)} & \gamma_{01}^{(n)} & \cdots & \gamma_{0n}^{(n)} \\ O & \gamma_{00}^{(n-1)} & \cdots & \gamma_{0,n-1}^{(n-1)} \\ \vdots & & \ddots & \vdots \\ O & O & \cdots & \gamma_{00}^{(0)} \end{bmatrix}$$

and

$$D_n = \mathrm{diag}\,\{\{\gamma_{00}^{(n)}\}^{-1}, \{\gamma_{00}^{(n-1)}\}^{-1}, \dots, \{\gamma_{00}^{(0)}\}^{-1}\}\,. \tag{12.20}$$

Exercise 12.23. Find formulas in terms of $\gamma_{ij}^{(k)}$ analogous to those given in the preceding exercise for the factors in a triangular factorization of the form

$$\Gamma_n = U_n D_n U_n^H \,, \tag{12.21}$$

where $T_n \succ O$, U_n is an upper triangular matrix and D_n is a diagonal matrix.

Positive definite Toeplitz matrices play a significant role in the theory of prediction of stationary stochastic sequences, which, when recast in the language of trigonometric approximation, focuses on evaluations of the following sort:

(12.22)

$$\min \left\{ \frac{1}{2\pi} \int_0^{2\pi} |e^{in\theta} - \sum_{j=0}^{n-1} c_j e^{ij\theta}|^2 f(e^{i\theta}) d\theta :\, c_0, \dots, c_{n-1} \in \mathbb{C} \right\} = \{\gamma_{nn}^{(n)}\}^{-1}$$

and

(12.23)

$$\min \left\{ \frac{1}{2\pi} \int_0^{2\pi} |1 - \sum_{j=1}^{n} c_j e^{ij\theta}|^2 f(e^{i\theta}) d\theta :\, c_1, \dots, c_n \in \mathbb{C} \right\} = \{\gamma_{00}^{(n)}\}^{-1} ,$$

where, for ease of exposition, we assume that $f(e^{i\theta})$ is a continuous function of θ on the interval $0 \le \theta \le 2\pi$ such that $f(e^{i\theta}) > 0$ on this interval. Let

$T_n = T_n(f)$ denote the Toeplitz matrix with entries

$$t_j = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})e^{-ij\theta}d\theta \quad \text{for} \quad j = 0, \pm 1, \pm 2, \dots . \tag{12.24}$$

Exercise 12.24. Show that

$$\text{if} \quad \mathbf{b} = \begin{bmatrix} b_0 \\ \vdots \\ b_n \end{bmatrix}, \quad \text{then} \quad \frac{1}{2\pi}\int_0^{2\pi} |\sum_{j=0}^{n} b_j e^{ij\theta}|^2 f(e^{i\theta})d\theta = \mathbf{b}^H T_n \mathbf{b}\,. \tag{12.25}$$

Exercise 12.25. Show that if $T_n \succ O$ and $\mathbf{u}^H = \begin{bmatrix} t_n & \cdots & t_1 \end{bmatrix}$, then

$$T_n = \begin{bmatrix} I_n & \mathbf{0} \\ \mathbf{u}^H T_{n-1}^{-1} & 1 \end{bmatrix} \begin{bmatrix} T_{n-1} & \mathbf{0} \\ \mathbf{0}^H & \{\gamma_{nn}^{(n)}\}^{-1} \end{bmatrix} \begin{bmatrix} I_n & T_{n-1}^{-1}\mathbf{u} \\ \mathbf{0}^H & 1 \end{bmatrix}. \tag{12.26}$$

Exercise 12.26. Show that if $T_n \succ O$ and $\mathbf{v}^H = \begin{bmatrix} t_1 & \cdots & t_n \end{bmatrix}$, then

$$T_n = \begin{bmatrix} 1 & \mathbf{v}^H T_{n-1}^{-1} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \{\gamma_{00}^{(n)}\}^{-1} & \mathbf{0}^H \\ \mathbf{0} & T_{n-1} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^H \\ T_{n-1}^{-1}\mathbf{v} & I_n \end{bmatrix}. \tag{12.27}$$

Exercise 12.27. Show that if $T_n \succ O$, then

$$\begin{aligned} \det T_n &= \{\gamma_{00}^{(n)}\}^{-1}\{\gamma_{00}^{(n-1)}\}^{-1}\cdots\{\gamma_{00}^{(0)}\}^{-1} \\ &= \{\gamma_{nn}^{(n)}\}^{-1}\{\gamma_{n-1,n-1}^{(n-1)}\}^{-1}\cdots\{\gamma_{00}^{(0)}\}^{-1}. \end{aligned} \tag{12.28}$$

Exercise 12.28. Verify formula (12.22). [HINT: Exploit formulas (12.25) and (12.26).]

Exercise 12.29. Verify formula (12.23). [HINT: Exploit formulas (12.25) and (12.27).]

Exercise 12.30. Use the formulas in Lemma 8.15 for calculating orthogonal projections to verify (12.22) and (12.23) another way.

Exercise 12.31. Show that if $T_n \succ O$, then $\gamma_{00}^{(n)} = \gamma_{nn}^{(n)}$ and

$$\gamma_{00}^{(n-1)} \geq \gamma_{00}^{(n)}\,. \tag{12.29}$$

[HINT: The monotonicity is an easy consequence of formula (12.23).]

Exercise 12.32. Show that if $T_n \succ O$, then the polynomials

$$q_k(\lambda) = \sum_{j=0}^{k} \gamma_{jk}^{(k)}\lambda^j \quad \text{for} \quad k = 0, \dots, n \tag{12.30}$$

are orthogonal with respect to the inner product

$$\langle q_j, q_k\rangle_f = \frac{1}{2\pi}\int_0^{2\pi} \overline{q_k(e^{i\theta})}f(e^{i\theta})q_j(e^{i\theta})d\theta\,, \quad \text{and} \quad \langle q_k, q_k\rangle_f = \gamma_{kk}^{(k)}\,. \tag{12.31}$$

Exercise 12.33. Use the orthogonal polynomials defined by formula (12.30) to give a new proof of formula (12.22). [HINT: Write $\zeta^n = \sum_{j=0}^n c_j q_j(\zeta)$.]

Exercise 12.34. Let $f(e^{i\theta}) = |h(e^{i\theta}|^2$, where $h(\zeta) = \sum_{j=0}^\infty h_j\zeta^j$, $\sum_{j=0}^\infty |h_j| < \infty$ and $|h(\zeta)| > 0$ for $|\zeta| \leq 1$. Granting that $1/h$ has the same properties as h (which follows from a theorem of Norbert Wiener), show that

$$\lim_{n\uparrow\infty}\{\gamma_{nn}^{(n)}\}^{-1} = |h_0|^2\,. \tag{12.32}$$

[HINT: $|1 - \sum_{j=1}^n c_j e^{ij\theta}|^2 f(e^{i\theta}) = |h(e^{\imath\theta}) - \sum_{j=1}^n c_j e^{ij\theta} h(e^{\imath\theta})|^2 = |h_0 + u(e^{i\theta})|^2$, where $u(e^{i\theta}) = \sum_{j=1}^\infty h_j e^{ij\theta} - \sum_{j=1}^n c_j e^{ij\theta} h(e^{i\theta})$ is orthogonal to h_0 with respect to the inner product of Exercise 8.3 adapted to $[0, 2\pi]$.]

The next lemma serves to guarantee that the conditions imposed on $f(e^{i\theta})$ in Exercise 12.34 are met if $f(\zeta) = a(\zeta)/b(\zeta)$, where $a(\zeta) = \sum_{j=-k}^k a_j\zeta^j$ and $b(\zeta) = \sum_{j=-\ell}^\ell b_j\zeta^j$ are trigonometric polynomials such that $a(\zeta) > 0$ and $b(\zeta) > 0$ when $|\zeta| = 1$.

Lemma 12.8. (Riesz-Fejér) *Let*

$$f(\zeta) = \sum_{j=-n}^n f_j\zeta^j \quad \textit{for} \quad |\zeta| = 1$$

be a trigonometric polynomial such that $|f(\zeta)| > 0$ for every point $\zeta \in \mathbb{C}$ with $|\zeta| = 1$ and $f_n \neq 0$. Then there exists a polynomial $\varphi_n(\zeta) = a(\zeta - \alpha_1)\cdots(\zeta - \alpha_n)$ such that

$$f(\zeta) = |\varphi_n(\zeta)|^2 \quad \textit{for} \quad |\zeta| = 1$$

and $|\alpha_j| > 1$ for $j = 1, \ldots, n$.

