Appendix A

Review of Matrices

A.0 INTRODUCTION

There are several excellent references on matrices, e.g., MacDuffee [1946], Gantmacher [1959], Bellman [1960], Franklin [1968], Halmos [1974], Horn and Johnson [1985], and Golub and Van Loan [1989], to name a few. Our aim here is to review those results from matrix theory which have direct relevance to this text. Most proofs can be found, or deduced from, the above texts. The material here is somewhat dense, as it is primarily meant to be a reference. Most of the deeper results mentioned here are required only in Chap. 13 and 14.

A.1 DEFINITIONS AND EXAMPLES

A $p \times r$ matrix $\mathbf{A}$ is a collection of pr elements (real or complex numbers) arranged in p rows and r columns. Thus

$$\mathbf{A} = \begin{bmatrix} A_{00} & A_{01} & \dots & A_{0,r-1} \\ A_{10} & A_{11} & \dots & A_{1,r-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p-1,0} & A_{p-1,1} & \dots & A_{p-1,r-1} \end{bmatrix}. \tag{A.1.1}$$

One often writes $\mathbf{A} = [A_{ij}]$ (or with a comma, as in $[A_{i,j}]$), and $A_{ij} = [\mathbf{A}]_{ij}$. So i is the row index and j the column index, both starting at zero. A $p \times 1$ matrix is said to be a p-vector or *column vector* (or just a vector). A $1 \times r$ matrix is called a *row vector*. Thus,

$$\underbrace{\begin{bmatrix} 1+j & -j \\ 2 & 3 \end{bmatrix}}_{2\times 2\ matrix}, \quad \underbrace{\begin{bmatrix} 1 & 2 \end{bmatrix}}_{row\ vector}, \quad \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{column\ vector}. \tag{A.1.2}$$

To save space, one often writes a column vector as the 'transpose' of a row vector. Thus $[1 \quad 2j]^T$ stands for the column vector

$$\begin{bmatrix} 1 \\ 2j \end{bmatrix}. \tag{A.1.3}$$

(See below for definition of transpose). The elements of a vector $\mathbf{v}$ are commonly denoted as v_i or $v(i)$. Whether $\mathbf{v}$ is a row or a column is usually clear from the context.

A square matrix is a matrix with $p = r$. Thus the first matrix in (A.1.2) is square. If a matrix is not square, it is said to be rectangular. A 1×1 matrix (i.e., just a single element) is said to be a scalar. A $p \times r$ null matrix, denoted $\mathbf{0}$, has all elements equal to zero. If $\mathbf{A} = \mathbf{0}$, $\mathbf{A}$ is said to be zero or null.

If two matrices have the same number of rows p and same number of columns r, they have the same size.

Diagonal matrices. The elements A_{ii} of a matrix $\mathbf{A}$ are called its diagonal elements. A matrix for which all elements are zero except possibly the diagonal elements is called a diagonal matrix. Examples are

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & j & 0 \end{bmatrix}. \tag{A.1.4}$$

The set of diagonal elements is sometimes called the main diagonal. Note that a diagonal matrix need not be square. A square diagonal matrix with all diagonal elements equal to unity is said to be the *identity matrix*, denoted as $\mathbf{I}$. Examples are

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{A.1.5}$$

which are 3×3 and 2×2 respectively. If the size of $\mathbf{I}$ is not clear from the context, a subscript will be used. The above examples represent, respectively, $\mathbf{I}_3$ and $\mathbf{I}_2$.

The notation

$$\text{diag}\,[d_0 \quad d_1 \quad \ldots \quad d_{N-1}]$$

stands for a $N \times N$ diagonal matrix $\mathbf{A}$ with diagonal elements $A_{ii} = d_i$.

Triangular matrices. A lower triangular matrix is one for which the elements above the main diagonal are equal to zero. An upper triangular matrix is one for which the elements below the main diagonal are equal to zero. Examples are:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 5 & 1 & 2 \\ 1 & 4 & 1 \end{bmatrix}}_{\textit{lower triangular}}, \quad \underbrace{\begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{\textit{upper triangular}}. \tag{A.1.6}$$

A.2 BASIC OPERATIONS

A number of operations, including arithmetic, can be performed with matrices.

Transpose and transpose-conjugate. Given $\mathbf{A} = [A_{ij}]$, we denote its transpose as $\mathbf{A}^T$. It is defined by $[\mathbf{A}^T]_{ij} = A_{ji}$. In other words, the (i, j) entry of the transpose is same as the (j, i) entry of $\mathbf{A}$. The transpose-conjugate of $\mathbf{A}$, denoted

$\mathbf{A}^\dagger$, is obtained by conjugating every element of $\mathbf{A}^T$. For example with $\mathbf{A}$ equal to the 2×2 matrix in (A.1.2), we have

$$\mathbf{A}^T = \begin{bmatrix} 1+j & 2 \\ -j & 3 \end{bmatrix}, \quad \mathbf{A}^\dagger = \begin{bmatrix} 1-j & 2 \\ j & 3 \end{bmatrix}. \tag{A.2.1}$$

Note that if $\mathbf{A}$ is $p \times r$ then $\mathbf{A}^T$ as well as $\mathbf{A}^\dagger$ are $r \times p$. Thus the transpose of a column vector is a row vector, and vice versa.

For a square matrix $\mathbf{A}$, the notation $\mathbf{A}^{-1}$ stands for its inverse (defined and discussed in Sec. A.4). The notation $\mathbf{A}^{-T}$ stands for $(\mathbf{A}^{-1})^T$. Similarly, $\mathbf{A}^{-\dagger}$ stands for $(\mathbf{A}^{-1})^\dagger$.

Submatrices. A submatrix of $\mathbf{A}$ is any matrix formed by deleting an arbitrary set of rows and an arbitrary set of columns.

Arithmetic Operations

Addition and scalar multiplication. Two matrices with the same size can be added or subtracted by adding or subtracting corresponding elements. The notation $c\mathbf{A}$ stands for a matrix which is obtained from $\mathbf{A}$ by multiplying each element A_{ij} with c. This operation is called scalar multiplication. Thus, $c\mathbf{A} = [cA_{ij}]$. Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size, the matrix $\mathbf{P} = c\mathbf{A} + d\mathbf{B}$ has elements $P_{ij} = cA_{ij} + dB_{ij}$. Matrix addition is evidently commutative, that is, $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. Note also that $[\mathbf{A} + \mathbf{B}]^T = \mathbf{A}^T + \mathbf{B}^T$.

Matrix multiplication. Given two matrices $\mathbf{A}$ and $\mathbf{B}$ with sizes $p \times m$ and $m \times r$ the product $\mathbf{C} = \mathbf{AB}$ is defined by defining the elements of $\mathbf{C}$ as

$$C_{ij} = \sum_{k=0}^{m-1} A_{ik} B_{kj}, \quad 0 \leq i \leq p-1, \; 0 \leq j \leq r-1. \tag{A.2.2}$$

Schematically,

$$\underbrace{\mathbf{C}}_{p \times r} = \underbrace{\mathbf{A}}_{p \times m} \underbrace{\mathbf{B}}_{m \times r}. \tag{A.2.3}$$

Note that the number of columns of $\mathbf{A}$ has to be the same as the number of rows of $\mathbf{B}$ (this is called compatibility requirement), and that $\mathbf{C}$ is $p \times r$. For example,

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 4 & 11 \\ 3 & 7 \end{bmatrix}.$$

Whenever we write $\mathbf{AB}$, the sizes of $\mathbf{A}$ and $\mathbf{B}$ are understood to be appropriate to make the product valid. The product $\mathbf{PQR}$ of three matrices is defined as $(\mathbf{PQ})\mathbf{R}$ provided that the sizes of the matrices are compatible. Note that matrix multiplications is associative, that is, $(\mathbf{PQ})\mathbf{R} = \mathbf{P}(\mathbf{QR})$. However, it is not commutative, i.e., in general $\mathbf{AB} \neq \mathbf{BA}$. For example if $\mathbf{A}$ is 2×3 and $\mathbf{B}$ is 3×4, then $\mathbf{AB}$ is well-defined but $\mathbf{BA}$ is not defined at all. It can be shown that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ and $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$. The notation $\mathbf{A}^n$ stands for the product $\mathbf{AAA}\ldots\mathbf{A}$ (n times).

