5

Linear Spaces

5.1 Introduction

Linear spaces are among the most important and widely used mathematical structures. Linear spaces consist of elements called *vectors* and are associated with a field (in most cases, the real field $\mathbb{R}$ or the complex field $\mathbb{C}$). The elements of this field are referred to as *scalars*. Two fundamental operations: vector addition and multiplication with scalars are defined such that certain axioms given are satisfied.

Linear spaces were introduced in their modern form by the Italian mathematician G. Peano in the second part of the 19^{th} century; precursor ideas can be traced to more than two centuries before in connection with analytic geometry problems.

Definition 5.1. *Let L be a nonempty set and let $\mathcal{F} = (F, \{0, +, -, \cdot, \})$ be a field whose carrier is a set F. An F-*linear space *is a triple $(L, +, \cdot)$ such that $(L, \{\boldsymbol{0}, +, -\})$ is an Abelian group and $\cdot : F \times L \longrightarrow L$ is an operation such that the following conditions are satisfied*

(i) $a \cdot (b \cdot \boldsymbol{x}) = (a \cdot b) \cdot \boldsymbol{x}$,
(ii) $1 \cdot \boldsymbol{x} = \boldsymbol{x}$,
(iii) $a \cdot (\boldsymbol{x} + \boldsymbol{y}) = a \cdot \boldsymbol{x} + a \cdot \boldsymbol{y}$, *and*
(iv) $(a + b) \cdot \boldsymbol{x} = a \cdot \boldsymbol{x} + b \cdot \boldsymbol{x}$

for every $a, b \in F$ and $\boldsymbol{x}, \boldsymbol{y} \in L$.

If $\mathcal{F}$ is the field of real numbers (the field of complex numbers), then we will refer to any F-linear space as a *real linear space* (*complex linear space*).

The commutative binary operation of L is denoted by the same symbol "+" as the corresponding operation of the field F. The multiplication by a scalar, $\cdot : F \times L \longrightarrow L$ is also referred to as an *external operation* since its two arguments belong to two different sets, F and L. Again, this operation is denoted by the same symbol used for denoting the multiplication on F; if there is no risk of confusion, we shall write $a\mathbf{x}$ instead of $a \cdot \mathbf{x}$.

The elements of the set L will be denoted using bold letters $\mathbf{x}, \mathbf{y}, \mathbf{z}$, etc. The members of the field will be denoted by small letters from the beginning of the alphabet.

The additive element $\mathbf{0}$ is a special element called the *zero element*; every F-linear space must contain at least this element.

Example 5.2. The set $\mathbb{R}^n$ of n-tuples of real numbers is a real linear space under the definitions

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \text{ and } a \cdot \mathbf{x} = \begin{pmatrix} a \cdot x_1 \\ \vdots \\ a \cdot x_n \end{pmatrix}$$

of the operations $+$ and $\cdot$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

In this linear space, the zero of the Abelian group is the n-tuple

$$\mathbf{0}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Similarly, the set $\mathbb{C}^n$ of n-tuples of complex numbers is a complex linear space under the same formal definitions of vector sum and scalar multiplication as $\mathbb{R}^n$, where $a \in \mathbb{C}$ in this case.

Example 5.3. The set of infinite sequences of complex numbers $\mathbf{Seq}_\infty(\mathbb{C})$ can be organized as a linear space by defining the addition of two sequences

$$\mathbf{x} = (x_0, x_1, \ldots) \text{ and } \mathbf{y} = (y_0, y_1, \ldots)$$

as $\mathbf{x} + \mathbf{y} = (x_0 + y_0, x_1 + y_1, \ldots)$, and the multiplication by $c\mathbf{x}$ as $c\mathbf{x} = (cx_0, cx_1, \ldots)$ for $c \in \mathbb{C}$.

Example 5.4. The set of complex-valued functions defined on a set S is a complex linear space. The addition of functions is given by $(f + g)(s) = f(s) + g(s)$, and the multiplication of a function with a complex number is defined by $(af)(s) = af(s)$ for $s \in S$ and $a \in \mathbb{C}$.

Example 5.5. Let C be the set of real-valued continuous functions defined on $\mathbb{R}$,

$$C = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f \text{ is continuous}\}.$$

Define $f + g$ by $(f + g)(x) = f(x) + g(x)$ and $(a \cdot f)(x) = a \cdot f(x)$ for $x \in \mathbb{R}$.

The triple $(C, +, \cdot)$ is a real linear space.

Definition 5.6. *Let L be an F-linear space. A* linear combination *of K (where $K \subseteq L$) is a member of L of the form $c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$, where $c_1, \ldots, c_n \in F$.*

A subset $K = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ of L is linearly independent *if $c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n = \boldsymbol{0}$ implies $c_1 = \cdots = c_n = 0$. If K is not linearly independent, we refer to K as a* linearly dependent *set.*

If $\mathbf{x} \neq \mathbf{0}$, then the set $\{\mathbf{x}\}$ is linearly independent. Of course, the set $\{\mathbf{0}\}$ is not linearly independent because $1\mathbf{0} = \mathbf{0}$. If K is a linearly independent subset of a linear space, then any subset of K is linearly independent.

Example 5.7. Let

$$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

be a vector that has a unique nonzero component equal to 1 in place i, where $1 \leqslant i \leqslant n$. The set $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is linearly independent. Indeed, suppose that $c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n = \mathbf{0}$. This is equivalent to

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

that is, with $c_1 = \cdots = c_n = 0$. Thus, E is linearly independent.

Theorem 5.8. *Let L be an F-linear space. A subset K of L is linearly independent if and only if for every $\boldsymbol{x} \in L$ there exists a linear combination of K, $\boldsymbol{x} = \sum_i c_i\boldsymbol{x}_i$ such that the coefficients c_i are uniquely determined.*

Proof. Suppose that $\mathbf{x} = c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n = c'_1\mathbf{x}_1 + \cdots + c'_n\mathbf{x}_n$ and there exists i such that $c_i \neq c'_i$. This implies $\sum_{i=1}^{n}(c_i - c'_i)\mathbf{x}_i = \mathbf{0}$, which contradicts the linear independence of K.

Definition 5.9. *A subset S of a linear space $(L, +, \cdot)$* spans *the space L (or S generates the linear space) if every $\boldsymbol{x} \in L$ is a linear combination of S.*

A basis *of the linear space $(L, +, \cdot)$ is a linearly independent subset that spans the linear space.*

In view of Theorem 5.8, a set B is a basis if every $\mathbf{x} \in L$ can be written uniquely as a linear combination of elements of B.

Definition 5.10. *A* subspace *of a F-linear space $(L, +, \cdot)$ is a nonempty subset U of L such that $\boldsymbol{x}, \boldsymbol{y} \in U$ implies $\boldsymbol{x} + \boldsymbol{y} \in U$ and $a \cdot \boldsymbol{x} \in U$ for every $a \in F$.*

If U is a subspace of an F-linear space, then U can be regarded as an F-linear space and various notions introduced for linear spaces are applicable to U.