Proof. Under the given assumptions, it is readily checked that $f_{-j} = \overline{f_j}$ for $j = 0, \ldots, n$ and hence, that $f(\beta) = \overline{f(1/\overline{\beta})}$ for every point $\beta \in \mathbb{C} \setminus \{0\}$. Moreover, since $g(\zeta) = \zeta^n f(\zeta) = f_{-n} + f_{1-n}\zeta + \cdots + f_n\zeta^{2n}$ is a polynomial of degree $2n$ with $g(0) = f_{-n} \neq 0$,

$$g(\zeta) = a(\zeta - \beta_1)\cdots(\zeta - \beta_{2n})$$

for some choice of points $a, \beta_1, \ldots, \beta_{2n} \in \mathbb{C} \setminus \{0\}$. However, in view of the preceding discussion, these roots can be indexed so that $|\beta_j| > 1$ and $\beta_{j+n} = 1/\overline{\beta_j}$ for $j = 1, \ldots, n$. Therefore,

$$\begin{aligned} f(\zeta) &= \zeta^{-n} a \prod_{j=1}^n (\zeta - \beta_j)(\zeta - 1/\overline{\beta_j}) \\ &= (-1)^n a(\overline{\beta_1}\cdots\overline{\beta_n})^{-1} \prod_{j=1}^n (\zeta - \beta_j)(\overline{\zeta} - \overline{\beta_j}) \quad \text{if} \quad |\zeta| = 1\,. \end{aligned}$$

The polynomial $\varphi_n(\zeta) = \sqrt{(-1)^n a(\overline{\beta_1}\cdots\overline{\beta_n})^{-1}}\prod_{j=1}^n(\zeta-\beta_j)$ meets the stated requirements of the lemma. □

12.6. Detour on block Toeplitz matrices

The interplay between the two Schur complements that was used to establish formula (12.10) is easily adapted to the more general setting of block Toeplitz matrices

$$T_k = \begin{bmatrix} t_0 & \cdots & t_{-k} \\ \vdots & \ddots & \vdots \\ t_k & \cdots & t_0 \end{bmatrix} \text{ with blocks } t_i \in \mathbb{C}^{p\times p},\ i=0,\ldots,k, \tag{12.33}$$

and their inverses

$$\Gamma_k = \begin{bmatrix} \gamma_{00}^{(k)} & \cdots & \gamma_{0k}^{(k)} \\ \vdots & & \vdots \\ \gamma_{k0}^{(k)} & \cdots & \gamma_{kk}^{(k)} \end{bmatrix} \text{ with blocks } \gamma_{ij}^{(k)} \in \mathbb{C}^{p\times p},\ i,j=0,\ldots,k, \tag{12.34}$$

when they exist.

Lemma 12.9. *Let T_k denote the block Toeplitz matrices defined by formula (12.33) and suppose that $n \geq 1$ and T_n is invertible. Let $\Gamma_n = T_n^{-1}$ be decomposed into blocks as in formula* (12.34)*. Then the following are equivalent:*

(1) *T_n and $\gamma_{00}^{(n)}$ are invertible.*

(2) *T_n and T_{n-1} are invertible.*

(3) *T_n and $\gamma_{nn}^{(n)}$ are invertible.*

Proof. The proof is an easy consequence of the two Schur decompositions used in the proof of Theorem 12.7 except that now block Schur decompositions are used and the vectors $\mathbf{x}^H$ and $\mathbf{y}^H$ are replaced by the block rows $\begin{bmatrix}\gamma_{01}^{(n)} & \cdots & \gamma_{0n}^{(n)}\end{bmatrix}$ and $\begin{bmatrix}\gamma_{n0}^{(n)} & \cdots & \gamma_{n,n-1}^{(n)}\end{bmatrix}$, respectively. The details are left to the reader. □

Exercise 12.35. Show that in the setting of Lemma 12.9,

$$\det \gamma_{00}^{(n)} = \det \gamma_{nn}^{(n)}.$$

Theorem 12.10. *Let T_n and T_{n-1} be invertible block Toeplitz matrices of the form* (12.33) *and let $\Gamma_n = T_n^{-1}$. Then the matrix polynomials*

$$P_n(\lambda) = \sum_{j=0}^n \lambda^j \gamma_{j0}^{(n)}, \quad P_n^\circ(\lambda) = \sum_{j=0}^n \lambda^j \gamma_{0j}^{(n)} \tag{12.35}$$

$$Q_n(\lambda) = \sum_{j=0}^n \lambda^j \gamma_{jn}^{(n)} \quad \textit{and} \quad Q_n^\circ(\lambda) = \sum_{j=0}^n \lambda^j \gamma_{nj}^{(n)} \tag{12.36}$$

are connected by the formula

$$(12.37)\quad \frac{P_n(\lambda)\{\gamma_{00}^{(n)}\}^{-1}P_n^\circ(\overline{\omega}) - \lambda\overline{\omega}Q_n(\lambda)\{\gamma_{nn}^{(n)}\}^{-1}Q_n^\circ(\overline{\omega})}{1-\lambda\overline{\omega}} = \sum_{i,j=0}^{n} \lambda^i \gamma_{ij}^{(n)} \overline{\omega}^j\,.$$

Proof. The proof is an almost exact paraphrase of the verification of formula (12.10), except that now block Schur decompositions are used and the vectors $\mathbf{x}^H$ and $\mathbf{y}^H$ are replaced by the block rows $\begin{bmatrix}\gamma_{01}^{(n)} & \cdots & \gamma_{0n}^{(n)}\end{bmatrix}$ and $\begin{bmatrix}\gamma_{n0}^{(n)} & \cdots & \gamma_{n,n-1}^{(n)}\end{bmatrix}$, respectively.

The well known Gohberg-Heinig formula for Γ_n can be obtained from formula (12.37) with the aid of the following evaluation:

Lemma 12.11. *Let*

$$X(\lambda) = \sum_{i=0}^{\ell} \lambda^i X_i \quad \text{and} \quad Y(\lambda) = \sum_{i=0}^{\ell} \lambda^i Y_i$$

be matrix polynomials with coefficients $X_i \in \mathbb{C}^{p\times q}$ and $Y_i \in \mathbb{C}^{q\times p}$ for $i = 0,\ldots,\ell$ such that

$$(12.38)\qquad X(e^{i\theta})Y(e^{i\theta})^H = O_{p\times p} \quad \text{for} \quad 0 \le \theta < 2\pi\,.$$

Then

$$\begin{aligned}
(12.39)\quad &\frac{X(\lambda)Y(\omega)^H}{1-\lambda\overline{\omega}} \\
&= -\Psi(\lambda)\begin{bmatrix} X_\ell & X_{\ell-1} & \cdots & X_1 \\ O & X_\ell & \cdots & X_2 \\ \vdots & \ddots & \ddots & \vdots \\ O & \cdots & O & X_\ell \end{bmatrix}\begin{bmatrix} Y_\ell^H & O & \cdots & O \\ Y_{\ell-1}^H & Y_\ell^H & \ddots & \vdots \\ \vdots & & \ddots & O \\ Y_1^H & Y_2^H & \cdots & Y_\ell^H \end{bmatrix}\Psi(\omega)^H \\
&= -\Psi(\lambda)\begin{bmatrix} X_1 & X_2 & \cdots & X_\ell \\ X_2 & X_3 & \cdots & O \\ \vdots & & & \vdots \\ X_\ell & O & \cdots & O \end{bmatrix}\begin{bmatrix} Y_1^H & Y_2^H & \cdots & Y_\ell^H \\ Y_2^H & Y_3^H & \cdots & O \\ \vdots & & & \vdots \\ Y_\ell^H & O & \cdots & O \end{bmatrix}\Psi(\omega)^H\,,
\end{aligned}$$

where

$$\Psi(\lambda) = \Psi_{\ell-1}(\lambda) = \begin{bmatrix} I_p & \lambda I_p & \cdots & \lambda^{\ell-1}\end{bmatrix}.$$

Proof. The condition (12.38) is equivalent to the condition

$$X(\lambda)Y(1/\overline{\lambda})^H = O_{p\times p} \quad \text{for} \quad \lambda \in \mathbb{C}\setminus\{0\}\,.$$

Thus, if $\omega \neq 0$ and $\mu = 1/\overline{\omega}$, then

$$\begin{aligned}
\frac{X(\lambda)Y(\omega)^H}{1-\lambda\overline{\omega}} &= \frac{\{X(\lambda)-X(1/\overline{\omega})\}Y(\omega)^H}{1-\lambda\overline{\omega}} \\
&= -\mu\sum_{j=0}^{\ell}\frac{\lambda^j-\mu^j}{\lambda-\mu}X_jY(\omega)^H \\
&= -\mu\sum_{j=1}^{\ell}\frac{\lambda^j-\mu^j}{\lambda-\mu}X_jY(\omega)^H \\
&= -\mu\sum_{j=1}^{\ell}\sum_{i=0}^{j-1}\lambda^i\mu^{j-1-i}X_jY(\omega)^H \\
&= -\mu\sum_{s=0}^{\ell-1}\sum_{i=0}^{s}\lambda^i\mu^{s-i}X_{s+1}Y(\omega)^H \\
&= -\mu\sum_{i=0}^{\ell-1}\lambda^i\sum_{s=i}^{\ell-1}\mu^{s-i}X_{s+1}Y(\omega)^H\,,
\end{aligned}$$

which serves to identify the coefficient of λ^i in the first formula on the right-hand side of (12.39) with

$$-\mu\sum_{s=i}^{\ell-1}\mu^{s-i}X_{s+1}Y(\omega)^H = -\sum_{s=i}^{\ell-1}\sum_{t=0}^{\ell}\overline{\omega}^{t-s+i-1}X_{s+1}Y_t^H\,.$$