The reversal matrix. Matrices of the form

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

are said to be reversal matrices. The general notation for an $M \times M$ reversal matrix is $\mathbf{J}_M$ with subscript omitted when obvious. The matrix $\mathbf{JA}$ is obtained from $\mathbf{A}$ by renumbering the rows in reverse order. Similarly $\mathbf{AJ}$ is obtained by renumbering the columns in reverse order. Given a diagonal matrix $\mathbf{\Lambda}$, the product $\mathbf{J\Lambda J}$ represents a new diagonal matrix with diagonal elements in reverse order. For example

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

Trace of a matrix. The trace of a square matrix $\mathbf{A}$, denoted $\text{Tr}(\mathbf{A})$ is defined to be the sum of the diagonal elements, i.e., $\sum_i A_{ii}$. It can be shown that $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$ as long as both products are meaningful.

Norms, Inner Products and Outer Products

Given two N-vectors $\mathbf{u}$ and $\mathbf{v}$, consider $\alpha = \mathbf{u}^\dagger\mathbf{v}$. This is a scalar quantity and is called the *inner product* of $\mathbf{u}$ with $\mathbf{v}$. The vectors are said to be (mutually) orthogonal if $\mathbf{u}^\dagger\mathbf{v} = 0$.

The inner product of $\mathbf{u}$ with $\mathbf{u}$, that is, $\mathbf{u}^\dagger\mathbf{u}$ is called the *energy* of $\mathbf{u}$. Denoting the elements of $\mathbf{u}$ as u_i, we see that $\mathbf{u}^\dagger\mathbf{u} = \sum_{i=0}^{N-1} |u_i|^2 \geq 0$. For example

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2j \end{bmatrix} \Rightarrow \mathbf{u}^\dagger\mathbf{u} = 5.$$

The quantity $\mathbf{u}^\dagger\mathbf{u}$ is nonzero (hence positive) unless $\mathbf{u} = \mathbf{0}$.

The norm $||\mathbf{u}||$ of $\mathbf{u}$ is defined as the positive square root of its energy, i.e.,

$$||\mathbf{u}|| = \sqrt{\mathbf{u}^\dagger\mathbf{u}} \qquad \text{(norm of } \mathbf{u}\text{)}. \qquad (A.2.4)$$

Sometimes this is also called the $\mathcal{L}_2$ *norm*, and denoted as $||\mathbf{u}||_2$.

Given a p-vector $\mathbf{u}$ and an r-vector $\mathbf{v}$, the quantity $\mathbf{A} = \mathbf{uv}^\dagger$ is a $p \times r$ matrix and is called the *outer product* of $\mathbf{u}$ with $\mathbf{v}$. This quantity is also called a *diadic* matrix. Example:

$$\mathbf{u} = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} v_0 \\ v_1 \\ v_2 \end{bmatrix}, \ \mathbf{uv}^\dagger = \begin{bmatrix} u_0v_0^* & u_0v_1^* & u_0v_2^* \\ u_1v_0^* & u_1v_1^* & u_1v_2^* \end{bmatrix}. \qquad (A.2.5)$$

Cauchy-Schwartz inequality. Given two column vectors $\mathbf{u}$ and $\mathbf{v}$, it can be shown that

$$|\mathbf{u}^\dagger\mathbf{v}| \leq ||\mathbf{u}||||\mathbf{v}||, \qquad (A.2.6)$$

with equality if and only if $\mathbf{v} = c\mathbf{u}$ for some scalar c. For example,

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow \mathbf{u}^{\dagger}\mathbf{u} = 2, \mathbf{v}^{\dagger}\mathbf{v} = 5, \mathbf{u}^{\dagger}\mathbf{v} = 3,$$

and (A.2.6) holds with strict inequality.

A.3 DETERMINANTS

There are several equivalent definitions of the determinant of a $p \times p$ matrix $\mathbf{A}$. We will conveniently define it recursively as follows:

$$\det \mathbf{A} = \sum_{k=0}^{p-1} (-1)^{k+m} a_{km} M_{km}, \tag{A.3.1}$$

where m is a fixed integer in $0 \leq m \leq p-1$. Here M_{km} is the determinant of the $(p-1) \times (p-1)$ submatrix obtained by deleting the kth row and mth column of $\mathbf{A}$.

Minors and cofactors. The quantity M_{km} is said to be the *minor* of the element a_{km}. The quantity $(-1)^{k+m} M_{km}$ is said to be the *cofactor* of a_{km}. In (A.3.1), the fixed column-index m is arbitrary.

In the above formula, the determinant has been computed by working with the mth column. Similarly, one can work with the mth row and obtain the determinant as

$$\det \mathbf{A} = \sum_{k=0}^{p-1} (-1)^{k+m} a_{mk} M_{mk}. \tag{A.3.2}$$

Whenever we write $[\det \mathbf{A}]$, it is implicit that $\mathbf{A}$ is square. The determinant of $\mathbf{A}$ is denoted either as $[\det \mathbf{A}]$ or as $|\mathbf{A}|$.

For 2×2 matrices the above formula is simplified to

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Here are some examples:

$$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1.$$

$$\begin{vmatrix} 1 & 4 & 5 \\ 2 & 1 & 1 \\ 3 & 1 & 2 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} - 2 \times \begin{vmatrix} 4 & 5 \\ 1 & 2 \end{vmatrix} + 3 \times \begin{vmatrix} 4 & 5 \\ 1 & 1 \end{vmatrix} = -8.$$

Determinants of block-diagonal matrices. Let $\mathbf{A}$ be a square matrix of the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}.$$

Then we can show that $[\det \mathbf{A}] = [\det \mathbf{P}][\det \mathbf{Q}]$.

Principal and leading-principal minors. In general, the determinant of any square submatrix of $\mathbf{A}$ is said to be a minor of $\mathbf{A}$. Let $\mathbf{A}$ be square. A principal minor is any minor whose diagonal elements are also the diagonal elements of $\mathbf{A}$. Thus for a 3×3 matrix $\mathbf{A}$ the principal minors are

$$a_{00}, a_{11}, a_{22}, \begin{vmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{vmatrix}, \begin{vmatrix} a_{00} & a_{02} \\ a_{20} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \text{and } [\det \mathbf{A}].$$

Next, a leading principal minor of $\mathbf{A}$ is a principal minor such that if a_{kk} is an element, then so is a_{ii} for all $i < k$. Thus, the leading principal minors for a 3×3 matrix are:

$$a_{00}, \begin{vmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{vmatrix}, \text{ and } [\det \mathbf{A}]. \tag{A.3.3}$$

Properties of Determinants

a) Let $\mathbf{C} = \mathbf{AB}$. Then $[\det \mathbf{C}] = [\det \mathbf{A}][\det \mathbf{B}]$.
b) If $\mathbf{B}$ is obtained from $\mathbf{A}$ by interchanging two rows (or two columns), then $[\det \mathbf{B}] = -[\det \mathbf{A}]$.
c) $[\det \mathbf{A}^T] = [\det \mathbf{A}]$.
d) For a $p \times p$ matrix $\mathbf{A}$, $[\det c\mathbf{A}] = c^p[\det \mathbf{A}]$, for any scalar c.
e) The determinant of a diagonal matrix is the product of its diagonal elements. The same is true for lower or upper triangular square matrices.
f) If any row is a scalar multiple of another row, the determinant is zero. If any row is zero, the determinant is zero. These statements also hold if 'row' is replaced with 'column' everywhere.

Singular and nonsingular matrices. A square matrix is said to be singular if $[\det \mathbf{A}] = 0$, and nonsingular if $[\det \mathbf{A}] \neq 0$. The product $\mathbf{AB}$ of two square matrices is nonsingular if and only if each of $\mathbf{A}$ and $\mathbf{B}$ is nonsingular (since the determinant of $\mathbf{AB}$ is the product of individual determinants).

A.4 LINEAR INDEPENDENCE, RANK, AND RELATED ISSUES

Let $\mathbf{v}_k, 0 \leq k \leq m-1$ be a set of m column vectors. A linear combination of these vectors is an expression of the form $\sum_{k=0}^{m-1} \alpha_k \mathbf{v}_k$, and is clearly a vector of the same size. (Here α_k are, in general, complex numbers.) The set $\mathcal{S}$ of all linear combinations of these vectors is said to be the *space* or *vector space* spanned by these vectors.

Another way to define a vector space in our context is this: a vector space is a collection of vectors of a given size such that every possible linear combination from this collection also belongs to this collection. In particular, the null vector (the vector with all components equal to zero) is a member of the vector space.