The set $\{\mathbf{0}\}$ is a subspace of any F-linear space $(L, +, \cdot)$ included in every subspace of L.

Example 5.11. We saw in Example 5.5 that the set of real-valued continuous functions C defined on $\mathbb{R}$ is a real linear space.

The set of even real-valued continuous functions defined on $\mathbb{R}$ given by $E = \{f \in C \mid f(x) = f(-x)\}$ is a subspace of C. Indeed, note that for $f, g \in E$ we have

$$(f + g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f + g)(x)$$

and $(af)(-x) = af(-x) = af(x) = (af)(x)$, so $f, g \in E$ implies $f + g \in E$ and $af \in E$. Similarly, it is possible to show that the set of odd real-valued continuous functions defined on $\mathbb{R}$ $D = \{f \in C \mid f(-x) = -f(x)\}$ is a subspace of C.

Example 5.12. In Example 5.2 we saw that $\mathbb{R}^n$ can be regarded as an $\mathbb{R}$-linear space, while $\mathbb{C}^n$ is a $\mathbb{C}$-linear space. Even though $\mathbb{R}^n \subseteq \mathbb{C}^n$, $\mathbb{R}^n$ is not a subspace of $\mathbb{C}^n$ because for $a \in \mathbb{C}$ and $\mathbf{x} \in \mathbb{R}^n$ we do not have $a\mathbf{x} \in \mathbb{R}^n$.

If $\{K_i \mid i \in I\}$ is a nonempty collection of subspaces of a linear space, then $\bigcap\{K_i \mid i \in I\}$ is also a linear subspace. Thus, the family of subspaces of a linear space is a closure system. If U is a subset of a linear space $(L, +, \cdot)$ and $\mathbf{K}$ is the corresponding closure operator for the closure system of linear spaces, then we say that U is *spanning* the subspace $\mathbf{K}(U)$ of L. The subspace $\mathbf{K}(U)$ is said to be spanned by U. We denote this subspace by $\langle U \rangle$.

Theorem 5.13. *Let $\mathcal{L} = (L, +, \cdot)$ be an F-linear space. The following statements are equivalent:*

(i) *The finite set $K = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is spanning the linear space $(L, +, \cdot)$ and K is minimal with this property.*
(ii) *K is a finite basis for $(L, +, \cdot)$.*
(iii) *The finite set K is linearly independent, and K is maximal with this property.*

Proof. (i) implies (ii): We need to prove that K is linearly independent. Suppose that this is not the case. Then, there exist $c_1, \ldots, c_n \in F$ such that $c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n = \mathbf{0}$ and at least one of $c_1, \ldots, c_n$, say c_i, is nonzero. Then,

$\mathbf{x}_i = -\frac{c_1}{c_i}\mathbf{x}_1 - \cdots - \frac{c_n}{c_i}\mathbf{x}_n$, and this implies that $K - \{\mathbf{x}_i\}$ also spans the linear space, thus contradicting the minimality of K.

(ii) implies (i): Let K be a finite basis. Suppose that K' is a proper subset of K that spans L. Then, if $\mathbf{z} \in K - K'$, z' is a linear combination of elements of K', which contradicts the fact that K is a basis.

We leave to the reader the proof of the equivalence between (ii) and (iii).

Corollary 5.14. *Every linear space that is spanned by a finite subset has a finite basis. Further, if B is a finite basis for an F-linear space $(L, +, \cdot)$, then each finite subset U of L such that $|U| = |B| + 1$ is linearly dependent.*

Proof. This statement follows directly from Theorem 5.13.

Corollary 5.15. *If B and B' are two finite bases for a linear space $(L, +, \cdot)$, then $|B| = |B'|$.*

Proof. If B is a finite basis, then $|B|$ is the maximum number of linearly independent elements in L. Thus, $|B'| \leqslant |B|$. Reversing the roles of B and B', we obtain $|B| \leqslant |B'|$, so $|B| = |B'|$.

Thus, the number of elements of a finite basis of L is a characteristic of L and does not depend on any particular basis.

Definition 5.16. *A linear space $(L, +, \cdot)$ is* n-dimensional *if there exists a basis of L such that $|B| = n$. The number n is the* dimension *of L and is denoted by* $\dim(L)$.

Theorem 5.17. *Let L be a finite-dimensional F-linear space and let $U = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be a linearly independent subset of L. There exists an extension of U that is a basis of L.*

Proof. If $\langle U \rangle = L$, then U is a basis of L. If this is not the case, let $\mathbf{w}_1 \in L - \langle U \rangle$. The set $U \cup \{\mathbf{w}_1\}$ is linearly independent and we have the strict inclusion $\langle U \rangle \subset \langle U \cup \{\mathbf{w}_1\} \rangle$. The subspace $\langle U \cup \{\mathbf{w}\} \rangle$ is $(k+1)$-dimensional. This argument can be repeated no more than $n - k$ times, where $n = \dim(L)$. Thus, $U \cup \{\mathbf{w}_1, \ldots, \mathbf{w}_{n-k}\}$ is a basis for L that extends U.

Definition 5.18. *Let L be an F-linear space and let U, V be subspaces of L. The* sum *of the subspaces U and V is the set $U + V$ defined by*

$$U + V = \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in U \text{ and } \boldsymbol{v} \in V\}.$$

It is easy to verify that $U + V$ is also a subspace of L.

Theorem 5.19. *Let U, V be two subspaces of the finite-dimensional F-linear space L. We have* $\dim(U + V) + \dim(U \cap V) = \dim(U) + \dim(V)$.

Proof. Suppose that $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ is a basis for $U \cap V$, where $k = \dim(U \cap V)$. This basis can be extended to a basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k, u_{k+1}, \ldots, u_p\}$ for U and to a basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k, v_{k+1}, \ldots, v_q\}$ for V.

Define $B = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, u_{k+1}, \ldots, u_p, v_{k+1}, \ldots, v_q\}$. It is clear that $\langle B \rangle = U + V$. Suppose that there exist $c_1, \ldots, c_{p+q-k}$ such that

$$c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k + c_{k+1}u_{k+1} + \cdots + c_pu_p + c_{p+1}v_{k+1} + \cdots + c_{p+q-k}v_q = \mathbf{0}.$$

The last equality implies

$$c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k + c_{k+1}u_{k+1} + \cdots + c_pu_p = -c_{p+1}v_{k+1} - \cdots - c_{p+q-k}v_q.$$

Therefore, $c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k + c_{k+1}u_{k+1} + \cdots + c_pu_p$ belongs to $U \cap V$, which implies $c_{k+1} = \cdots = c_p = 0$. Since

$$c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k + c_{p+1}v_{k+1} + \cdots + c_{p+q-k}v_q = \mathbf{0},$$

and $\{\mathbf{w}_1, \ldots, \mathbf{w}_k, v_{k+1}, \ldots, v_q\}$ is a basis for V, it follows that $c_1 = \cdots = c_k = c_{p+1} = \cdots = c_{p+q-k} = 0$.

This allows to conclude that $\dim(U + V) = p + q - k$ and this implies the equality of the theorem.