Moreover, since the right-hand side of the last formula is a polynomial in $\overline{\omega}$, it can be reexpressed as

$$\begin{aligned}
&-\left\{X_{i+1}\sum_{t=0}^{\ell}\overline{\omega}^{t-1}Y_t^H + X_{i+2}\sum_{t=0}^{\ell}\overline{\omega}^{t-2}Y_t^H + \cdots + X_\ell\sum_{t=0}^{\ell}\overline{\omega}^{t-\ell+i}Y_t^H\right\} \\
= \quad &-\left\{X_{i+1}\sum_{t=1}^{\ell}\overline{\omega}^{t-1}Y_t^H + X_{i+2}\sum_{t=2}^{\ell}\overline{\omega}^{t-2}Y_t^H + \cdots + X_\ell\sum_{t=\ell-i}^{\ell}\overline{\omega}^{t-\ell+i}Y_t^H\right\} \\
= \quad &-\left\{X_{i+1}\sum_{t=0}^{\ell-1}\overline{\omega}^{t}Y_{t+1}^H + X_{i+2}\sum_{t=0}^{\ell-2}\overline{\omega}^{t}Y_{t+2}^H + \cdots + X_\ell\sum_{t=0}^{i}\overline{\omega}^{t}Y_{t+\ell-i}^H\right\}.
\end{aligned}$$

Therefore, the coefficient of $\lambda^i\overline{\omega}^j$ in the first formula on the right-hand side of (12.39) is equal to

$$-\left\{X_{i+1}Y_{j+1}^H + X_{i+2}Y_{j+2}^H + \cdots + X_\ell Y_{\ell-i+j}^H\right\} \quad \text{if} \quad i \geq j$$

and to

$$-\left\{X_{i+1}Y_{j+1}^H + X_{i+2}Y_{j+2}^H + \cdots + X_{\ell-i+j}Y_{\ell}^H\right\} \quad \text{if} \quad i \leq j\,,$$

which yields the first formula in (12.39). The second formula follows easily from the first upon noting that

$$\begin{bmatrix} X_1 & X_2 & \cdots & X_\ell \\ X_2 & X_3 & \cdots & O \\ \vdots & & & \vdots \\ X_\ell & O & \cdots & O \end{bmatrix} = \begin{bmatrix} X_\ell & X_{\ell-1} & \cdots & X_1 \\ O & X_\ell & \cdots & X_2 \\ \vdots & & & \vdots \\ O & O & \cdots & X_\ell \end{bmatrix} \begin{bmatrix} O & \cdots & O & I_p \\ O & \cdots & I_p & O \\ \vdots & & & \vdots \\ I_p & \cdots & O & O \end{bmatrix}.$$

□

Theorem 12.12. (Gohberg-Heinig) *In the setting of Theorem 12.10,*

$$(12.40)\quad T_n^{-1} = \begin{bmatrix} \gamma_{nn}^{(n)} & \gamma_{n-1,n}^{(n)} & \cdots & \gamma_{0n}^{(n)} \\ O & \gamma_{nn}^{(n)} & \cdots & \gamma_{1n}^{(n)} \\ \vdots & & \ddots & \vdots \\ O & O & \cdots & \gamma_{nn}^{(n)} \end{bmatrix} D_n^{-1} \begin{bmatrix} \gamma_{nn}^{(n)} & O & \cdots & O \\ \gamma_{n,n-1}^{(n)} & \gamma_{nn}^{(n)} & \cdots & O \\ \vdots & & \ddots & \vdots \\ \gamma_{n0}^{(n)} & \gamma_{n1}^{(n)} & \cdots & \gamma_{nn}^{(n)} \end{bmatrix}$$

$$- \begin{bmatrix} O & \gamma_{n0}^{(n)} & \gamma_{n-1,0}^{(n)} & \cdots & \gamma_{10}^{(n)} \\ O & O & \gamma_{n0}^{(n)} & \cdots & \gamma_{20}^{(n)} \\ \vdots & & \cdots & \ddots & \vdots \\ O & O & O & \cdots & \gamma_{n0}^{(n)} \\ O & O & O & \cdots & O \end{bmatrix} D_0^{-1} \begin{bmatrix} O & O & \cdots & O \\ \gamma_{0n}^{(n)} & O & \cdots & O \\ \gamma_{0,n-1}^{(n)} & \gamma_{0n}^{(n)} & \cdots & O \\ \vdots & & \ddots & \vdots \\ \gamma_{01}^{(n)} & \gamma_{02}^{(n)} & \cdots & \gamma_{0n}^{(n)} \end{bmatrix},$$

where

$$D_0 = \operatorname{diag}\{\gamma_{00}^{(n)}, \ldots, \gamma_{00}^{(n)}\} \quad \textit{and} \quad \operatorname{diag}\{\gamma_{nn}^{(n)}, \ldots, \gamma_{nn}^{(n)}\}.$$

Proof. Let

$$X(\lambda) = \left[P_n(\lambda)\{\gamma_{00}^{(n)}\}^{-1} \quad \lambda Q_n(\lambda)\{\gamma_{nn}^{(n)}\}^{-1} \right]$$

and

$$Y(\lambda) = \left[P_n^\circ(\overline{\lambda})^H \quad -\lambda Q_n^\circ(\overline{\lambda})^H \right].$$

Then

$$X_j = \begin{cases} \left[\gamma_{j0}^{(n)}\{\gamma_{00}^{(n)}\}^{-1} \quad \gamma_{j-1,n}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1} \right] & \text{for} \quad j = 1, \ldots, n \\ \left[O \quad I_p \right] \quad \text{for} \quad j = n+1 \end{cases}$$

and

$$Y_j = \begin{cases} \left[\{\gamma_{0j}^{(n)}\}^H \quad -\{\gamma_{n,j-1}^{(n)}\}^H \right] & \text{for} \quad j = 1, \ldots, n \\ \left[O \quad -\{\gamma_{nn}^{(n)}\}^H \right] \quad \text{for} \quad j = n+1 \end{cases}$$

and the formula emerges from formula (12.39) upon making the requisite substitutions. □

12.7. A maximum entropy matrix completion problem

In this section we shall consider the problem of completing a matrix in $\mathbb{C}^{n\times n}$ that belongs to the class

$$\mathbb{C}_{\succ}^{n\times n} = \{A \in \mathbb{C}^{n\times n} : A \succ O\}$$

when only the entries in the $2m+1$ central diagonals of A, i.e., the entries a_{ij} with indices in the set

$$\Lambda_m = \{(i,j) : i,j = 1,\ldots,n \quad \text{and} \quad |i-j| \le m\}, \tag{12.41}$$

are given. It is tempting to set the unknown entries equal to zero. However, the matrix that is obtained this way is not necessarily positive definite; see Exercise 12.18 for a simple example. A remarkable fact is that there exists exactly one completion $\widetilde{A} \in \mathbb{C}_{\succ}^{n\times n}$ of the partially specified A such that $\mathbf{e}_i^T(\widetilde{A})^{-1}\mathbf{e}_j = 0$ for $(i,j) \notin \Lambda_m$. We shall sketch an algorithm for obtaining this particular completion that is based on factorization and shall show that $\widetilde{A}$ can also be characterized as the completion which maximizes the determinant. Because of this property $\widetilde{A}$ is commonly referred to as the **maximum entropy completion**.

Theorem 12.13. *Let m be an integer such that $0 \le m \le n-1$ and let*

$$\{b_{ij} : (i,j) \in \Lambda_m\}$$

be a given set of complex numbers. Then there exists a matrix $A \in \mathbb{C}_{\succ}^{n\times n}$ such that

$$a_{ij} = b_{ij} \quad \textit{for} \quad (i,j) \in \Lambda_m \tag{12.42}$$

if and only if

$$\begin{bmatrix} b_{jj} & \cdots & b_{j,j+m} \\ \vdots & & \vdots \\ b_{j+m,j} & \cdots & b_{j+m,j+m} \end{bmatrix} \succ O \quad \textit{for} \quad j = 1,\ldots,n-m\,. \tag{12.43}$$

Discussion. The proof of necessity is easy and is left to the reader as an exercise. The verification of sufficiency is by construction and is most easily understood by example. To this end, let $n = 5$ and $m = 2$, and let $X \in \mathbb{C}^{5\times 5}$ denote the lower triangular matrix with entries x_{ij} that are set equal to zero for $i > j+2$ (i.e., $x_{41} = x_{51} = x_{52} = 0$) and are determined by the following equations when $j \le i \le j+2$:

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} b_{22} & b_{23} & b_{24} \\ b_{32} & b_{33} & b_{34} \\ b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} x_{22} \\ x_{32} \\ x_{42} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{12.44}$$

(12.45)

$$\begin{bmatrix} b_{33} & b_{34} & b_{35} \\ b_{43} & b_{44} & b_{45} \\ b_{53} & b_{54} & b_{55} \end{bmatrix} \begin{bmatrix} x_{33} \\ x_{43} \\ x_{53} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} b_{44} & b_{45} \\ b_{54} & b_{55} \end{bmatrix} \begin{bmatrix} x_{44} \\ x_{54} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad b_{55}x_{55} = 1\,.$$

Next, let $D = \mathrm{diag}\,\{x_{11}, \ldots, x_{55}\}$. Since $x_{jj} > 0$ for $j = 1, \ldots, 5$, $D \succ O$ and the matrix $L = XD^{-1}$ is lower triangular with ones on the diagonal. Now set

$$A = (L^H)^{-1}D^{-1}L^{-1}\,. \tag{12.46}$$

Then clearly $A \in \mathbb{C}_{\succ}^{5\times 5}$ and equations (12.44) and (12.45) are in force, but with a_{ij} in place of b_{ij}. Therefore, the numbers $c_{ij} = b_{ij} - a_{ij}$ are solutions of the equations

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} c_{22} & c_{23} & c_{24} \\ c_{32} & c_{33} & c_{34} \\ c_{42} & c_{43} & c_{44} \end{bmatrix} \begin{bmatrix} x_{22} \\ x_{32} \\ x_{42} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{12.47}$$

(12.48)

$$\begin{bmatrix} c_{33} & c_{34} & c_{35} \\ c_{43} & c_{44} & c_{45} \\ c_{53} & c_{54} & c_{55} \end{bmatrix} \begin{bmatrix} x_{33} \\ x_{43} \\ x_{53} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} c_{44} & c_{45} \\ c_{54} & c_{55} \end{bmatrix} \begin{bmatrix} x_{44} \\ x_{54} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad c_{55}x_{55} = 0\,.$$

But, since the x_{jj} are positive and each of the five submatrices are Hermitian, it is readily seen that $c_{ij} = 0$ for all the indicated entries; i.e., $a_{ij} = b_{ij}$ for $|i - j| \leq 2$. Thus, the matrix A constructed above is a positive definite completion. □

At first glance it might seem that the missing entries in the partially specified matrix should be set equal to zero. However, as we have already noted, Exercise 12.18 shows that the matrix that arises this way is not necessarily positive definite over $\mathbb{C}^n$.