Linear independence. We say that the vectors $\mathbf{v}_k$, $0 \leq k \leq m-1$ are linearly dependent if there exists a set of m scalars α_k, not all zero, such that $\sum_{k=0}^{m-1} \alpha_k \mathbf{v}_k = \mathbf{0}$. The set of vectors is linearly independent if they are not linearly dependent. For row vectors, we have an identical definition.

Basis vectors. The set of all vectors of the form $\sum_{k=0}^{m-1} \alpha_k \mathbf{v}_k$ is said to be the space spanned by the m column-vectors $\mathbf{v}_k$. If $\mathcal{S}$ is the space spanned by a set of *linearly independent* vectors, then these vectors are said to form a basis for this space. The minimum number of basis vectors required to span the space under question is called the *dimension* of the space. The basis set is not unique. For example, either of the following two sets of vectors spans the entire space of two-component vectors:

$$\underbrace{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}}_{set\ 1}, \quad \underbrace{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}}_{set\ 2} \tag{A.4.1}$$

Rank of a Matrix

There are several equivalent ways to define rank (even though the equivalence is not obvious). We define the rank of a $p \times r$ matrix $\mathbf{A}$ to be an integer $\rho(\mathbf{A})$ (denoted just ρ when there is no confusion) such that there exists a $\rho \times \rho$ nonsingular submatrix of $\mathbf{A}$ but there does not exist a larger nonsingular submatrix. If ρ_1 denotes the largest number of columns of $\mathbf{A}$ that form a linearly independent set, we say that ρ_1 is the column rank of $\mathbf{A}$. Evidently $\rho_1 \leq r$, and if $\rho_1 = r =$ number of columns, we say that $\mathbf{A}$ has full column rank. Similarly we can define row rank ρ_2, and full row-rank matrices. It turns out that, for any matrix, $\rho = \rho_1 = \rho_2$. It is also clear from the definition that a $p \times p$ matrix is nonsingular if and only if it has full rank p.

♠**Fact A.4.1. Important properties of rank.** We now list the key features of rank.

a) $\rho = \rho_1 = \rho_2$, as stated above.
b) A $p \times p$ matrix is nonsingular if and only if the rank $\rho = p$.
c) $\rho(\mathbf{AB}) \leq \min(\rho(\mathbf{A}), \rho(\mathbf{B}))$.
d) If $\rho(\mathbf{A}) = 0$ then $\mathbf{A} = \mathbf{0}$.
e) Suppose $\mathbf{A}$ is $N \times N$ with rank $\rho < N$. This means that there are ρ linearly independent columns, from which all columns can be generated by linear combination. As a result, $\mathbf{A}$ can be written as

$$\mathbf{A} = \underbrace{\mathbf{B}}_{N\times\rho}\underbrace{\mathbf{C}}_{\rho\times N}. \tag{A.4.2}$$

Any $N \times N$ matrix $\mathbf{A}$ with rank ρ can be factorized like this.

f) *Sylvester's inequality.* Let $\mathbf{P}$ and $\mathbf{Q}$ be $M \times N$ and $N \times K$ matrices with ranks ρ_p and ρ_q. Let ρ_{pq} be the rank of $\mathbf{PQ}$. Then,

$$\rho_p + \rho_q - N \leq \rho_{pq} \leq \min(\rho_p, \rho_q).$$

g) Given two square matrices $\mathbf{A}$ and $\mathbf{B}$, the products $\mathbf{AB}$ and $\mathbf{BA}$ may not have the same rank. However the matrices $\mathbf{I} - \mathbf{AB}$ and $\mathbf{I} - \mathbf{BA}$ do have the same rank. See Problem A.13. ◇

Diadic matrices. If $\rho(\mathbf{A}) = 1$, then every column of $\mathbf{A}$ is a scalar multiple of every other column. The same is true of the rows. In this case we can write $\mathbf{A} = \mathbf{u}\mathbf{v}^\dagger$ where $\mathbf{u}$ and $\mathbf{v}$ are column vectors (so that $\mathbf{v}^\dagger$ is a row vector). In other words, any rank-one matrix is an outer product of two column vectors, and is sometimes called a *diadic* matrix. Here is an example:

$$\mathbf{A} = \begin{bmatrix}1 & 2 & 3j\\2 & 4 & 6j\\1 & 2 & 3j\\4 & 8 & 12j\end{bmatrix} = \underbrace{\begin{bmatrix}1\\2\\1\\4\end{bmatrix}}_{\mathbf{u}}\underbrace{\begin{bmatrix}1 & 2 & 3j\end{bmatrix}}_{\mathbf{v}^\dagger}, \tag{A.4.3}$$

so that

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ -3j \end{bmatrix} \tag{A.4.4}$$

in this example.

Range space and null space. Given a $p \times r$ matrix $\mathbf{A}$, the space spanned by its columns is said to be the range space (or *column* space) of $\mathbf{A}$. This space is the set of all vectors of the form $\mathbf{Ax}$ where $\mathbf{x}$ is any r-component column vector. The dimension of the range space is equal to the rank of $\mathbf{A}$. Note that the elements of the range space are p-vectors. The null space of $\mathbf{A}$ is the set of all r-vectors $\mathbf{y}$ such that $\mathbf{Ay} = \mathbf{0}$. It can be shown that the set of all linear combinations from the range space of $\mathbf{A}$ and the null space of $\mathbf{A}^\dagger$ is equal to the complete space of all vectors of size p.

Orthogonal complements. Let $\mathbf{t}_0, \mathbf{t}_1 \ldots \mathbf{t}_{\rho-1}$ be a set of linearly independent N-vectors and let $\mathcal{V}$ be the vector space spanned by them. (Clearly, $\rho \leq N$). Now consider the set $\mathcal{V}^\perp$ of all N-vectors orthogonal to all the vectors in $\mathcal{V}$ (i.e., orthogonal to all the above $\mathbf{t}_i$). The set $\mathcal{V}^\perp$ is itself a vector space, and is called the *orthogonal complement* of $\mathcal{V}$. It has dimension $N - \rho$. Any N-vector $\mathbf{x}$ can be expressed as a linear combination of one vector in $\mathcal{V}$ and one in $\mathcal{V}^\perp$. That is,

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1, \quad \mathbf{x}_0 \in \mathcal{V},\ \mathbf{x}_1 \in \mathcal{V}^\perp$$

Moreover, for a given $\mathbf{x}$, the components $\mathbf{x}_0$ and $\mathbf{x}_1$ are unique. Letting $\mathbf{t}_\rho \ldots \mathbf{t}_{N-1}$ denote a basis for $\mathcal{V}^\perp$, the matrix

$$[\mathbf{t}_0 \quad \mathbf{t}_1 \quad \ldots \quad \mathbf{t}_{N-1}]$$

is $N \times N$ nonsingular. Its columns span the space of all N-vectors.

The annihilating vector. Given a matrix $\mathbf{A}$, any vector $\mathbf{y}$ such that $\mathbf{Ay} = \mathbf{0}$ is called an annihilating vector for $\mathbf{A}$. Evidently, the set of all annihilating vectors is equal to the null space of $\mathbf{A}$ defined above. A nonzero annihilating vector exists whenever the column rank of $\mathbf{A}$ is less than full, so that a linear combination of the columns can be made zero. For square matrices, this is equivalent to the condition that the determinant be zero.

Inverse of a Square Matrix

Given a $p \times r$ matrix $\mathbf{A}$, we say that the $r \times p$ matrix $\mathbf{L}$ is a left inverse if $\mathbf{LA} = \mathbf{I}_r$. We say that the $r \times p$ matrix $\mathbf{R}$ is a right inverse if $\mathbf{AR} = \mathbf{I}_p$. Inverses may or may not exist, and in general are not unique. For the case of square matrices, however the following are true: (a) An inverse exists if and only if $\mathbf{A}$ has full rank, i.e., $\mathbf{A}$ is nonsingular, and (b) when they exist, the left and right inverses are the same, and unique.

There is a closed form expression for the inverse of a nonsingular square matrix, given by

$$\mathbf{A}^{-1} = \frac{\text{Adj } \mathbf{A}}{\det \mathbf{A}}, \tag{A.4.5}$$

where [Adj $\mathbf{A}$] is the *adjugate* of $\mathbf{A}$ (referred to as adjoint in some texts), defined as

$$[\text{Adj } \mathbf{A}]_{ij} = \text{cofactor of } A_{ji}. \tag{A.4.6}$$

In other words, the (i,j) element of the adjugate is equal to the cofactor of the (j,i) element of $\mathbf{A}$.