5.2 Linear Mappings

Linear mappings between linear spaces are functions that are compatible with the algebraic operations of linear spaces.

Definition 5.20. *Let L and K be two F-linear spaces. A* linear mapping *is a function $h : L \longrightarrow K$ such that $h(a\boldsymbol{x} + b\boldsymbol{y}) = ah(\boldsymbol{x}) + bh(\boldsymbol{y})$ for every $a, b \in F$ and $\boldsymbol{x}, \boldsymbol{y} \in L$.*

An affine mapping *is a function $f : L \longrightarrow K$ such that there exists a linear mapping $h : L \longrightarrow K$ and $\boldsymbol{b} \in K$ such that $f(\boldsymbol{x}) = h(\boldsymbol{x}) + \boldsymbol{b}$ for $\boldsymbol{x} \in L$.*

Linear mappings are also referred to as *homomorphisms*, as *morphisms*, or *linear operators* and it is the latter term that we usually use. The set of morphisms between two F-linear spaces L and K is denoted by $\mathsf{Hom}(L, K)$. The set of affine mappings between two F-linear spaces L and K is denoted by $\mathsf{Aff}(L, K)$.

A linear mapping $h : L \longrightarrow K$ is a morphism between the Abelian additive groups of the linear spaces; therefore, $h(\mathbf{0}_L) = \mathbf{0}_K$ and $h(-\mathbf{x}) = -h(\mathbf{x})$ for $\mathbf{x} \in L$.

Theorem 5.21. *Let L and K be two F-linear spaces having $\boldsymbol{0}_L$ and $\boldsymbol{0}_K$ as their zero elements, respectively. A morphism $h \in \mathsf{Hom}(L, K)$ is injective if and only if $h(\boldsymbol{x}) = \boldsymbol{0}_K$ implies $\boldsymbol{x} = \boldsymbol{0}_L$.*

Proof. Let h be a morphism such that $h(\mathbf{x}) = \mathbf{0}_K$ implies $\mathbf{x} = \mathbf{0}_L$. If $h(\mathbf{x}) = h(\mathbf{y})$, by the linearity of h we have $h(\mathbf{x}-\mathbf{y}) = \mathbf{0}_K$, which implies $\mathbf{x}-\mathbf{y} = \mathbf{0}_L$, that is, $\mathbf{x} = \mathbf{y}$. Thus, h is injective.

Conversely, suppose that h is injective. If $\mathbf{x} \neq \mathbf{0}_L$, then $h(\mathbf{x}) \neq h(\mathbf{0}_L) = \mathbf{0}_K$. Thus, $h(\mathbf{x}) = \mathbf{0}_K$ implies $\mathbf{x} = \mathbf{0}_L$.

An *endomorphism* of an F-linear space L is a morphism $h : L \longrightarrow L$. The set of endomorphisms of L is denoted by $\mathsf{Endo}(L)$. Often, we refer to endomorphisms of L as *linear operators* on L.

The term *linear form* is reserved for linear mappings between an F-linear space and the field F itself, where F is considered as an F-linear space.

Example 5.22. For $a \in \mathbb{R}$ define the mapping $\mathsf{h}_a : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by $\mathsf{h}_a(\mathbf{x}) = a\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$. It is easy to verify that h_a is a linear operator on $\mathbb{R}^n$. This mapping is known as a *homotety on* $\mathbb{R}^n$.

If $a = 1$, then h_1 is given by $\mathsf{h}_1(\mathbf{x}) = \mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$; this is the identity morphism of $\mathbb{R}^n$, which is usually denoted by $1_{\mathbb{R}^n}$.

For $a = 0$ we obtain the *zero morphism of* $\mathbb{R}^n$ given by $\mathsf{h}_0(\mathbf{x}) = \mathbf{0}$ for $\mathbf{x} \in \mathbb{R}^n$.

Example 5.23. The *translation* generated by $\mathbf{z} \in \mathbb{R}^n$ is the mapping $\mathsf{t}_{\mathbf{z}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ defined by $\mathsf{t}_{\mathbf{z}}(\mathbf{x}) = \mathbf{x} + \mathbf{z}$ is a bijection but not a morphism unless $\mathbf{z} = \mathbf{0}$. Its inverse is $\mathsf{t}_{-\mathbf{z}}$.

Definition 5.24. *Let U and V be two subsets of $\mathbb{R}^n$. We define the subset $U+V$ of $\mathbb{R}^n$ as*

$$U + V = \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \in U \text{ and } \boldsymbol{v} \in V\}.$$

For $a \in \mathbb{R}$, the set aU is

$$aU = \{a\boldsymbol{u} \mid \boldsymbol{u} \in U\}.$$

If L, K are two linear spaces, then the set $\mathsf{Hom}(L, K)$ is never empty because the zero morphism $h_0 : L \longrightarrow K$ given by $h_0(\mathbf{x}) = \mathbf{0}_L$ for $\mathbf{x} \in K$ is always an element of $\mathsf{Hom}(L, K)$.

Definition 5.25. *Let L and K be two F-linear spaces. If $f, g \in \mathsf{Hom}(L, K)$, the* sum *$f + g$ is defined by $(f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$ for $\boldsymbol{x} \in L$.*

The sum of two linear mappings is also a linear mapping because

$$\begin{aligned}(f+g)(a\mathbf{x} + b\mathbf{y}) &= f(a\mathbf{x} + b\mathbf{y}) + g(a\mathbf{x} + b\mathbf{y})\\ &= af(\mathbf{x}) + bf(\mathbf{y}) + ag(\mathbf{x}) + bg(\mathbf{y})\\ &= f(a\mathbf{x} + b\mathbf{y}) + g(a\mathbf{x} + b\mathbf{y}),\end{aligned}$$

for all $a, b \in F$ and $\mathbf{x}, \mathbf{y} \in L$.

Theorem 5.26. *Let M, P, Q be three F-linear spaces. The following properties of compositions of linear mappings hold:*
(i) *If $f \in \mathsf{Hom}(M, P)$ and $g \in \mathsf{Hom}(P, Q)$, then $gf \in \mathsf{Hom}(M, Q)$.*
(ii) *If $f \in \mathsf{Hom}(M, P)$ and $g_0, g_1 \in \mathsf{Hom}(P, Q)$, then*

$$f(g_0 + g_1) = fg_0 + fg_1.$$

(iii) *If $f_0, f_1 \in \mathsf{Hom}(M, P)$ and $g \in \mathsf{Hom}(P, Q)$, then*

$$(f_0 + f_1)g = f_0 g + f_1 g.$$

Proof. We prove only the second part of the theorem and leave the proofs of the remaining parts to the reader.

Let $\mathbf{x} \in M$. Then, $f(g_0 + g_1)(\mathbf{x}) = f((g_0 + g_1)(\mathbf{x})) = f(g_0(\mathbf{x}) + g_1(\mathbf{x})) = f(g_0(\mathbf{x})) + f(g_1(\mathbf{x}))$ for $\mathbf{x} \in M$, which yields the desired equality.