The matrix A that is constructed in Theorem 12.13 inherits special properties from the construction.

Lemma 12.14. *If $A = XX^H$ and $X \in \mathbb{C}^{n\times n}$ is a lower triangular invertible matrix, then*

$$a_{ij} = 0 \quad \textit{for} \quad i - j \geq k \Longleftrightarrow x_{ij} = 0 \quad \textit{for} \quad i - j \geq k\,. \tag{12.49}$$

Discussion. The verification of (12.49) becomes transparent if the calculations are organized properly. The underlying ideas are best conveyed by example. Let $A \in \mathbb{C}^{7\times 7}$ and suppose that $k = 3$. Then the entries a_{ij} in A that meet the constraint $i - j \geq k$ with $k = 3$ can be expressed in terms of the corresponding entries x_{ij} in the lower triangular matrix X by means of the formulas

$$\begin{bmatrix} a_{41} \\ a_{51} \\ a_{61} \\ a_{71} \end{bmatrix} = \begin{bmatrix} x_{41} \\ x_{51} \\ x_{61} \\ x_{71} \end{bmatrix} \overline{x_{11}}\,, \quad \begin{bmatrix} a_{52} \\ a_{62} \\ a_{72} \end{bmatrix} = \begin{bmatrix} x_{51} & x_{52} \\ x_{61} & x_{62} \\ x_{71} & x_{72} \end{bmatrix} \begin{bmatrix} \overline{x_{21}} \\ \overline{x_{22}} \end{bmatrix},$$

(12.50)

$$\begin{bmatrix} a_{63} \\ a_{73} \end{bmatrix} = \begin{bmatrix} x_{61} & x_{62} & x_{63} \\ x_{71} & x_{72} & x_{73} \end{bmatrix} \begin{bmatrix} \overline{x_{31}} \\ \overline{x_{32}} \\ \overline{x_{33}} \end{bmatrix} \quad \text{and} \quad a_{74} = \sum_{j=1}^{4} x_{7j}\overline{x_{4j}}\,.$$

Thus, as the diagonal entries of X are all nonzero, assertion (12.71) is easily verified for this special case. The general case may be established in just the same way.

There is a companion result, which we state without proof:

Lemma 12.15. *If $A = YY^H$ and $Y \in \mathbb{C}^{n\times n}$ is an upper triangular invertible matrix, then*

(12.51) $$a_{ij} = 0 \quad \text{for} \quad i - j \le -k \Longleftrightarrow y_{ij} = 0 \quad \text{for} \quad i - j \le -k\,.$$

Exercise 12.36. Verify Lemma 12.15 if $n = 7$ and $k = 3$.

Theorem 12.16. *Let m be an integer such that $0 \le m \le n-1$ and let*

$$\{b_{ij} : (i,j) \in \Lambda_m\}$$

be a given set of complex numbers such that the conditions (12.43) *are in force. Then there exists exactly one matrix $A \in \mathbb{C}_{\succ}^{n\times n}$ such that*

(12.52) $$a_{ij} = b_{ij} \quad \text{for} \quad (i,j) \in \Lambda_m$$

and

(12.53) $$\mathbf{e}_i^T A^{-1} \mathbf{e}_j = 0 \quad \text{for} \quad (i,j) \notin \Lambda_m\,.$$

Proof. In view of Lemma 12.14, the matrix $A = (L^H)^{-1}D^{-1}L^{-1}$ that was constructed in the discussion of Theorem 12.13 meets both of the stated conditions. Suppose next that $A_1 \in \mathbb{C}_{\succ}^{n\times n}$ is a second matrix that meets the conditions (12.52) and (12.53). Then, by Theorem 12.4, A_1 admits a factorization of the form

$$A_1^{-1} = L_1 D_1 L_1^H$$

for some lower triangular matrix $L_1 \in \mathbb{C}^{n\times n}$ with ones on the diagonal and some diagonal matrix $D_1 \in \mathbb{C}_{\succ}^{n\times n}$. Moreover, by Lemma 12.14 the entries z_{ij} in the lower triangular matrix $Z = L_1 D$ are equal to zero for $i > j + m$. Consequently, the entries z_{ij} with $j \le i \le j+m$ are determined by the same equations as the x_{ij} for $j \le i \le j+m$, i.e., by equations (12.44) and (12.45) if $n = 5$ and $m = 2$, or, in general, by the equations

$$(12.54) \qquad B_{[j,j+m]} \begin{bmatrix} x_{jj} \\ \vdots \\ x_{j+m,j} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{for} \quad j = 1, \ldots, n-m$$

and

$$(12.55) \qquad B_{[j,n]} \begin{bmatrix} x_{jj} \\ \vdots \\ x_{n,j} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix} \quad \text{for} \quad j = n-m+1, \ldots, n\,.$$

Thus, $z_{ij} = x_{ij}$ for $i, j = 1, \ldots, n$ and hence $A_1 = A$; i.e., the proof of uniqueness is complete. □

Theorem 12.17. *Let m be an integer such that $0 \le m \le n-1$ and let*

$$\{b_{ij} : (i,j) \in \Lambda_m\}$$

be a given set of complex numbers such that the conditions (12.43) *are in force. Let $A \in \mathbb{C}_{\succ}^{n\times n}$ meet conditions* (12.52) *and* (12.53) *and let $C \in \mathbb{C}_{\succ}^{n\times n}$ meet condition* (12.52). *Then*

(1) $\det A \ge \det C$, *with equality if and only if $A = C$.*
(2) *If $A = L_A D_A L_A^H$ and $C = L_C D_C L_C^H$, in which L_A and L_C are lower triangular with ones on the diagonal and D_A and D_C are $n \times n$ diagonal matrices, then $D_A \succeq D_C$, with equality if and only if $A = C$.*

Proof. In view of Theorem 12.4,

$$A = (X^H)^{-1} D X^{-1} \quad \text{and} \quad C = (Y^H)^{-1} G Y^{-1}\,,$$

where $X \in \mathbb{C}^{n\times n}$ and $Y \in \mathbb{C}^{n\times n}$ are lower triangular matrices with ones on the diagonal, $D \in \mathbb{C}_{\succ}^{n\times n}$ and $G \in \mathbb{C}_{\succ}^{n\times n}$ are diagonal matrices and $x_{ij} = 0$ for $i \ge m+j$. Therefore, the formulas

$$C = A + (C - A) \quad \text{and} \quad Z = Y^{-1}X$$

imply that

$$Z^H G Z = D + X^H (C - A) X\,.$$

Thus, as Z is lower triangular with ones on the diagonal and the diagonal entries of $X^H(C-A)X$ are all equal to zero,

$$d_{jj} = \sum_{s=j}^{n} g_{ss}|z_{sj}|^2 = g_{jj} + \sum_{s>j} g_{ss}|z_{sj}|^2 \ge g_{jj}$$

with strict inequality unless $z_{sj} = 0$ for $s > j$, i.e., unless $Z = I_n$. This completes the proof of (1). Much the same argument serves to justify (2). □

Remark 12.18. Theorem 12.16 can also be expressed in terms of the orthogonal projection P_{Λ_m} that is defined by the formula

$$P_{\Lambda_m} A = \sum_{(i,j)\in\Lambda_m} \langle A, \mathbf{e}_i\mathbf{e}_j^T\rangle \mathbf{e}_i\mathbf{e}_j^T$$

on the inner product space $\mathbb{C}^{n\times n}$ with inner product $\langle A, B\rangle = \text{trace}\,\{B^H A\}$: If the conditions of Theorem 12.16 are met and if $Q \in \mathbb{C}^{n\times n}$ with $q_{ij} = b_{ij}$ for $(i,j) \in \Lambda_m$, then there exists exactly one matrix $A \in \mathbb{C}_{\succ}^{n\times n}$ such that

$$P_{\Lambda_m} A = Q \quad \text{and} \quad (I_n - P_{\Lambda_m})A^{-1} = O\,.$$

This formulation suggests that results analogous to those discussed above can be obtained in other algebras, which is indeed the case.