The matrix-inversion lemma. The following inversion formula, which holds whenever $\mathbf{P}$ and $\mathbf{R}$ are nonsingular, is very useful in system theoretic work. ($\mathbf{Q}$ and $\mathbf{S}$ need not be square).

$$(\mathbf{P}+\mathbf{QRS})^{-1} = \mathbf{P}^{-1} - \mathbf{P}^{-1}\mathbf{Q}(\mathbf{SP}^{-1}\mathbf{Q}+\mathbf{R}^{-1})^{-1}\mathbf{SP}^{-1}, \tag{A.4.7}$$

Linear Equations

Consider an equation of the form $\mathbf{Ax}=\mathbf{b}$, where $\mathbf{A}$ is $N\times N$, and $\mathbf{b}$ and $\mathbf{x}$ are N-vectors. Given the quantities $\mathbf{A}$ and $\mathbf{b}$, we wish to find $\mathbf{x}$ satisfying this equation. Basically we have a set of N linear equations in N unknowns (the elements of $\mathbf{x}$). Given $\mathbf{A}$ and $\mathbf{b}$, there are several possibilities:

1. If $\mathbf{A}$ is nonsingular, there exists a unique solution given by $\mathbf{x}=\mathbf{A}^{-1}\mathbf{b}$.
2. If $\mathbf{A}$ is singular, then there are two possibilities: either there does not exist a solution, or there is an infinite number of solutions. (Since $\mathbf{A}$ is singular, $\mathbf{Av}=\mathbf{0}$ for some $\mathbf{v}\neq\mathbf{0}$ which shows that if $\mathbf{x}$ is a solution, then $\mathbf{x}+c\mathbf{v}$ is also a solution for *any* scalar c.)

More generally, consider the equation

$$\mathbf{Ax}=\mathbf{b}, \tag{A.4.8}$$

where $\mathbf{A}$ is $N\times M$. The following are true.

1. If $\mathbf{A}$ has rank N (which implies $N\leq M$), then there exists $\mathbf{x}$ such that (A.4.8) holds.
2. If the rank of $\mathbf{A}$ is less than N, then depending on $\mathbf{b}$ there may or may not exist $\mathbf{x}$ satisfying (A.4.8).
3. In any case, if a solution exists, it is *unique* if and only if the rank of the matrix $\mathbf{A}$ equals the number of columns M.

A.5 EIGENVALUES AND EIGENVECTORS

Given a $N\times N$ square matrix $\mathbf{A}$, consider $D(s)=\det[s\mathbf{I}-\mathbf{A}]$. This is a polynomial in s with order N, called the *characteristic polynomial.* The N roots of $D(s)$ are said to be the N eigenvalues of $\mathbf{A}$. If a particular root λ_0 has multiplicity K, that is, if $D(s)$ has the factor $(s-\lambda_0)^K$, then the eigenvalue λ_0 is said to have multiplicity K. From the above definition, one can verify that λ is an eigenvalue of $\mathbf{A}$ if and only if there exists a nonzero vector $\mathbf{v}$ such that

$$\mathbf{Av}=\lambda\mathbf{v}. \tag{A.5.1}$$

The vector $\mathbf{v}$ is said to be an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ.

For example, let $\mathbf{A}=\begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}$. Then

$$D(s)=\begin{vmatrix}s-3 & -1\\ -1 & s-3\end{vmatrix} = s^2-6s+8=(s-4)(s-2),$$

so that the eigenvalues are $\lambda_0 = 4$ and $\lambda_1 = 2$. Furthermore

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

so that the corresponding eigenvectors are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Properties of Eigenvalues and Eigenvectors

a) If $\mathbf{v}$ is an eigenvector of $\mathbf{A}$, then so is $c\mathbf{v}$ for any scalar $c \neq 0$.

b) If $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors corresponding to distinct eigenvalues λ_1 and λ_2, then $\mathbf{v}_1$ and $\mathbf{v}_2$ are "distinct". More accurately, we cannot write $\mathbf{v}_1 = \alpha\mathbf{v}_2$ for any scalar α.

c) Let $\mathbf{A}$ be $N \times N$ with N distinct eigenvalues λ_k, $0 \leq k \leq N-1$. In other words, none of the eigenvalues is a multiple root of $D(s)$. Then the corresponding eigenvectors $\mathbf{v}_k$, $0 \leq k \leq N-1$ are linearly independent. Also, each eigenvector $\mathbf{v}_k$ is unique (except, of course, for a scale factor). Notice, in general, that if $\mathbf{A}$ has less than N distinct eigenvalues, then there may or may not exist a set of N linearly independent eigenvectors.

d) Suppose λ is a complex eigenvalue of a real matrix $\mathbf{A}$. Then its conjugate λ^* is also an eigenvalue. Also if an eigenvalue of a real matrix is complex then the corresponding eigenvector is necessarily complex.

e) $\mathbf{A}$ has an eigenvalue equal to zero if and only if it is singular.

f) The eigenvalues of $\mathbf{A}$ are same as those of $\mathbf{A}^T$.

g) For a square (lower or upper) triangular matrix, the eigenvalues are equal to the diagonal elements. (The same is true of a diagonal matrix, which is a special case.) This follows by noting that, in these cases, $[\det(s\mathbf{I} - \mathbf{A})] = \prod(s - a_{ii})$.

h) Let $\mathbf{A}$ be $N \times N$ with eigenvalues λ_i. Then the determinant and trace can be expressed as

$$\det \mathbf{A} = \prod_{i=0}^{N-1} \lambda_i, \quad \mathrm{Tr}\,(\mathbf{A}) = \sum_{i=0}^{N-1} \lambda_i.$$

i) For nonsingular $\mathbf{A}$, the eigenvalues of $\mathbf{A}^{-1}$ are reciprocals of those of $\mathbf{A}$.

j) If λ_k are the eigenvalues of $\mathbf{A}$, the eigenvalues of $\mathbf{A} + \sigma\mathbf{I}$ are equal to $\lambda_k + \sigma$. *Proof.* Let $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, then $(\mathbf{A} + \sigma\mathbf{I})\mathbf{v} = \mathbf{A}\mathbf{v} + \sigma\mathbf{v} = (\lambda + \sigma)\mathbf{v}$.

It is possible for all eigenvalues to be equal to zero, even if $\mathbf{A} \neq \mathbf{0}$. An example is a triangular matrix with all diagonal elements equal to zero.

This appears to be a good place to summarize the various ways in which singularity of a matrix can manifest:

♠**Fact A.5.1. On singularity.** Let $\mathbf{A}$ be $N \times N$. Then the following statements are equivalent:

a) $\mathbf{A}$ is singular.
b) $[\det \mathbf{A}] = 0$.
c) There exists an eigenvalue of $\mathbf{A}$ equal to zero.
d) There exists a nonzero vector $\mathbf{v}$ such that $\mathbf{A}\mathbf{v} = \mathbf{0}$.
e) The rank of $\mathbf{A}$ is less than N.
f) The N columns (and rows) of $\mathbf{A}$ are not linearly independent.
g) $\mathbf{A}$ has no inverse.

h) The equation $\mathbf{Ax} = \mathbf{b}$ has no unique solution $\mathbf{x}$ (i.e., it has either no solutions, or an infinite number of them). ◇

Eigenspaces. Suppose the N eigenvalues of $\mathbf{A}$ are not distinct. It is then conceivable that an eigenvalue, say λ_0, has more than one eigenvector. Suppose, for example, that $\{\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2\}$ is a set of three linearly independent eigenvectors corresponding to λ_0, i.e., $\mathbf{A}\mathbf{v}_k = \lambda_0 \mathbf{v}_k$ for $k = 0, 1, 2$. Then any linear combination of $\mathbf{v}_0, \mathbf{v}_1$ and $\mathbf{v}_2$ (i.e., any vector in the space spanned by $\mathbf{v}_0, \mathbf{v}_1$ and $\mathbf{v}_2$) is an eigenvector for λ_0. This vector space spanned by $\mathbf{v}_0, \mathbf{v}_1$, and $\mathbf{v}_2$ is the *eigenspace* corresponding to λ_0.

Similarity transformations. Given a square matrix $\mathbf{A}$, suppose we define $\mathbf{A}_1 = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ where $\mathbf{T}$ is some nonsingular matrix. It turns out that $\mathbf{A}_1$ and $\mathbf{A}$ have the same set of eigenvalues. (*Proof:* $\mathbf{Av} = \lambda\mathbf{v} \Rightarrow \mathbf{T}^{-1}\mathbf{AT}(\mathbf{T}^{-1}\mathbf{v}) = \lambda(\mathbf{T}^{-1}\mathbf{v})$. This transformation of $\mathbf{A}$ to $\mathbf{A}_1$ is said to be a similarity transformation.