We leave to the reader to verify that for any F-linear spaces M and P the algebra $(\mathsf{Hom}(M, P), \{h_0, +, -\})$ is an Abelian group that has the zero morphism h_0 as its zero-ary operations and the addition of linear mappings as its binary operation; the opposite of a linear mapping h is the mapping $-h$.

Moreover, $(\mathsf{Endo}(M), \{h_0, 1_M, +, -, \cdot\})$ is a unitary ring, where the multiplication is defined as the composition of linear mappings.

If M and P are F-linear spaces, $\mathsf{Hom}(M, P)$ is itself an linear space, where the multiplication of a morphism h by a scalar c is the morphism ch defined by $(ch)(\mathbf{x}) = c \cdot h(\mathbf{x})$. Indeed, the mapping ch is linear because

$$\begin{aligned}(ch)(a\mathbf{x} + b\mathbf{y}) &= c(ah(\mathbf{x}) + bh(\mathbf{y})) = cah(\mathbf{x}) + cbh(\mathbf{y})) \\ &= ach(\mathbf{x}) + bch(\mathbf{y})) = a(ch)(\mathbf{x}) + b(ch)(\mathbf{y})),\end{aligned}$$

for every $a, b, c \in F$ and $\mathbf{x}, \mathbf{y} \in M$.

Definition 5.27. *Let h be an endomorphism of a linear space M. The m^{th}* iteration *of h (for $m \in \mathbb{N}$) is defined as*
(i) $h^0 = 1_M$;
(ii) $h^{m+1}(\boldsymbol{x}) = h(h^m(\boldsymbol{x}))$ *for $m \in \mathbb{N}$.*

For every $m \geqslant 1$, h^m is an endomorphism of M; this can be shown by a straightforward proof by induction on m.

Theorem 5.28. *Let L, M be two F-linear spaces and let $h : L \longrightarrow M$ be a morphism. Then, the sets*

$$\begin{aligned}\mathit{Ker}(h) &= \{\boldsymbol{x} \in L \mid h(\boldsymbol{x}) = \boldsymbol{0}_M\}, \\ \mathit{Img}(h) &= \{\boldsymbol{y} \in M \mid \boldsymbol{y} = h(\boldsymbol{x}) \text{ for some } \boldsymbol{x} \in L\}\end{aligned}$$

are subspaces of L and M, respectively.

Proof. Let $\mathbf{u}, \mathbf{v} \in \mathsf{Ker}(h)$. Since $h(\mathbf{u}) = h(\mathbf{v}) = \mathbf{0}_M$ it follows that

$$h(a\mathbf{u} + b\mathbf{v}) = ah(\mathbf{u}) + bh(\mathbf{v}) = \mathbf{0}_M,$$

for $a, b \in F$, so $a\mathbf{u} + b\mathbf{v} \in \mathsf{Ker}(h)$. This shows that $\mathsf{Ker}(h)$ is indeed a subspace of L.

Let now $\mathbf{s}, \mathbf{t} \in \mathsf{Img}(h)$. There exist $\mathbf{x}, \mathbf{y} \in L$ such that $\mathbf{s} = h(\mathbf{x})$ and $\mathbf{t} = h(\mathbf{y})$. Therefore, $a\mathbf{s} + b\mathbf{t} = ah(\mathbf{x}) + bh(\mathbf{y}) = h(a\mathbf{x} + b\mathbf{y})$, hence $a\mathbf{s} + b\mathbf{t} \in \mathsf{Img}(h)$. This implies that $\mathsf{Img}(h)$ is a subspace of M.

Definition 5.29. *Let L be a F-linear space and let $M_1, \ldots, M_p$ be p linear subspaces of L. L is the* direct sum *of $M_1, \ldots, M_p$ if for every $\boldsymbol{x}$ of L there exists a unique sequence $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_p)$ such that $\boldsymbol{y}_i \in M_i$ for $1 \leqslant i \leqslant p$ and $\boldsymbol{x} = \boldsymbol{y}_1 + \cdots + \boldsymbol{y}_p$. This is denoted by $L = M_1 \boxplus M_2 \boxplus \cdots \boxplus M_p$.*

Observe that if $L = M_1 \boxplus M_2 \boxplus \cdots \boxplus M_p$, then the function $h_i : L \longrightarrow M_i$ given by $h_i(\mathbf{x}) = \mathbf{y}_i$ is well-defined due to the uniqueness of the sequence $(\mathbf{y}_1, \ldots, \mathbf{y}_p)$ for $1 \leqslant i \leqslant p$.

It is easy to verify that each h_i is a linear mapping. Indeed, if $a, b \in F$ and $\mathbf{u}, \mathbf{v} \in L$ can be uniquely written as $\mathbf{u} = \mathbf{y}_1 + \cdots + \mathbf{y}_p$ and $\mathbf{v} = \mathbf{z}_1 + \cdots + \mathbf{z}_p$, where $\mathbf{y}_i, \mathbf{z}_i \in M_i$ for $1 \leqslant i \leqslant p$, then $a\mathbf{u} + b\mathbf{v} = (a\mathbf{y}_1 + b\mathbf{x}_1) + \cdots + (a\mathbf{y}_p + b\mathbf{z}_p)$. Since each M_i is a subspace, $a\mathbf{y}_i + b\mathbf{z}_i \in M_i$ for $1 \leqslant i \leqslant p$ and the uniqueness of the decomposition implies that $h_i(a\mathbf{u} + b\mathbf{v}) = ah_i(\mathbf{u}) + bh_i(\mathbf{v})$.

Each morphism h_i is idempotent, that is, $h_i(h_i(\mathbf{x})) = h_i(\mathbf{x})$ for $\mathbf{x} \in L$. Indeed, $h_i(\mathbf{x}) \in M$ and applying the uniqueness of the decomposition to $h_i(\mathbf{x})$ we obtain $h_i(h_i(\mathbf{x})) = h_i(\mathbf{x})$.

Theorem 5.30. *Let L be a linear space that is the direct sum $L = M_1 \boxplus M_2 \boxplus \cdots \boxplus M_p$. Then $M_i \cap M_j = \{\boldsymbol{0}\}$ for every i, j such that $i \neq j$ and $1 \leqslant i, j \leqslant p$.*

Proof. Suppose that $\mathbf{t} \neq \mathbf{0}$ belongs to $M_i \cap M_j$ for some i, j such that $i \neq j$ and $1 \leqslant i, j \leqslant p$. Let $\mathbf{x} = \mathbf{y}_1 + \cdots + \mathbf{y}_p$ be the unique decomposition of $\mathbf{x} \in L$ as a sum of vectors from the subspaces M_i. Since $\mathbf{t} \neq 0$ and $\mathbf{t} \in M_i \cap M_j$, we would have the distinct decompositions

$$\begin{aligned}
\mathbf{x} &= \mathbf{y}_1 + \cdots + (\mathbf{y}_i + \mathbf{t}) + \cdots + (\mathbf{y}_j - \mathbf{t}) + \cdots + \mathbf{y}_p \\
\mathbf{x} &= \mathbf{y}_1 + \cdots + (\mathbf{y}_i - \mathbf{t}) + \cdots + (\mathbf{y}_j + \mathbf{t}) + \cdots + \mathbf{y}_p,
\end{aligned}$$

contradicting the uniqueness of the decomposition of $\mathbf{x}$.