12.8. Schur complements for semidefinite matrices

In this section we shall show that if $E \succeq O$, then analogues of the Schur complement formulas hold even if neither of the block diagonal entries are invertible. (Similar formulas hold if $E \preceq O$.)

Lemma 12.19. *Let $A \in \mathbb{C}^{p\times p}$, $D \in \mathbb{C}^{q\times q}$, $n = p + q$, and let*

$$E = \begin{bmatrix} A & B \\ B^H & D \end{bmatrix}$$

be positive semidefinite over $\mathbb{C}^n$. Then:

(1) *$\mathcal{N}_A \subseteq \mathcal{N}_{B^H}$ and $\mathcal{N}_D \subseteq \mathcal{N}_B$.*

(2) *$\mathcal{R}_B \subseteq \mathcal{R}_A$ and $\mathcal{R}_{B^H} \subseteq \mathcal{R}_D$.*

(3) *$AA^\dagger B = B$ and $DD^\dagger B^H = B^H$.*

(4) *The matrix E admits the (lower-upper) factorization*

$$E = \begin{bmatrix} I_p & O \\ B^H A^\dagger & I_q \end{bmatrix} \begin{bmatrix} A & O \\ O & D - B^H A^\dagger B \end{bmatrix} \begin{bmatrix} I_p & A^\dagger B \\ O & I_q \end{bmatrix}, \tag{12.56}$$

where $A^\dagger$ denotes the Moore-Penrose inverse of A.

(5) *The matrix E admits the (upper-lower) factorization*

$$E = \begin{bmatrix} I_p & BD^\dagger \\ O & I_q \end{bmatrix} \begin{bmatrix} A - BD^\dagger B^H & O \\ O & D \end{bmatrix} \begin{bmatrix} I_p & O \\ D^\dagger B^H & I_q \end{bmatrix}, \tag{12.57}$$

where $D^\dagger$ denotes the Moore-Penrose inverse of D.

Proof. Since E is presumed to be positive semidefinite, the inequality

$$\mathbf{x}^H(A\mathbf{x} + B\mathbf{y}) + \mathbf{y}^H(B^H\mathbf{x} + D\mathbf{y}) \geq 0$$

must be in force for every choice of $\mathbf{x} \in \mathbb{C}^p$ and $\mathbf{y} \in \mathbb{C}^q$. If, in particular, $\mathbf{x} \in \mathcal{N}_A$, then this reduces to

$$\mathbf{x}^H B\mathbf{y} + \mathbf{y}^H(B^H\mathbf{x} + D\mathbf{y}) \geq 0$$

for every choice of $\mathbf{y} \in \mathbb{C}^q$ and hence, upon replacing $\mathbf{y}$ by $\varepsilon\mathbf{y}$, to

$$\varepsilon\mathbf{x}^H B\mathbf{y} + \varepsilon\mathbf{y}^H B^H\mathbf{x} + \varepsilon^2\mathbf{y}^H D\mathbf{y} \geq 0$$

for every choice of $\varepsilon > 0$ as well. Consequently, upon dividing through by ε and then letting $\varepsilon \downarrow 0$, it follows that

$$\mathbf{x}^H B\mathbf{y} + \mathbf{y}^H B^H\mathbf{x} \geq 0$$

for every choice of $\mathbf{y} \in \mathbb{C}^q$. But if $\mathbf{y} = -B^H\mathbf{x}$, then the last inequality implies that

$$-2\|B^H\mathbf{x}\|^2 = -\mathbf{x}^H BB^H\mathbf{x} - \mathbf{x}^H BB^H\mathbf{x} \geq 0\,.$$

Therefore,

$$B^H\mathbf{x} = \mathbf{0}\,,$$

which serves to complete the proof of the first statement in (1) and, since the orthogonal complements of the indicated sets satisfy the opposite inclusion, implies that

$$\mathcal{R}_{A^H} = (\mathcal{N}_A)^\perp \supseteq (\mathcal{N}_{B^H})^\perp = \mathcal{R}_B\,.$$

Since $A = A^H$, this verifies the first assertion in (2); the proofs of the second assertions in (1) and (2) are similar. The fourth assertion is a straightforward consequence of the formula $AA^\dagger A = A$ and the fact that

$$A = A^H \Longrightarrow (A^\dagger)^H = A^\dagger\,.$$

Items (3) and (5) are left to the reader. □

Exercise 12.37. Verify items (3) and (5) of Lemma 12.19.

Theorem 12.20. *If $A \in \mathbb{C}^{n\times n}$ and $A \succeq O$, then A admits factorizations of the form*

$$A = LD_1L^H \quad \textit{and} \quad A = UD_2U^H\,, \tag{12.58}$$

where L is lower triangular with ones on the diagonal, U is upper triangular with ones on the diagonal, and D_1 and D_2 are $n \times n$ diagonal matrices with nonnegative entries.

Proof. Since $A_{[1,k]} \succeq O$ for $k = 1,\ldots,n$, formula (12.56) implies that

$$A_{[1,k]} = \widetilde{L}_k \begin{bmatrix} A_{[1,k-1]} & O \\ O & \alpha_k \end{bmatrix} \widetilde{L}_k^H \quad \text{for} \quad k = 2,\ldots,n\,,$$

where $\widetilde{L}_k$ is a $k \times k$ lower triangular matrix with ones on the diagonal and $\alpha_k \geq 0$:

$$A = \widetilde{L}_n \begin{bmatrix} A_{[1,n-1]} & O \\ O & \alpha_n \end{bmatrix} \widetilde{L}_n^H\,,\ldots, A_{[1,2]} = \widetilde{L}_2 \begin{bmatrix} A_{[1,1]} & O \\ O & \alpha_2 \end{bmatrix} \widetilde{L}_2^H\,.$$

The first formula in (12.58) is obtained by setting $L_k = \operatorname{diag}\{\widetilde{L}_k, I_{n-k}\}$ and writing

$$A = A_{[1,n]} = L_n L_{n-1} \cdots L_2 \operatorname{diag}\{a_{11}, \alpha_2, \ldots, \alpha_n\} L_2^H \cdots L_{n-1}^H L_n^H .$$

The second formula in (12.57) is verified on the basis of (12.58) in much the same way. □

Exercise 12.38. Let

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} c & b \\ b & a \end{bmatrix} .$$

Show that if $a > b > c > 0$ and $ac > b^2$, then:

(1) The matrices A and B are both positive definite over $\mathbb{C}^2$.

(2) The matrix AB is not positive definite over $\mathbb{C}^2$.

(3) The matrix $AB + BA$ is not positive definite over $\mathbb{C}^2$.

Exercise 12.39. Show that the matrix AB considered in Exercise 12.38 is not positive definite over $\mathbb{R}^2$.

Exercise 12.40. Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ both be positive semidefinite over $\mathbb{C}^n$. Show that $A^2B^2 + B^2A^2$ need not be positive semidefinite over $\mathbb{C}^n$. [HINT: See Exercise 12.38.]

Exercise 12.41. Let $A \in \mathbb{C}^{n\times n}$ be expressed in block form as

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

with square blocks A_{11} and A_{22} and suppose that $A \succeq O$. Show that:

(1) There exists a matrix $K \in \mathbb{C}^{p\times q}$ such that $A_{12} = A_{11}K$.

(2) $A = \begin{bmatrix} I_p & O \\ K^H & I_q \end{bmatrix} \begin{bmatrix} A_{11} & O \\ O & A_{22} - K^H A_{11} K \end{bmatrix} \begin{bmatrix} I_p & K \\ O & I_q \end{bmatrix}$.

(3) $K^H A_{11} K = A_{21} A_{11}^{\dagger} A_{12}$.

Exercise 12.42. Show that in the setting of Exercise 12.41

(1) There exists a matrix $K \in \mathbb{C}^{q\times p}$ such that $A_{21} = A_{22}K$.

(2) $A = \begin{bmatrix} I_p & K^H \\ O & I_q \end{bmatrix} \begin{bmatrix} A_{11} - K A_{22} K^H & O \\ O & A_{22} \end{bmatrix} \begin{bmatrix} I_p & O \\ K & I_q \end{bmatrix}$.

(3) $K A_{22} K^H = A_{12} A_{22}^{\dagger} A_{21}$.

Exercise 12.43. Let $A = BB^H$, where $B \in \mathbb{C}^{n\times k}$ and $\operatorname{rank} B = k$; let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal basis for $\mathcal{R}_{B^H}$; and let $A_\ell = \sum_{j=1}^{\ell} B\mathbf{u}_j\mathbf{u}_j^H B^H$ for $\ell = 1, \ldots, k$. Show that $A = A_k$ and that $A - A_\ell$ is a positive semidefinite matrix of rank $k - \ell$ for $\ell = 1, \ldots, k-1$.

Exercise 12.44. Show that if $A \in \mathbb{C}^{n\times n}$, then

$$A \succeq O \Longrightarrow A = A^H \quad \text{and} \quad \det A_{[1,k]} \geq 0 \quad \text{for} \quad k = 1, \ldots, n\,, \tag{12.59}$$

but the converse implication is false.