Diagonalization

Suppose $\mathbf{A}$ is $N \times N$, and assume that it has N linearly independent eigenvectors $\mathbf{t}_k$. (This does not necessarily mean that there are N distinct eigenvalues.) We can write $\mathbf{A}\mathbf{t}_k = \lambda_k \mathbf{t}_k$, $0 \le k \le N-1$. We can compactly write these as one matrix equation:

$$\mathbf{A}\underbrace{[\,\mathbf{t}_0 \quad \mathbf{t}_1 \quad \ldots \quad \mathbf{t}_{N-1}\,]}_{\mathbf{T}} = \underbrace{[\,\mathbf{t}_0 \quad \mathbf{t}_1 \quad \ldots \quad \mathbf{t}_{N-1}\,]}_{\mathbf{T}}\boldsymbol{\Lambda}, \tag{A.5.2}$$

where $\boldsymbol{\Lambda}$ is an $N \times N$ diagonal matrix with kth diagonal element equal to λ_k. We can rearrange this as

$$\mathbf{T}^{-1}\mathbf{AT} = \boldsymbol{\Lambda}, \quad i.e., \quad \mathbf{A} = \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{-1}. \tag{A.5.3}$$

This shows that if there exist N linearly independent eigenvectors, then we can diagonalize $\mathbf{A}$ by applying a similarity transformation. Conversely, whenever we can find $\mathbf{T}$ such that $\mathbf{T}^{-1}\mathbf{AT}$ is diagonal, the columns of $\mathbf{T}$ are eigenvectors of $\mathbf{A}$ with corresponding eigenvalues appearing on the diagonals of $\mathbf{T}^{-1}\mathbf{AT}$. For example, let $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. We have already computed the eigenvalues and eigenvectors above. From these we obtain

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}}_{\mathbf{T}^{-1}} \underbrace{\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{T}} = \underbrace{\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}}_{\boldsymbol{\Lambda}}.$$

An $N \times N$ matrix $\mathbf{A}$ is said to be *diagonalizable* if it can be written as in (A.5.3) for some diagonal $\boldsymbol{\Lambda}$ and nonsingular $\mathbf{T}$. The following points are worth noting.

1. Not every $N \times N$ matrix is diagonalizable. For example suppose $\mathbf{A}$ is a nonzero matrix such that the only possible eigenvalue is $\lambda = 0$. (An example is $\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.) If $\mathbf{A}$ is diagonalizable, then $\mathbf{A} = \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{-1}$ with $\boldsymbol{\Lambda} = \mathbf{0}$ so that $\mathbf{A} = \mathbf{0}$, which is a contradiction.
2. Every matrix with N distinct eigenvalues is diagonalizable because there are N linearly independent eigenvectors.

3. There is a class of matrices called normal matrices (defined below), which are diagonalizable even if the eigenvalues are not distinct.

Cayley-Hamilton Theorem

Recall that the characteristic polynomial $D(s)$ is defined as [det $(s\mathbf{I}-\mathbf{A})$] and has the form $D(s) = s^N + d_{N-1}s^{N-1} + \ldots + d_0$. The equation $D(s) = 0$ is called the *characteristic equation* (its solutions being the eigenvalues). It turns out that the $N \times N$ matrix $\mathbf{A}$ satisfies this equation, that is,

$$\mathbf{A}^N + d_{N-1}\mathbf{A}^{N-1} + \ldots + d_1\mathbf{A} + d_0\mathbf{I} = \mathbf{0}. \quad (A.5.4)$$

This result is called the *Cayley-Hamilton theorem.* This says that the matrix $\mathbf{A}^N$ can be expressed as a linear combination of the lower powers.

A.6 SPECIAL TYPES OF MATRICES

We now discuss a number of special types of matrices which arise in our discussions throughout the text.

Hermitian matrix. $\mathbf{H}$ is said to be Hermitian if $\mathbf{H}^\dagger = \mathbf{H}$. This implies $H_{ij} = H_{ji}^*$. Note that a Hermitian matrix, by definition, is square. A real Hermitian matrix is *symmetric* (i.e., $\mathbf{H}^T = \mathbf{H}$). Variations of this class are the skew-Hermitian matrix ($\mathbf{H}^\dagger = -\mathbf{H}$), and antisymmetric matrix ($\mathbf{H}^T = -\mathbf{H}$).

Any matrix $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{A}_h + \mathbf{A}_s$$

where $\mathbf{A}_h$ is Hermitian and $\mathbf{A}_s$ is skew-Hermitian. For this, just define

$$\mathbf{A}_h = \frac{\mathbf{A}+\mathbf{A}^\dagger}{2}, \quad \mathbf{A}_s = \frac{\mathbf{A}-\mathbf{A}^\dagger}{2}$$

Unitary matrix. $\mathbf{U}$ is said to be unitary if $\mathbf{U}^\dagger\mathbf{U} = c\mathbf{I}$ for some $c > 0$. This means that every pair of columns is mutually orthogonal, and that all columns have the same norm $\sqrt{c}$. If the unitary matrix is square, then $\mathbf{U}^T$ as well as $\mathbf{U}^\dagger$ are unitary. Thus, for a square unitary matrix $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = c\mathbf{I}$. If $c = 1$, $\mathbf{U}$ is *normalized* unitary. A real unitary matrix is usually said to be an orthogonal matrix (ortho*normal* if $\mathbf{U}^T\mathbf{U} = \mathbf{I}$). In Chap. 14 one can find more details about unitary matrices, planar rotations, Householder forms, and factorizations.

Let $\mathbf{y} = \mathbf{U}\mathbf{x}$. If $\mathbf{U}$ is unitary, it is clear that $\mathbf{y}^\dagger\mathbf{y} = c\mathbf{x}^\dagger\mathbf{x}$, for any choice of $\mathbf{x}$. So a unitary matrix changes the norms of all vectors by the same factor $\sqrt{c}$. Conversely, suppose a matrix $\mathbf{U}$ is such that $\mathbf{y}^\dagger\mathbf{y} = c\mathbf{x}^\dagger\mathbf{x}$ for all vectors $\mathbf{x}$. Then $\mathbf{U}$ is unitary (Problem A.17).

Circulant matrices. A square matrix is right circulant if each row is obtained by a right circular shift of the previous row. Example:

$$\mathbf{A} = \begin{bmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{bmatrix}. \quad (A.6.1)$$

The left-circulant property is similarly defined. Unless mentioned otherwise, 'circulant' denotes right circulants.

Normal matrix. $\mathbf{A}$ is said to be normal if $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}$. By definition, $\mathbf{A}$ has to be a square matrix. It can be verified that the following matrices are normal: (a) Hermitian and skew-Hermitian matrices, (b) square unitary matrices and (c) circulants.

The DFT and IDFT Matrices

A matrix of special interest in digital signal processing is the Discrete Fourier Transform (DFT) matrix. This is a $N \times N$ matrix defined as $\mathbf{W}_N = [W_N^{km}]$ where $W_N = e^{-j2\pi/N}$. In other words, the entry at the kth row and mth column is equal to $e^{-j2\pi km/N}$. Evidently this is a symmetric (but complex) matrix. Examples are:

$$\mathbf{W}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{W}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix}.$$

The subscripts N on $\mathbf{W}$ and W are usually omitted if they are clear from the context. The matrix $\mathbf{W}$ satisfies the property $\mathbf{W}^\dagger\mathbf{W} = N\mathbf{I}$ so that it is unitary. Given a finite length sequence $x(n)$, $0 \le n \le N-1$, suppose we define the vector $\mathbf{x} = [x(0) \quad x(1) \dots \quad x(N-1)]^T$, and compute the vector $\mathbf{X} = \mathbf{W}\mathbf{x}$. Then the components of $\mathbf{X}$, viz., $X(k), 0 \le k \le N-1$ are said to form the DFT coefficients of the sequence $x(n)$. The sequence $x(n)$ is the inverse DFT (abbreviated IDFT) of the sequence $X(k)$. The matrix $\mathbf{W}^{-1}$ (which is equal to $\mathbf{W}^\dagger/N$) is called the *IDFT matrix.* Notice that $\mathbf{W}$ is symmetric, that is, $\mathbf{W}^T = \mathbf{W}$ so that $\mathbf{W}^\dagger = \mathbf{W}^*$.