Theorem 5.31. *Let L be a linear space. If $B_1, \ldots, B_p$ be bases in the subspaces $M_1, \ldots, M_p$, then $L = M_1 \boxplus M_2 \boxplus \cdots \boxplus M_p$ if and only if $B = \bigcup_{i=1}^{p}$ is a basis for L.*

Proof. Suppose that $L = M_1 \boxplus M_2 \boxplus \cdots \boxplus M_p$. Each $\mathbf{x} \in L$ can be uniquely written as a sum $\mathbf{x} = \mathbf{y}_1 + \cdots + \mathbf{y}_p$, where $\mathbf{y}_i \in M_i$ for $1 \leqslant i \leqslant p$. It is clear that B spans L, so we need to show only that B is linearly independent.

Let $B_i = \{\mathbf{y}^i_1, \ldots, \mathbf{y}^i_{k_i}\}$ for $1 \leqslant i \leqslant p$. Suppose that

$$\sum_{i=1}^{p} \sum_{j=1}^{k_i} c^i_j y^i_j = \mathbf{0}.$$

Since $\mathbf{0}$ can be regarded as a sum of p copies of itself (where each copy is in M_i for $1 \leqslant i \leqslant p$, we have $\sum_{j=1}^{k_i} c^i_j y^i_j = \mathbf{0}$, so $c^i_j = 0$ for $1 \leqslant j \leqslant k_i$ and for $1 \leqslant i \leqslant p$. Thus, B is linearly independent and, therefore, it is a basis.

We leave to the reader the proof of the reverse implication.

If U, V are subspaces of L such that $L = U \boxplus V$, then U, V are said to be *complementary subspaces* of L.

Theorem 5.32. *Let $h : L \longrightarrow L$ be an idempotent endomorphism of the F-linear space L. Then, $L = \mathsf{Ker}(h) \boxplus \mathsf{Img}(h)$.*

Proof. If $\mathbf{t} \in \mathsf{Ker}(h) \cap \mathsf{Img}(h)$, we have $h(\mathbf{t}) = \mathbf{0}$ and $t = h(\mathbf{z})$ for some $\mathbf{z} \in L$. Thus, $\mathbf{t} = h(\mathbf{z}) = h(h(\mathbf{z})) = h(\mathbf{t}) = \mathbf{0}$, which implies $\mathsf{Ker}(h) \cap \mathsf{Img}(h) = \{\mathbf{0}\}$.

Since $h(\mathbf{x}) = h(h(\mathbf{x}))$ for every $\mathbf{x} \in L$, it follows that $h(\mathbf{x} - h(\mathbf{x})) = \mathbf{0}$, so $\mathbf{y} = \mathbf{x} - h(\mathbf{x}) \in \mathsf{Ker}(h)$. This allows us to write $\mathbf{x} = \mathbf{y} + \mathbf{z}$, where $\mathbf{z} = h(\mathbf{x}) \in \mathsf{Img}(h)$, which shows that every element $\mathbf{x}$ of L can be written as a sum of an element in $\mathsf{Ker}(h)$ and an element in $\mathsf{Img}(h)$.

Suppose now that $\mathbf{x} = \tilde{\mathbf{y}} + \tilde{\mathbf{z}}$, where $\tilde{\mathbf{y}} \in \mathsf{Ker}(h)$ and $\tilde{\mathbf{z}} \in \mathsf{Img}(h)$. Since $\mathbf{y} - \tilde{\mathbf{y}} = \tilde{\mathbf{z}} - \mathbf{z} = \mathbf{0}$, it follows that the expression of $\mathbf{t}$ as a sum of two vectors in $\mathsf{Ker}(h)$ and $\mathsf{Img}(h)$ is unique, so $L = \mathsf{Ker}(h) \boxplus \mathsf{Img}(h)$.

Theorem 5.33. *Let U and V be two subspaces of an F-linear space L. If $L = U \boxplus V$, then there exists an idempotent endomorphism h of L such that $U = \mathsf{Ker}(h)$ and $V = \mathsf{Img}(h)$.*

Proof. Let $\mathbf{x}$ be a vector in L and let $\mathbf{x} = \mathbf{u} + \mathbf{v}$ be the decomposition of $\mathbf{x}$, where $\mathbf{u} \in U$ and $\mathbf{v} \in V$. Define the mapping $h : L \longrightarrow L$ as $h(\mathbf{x}) = \mathbf{v}$. The uniqueness of the decomposition of h implies that h is well-defined.

Note that a vector $\mathbf{u} \in U$ has the decomposition $\mathbf{u} = \mathbf{u} + \mathbf{0}$, so $h(\mathbf{u}) = \mathbf{0}$. Thus, $U = \mathsf{Ker}(h)$. Since $h(h(\mathbf{x})) = h(\mathbf{v}) = h(\mathbf{u} + \mathbf{v}) = h(\mathbf{x})$, it follows that h is idempotent.

If $\mathbf{w} \in W$, its decomposition is $\mathbf{w} = \mathbf{0} + \mathbf{w}$, so $h(\mathbf{w}) = \mathbf{w}$. Therefore, $W = \mathsf{Img}(h)$.

5.3 Matrices

Definition 5.34. *Let F be a field. A* matrix on F *is a function*

$$A : \{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow F.$$

The pair (m, n) is the format *of the matrix A.*

If $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow F$ is a matrix on F, we say that A is an $(m \times n)$-matrix on F. The set of all such matrices will be denoted by $F^{m\times n}$.

A matrix $A \in F^{m\times n}$ can be written as

$$\begin{pmatrix} A(1,1) & A(1,2) & \ldots & A(1,n) \\ A(2,1) & A(2,2) & \ldots & A(2,n) \\ \vdots & \vdots & \ldots & \vdots \\ A(m,1) & A(m,2) & \ldots & A(m,n) \end{pmatrix}.$$

Alternatively, a matrix $A \in F^{m\times n}$ can be regarded as consisting of m rows, where each row is a sequence of the form

$$(A(i,1), A(i,2), \ldots, A(i,n)),$$

for $1 \leqslant i \leqslant n$, or as a collection of n columns of the form

$$\begin{pmatrix} A(1,j) \\ A(2,j) \\ \vdots \\ A(m,j) \end{pmatrix},$$

where $1 \leqslant j \leqslant m$.

Example 5.35. Let $F = \{0, 1\}$. The matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

is a (3×2)-matrix on the set F.

The element $A(i, j)$ of the matrix A will be denoted by a_{ij} or by A_{ij} and the matrix itself will be writen as $A = (a_{ij})$.