Exercise 12.45. Let $A \in \mathbb{C}^{n\times n}$. Show that

$$A = A^H \quad \text{and} \quad \sigma(A) \subset [0, \infty) \Longleftrightarrow A \succeq O\,. \tag{12.60}$$

12.9. Square roots

Theorem 12.21. *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, say B_1 and B_2. Then, since B_1 and B_2 are both positive semidefinite over $\mathbb{C}^n$ and hence Hermitian, there exist a pair of unitary matrices U_1 and U_2 and a pair of diagonal matrices $D_1 \succeq O$ and $D_2 \succeq O$ 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 that

$$(U_2^HU_1D_1 - D_2U_2^HU_1)D_1 + D_2(U_2^HU_1D_1 - D_2U_2^HU_1) = O\,.$$

But this in turn implies that the matrix

$$X = U_2^HU_1D_1 - D_2U_2^HU_1$$

is a solution of the equation

$$XD_1 + D_2X = O.$$

The next step is to show that $X = O$ is the only solution of this equation. Upon writing

$$D_1 = \text{diag}\,\{d_{11}^{(1)}, \ldots, d_{nn}^{(1)}\} \text{ and } D_2 = \text{diag}\,\{d_{11}^{(2)}, \ldots, d_{nn}^{(2)}\}\,,$$

one can readily check that x_{ij}, the ij entry of the matrix X, is a solution of the equation

$$x_{ij}d_{jj}^{(1)} + d_{ii}^{(2)}x_{ij} = 0.$$

Thus, if $d_{jj}^{(1)} + d_{ii}^{(2)} > 0$, then $x_{ij} = 0$. On the other hand, if $d_{jj}^{(1)} + d_{ii}^{(2)} = 0$, then $d_{jj}^{(1)} = d_{ii}^{(2)} = 0$ and, as follows from the definition of X, $x_{ij} = 0$ in this case too. 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 insure uniqueness. Thus, for example, if A is Hermitian, then the formula

$$\begin{bmatrix} A & O \\ C & -A \end{bmatrix} \begin{bmatrix} A & O \\ C & -A \end{bmatrix} = \begin{bmatrix} A^2 & O \\ CA - AC & A^2 \end{bmatrix}$$

exhibits the matrix

$$\begin{bmatrix} A & O \\ C & -A \end{bmatrix} \quad \text{as a square root of} \quad \begin{bmatrix} A^2 & O \\ O & A^2 \end{bmatrix}$$

for every choice of C that commutes with A. In particular,

$$\begin{bmatrix} I_k & O \\ C & -I_k \end{bmatrix} \begin{bmatrix} I_k & O \\ C & -I_k \end{bmatrix} = \begin{bmatrix} I_k & O \\ O & I_k \end{bmatrix} \quad \text{for every} \quad C \in \mathbb{C}^{k \times k}\,.$$

Exercise 12.46. Show that if $A, B \in \mathbb{C}^{n \times n}$ and if $A \succ O$ and $B = B^H$, then there exists a matrix $V \in \mathbb{C}^{n \times n}$ such that

$$V^H A V = I_n \quad \text{and} \quad V^H B V = D = \text{diag}\,\{\lambda_1, \ldots, \lambda_n\}\,.$$

[HINT: Reexpress the problem in terms of $U = A^{1/2}V$.]

Exercise 12.47. 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 12.48. 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 $AB \not\succeq O$).

12.10. Polar forms

If $A \in \mathbb{C}^{p\times q}$ and $r = \operatorname{rank} A \geq 1$, then the formula $A = V_1DU_1^H$ that was obtained in Corollary 10.2 on the basis of the singular value decomposition of A can be reexpressed in **polar form**:

$$A = V_1U_1^H(U_1DU_1^H) \quad \text{and} \quad A = (V_1DV_1^H)V_1U_1^H\,, \tag{12.61}$$

where $V_1U_1^H$ maps $\mathcal{R}_{A^H}$ isometrically onto $\mathcal{R}_A$, $U_1DU_1^H = \{A^HA\}^{1/2}$ is positive definite on $\mathcal{R}_{A^H}$ and $V_1DV_1^H = \{AA^H\}^{1/2}$ is positive definite on $\mathcal{R}_A$. These formulas are matrix analogues of the polar decomposition of a complex number.

Theorem 12.22. *Let $A \in \mathbb{C}^{p\times q}$. Then*

(1) $\operatorname{rank} A = q$ *if and only if A admits a factorization of the form $A = V_1P_1$, where $V_1 \in \mathbb{C}^{p\times q}$ is* **isometric***; i.e., $V_1^HV_1 = I_q$, and $P_1 \in \mathbb{C}^{q\times q}$ is positive definite over $\mathbb{C}^q$.*

(2) $\operatorname{rank} A = p$ *if and only if A admits a factorization of the form $A = P_2U_2$, where $U_2 \in \mathbb{C}^{p\times q}$ is* **coisometric***; i.e., $U_2U_2^H = I_p$, and $P_2 \in \mathbb{C}^{p\times p}$ is positive definite over $\mathbb{C}^p$.*

Proof. If $\operatorname{rank} A = q$, then $p \geq q$ and, by Theorem 10.1, A admits a factorization of the form

$$A = V\begin{bmatrix} D \\ O \end{bmatrix} U = V\begin{bmatrix} I_q \\ O \end{bmatrix} DU\,,$$

where V and U are unitary matrices of sizes $p \times p$ and $q \times q$, respectively, and $D \in \mathbb{C}^{q\times q}$ is positive definite over $\mathbb{C}^q$. But this yields a factorization of the asserted form with $V_1 = V\begin{bmatrix} I_q \\ O \end{bmatrix}U$ and $P_1 = U^HDU$. Conversely, if A admits a factorization of this form, it is easily seen that $\operatorname{rank} A = q$. The details are left to the reader.

Assertion (2) may be established in much the same way or by invoking (1) and passing to transposes. The details are left to the reader. □

Exercise 12.49. Complete the proof of assertion (1) in Theorem 12.22.

Exercise 12.50. Verify assertion (2) in Theorem 12.22.

Exercise 12.51. Show that if $UU^H = VV^H$ for a pair of matrices $U, V \in \mathbb{C}^{n\times d}$ with $\operatorname{rank} U = \operatorname{rank} V = d$, then $U = VK$ for some unitary matrix $K \in \mathbb{C}^{d\times d}$.

Exercise 12.52. Find an isometric matrix V_1 and a matrix $P_1 \succ O$ such that $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = V_1P_1$.

12.11. Matrix inequalities

Lemma 12.23. *If $F \in \mathbb{C}^{p\times q}$, $G \in \mathbb{C}^{r\times q}$ and $F^HF - G^HG \succeq O$, then there exists exactly one matrix $K \in \mathbb{C}^{r\times p}$ such that*

$$G = KF \quad \text{and} \quad K\mathbf{u} = \mathbf{0} \quad \text{for every} \quad \mathbf{u} \in \mathcal{N}_{F^H}\,. \tag{12.62}$$

Moreover, this matrix K is contractive: $\|K\| \le 1$.

Proof. The given conditions imply that

$$\langle F^HF\mathbf{x}, \mathbf{x}\rangle \ge \langle G^HG\mathbf{x}, \mathbf{x}\rangle \quad \text{for every} \quad \mathbf{x} \in \mathbb{C}^q\,.$$

Thus,

$$F\mathbf{x} = \mathbf{0} \Longrightarrow \|G\mathbf{x}\| = 0 \Longrightarrow G\mathbf{x} = \mathbf{0}\,;$$

i.e., $\mathcal{N}_F \subseteq \mathcal{N}_G$ and hence $\mathcal{R}_{G^H} \subseteq \mathcal{R}_{F^H}$. Therefore, there exists a matrix $K_1^H \in \mathbb{C}^{p\times r}$ such that $G^H = F^HK_1^H$.

If $\mathcal{N}_{F^H} = \{\mathbf{0}\}$, then the matrix $K = K_1$ meets both of the conditions in (12.62).

If $\mathcal{N}_{F^H} \ne \{\mathbf{0}\}$ and $V \in \mathbb{C}^{p\times\ell}$ is a matrix whose columns form a basis for $\mathcal{N}_{F^H}$, then

$$F^H(K_1^H + VL) = F^HK_1^H = G^H$$

for every choice of $L \in \mathbb{C}^{\ell\times r}$. Moreover,

$$\begin{aligned}(K_1 + L^HV^H)V = O \quad &\Longleftrightarrow \quad L^H = -K_1V(V^HV)^{-1}\\ &\Longleftrightarrow \quad K_1 + L^HV^H = K_1(I_p - V(V^HV)^{-1}V^H)\,.\end{aligned}$$

Thus, the matrix $K = K_1(I_p - V(V^HV)^{-1}V)$ meets the two conditions stated in (12.62). This is eminently reasonable, since $I_p - V(V^HV)^{-1}V^H$ is the formula for the orthogonal projection of $\mathbb{C}^p$ onto $\mathcal{R}_F$.

It is readily checked that if $\widetilde{K} \in \mathbb{C}^{r\times p}$ is a second matrix that meets the two conditions in (12.62), then $\widetilde{K} = K$. The details are left to the reader.