The DFT and IDFT relations are more commonly written as

$$X(k) = \sum_{m=0}^{N-1} x(m)W^{km}, \quad x(m) = \frac{1}{N}\sum_{k=0}^{N-1} X(k)W^{-km} \tag{A.6.2}$$

Toeplitz matrices. An $N \times N$ matrix $\mathbf{A}$ is said to be Toeplitz if the elements A_{ij} are determined completely by the difference $i-j$. For example,

$$\mathbf{A} = \begin{bmatrix} a_0 & a_1 & a_2 \\ a_3 & a_0 & a_1 \\ a_4 & a_3 & a_0 \end{bmatrix}, \tag{A.6.3}$$

is Toeplitz. Pictorially, if we draw a line parallel to the main diagonal, then all elements on this line are equal. Thus, a Toeplitz matrix is completely determined by the 0th row and 0th column, that is, by $2N-1$ elements. Notice that circulants are Toeplitz.

If we replace each of the $2N-1$ elements in a Toeplitz matrix by a (possibly rectangular) matrix, we obtain a block-Toeplitz matriz. An example is

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_0 & \mathbf{a}_1 & \mathbf{a}_2 \\ \mathbf{a}_3 & \mathbf{a}_0 & \mathbf{a}_1 \\ \mathbf{a}_4 & \mathbf{a}_3 & \mathbf{a}_0 \end{bmatrix}, \tag{A.6.4}$$

where $\mathbf{a}_i$ are themselves 2×2 matrices.

Vandermonde matrices. An $N \times N$ matrix, each of whose rows has the form

$$[1 \quad a_i \quad a_i^2 \quad \ldots \quad a_i^{N-1}] \tag{A.6.5}$$

is a Vandermonde matrix. Example:

$$\mathbf{A} = \begin{bmatrix} 1 & a_0 & a_0^2 \\ 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \end{bmatrix}. \tag{A.6.6}$$

The transpose of a Vandermonde matrix, for example,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ a_0 & a_1 & a_2 \\ a_0^2 & a_1^2 & a_2^2 \end{bmatrix}, \tag{A.6.7}$$

is also said to be a Vandermonde matrix. Note that the DFT matrix is Vandermonde. The determinant of a Vandermonde matrix is given by

$$\det \mathbf{A} = \prod_{i>j}(a_i - a_j). \tag{A.6.8}$$

For example, if $\mathbf{A}$ is as in (A.6.6),

$$\det \mathbf{A} = (a_1 - a_0)(a_2 - a_0)(a_2 - a_1). \tag{A.6.9}$$

It follows that a Vandermonde matrix is nonsingular if and only if the a_i's are distinct.

Eigenstructures of Special Matrices

Some of the above mentioned special matrices satisfy special properties related to eigenvalues and eigenvectors.

♠**Fact A.6.1. Normal matrices.** The $N \times N$ matrix $\mathbf{A}$ is normal if and only if we can find $N \times N$ unitary $\mathbf{U}$ such that $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ is diagonal, that is, if and only if we can write

$$\mathbf{AU} = \mathbf{U\Lambda} \tag{A.6.10}$$

for diagonal $\mathbf{\Lambda}$ and unitary $\mathbf{U}$. This means that normal matrices are precisely those for which there exists a complete set of *mutually orthogonal eigenvectors* (i.e., unitary diagonalization is possible). Without loss of generality we can assume the columns of $\mathbf{U}$ to have unit norm. Then (A.6.10) is the same as

$$\mathbf{A} = \mathbf{U\Lambda U}^\dagger, \quad \text{where} \quad \mathbf{U}^\dagger\mathbf{U} = \mathbf{UU}^\dagger = \mathbf{I}. \tag{A.6.11}$$

This is identical to (A.5.3) with $\mathbf{T} = \mathbf{U}$. ◇

Notice, as a corollary, that if all the eigenvalues of a normal matrix are identical, then it has the form $\mathbf{A} = \lambda\mathbf{I}$ (where λ is this common eigenvalue). The same is not true for arbitary matrices, for example, $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$.

♠**Fact A.6.2. Special normal matrices.** Since Hermitian, unitary and circulant square matrices are normal, they can be written as in (A.6.11). In addition, the following are true.

a) If $\mathbf{A}$ is Hermitian, all eigenvalues are real. Moreover, $\mathbf{v}^\dagger\mathbf{A}\mathbf{v}$ is real for all vectors $\mathbf{v}$.
b) If $\mathbf{A}$ is unitary ($\mathbf{A}^\dagger\mathbf{A} = c\mathbf{I}, c > 0$), then all the eigenvalues have magnitude $\sqrt{c}$.
c) If $\mathbf{A}$ is $M \times M$ circulant, then we can write (A.6.10) with $\mathbf{U} = \mathbf{W}/\sqrt{M}$, where $\mathbf{W}$ is the DFT matrix. The eigenvalues of $\mathbf{A}$ are the DFT coefficients of the 0th row of $\mathbf{A}$. The eigenvectors are the columns of $\mathbf{W}/\sqrt{M}$.

Quadratic Forms and Positive Definite Matrices

For any $N \times N$ matrix $\mathbf{P}$, the scalar $\mathbf{v}^\dagger\mathbf{P}\mathbf{v}$ is said to be a quadratic form. In particular when $\mathbf{P}$ is Hermitian we know that $\mathbf{v}^\dagger\mathbf{P}\mathbf{v}$ is real. If this is positive for *all nonzero* $\mathbf{v}$, we say that $\mathbf{P}$ is positive definite. Notice that this property is defined only for Hermitian matrices.

Based on the properties of $\mathbf{v}^\dagger\mathbf{P}\mathbf{v}$ we can in fact identify a number of definitions as follows:

$$\mathbf{v}^\dagger\mathbf{P}\mathbf{v}\begin{cases} > 0, \ \forall\mathbf{v} \neq \mathbf{0} & \text{(positive definite)} \\ \geq 0, \ \forall\mathbf{v} & \text{(positive semidefinite or nonnegative definite)} \\ < 0, \ \forall\mathbf{v} \neq \mathbf{0} & \text{(negative definite)} \\ \leq 0, \ \forall\mathbf{v} & \text{(negative semidefinite).} \end{cases} \tag{A.6.12}$$

Here are some examples:

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}_{\text{positive definite}} \quad \underbrace{\begin{bmatrix} 2 & j \\ -j & 2 \end{bmatrix}}_{\text{positive definite}} \quad \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}}_{\text{positive semidefinite}}$$

(See test for positive (semi)definiteness below). By definition, a positive definite matrix is also positive semidefinite. Notice that $\mathbf{P}$ is positive definite (semidefinite) if and only if $-\mathbf{P}$ is negative definite (semidefinite). It is possible that $\mathbf{P}$ does not belong to any of these categories. (Example: $\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.) In other words, it is possible that $\mathbf{v}^\dagger\mathbf{P}\mathbf{v}$ has different signs for different $\mathbf{v}$. In that case, $\mathbf{P}$ is said to be indefinite.

Matrix inequalities. If $\mathbf{P}$ is positive definite, we indicate it as $\mathbf{P} > 0$ ($\mathbf{P} \geq 0$ for semidefinite). Given two Hermitian matrices $\mathbf{P}$ and $\mathbf{Q}$ of the same size, we write $\mathbf{P} > \mathbf{Q}$ if $\mathbf{P} - \mathbf{Q}$ is positive definite ($\mathbf{P} \geq \mathbf{Q}$ if $\mathbf{P} - \mathbf{Q}$ is positive semidefinite). Notice however that, in general, the difference matrix $\mathbf{P} - \mathbf{Q}$ can be indefinite, even though $\mathbf{P}$ and $\mathbf{Q}$ are definite.