Definition 5.36. *A* square matrix *on F is an $(n \times n)$-matrix on F for some $n \geqslant 1$.*

Let $A \in F^{n\times n}$. The *main diagonal* of the matrix A is the sequence $(a_{11}, \ldots, a_{nn})$. The set $\{a_{ij} \mid 1 \leqslant i, j \leqslant n \text{ and } i - j = k\}$ consists of elements located on the k^{th} diagonal above the main diagonal, while $\{a_{ij} \mid j - i = k\}$ consists of elements located on the k^{th} diagonal blow the main diagonal for $1 \leqslant k \leqslant n - 1$.

A square matrix $A \in F^{n\times n}$ is *diagonal* if $i \neq j$ implies $a_{ij} = 0$ for $1 \leqslant i, j \leqslant n$. A diagonal matrix $A \in F^{n\times n}$ having the diagonal elements $d_1, \ldots, d_n$ is be denoted as $A = \mathsf{diag}(d_1, \ldots, d_n)$.

A matrix $A \in F^{n\times n}$ is *upper triangular* (*lower triangular*) if $i > j$ implies $a_{ij} = 0$ ($j > i$ implies $a_{ij} = 0$).

Example 5.37. The matrices $A, B \in F^{4\times 4}$ defined by

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 \\ b_{41} & b_{42} & b_{43} & a_{44} \end{pmatrix}$$

are upper triangular and lower triangular, respectively.

Definition 5.38. *Let $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ be a matrix on the set of real numbers. Its* transpose *is the matrix $A' \in \mathbb{R}^{n\times m}$ defined by $(A')_{ij} = a_{ji}$ for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$.*

If $A' = A$, we say that A is a symmetric matrix.

A matrix A is skew-symmetric *if $A' = -A$.*

It is easy to verify that $(A')' = A$.

A similar notion exists for complex matrices.

Definition 5.39. *Let $A = (a_{ij}) \in \mathbb{C}^{m\times n}$. Its* Hermitian conjugate *is the matrix $A^{\mathsf{H}} \in \mathbb{C}^{n\times m}$ defined by $(A^{\mathsf{H}})_{ij} = \overline{a_{ji}}$ for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$.*

If $A^{\mathsf{H}} = A$, we say that A is a Hermitian matrix.

A is skew-Hermitian *if $A^{\mathsf{H}} = -A$.*

Example 5.40. The transpose of the matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 3}$$

is the matrix

$$A' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

The Hermitian conjugate of the matrix

$$L = \begin{pmatrix} 1+i & 0 & 2-3i \\ 2 & 1-i & i \end{pmatrix} \in \mathbb{C}^{2\times 3}$$

is the matrix

$$L^{\mathsf{H}} = \begin{pmatrix} 1-i & 2 \\ 0 & 1+i \\ 2+3i & -i \end{pmatrix} \in \mathbb{C}^{3\times 2}.$$

A complex matrix having real entries is symmetric if and only if it is Hermitian.

Dissimilarities defined on finite sets can be represented by matrices. If $S = \{x_1, \ldots, x_n\}$ is a finite set and $d : S \times S \longrightarrow \mathbb{R}_{\geqslant 0}$ is a dissimilarity, let $M_d \in (\mathbb{R}_{\geqslant 0})^{n\times n}$ be the matrix defined by $M_{ij} = d(x_i, x_j)$ for $1 \leqslant i, j \leqslant n$. Clearly, all main diagonal elements of M_d are 0 and the matrix M is symmetric.

Example 5.41. Let S be the set $\{x_1, x_2, x_3, x_4\}$. The discrete metric on S is represented by the 4×4-matrix

$$M_d = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

If $x_1, x_2, x_3 \in \mathbb{R}$ are three real numbers the matrix that represents the distance $e(x_i, x_j) = |x_i - x_j|$ measured on the real line is

$$M_e = \begin{pmatrix} 0 & |x_1 - x_2| & |x_1 - x_3| \\ |x_1 - x_2| & 0 & |x_2 - x_3| \\ |x_1 - x_3| & |x_2 - x_3| & 0 \end{pmatrix}.$$

Definition 5.42. *The* $(n \times n)$-unit matrix *on the field* $\mathcal{F} = (F, \{0, +, -, \cdot\})$ *is the square matrix* $I_n \in F^{n \times n}$ *given by*

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

whose entries located outside its main diagonal are 0*s.*

The $(m \times n)$-zero matrix *is the* $(m \times n)$*-matrix* $O_{m,n} \in F^{n \times n}$ *given by*

$$O_{m,n} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Definition 5.43. *Let* $A, B \in F^{m \times n}$ *be two matrices that have the same format. The* sum *of the matrices* A *and* B *is the matrix* $A + B$ *having the same format and defined by*

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

for $1 \leqslant i \leqslant m$ *and* $1 \leqslant j \leqslant n$*.*

Example 5.44. The sum of the matrices $A, B \in \mathbb{R}^{2 \times 3}$ given by

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & 2 & 3 \\ 1 & 4 & 2 \end{pmatrix}$$

is the matrix

$$A + B = \begin{pmatrix} 0 & 0 & 6 \\ 1 & 6 & 1 \end{pmatrix}.$$

It is easy to verify that the matrix sum is an associative and commutative operation on $F^{m \times n}$; that is,

$$\begin{gathered} A + (B + C) = (A + B) + C, \\ A + B = B + A, \end{gathered}$$

for all $A, B, C \in F^{m \times n}$.

The zero matrix $O_{m,n}$ acts as an additive unit on the set $F^{m \times n}$; that is,

$$A + O_{m,n} = O_{m,n} + A$$

for every $A \in S^{m \times n}$.

The additive inverse, or the opposite of a matrix $A \in F^{m \times n}$, is the matrix $-A$ given by $(-A)_{ij} = -A_{ij}$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

Example 5.45. The opposite of $A \in \mathbb{R}^{2 \times 3}$, given by

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \end{pmatrix}$$

is the matrix

$$-A = \begin{pmatrix} -1 & 2 & -3 \\ 0 & -2 & 1 \end{pmatrix}.$$

It is immediate that $A + (-A) = O_{2,3}$.

The set of matrices $F^{m \times n}$ is an F-linear space. Furthermore, it is easy to see that the sets of symmetric matrices and skew-symmetric matrices are subspaces of the linear space of square matrices $\mathbb{R}^{n \times n}$ and the sets of Hermitian and skew-Hermitian matrices are subspaces of $\mathbb{C}^{n \times n}$.

Definition 5.46. *Let $A \in F^{m \times n}$ and $B \in F^{n \times p}$ be two matrices. The* product *of the matrices A, B is the matrix $C \in F^{m \times p}$ defined by $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$, where $1 \leqslant i \leqslant m$ and $1 \leqslant k \leqslant p$. The product of the matrices A, B will be denoted by AB.*

The matrix product is a partial operation because in order to multiply two matrices A and B, they must have the formats $m \times n$ and $n \times p$, respectively. In other words, the number of columns of the first matrix must equal the number of rows of the second matrix.

Theorem 5.47. *Matrix multiplication is associative.*

Proof. Let $A \in F^{m \times n}$, $B \in F^{n \times p}$, and $C \in F^{p \times r}$ be three matrices. We prove that $(AB)C = A(BC)$.