It remains to check that K is contractive. Since

$$\mathbb{C}^p = \mathcal{R}_F \oplus \mathcal{N}_{F^H}\,,$$

every vector $\mathbf{u} \in \mathbb{C}^p$ can be expressed as $\mathbf{u} = F\mathbf{x} + V\mathbf{y}$ for some choice of $\mathbf{x} \in \mathbb{C}^q$ and $\mathbf{y} \in \mathbb{C}^\ell$. Correspondingly,

$$\begin{aligned}\langle K\mathbf{u}, K\mathbf{u}\rangle &= \langle K(F\mathbf{x} + V\mathbf{y}), K(F\mathbf{x} + V\mathbf{y})\rangle = \langle KF\mathbf{x}, KF\mathbf{x}\rangle\\ &= \langle G\mathbf{x}, G\mathbf{x}\rangle \le \langle F\mathbf{x}, F\mathbf{x}\rangle\\ &\le \langle F\mathbf{x}, F\mathbf{x}\rangle + \langle V\mathbf{y}, V\mathbf{y}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle\,.\end{aligned}$$

□

Exercise 12.53. Show that if $K \in \mathbb{C}^{r\times p}$ and $\widetilde{K} \in \mathbb{C}^{r\times p}$ both meet the two conditions in (12.62), then $K = \widetilde{K}$ and hence that K is uniquely specified in terms of the Moore-Penrose inverse $F^\dagger$ of F by the formula $K = GFF^\dagger$.

Corollary 12.24. If, in the setting of Lemma 12.23, $F^H F = G^H G$, then the unique matrix K that meets the two conditions in (12.62) is an isometry on $\mathcal{R}_F$.

Proof. This is immediate from the identity

$$\langle KF\mathbf{x}, KF\mathbf{x}\rangle = \langle G\mathbf{x}, G\mathbf{x}\rangle = \langle F\mathbf{x}, F\mathbf{x}\rangle\,,$$

which is valid for every $\mathbf{x} \in \mathbb{C}^q$. □

Lemma 12.25. *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$, and suppose that $A \succeq B \succeq O$. Then:*

(1) *There exists a matrix $K \in \mathbb{C}^{n\times n}$ such that $B = K^H AK$ and $I_n - K^H K \succeq O$.*
(2) $A^{1/2} \succeq B^{1/2}$.
(3) $\det\, A \geq \det\, B \geq O$.

Moreover, if $A \succ O$, then

(4) $\det\, A = \det\, B$ *if and only if* $A = B$.

Proof. Lemma 12.23 with $F = A^{1/2}$ and $G = B^{1/2}$ guarantees the existence of a contractive matrix $K \in \mathbb{C}^{n\times n}$ such that $KA^{1/2} = B^{1/2}$. Therefore, since $B^{1/2} = (B^{1/2})^H = A^{1/2}K^H$,

$$B = B^{1/2}B^{1/2} = KA^{1/2}(KA^{1/2})^H = KAK^H\,.$$

Next, in view of Exercise 20.1, it suffices to show that all the eigenvalues of the Hermitian matrix $A^{1/2} - B^{1/2}$ are nonnegative in order to verify (2). To this end, let $(A^{1/2} - B^{1/2})\mathbf{u} = \lambda\mathbf{u}$ for some nonzero vector $\mathbf{u}$. Then

$$\begin{aligned}\lambda\langle (A^{1/2} + B^{1/2})\mathbf{u}, \mathbf{u}\rangle &= \langle (A^{1/2} + B^{1/2})(A^{1/2} - B^{1/2})\mathbf{u}, \mathbf{u}\rangle \\ &= \langle (A + B^{1/2}A^{1/2} - A^{1/2}B^{1/2} - B)\mathbf{u}, \mathbf{u}\rangle \\ &= \langle (A - B)\mathbf{u}, \mathbf{u}\rangle \geq 0\,,\end{aligned}$$

since

$$\langle B^{1/2}A^{1/2}\mathbf{u}, \mathbf{u}\rangle = \langle B^{1/2}\mathbf{u}, B^{1/2}\mathbf{u}\rangle + \lambda\langle \mathbf{u}, \mathbf{u}\rangle = \langle A^{1/2}B^{1/2}\mathbf{u}, \mathbf{u}\rangle\,.$$

The last inequality implies that $\lambda \geq 0$ if $\langle (A^{1/2} + B^{1/2})\mathbf{u}, \mathbf{u}\rangle > 0$. On the other hand, if $\langle (A^{1/2} + B^{1/2})\mathbf{u}, \mathbf{u}\rangle = 0$, then $\langle A^{1/2}\mathbf{u}, \mathbf{u}\rangle = \langle B^{1/2}\mathbf{u}, \mathbf{u}\rangle = 0$ and hence $\lambda = 0$.

To obtain (3), observe first that in view of (1), the eigenvalues $\mu_1, \dots, \mu_n$ of $K^H K$ are subject to the bounds $0 \leq \mu_j \leq 1$ for $j = 1, \dots, n$. Therefore,

$$\det\, B = \det\,(KAK^H) = \det\,(K^H K)\,\det\, A = (\mu_1 \cdots \mu_n)\det\, A \leq \det\, A\,.$$

Moreover, if $\det B = \det A$ and A is invertible, then $\mu_1 = \cdots = \mu_n = 1$, i.e., $K^H K = I_n$. Therefore, since $KA^{1/2} = B^{1/2} = (B^{1/2})^H = A^{1/2}K^H$,

$$B = A^{1/2}K^H K A^{1/2} = A^{1/2} I_n A^{1/2} = A\,,$$

which justifies (4) and completes the proof. □

Exercise 12.54. Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Show that $A - B \succ O$, but $A^2 - B^2$ has one positive eigenvalue and one negative eigenvalue.

Theorem 12.26. *If $A_1 \in \mathbb{C}^{n\times s}$, $A_2 \in \mathbb{C}^{n\times t}$, $\operatorname{rank} A_1 = s$, $\operatorname{rank} A_2 = t$ and $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$, then*

$$\det(A^H A) \le \det(A_1^H A_1)\,\det(A_2^H A_2)\,,$$

with equality if and only if $A_1^H A_2 = O$.

Proof. Clearly

$$A^H A = \begin{bmatrix} A_1^H A_1 & A_1^H A_2 \\ A_2^H A_1 & A_2^H A_2 \end{bmatrix}.$$

Therefore, since $A_1^H A_1$ is invertible by Exercise 12.17, it follows from the Schur complement formulas that

$$\det(A^H A) = \det(A_1^H A_1)\,\det(A_2^H A_2 - A_2^H A_1 (A_1^H A_1)^{-1} A_1^H A_2)\,.$$

Thus, as

$$A_2^H A_2 - A_2^H A_1 (A_1^H A_1)^{-1} A_1^H A_2 \preceq A_2^H A_2\,,$$

Lemma 12.25 guarantees that

$$\det(A_2^H A_2 - A_2^H A_1 (A_1^H A_1)^{-1} A_1^H A_2) \le \det(A_2^H A_2)\,,$$

with equality if and only if

$$A_2^H A_2 - A_2^H A_1 (A_1^H A_1)^{-1} A_1^H A_2 = A_2^H A_2\,.$$

This serves to complete the proof, since the last equality holds if and only if $A_1^H A_2 = O$. □

The lemma leads to another inequality (12.63) that is also credited to Hadamard. This inequality is sharper than the inequality (9.13).

Corollary 12.27. Let $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ be an $n \times n$ matrix with columns $\mathbf{a}_j \in \mathbb{C}^n$ for $j = 1, \ldots, n$. Then

$$|\det A|^2 \le \prod_{j=1}^{n} \mathbf{a}_j^T \mathbf{a}_j\,. \tag{12.63}$$

Moreover, if A is invertible, then equality holds in (12.63) if and only if the columns of A are orthogonal.

Proof. The basic strategy is to iterate Theorem 12.26. The details are left to the reader. □

Exercise 12.55. Complete the proof of Corollary 12.27.

Exercise 12.56. Show that if $U,\, V \in \mathbb{C}^{n\times d}$ and $\operatorname{rank} U = \operatorname{rank} V = d$, then

$$UU^H = VV^H \Longleftrightarrow U = VK \quad \text{for some unitary matrix} \quad K \in \mathbb{C}^{d\times d}.$$

Exercise 12.57. Show that if $A \in \mathbb{C}^{n\times n}$ and $O \preceq A \preceq I_n$, then a vector $\mathbf{x} \in \mathcal{R}_{(I_n - A)}$ if and only if

$$\lim_{\delta\uparrow 1}\langle (I_n - \delta A)^{-1}\mathbf{x}, \mathbf{x}\rangle < \infty\,.$$

[HINT: The result is transparent if A is diagonal.]

Exercise 12.58. Show that if $A,\, B \in \mathbb{C}^{n\times n}$ and if $A \succ O$ and $B \succ O$, then

$$\sqrt{\det\, A\, \det\, B} \le \det\, \frac{A+B}{2}\,. \tag{12.64}$$

Exercise 12.59. Show that if $A,\, B \in \mathbb{C}^{n\times n}$ and if $AB = O$ but $A+B \succ O$, then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$U^H AU = \begin{bmatrix} A_{11} & O \\ O & O \end{bmatrix} \quad \text{and} \quad U^H BU = \begin{bmatrix} O & O \\ O & B_{22} \end{bmatrix}\,,$$

where $A_{11} \succ O$ and $B_{22} \succ O$.

12.12. A minimal norm completion problem

The next result, which is usually referred to as Parrott's lemma, is a nice application of the preceding circle of ideas.

Lemma 12.28. *Let $A \in \mathbb{C}^{p\times q}$, $B \in \mathbb{C}^{p\times r}$ and $C \in \mathbb{C}^{s\times q}$. Then*

(12.65)

$$\min\left\{\left\|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right\| : D \in \mathbb{C}^{s\times r}\right\} = \max\left\{\left\|\begin{bmatrix} A \\ C \end{bmatrix}\right\|,\ \|\begin{bmatrix} A & B \end{bmatrix}\|\right\}\,.$$

The proof will be developed in a sequence of auxiliary lemmas, most of which will be left to the reader to verify.