Properties of Positive Definite Matrices

For convenience of reference we now list a number of properties of positive (semi)definite matrices. We encourage the reader to verify these for the examples shown above.

a) All diagonal elements of a positive definite (semidefinite) matrix are positive (nonnegative).

b) The Hermitian matrix $\mathbf{P}$ is positive definite (semidefinite) if and only if all the eigenvalues are positive (nonnegative).
c) *Test for positive definiteness.* The Hermitian matrix $\mathbf{P}$ is positive definite if and only if all leading principal minors of $\mathbf{P}$ are positive, and positive semidefinite if and only if all principal minors are nonnegative.
d) If $\mathbf{P} \geq \mathbf{0}$ and $\mathbf{Q} \geq \mathbf{0}$, then $\mathbf{P} + \mathbf{Q} \geq \mathbf{0}$. If in addition $\mathbf{Q} > \mathbf{0}$ (or $\mathbf{P} > \mathbf{0}$), then $\mathbf{P} + \mathbf{Q} > \mathbf{0}$.
e) *Square roots.* Given a positive number a, we know that we can find a real square root. The beauty of positive definite matrices is that we can define square roots in a similar way. Given a Hermitian matrix $\mathbf{P}$, if we can factorize it as $\mathbf{P} = \mathbf{Q}^\dagger\mathbf{Q}$ for some $\mathbf{Q}$ (possibly rectangular), we say that $\mathbf{Q}$ is a square root. For example, consider $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. This is Hermitian, and we have already calculated its eigenvalues to be 4 and 2, so that it is positive definite. By using the diagonalization result for this, we obtain

$$\begin{aligned}\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{\mathbf{Q}^\dagger}\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{Q}}.\end{aligned}$$

Square roots are not unique. For example, suppose $\mathbf{Q}$ is a square root, then $\mathbf{UQ}$ is also a square root for any normalized unitary $\mathbf{U}$. It is clear that if there exists a square root for $\mathbf{P}$, then $\mathbf{P}$ is nonnegative definite because $\mathbf{v}^\dagger\mathbf{P}\mathbf{v} = \mathbf{v}^\dagger\mathbf{Q}^\dagger\mathbf{Q}\mathbf{v} = \mathbf{w}^\dagger\mathbf{w} \geq 0$. Conversely, it can be shown that any $N \times N$ nonnegative definite $\mathbf{P}$ with rank ρ can be factorized as $\mathbf{Q}^\dagger\mathbf{Q}$ where $\mathbf{Q}$ is $\rho \times N$. One technique to find such a factor $\mathbf{Q}$ is called *Cholesky decomposition* [Golub and Van Loan, 1989], which produces a lower triangular square root.
f) Suppose $\mathbf{Q}$ is $p \times r$, with $p \geq r$. Evidently the rank of $\mathbf{Q} \leq r$. Define the $r \times r$ positive semidefinite matrix $\mathbf{P} = \mathbf{Q}^\dagger\mathbf{Q}$. This is nonsingular (hence positive definite) if and only if $\mathbf{Q}$ has full rank r.
g) *Determinant and diagonal elements.* Let $\mathbf{P}$ be $N \times N$ Hermitian positive definite, and let P_{ii} denote its diagonal elements. Then

$$\det \mathbf{P} \leq \prod_{i=0}^{N-1} P_{ii},$$

with equality if and only if $\mathbf{P}$ is diagonal. See Problem A.19 for a proof.

♠**Fact A.6.3. Positive definite matrices.** Let $\mathbf{P}$ be $N \times N$ Hermitian. Then the following statements are equivalent.

a) $\mathbf{P}$ is positive definite (i.e., $\mathbf{v}^\dagger\mathbf{P}\mathbf{v} > 0$ for all vectors $\mathbf{v} \neq \mathbf{0}$.)
b) All eigenvalues of $\mathbf{P}$ are positive.
c) There exists an $N \times N$ nonsingular square root $\mathbf{Q}$.
d) There exists an $N \times N$ nonsingular *lower triangular* square root $\mathbf{\Delta}_\ell$.
e) There exists an $N \times N$ nonsingular *upper triangular* square root $\mathbf{\Delta}_u$.

f) All leading principal minors of $\mathbf{P}$ are positive. $\diamond$

A.7 UNITARY TRIANGULARIZATION

It should be noticed that an arbitrary square matrix may not be diagonalizable [i.e., expressible in the form (A.5.3)]. The class of matrices which can be diagonalized by unitary matrices [i.e., which can be expressed as in (A.6.11)] is even smaller (namely *normal* matrices). However, *every* square matrix can be *triangularized* by a unitary transformation, as stated next.

♠**Fact A.7.1.** Let $\mathbf{A}$ be an arbitrary $N \times N$ matrix. Then, we can always write it in the form

$$\mathbf{A} = \mathbf{U}\mathbf{\Delta}\mathbf{U}^{\dagger}, \tag{A.7.1}$$

where $\mathbf{U}$ is $N \times N$ normalized unitary (i.e., $\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}$), and $\mathbf{\Delta}$ is lower triangular. (A.7.1) can be regarded as a similarity transformation of $\mathbf{A}$ into $\mathbf{\Delta}$. So the eigenvalues of $\mathbf{A}$ are the same as those of $\mathbf{\Delta}$, which in turn are the diagonal elements of $\mathbf{\Delta}$. Note that the columns of $\mathbf{U}$ are not necessarily the eigenvectors of $\mathbf{A}$, unlike in diagonalization. $\diamond$

This result, due to Schur, is of great importance. As an example,

$$\begin{bmatrix} 2+j & j \\ -j & 2-j \end{bmatrix} = \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} 2 & 0 \\ 2j & 2 \end{bmatrix}}_{\mathbf{\Delta}} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{U}^{\dagger}}.$$

Since triangular matrices play an important role in many applications, it is useful to summarize some of their properties.

♠**Fact. A.7.2. Properties of triangular matrices.** Let $\mathbf{A}$ and $\mathbf{B}$ be $N \times N$ lower triangular. Then,

a) The product $\mathbf{AB}$ is lower triangular.
b) [det $\mathbf{A}$] is equal to the product of diagonal elements A_{ii}.
c) The eigenvalues of $\mathbf{A}$ are equal to the diagonal elements A_{ii}.
d) If all diagonal elements are such that $|A_{ii}| < 1$, then $\mathbf{A}^n \to \mathbf{0}$ as $n \to \infty$.
e) If all diagonal elements are equal to zero then $\mathbf{A}^N = \mathbf{0}$. $\diamond$

Property (d) above finds application in stability analysis (Chap. 13). Property (e) is useful in the study of FIR systems. The above results hold if "lower triangular" is replaced with "upper triangular" everywhere.

A.8 MAXIMIZATION AND MINIMIZATION

Suppose $\mathbf{A}$ is Hermitian. Consider the quadratic form $\mathbf{v}^{\dagger}\mathbf{A}\mathbf{v}$. If we constrain $\mathbf{v}$ to be a unit norm vector, then this quadratic form cannot take arbitrarily large or small values. The extreme values are determined by the eigenvalues of $\mathbf{A}$ as summarized in the following result.

♠**Fact A.8.1. Rayleigh's principle.** Let $\mathbf{A}$ be $N \times N$ Hermitian. We know that the eigenvalues are real. Let λ_{min} and λ_{max} be the smallest and largest eigenvalues. Then the maximum value of $\mathbf{v}^{\dagger}\mathbf{A}\mathbf{v}$, as we vary $\mathbf{v}$ over all unit norm vectors, is equal to λ_{max} and occurs if and only if $\mathbf{v}$ is an eigenvector corresponding to λ_{max}. Similarly the minimum value of $\mathbf{v}^\dagger \mathbf{A}\mathbf{v}$ over unit norm $\mathbf{v}$ is equal to λ_{min} and occurs if, and only if, $\mathbf{v}$ is an eigenvector corresponding to λ_{min}. $\diamond$

Note that λ_{max} may have multiplicity > 1, in which case the eigenvector which maximizes the quadratic form is any vector from the corresponding eigenspace.

The "power method" for computing λ_{max} and its eigenvector

Let $\mathbf{A}$ be Hermitian positive semidefinite. So $\lambda_{min} \geq 0$ and $\lambda_{max} \geq 0$. Suppose we perform the following iteration

$$\mathbf{v}_k = \mathbf{A}\mathbf{v}_{k-1}, \tag{A.8.1}$$

with the initial vector $\mathbf{v}_0$ chosen arbitrarily. Unless $\mathbf{v}_0$ is orthogonal to every vector in the eigenspace of λ_{max}, this iteration eventually converges to an eigenvector corresponding to λ_{max}. This technique is called the *power method.* Once an eigenvector $\mathbf{v}$ is so computed, we compute λ_{max} from $\mathbf{A}\mathbf{v} = \lambda_{max}\mathbf{v}$.

If we are interested in computing λ_{min} and a corresponding eigenvector, there are several tricks we can use. If $\mathbf{A}$ is nonsingular, we can invert it (so that $1/\lambda_{min}$ is its largest eigenvalue. If we wish to avoid inversion (which is time consuming) we can first compute λ_{max} and define $\mathbf{B} = \lambda_{max}\mathbf{I} - \mathbf{A}$. This is positive semidefinite with largest eigenvalue $\lambda_{max} - \lambda_{min}$ which can be computed by the power method. So λ_{min} can be found.