By applying the definition of the matrix product, we have

$$
\begin{aligned}
((AB)C)_{i\ell} &= \sum_{k=1}^{p}(AB)_{ik}C_{k\ell} = \sum_{k=1}^{p}\left(\sum_{j=1}^{n}A_{ij}B_{jk}\right)C_{k\ell} \\
&= \sum_{j=1}^{n}A_{ij}\sum_{k=1}^{p}B_{jk}C_{k\ell} = \sum_{j=1}^{n}A_{ij}(BC)_{j\ell} = (A(BC))_{i\ell}
\end{aligned}
$$

for $1 \leqslant i \leqslant m$ and $1 \leqslant \ell \leqslant r$, which shows that matrix multiplication is indeed associative.

Theorem 5.48. *If $A \in F^{m\times n}$, then $I_m A = A I_n = A$.*

Proof. The statement follows immediately from the definition of a matrix product.

Note that if $A \in F^{n\times n}$, then $I_n A = A I_n = A$, so I_n is a unit relative to matrix multiplication considered as an operation on the set of square matrices $F^{n\times n}$.

The product of matrices is *not* commutative. Indeed, consider the matrices $A, B \in \mathbb{R}^{2\times 2}$ defined by

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}.$$

We have

$$AB = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \text{ and } BA = \begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix},$$

so $AB \neq BA$.

For $A \in \mathbb{C}^{n\times n}$ the power A^n, where $n \in \mathbb{N}$ is defined inductively by $A^0 = I_n$ and $A^{n+1} = A^n\ A$. It is immediate that $A^1 = A$. This allows us to define the matrix $f(A)$, where f is a polynomial with complex coefficients, $f(x) = a^n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, as

$$f(A) = a^n A^n + a_{n-1}A^{n-1} + \cdots + a_0 I_n.$$

The product of two lower (upper) triangular matrices lower (upper) triangular matrix. Therefore, any power of a lower (upper) triangular matrix is a triangular matrix.

Theorem 5.49. *If $T \in \mathbb{C}^{m\times m}$ is an upper (a lower) triangular matrix and f is a polynomial, then $f(T)$ is an upper (a lower) triangular matrix. Furthermore, if the diagonal elements of T are $t_{11}, t_{22}, \ldots, t_{mm}$, then the diagonal elements of $f(T)$ are $f(t_{11}), f(t_{22}), \ldots, f(t_{mm})$, respectively.*

Proof. Since every power T^k of T is an upper (a lower) triangular matrix, and the sum of upper (lower) triangular matrices is upper (lower) triangular, if follows that $f(T)$ is an upper triangular (a lower triangular) matrix.

An easy argument by induction on k (left to the reader) shows that if the diagonal elements of T are $t_{11}, t_{22}, \ldots, t_{mm}$, then the diagonal elements of T^k are $t_{11}^k, t_{22}^k, \ldots, t_{mm}^k$. The second part of the theorem follows immediately.

Definition 5.50. *Let* $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ *be a square matrix. The* trace *of* A *is the number* $\mathit{trace}(A)$ *given by*

$$\mathit{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Theorem 5.51. *Let* A *and* B *be two square matrices in* $\mathbb{C}^{n \times n}$. *We have:*
(i) $\mathit{trace}(aA) = a\ \mathit{trace}(A)$;
(ii) $\mathit{trace}(A + B) = \mathit{trace}(A) + \mathit{trace}(B)$, *and*
(iii) $\mathit{trace}(AB) = \mathit{trace}(BA)$.

Proof. The first two parts are direct consequences of the definition of the trace. For the last part we can write:

$$\mathit{trace}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji}.$$

Exchanging the subscripts i and j and, then the order of the summations, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} b_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} a_{ji} = \sum_{i=1}^{n} (BA)_{ii},$$

which proves the desired equality.

Let A, B, C be three matrices in $\mathbb{C}^{n \times n}$. We have

$$\mathit{trace}(ABC) = \mathit{trace}((AB)C) = \mathit{trace}(C(AB)) = \mathit{trace}(CAB),$$

and

$$\mathit{trace}(ABC) = \mathit{trace}(A(BC)) = \mathit{trace}((BC)A) = \mathit{trace}(BCA).$$

However, it is important to notice that the third part of Theorem 5.51 *cannot* be extended to arbitary permutations of a product of matrices. Consider, for example the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We have

$$ABC = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \text{ and } ACB = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix},$$

so $\mathit{trace}(ABC) = 4$ and $\mathit{trace}(ACB) = 3$.

Definition 5.52. *A matrix $A \in \mathbb{C}^{m \times n}$ is* non-negative *if all its entries a_{ij} are real numbers and $a_{ij} \geqslant 0$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. This is denoted by $A \geqslant O_{m,n}$.*

A is positive *if all its entries are real numbers, and $a_{ij} > 0$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. This is denoted by $A > 0_{m,n}$.*

If $B, C \in \mathbb{R}^{m \times n}$ we write $B \geqslant C$ $(B > C)$ if $B - C \geq O_{m,n}$ $(B - C > O_{m,n}$, respectively).

The sets of non-negative (non-positive, positive, negative) $m \times n$-matrices is denoted by $\mathbb{R}^{m \times n}_{\geqslant 0}$ ($\mathbb{R}^{m \times n}_{\leqslant 0}$, $\mathbb{R}^{m \times n}_{>0}$, $\mathbb{R}^{m \times n}_{<0}$, respectively).

Example 5.53. The diagonal matrix I_n is non-negative but not positive.

Definition 5.54. *A matrix $A \in S^{n \times n}$ is* nilpotent *if there is $m \in \mathbb{N}$ such that $A^m = O_{n,n}$. The* nilpotency *of A is the number $\mathsf{nilp}(A) = \min\{m \in \mathbb{N} \mid A^m = O_{n,n}\}$.*

In other words, if $A \in S^{n \times n}$ is a nilpotent matrix, we have $\mathsf{nilp}(A) = m$ if and only if $A^m = O_{n,n}$ but $A^{m-1} \neq O_{n,n}$.

Example 5.55. Let a and b be two positive numbers in $\mathbb{R}$. The matrix $A \in \mathbb{R}^{3 \times 3}$ given by

$$A = \begin{pmatrix} 0 & a & 0 \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}$$

is nilpotent because

$$A^2 = \begin{pmatrix} 0 & 0 & ab \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, $\mathsf{nilp}(A) = 3$.

Definition 5.56. *A matrix $A \in S^{n \times n}$ is* idempotent *if $A^2 = A$.*

Example 5.57. The matrix

$$A = \begin{pmatrix} 0.5 & 1 \\ 0.25 & 0.5 \end{pmatrix}$$

is idempotent, as the reader can easily verify.