Lemma 12.29. *Let $A \in \mathbb{C}^{p\times q}$. Then*

$$\|A\| \le \gamma \Longleftrightarrow \gamma^2 I_q - A^H A \succeq O \Longleftrightarrow \gamma^2 I_p - AA^H \succeq O\,.$$

Proof. The proof is easily extracted from the inequalities in Exercise 12.9. □

Lemma 12.30. *Let $A \in \mathbb{C}^{p\times q}$, $B \in \mathbb{C}^{p\times r}$ and $C \in \mathbb{C}^{s\times q}$. Then*

(1) $\gamma \ge \left\|\begin{bmatrix} A \\ C \end{bmatrix}\right\| \Longleftrightarrow \gamma^2 I_q - A^H A \succeq C^H C.$

(2) $\gamma \geq \|[\ A \quad B\ ]\| \Longleftrightarrow \gamma^2 I_p - AA^H \succeq BB^H$.

Proof. This is an easy consequence of the preceding lemma and the fact that $\|E\| = \|E^H\|$. □

Lemma 12.31. *If $A \in \mathbb{C}^{p\times q}$ and $\|A\| \leq \gamma$, then:*

$$(\gamma^2 I_q - A^H A)^{1/2} A^H = A^H (\gamma^2 I_p - AA^H)^{1/2} \tag{12.66}$$

and

$$(\gamma^2 I_p - AA^H)^{1/2} A = A(\gamma^2 I_q - A^H A)^{1/2}\,. \tag{12.67}$$

Proof. These formulas may also be established with the aid of the singular value decomposition of A. □

Lemma 12.32. *If $A \in \mathbb{C}^{p\times q}$, $p+q=n$ and $\|A\| \leq \gamma$, then the matrix*

$$E = \begin{bmatrix} A & (\gamma^2 I_p - AA^H)^{1/2} \\ (\gamma^2 I_q - A^H A)^{1/2} & -A^H \end{bmatrix} \tag{12.68}$$

satisfies the identity

$$EE^H = \begin{bmatrix} \gamma^2 I_p & O \\ O & \gamma^2 I_q \end{bmatrix} = \gamma^2 I_n\,.$$

Proof. This is a straightforward multiplication, thanks to Lemma 12.31. □

Lemma 12.33. *Let $A \in \mathbb{C}^{p\times q}$, $B \in \mathbb{C}^{p\times r}$ and $C \in \mathbb{C}^{s\times q}$ and suppose that*

$$\gamma \geq \max\left\{ \left\| \begin{bmatrix} A \\ C \end{bmatrix} \right\|,\ \|[\ A \quad B\ ]\| \right\}.$$

Then there exists a matrix $D \in \mathbb{C}^{s\times r}$ such that

$$\left\| \begin{bmatrix} A & B \\ C & D \end{bmatrix} \right\| \leq \gamma\,.$$

Proof. The given inequality implies that

$$\gamma^2 I_q - A^H A \succeq C^H C \quad \text{and} \quad \gamma^2 I_p - AA^H \succeq BB^H\,.$$

Therefore, by Lemma 12.23,

$$B = (\gamma^2 I_p - AA^H)^{1/2} X \quad \text{and} \quad C = Y(\gamma^2 I_q - A^H A)^{1/2} \tag{12.69}$$

for some choice of $X \in \mathbb{C}^{p\times r}$ and $Y \in \mathbb{C}^{s\times q}$ with $\|X\| \leq 1$ and $\|Y\| \leq 1$. Thus, upon setting $D = -YA^H X$, it is readily seen that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_p & O \\ O & Y \end{bmatrix} E \begin{bmatrix} I_q & O \\ O & X \end{bmatrix},$$

where E is given by formula (12.68). But this does the trick, since $EE^H = \gamma^2 I_n$ by Lemma 12.32 and the norm of each of the two outside factors on the right is equal to one. □

12.13. A description of all solutions to the minimal norm completion problem

Theorem 12.34. *A matrix $D \in \mathbb{C}^{s\times r}$ achieves the minimum in* (12.65) *if and only if it can be expressed in the form*

$$D = -YA^HX + (I_s - YY^H)^{1/2}Z(I_r - X^HX)^{1/2}\,, \tag{12.70}$$

where

$$X = \left\{\gamma^2 I_p - AA^H)^{1/2}\right\}^{\dagger} B\,, \quad Y = C\left\{\gamma^2 I_q - A^HA)^{1/2}\right\}^{\dagger} \tag{12.71}$$

and

$$Z \quad \textit{is any matrix in } \mathbb{C}^{s\times r} \textit{ such that } \quad Z^HZ \le \gamma^2 I_r\,. \tag{12.72}$$

Discussion. We shall outline the main steps in the proof:

1.

$$\begin{bmatrix} A^H & C^H \\ B^H & D^H \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix} \preceq \gamma^2 I_{q+r}$$

if and only if

$$\begin{bmatrix} C^H \\ D^H \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix} \preceq \gamma^2 I_{q+r} - \begin{bmatrix} A^H \\ B^H \end{bmatrix}\begin{bmatrix} A & B \end{bmatrix}. \tag{12.73}$$

2. In view of Lemma 12.31 and the formulas in (12.69),

$$\gamma^2 I_{q+r} - \begin{bmatrix} A^H \\ B^H \end{bmatrix}\begin{bmatrix} A & B \end{bmatrix} = M^HM\,,$$

where

$$M = \begin{bmatrix} (\gamma^2 I_q - A^HA)^{1/2} & -A^HX \\ O & \gamma(I_r - X^HX)^{1/2} \end{bmatrix}.$$

3. In view of (12.73), the identity in Step 2 and Lemma 12.23, there exists a unique matrix $\begin{bmatrix} K_1 & K_2 \end{bmatrix}$ with components $K_1 \in \mathbb{C}^{s\times q}$ and $K_2 \in \mathbb{C}^{s\times r}$ such that

$$\begin{aligned} \begin{bmatrix} C & D \end{bmatrix} &= \begin{bmatrix} K_1 & K_2 \end{bmatrix} M \\ &= \begin{bmatrix} K_1(\gamma^2 I_q - A^HA)^{1/2} & -K_1A^HX + K_2\gamma(I_r - X^HX)^{1/2} \end{bmatrix} \end{aligned}$$

and

$$K_1\mathbf{u}_1 + K_2\mathbf{u}_2 = \mathbf{0} \quad \text{if} \quad M^H\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \mathbf{0}\,.$$

4. $K_1 = Y$, since

$$\begin{aligned} M^H\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \mathbf{0} \iff & (\gamma^2 I_q - A^HA)^{1/2}\mathbf{u}_1 = \mathbf{0} \quad \text{and} \\ & -X^HA\mathbf{u}_1 + \gamma(I_r - X^HX)^{1/2}\mathbf{u}_2 = \mathbf{0} \\ \iff & (\gamma^2 I_q - A^HA)^{1/2}\mathbf{u}_1 = \mathbf{0} \text{ and } (I_r - X^HX)^{1/2}\mathbf{u}_2 = \mathbf{0}\,, \end{aligned}$$

because

$$X^H A = B^H \left\{(\gamma^2 I_q - A^H A)^{1/2}\right\}^\dagger A = B^H A \left\{(\gamma^2 I_p - AA^H)^{1/2}\right\}^\dagger$$

and $\mathcal{N}_{W^H} = \mathcal{N}_{W^\dagger}$ for any matrix $W \in \mathbb{C}^{k\times k}$.

5. Extract the formula

$$D = -K_1 A^H X + \gamma K_2 (I_q - XX^H)^{1/2}$$

from Step 3 and then, taking note of the fact that $K_1K_1^H + K_2K_2^H \preceq I_s$, replace K_1 by Y and γK_2 by $(I_s - YY^H)^{1/2}Z$.

12.14. Bibliographical notes

The section on maximum entropy interpolants is adapted from the paper [**24**]. It is included here to illustrate the power of factorization methods. The underlying algebraic structure is clarified in [**25**]; see also [**34**] for further generalizations. A description of all completions of the problem considered in Section 12.7 may be found e.g., in Chapter 10 of [**21**]. Formulas (12.32) and (12.28) imply that

$$\lim_{n\uparrow\infty} \frac{\ln\det T_n(f)}{n} = \ln|h_0|^2 = \frac{1}{2\pi}\int_0^{2\pi} \ln f(e^{i\theta})d\theta \qquad (12.74)$$

for the Toeplitz matrix $T_n(f)$ based on the Fourier coefficients of the considered function f. This is a special case of a theorem that was proved by Szegö in 1915 and is still the subject of active research today; see e.g., [**9**] and [**65**] for two recent expository articles on the subject; [**64**] for additional background material; and the references cited in all three. Lemma 12.8 is due to Fejér and Riesz. Formula (12.40) is one way of writing a formula due to Gohberg and Heinig. Other variants may be obtained by invoking appropriate generalizations of the observation

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ b & a & 0 \\ c & b & a \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{bmatrix}.$$

The minimal norm completion problem is adapted from [**31**] and [**74**], both of which cite [**18**] as a basic reference for this problem. Exercises 12.58 and 12.59 are adapted from [**69**] and [**26**], respectively.