A.9 PROPERTIES PRESERVED IN MATRIX PRODUCTS

Let $\mathbf{A}$ and $\mathbf{B}$ be $N \times N$ matrices satisfying a given property. It is often important to know whether the product $\mathbf{C} = \mathbf{A}\mathbf{B}$ satisfies the same property. For example, the product of two unitary matrices is unitary, but the product of two Hermitian matrices is not necessarily Hermitian.

Properties preserved under matrix multiplication. (a) Unitariness, (b) circulant property, (c) nonsingularity, (d) lower (hence upper) triangular property.

Properties not necessarily preserved. (a) Hermitian property, (b) positive definiteness, (c) Toeplitz property, (d) Vandermonde property, (e) *normal* property, and (f) stability (i.e., all eigenvalues inside the unit circle). See Problem A.16.

PROBLEMS

A.1. We know that matrix products do not commute, that is, in general $\mathbf{AB} \neq \mathbf{BA}$.

a) Demonstrate this with an example when $\mathbf{A}$ and $\mathbf{B}$ are (i) 2×2 and (ii) 3×3.

b) Find examples of 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA}$. (To avoid trivial answers, make sure the matrices are non diagonal.)

A.2. Which of the following matrices is diagonalizable?

$$\begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}, \begin{bmatrix} 0 & j \\ -j & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

A.3. Consider the Toeplitz matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

a) Compute the quantity $\mathbf{A}^2$ and verify that it is not Toeplitz. This shows that the product of Toeplitz matrices may not be Toeplitz.

b) Compute the determinant and verify that this is nonsingular. Find the inverse.

A.4. Verify Cayley-Hamilton theorem (A.5.4), for $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

A.5. Let $\mathbf{A}$ and $\mathbf{B}$ be $N \times N$ lower triangular matrices. Thus, $A_{ij} = 0$ for $j > i$ and $B_{jk} = 0$ for $k > j$. Using these prove that $\mathbf{AB}$ is lower triangular.

A.6. Check whether each of the following matrices has any of these properties: (a) Hermitian, (b) positive definiteness, (c) unitariness, and (d) normal property.

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ -2 & 2 \end{bmatrix}, \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}.$$

A.7. Find the eigenvalues and eigenvectors of the matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Is the matrix positive definite?

A.8. Evaluate all the leading principal minors of the matrices

$$\begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Which of these matrices are positive definite?

A.9. Let $\mathbf{A} = \mathbf{v}\mathbf{v}^\dagger$ where $\mathbf{v}$ is an $N \times 1$ matrix, i.e., a column vector.

a) Show that $\mathbf{v}^\dagger\mathbf{v}$ is an eigenvalue of $\mathbf{A}$, with corresponding eigenvector $\mathbf{v}$.

b) Show that the remaining $N-1$ eigenvalues are equal to zero. Find a set of N independent eigenvectors for $\mathbf{A}$.

A.10. Let $\mathbf{H}$ be Hermitian and $\mathbf{S}$ skew-Hermitian. Show that $\mathbf{v}^\dagger\mathbf{H}\mathbf{v}$ is real and $\mathbf{v}^\dagger\mathbf{S}\mathbf{v}$ imaginary for any choice of $\mathbf{v}$.

A.11.

a) It is possible for a nonzero matrix $\mathbf{A}$ to be such that $\mathbf{v}^T\mathbf{A}\mathbf{v} = 0$ for all $\mathbf{v}$. Show that $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is such an example.

b) Next show that if $\mathbf{v}^\dagger\mathbf{A}\mathbf{v} = 0$ for all vectors $\mathbf{v}$, then $\mathbf{A} = \mathbf{0}$. (*Hint:* Write $\mathbf{A} = \mathbf{A}_h + \mathbf{A}_s$ where $\mathbf{A}_h$ is Hermitian and $\mathbf{A}_s$ is skew Hermitian.)

A.12. Show that a lower triangular matrix cannot be unitary unless it is diagonal.

A.13. Let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$.

a) Compute $\mathbf{AB}$ and $\mathbf{BA}$ and verify that they do not have the same rank.
b) Compute $\mathbf{I}-\mathbf{AB}$ and $\mathbf{I}-\mathbf{BA}$ and verify that these have the same rank.
c) (This is tricky.) More generally, let $\mathbf{A}$ and $\mathbf{B}$ be arbitrary square matrices of the same size. Show that $\mathbf{I}-\mathbf{AB}$ and $\mathbf{I}-\mathbf{BA}$ have the same rank.

A.14. Find examples of Hermitian positive definite matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{P}-\mathbf{Q}$ is indefinite. (Avoid trivial answers by finding *non diagonal* examples!)

A.15. The product of two right-circulant matrices is right-circulant. Prove this for the 3×3 case.

A.16. Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices with a certain property in common (e.g., unitary, circulant, and so on.) If the product $\mathbf{AB}$ also has this property, we say that the property is preserved under multiplication. Prove by examples that the following properties are not necessarily preserved under multiplication: (i) Hermitian, (ii) Vandermonde, (iii) normal property, and (iv) stability (i.e., all eigenvalues of $\mathbf{A}$ have magnitude less than unity).

A.17. Let $\mathbf{U}$ be a square matrix, and let $\mathbf{y} = \mathbf{Ux}$. Let $\mathbf{U}$ be such that $\mathbf{y}^\dagger\mathbf{y} = \mathbf{x}^\dagger\mathbf{x}$ for all vectors $\mathbf{x}$. Show that $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$.

A.18. Prove the following matrix identity

$$\underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}}_{\mathbf{R}} = \begin{bmatrix} \mathbf{I} & \mathbf{BD}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}-\mathbf{BD}^{-1}\mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I} \end{bmatrix},$$

where $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ are matrices of appropriate dimensions, $\mathbf{A}$ and $\mathbf{D}$ are square (hence $\mathbf{R}$ is square), and $\mathbf{D}$ is nonsingular. Hence prove that

$$\det \mathbf{R} = [\det \mathbf{D}] \times \det [\mathbf{A}-\mathbf{BD}^{-1}\mathbf{C}].$$

A.19. Let $\mathbf{P}$ be $N \times N$ Hermitian positive definite. Partition it as

$$\mathbf{P} = \begin{bmatrix} P_{00} & \mathbf{p}_{10}^\dagger \\ \mathbf{p}_{10} & \mathbf{P}_{11} \end{bmatrix}. \qquad (PA.19a)$$

Here P_{00} is scalar, whereas $\mathbf{P}_{11}$ is $(N-1) \times (N-1)$. Evidently $\mathbf{p}_{10}$ is a column vector.

a) Using the definition of positive definiteness, show that P_{00} is real and positive. Also show that $\mathbf{P}_{11}$ is Hermitian positive definite.

b) Using the previous problem, show that

$$\det \mathbf{P} = \left(P_{00} - \mathbf{p}_{10}^{\dagger}\mathbf{P}_{11}^{-1}\mathbf{p}_{10}\right)\det \mathbf{P}_{11}. \qquad (PA.19b)$$

c) Using some or all of the results proved above, show that

$$\det \mathbf{P} \leq P_{00}[\det \mathbf{P}_{11}], \qquad (PA.19c)$$

with equality if and only if $\mathbf{p}_{10} = \mathbf{0}$.

d) Let P_{ii} denote the diagonal elements of $\mathbf{P}$. By repeated application of the above result, prove that

$$\det \mathbf{P} \leq \prod_{i=0}^{N-1} P_{ii}, \qquad (PA.19d)$$

with equality if and only if $\mathbf{P}$ is diagonal.

A.20. Let $\mathbf{P}$ be as in Problem A.19. We now provide a second proof of the inequality in (PA.19d).

a) Prove that $\mathbf{P}$ can be written as $\mathbf{P} = \mathbf{D}\mathbf{Q}\mathbf{D}^{\dagger}$ where $\mathbf{D}$ is a diagonal matrix of positive elements, and $Q_{ii} = 1$ for all i.

b) Hence show that $[\det \mathbf{P}] = [\det \mathbf{Q}] \times [\prod_{i=0}^{N-1} P_{ii}]$.

c) Now show that $\det \mathbf{Q} \leq 1$, with equality if and only if $\mathbf{Q} = \mathbf{I}$. This establishes the desired result (PA.19d). [*Hint.* The determinant is the product of eigenvalues and the trace is the sum of eigenvalues. Use this along with the arithmetic/geometric mean inequality, which can be found in Appendix C; see discussion around Eq. (C.2.3).]