Let $A \in F^{m \times n}$ be a matrix and suppose that $m = m_1 + \cdots + m_p$ and $n = n_1 + \cdots + n_q$, where F is the real or the complex field. A *partitioning of* A is a collection of matrices $A_{hk} \in F^{m_h \times n_k}$ such that A_{hk} is the contiguous submatrix

$$A\begin{bmatrix} m_1 + \cdots + m_{h-1} + 1, \ldots, m_1 + \cdots + m_{h-1} + m_h \\ n_1 + \cdots + n_{k-1} + 1, \ldots, n_1 + \cdots + n_k \end{bmatrix},$$

for $1 \leqslant h \leqslant p$ and $1 \leqslant k \leqslant q$.

If $\{A_{hk} \mid 1 \leqslant h \leqslant p \text{ and } 1 \leqslant k \leqslant q\}$ is a partitioning of A, A is written as

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \cdots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix}.$$

The matrices A_{hk} are referred to as the *blocks* of the partitioning. All blocks located in a column must have the number of columns; all blocks located in a row must have the same number of rows.

Example 5.58. The matrix $A \in F^{5\times 6}$ given by

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{pmatrix}$$

can be partitioned as

$$\left(\begin{array}{ccc|c|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{array}\right)$$

Thus, if we introduce the matrices

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, A_{12} = \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix}, A_{13} = \begin{pmatrix} a_{15} & a_{16} \\ a_{25} & a_{26} \\ a_{35} & a_{36} \end{pmatrix},$$
$$A_{21} = \begin{pmatrix} a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{pmatrix}, A_{22} = \begin{pmatrix} a_{45} \\ a_{55} \end{pmatrix}, A_{23} = \begin{pmatrix} a_{45} & a_{46} \\ a_{55} & a_{56} \end{pmatrix},$$

the matrix A can be written as

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}.$$

Definition 5.59. *A matrix is* $A \in \mathbb{C}^{n\times n}$ *is* normal *if* $A^{\mathsf{H}}A = AA^{\mathsf{H}}$ *and is* unitary *if* $A^{\mathsf{H}}A = AA^{\mathsf{H}} = I_n$. *Every unitary matrix is normal.*

Theorem 5.60. *A matrix $A \in \mathbb{C}^{n\times n}$ is normal and upper triangular (or lower triangular) if and only if A is a diagonal matrix.*

Proof. Suppose that A is both normal and upper triangular. The normality of A implies $(A^{\mathsf{H}}A)_{pp} = (AA^{\mathsf{H}})_{pp}$ for $1 \leqslant p \leqslant n$. We show, by induction on p that all non-diagonal elements of A are 0.

For the base step $p = 1$ we have $\bar{a}_{11}a_{11} = \sum_{j=1}^{n} a_{1j}\bar{a}_{j1} = a_{11}\bar{a}_{11} + \sum_{j=2}^{n} a_{1j}\bar{a}_{j1}$. Since $\bar{a}_{11}a_{11} = a_{11}\bar{a}_{11} = |a_{11}|^2$, it follows that $\sum_{j=2}^{n} a_{1j}\bar{a}_{j1} = \sum_{j=2}^{n} |a_{1j}|^2 = 0$, so $a_{1j} = 0$ for $2 \leqslant j \leqslant n$, which implies that all non-diagonal elements of the first line of A are 0.

For the inductive step suppose that all non-diagonal elements of the first $p-1$ rows are 0. Then

$$(A^{\mathsf{H}}A)_{pp} = \sum_{i=1}^{n} \bar{a}_{ip}a_{ip} = \sum_{i=p}^{n} \bar{a}_{ip}a_{ip} = \bar{a}_{pp}a_{pp},$$

by the inductive hypothesis and the fact that A is upper diagonal. Therefore,

$$\bar{a}_{pp}a_{pp} = \sum_{j=p}^{n} a_{pj}\bar{a}_{pj} = a_{pp}\bar{a}_{pp} + \sum_{j=p+1}^{n} a_{pj}\bar{a}_{pj},$$

which implies $\sum_{j=p+1}^{n} a_{pj}\bar{a}_{pj} = \sum_{j=p+1}^{n} |a_{pj}|^2 = 0$. This, in turn, yields $a_{p\,p+1} = \cdots = a_{pn} = 0$.

The argument is similar for the lower diagonal case.

Clearly, any diagonal matrix is normal and both upper triangular and lower triangular.

Let $A \in \mathbb{C}^{m\times n}$ be a matrix. The *matrix of the absolute values of A* is the matrix $\mathsf{abs}(A) \in \mathbb{R}^{m\times n}$ defined by

$$(\mathsf{abs}(A))_{ij} = |a_{ij}|$$

for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. In particular, if $\mathbf{x} \in \mathbb{C}^n$, we have $(\mathsf{abs}(\mathbf{x}))_j = |x_j|$.

Theorem 5.61. *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times p}$ be two matrices. We have $\mathsf{abs}(AB) \leqslant \mathsf{abs}(A)\mathsf{abs}(B)$.*

Proof. Since $(AB)_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$, it follows that

$$|(AB)_{ik}| = \left|\sum_{j=1}^{n} a_{ij}b_{jk}\right| \leqslant \sum_{j=1}^{n} |a_{ij}b_{jk}| = \sum_{j=1}^{n} |a_{ij}|\,|b_{jk}|,$$

for $1 \leqslant i \leqslant m$ and $1 \leqslant k \leqslant p$. This amounts to $\mathsf{abs}(AB) \leqslant \mathsf{abs}(A)\mathsf{abs}(B)$.

Theorem 5.62. *For $A \in \mathbb{C}^{n\times n}$ we have $\mathsf{abs}(A^k) \leqslant (\mathsf{abs}(A))^k$ for every $k \in \mathbb{N}$.*

Proof. The proof is by induction on k. The base case, $k = 0$, is immediate. Suppose that the inequality holds for k. We have

$$\begin{aligned}\mathsf{abs}(A^{k+1}) &= \mathsf{abs}(A^k\ A)\\ &\leqslant \mathsf{abs}(A^k)\mathsf{abs}(A)\\ &\quad \text{(by Theorem 5.61)}\\ &\leqslant (\mathsf{abs}(A))^k\mathsf{abs}(A)\\ &\quad \text{(by the inductive hypothesis)}\\ &= (\mathsf{abs}(A))^{k+1},\end{aligned}$$

which completes the induction case.

Partitioning matrices is useful because matrix operations can be performed on block submatrices in a manner similar to scalar operations as we show next.

Theorem 5.63. *Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$ be two matrices. Suppose that the matrices A, B are partitioned as*

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1k}\\ \vdots & \cdots & \vdots\\ A_{h1} & \cdots & A_{hk}\end{pmatrix} \text{ and } B = \begin{pmatrix} B_{11} & \cdots & B_{1\ell}\\ \vdots & \cdots & \vdots\\ B_{k1} & \cdots & B_{k\ell}\end{pmatrix},$$

where $A_{rs} \in F^{m_r\times n_s}$, $B_{st} \in F^{n_s\times p_t}$ for $1 \leqslant r \leqslant h$, $1 \leqslant s \leqslant k$ and $1 \leqslant t \leqslant \ell$. Then, the product $C = AB$ can be partitioned as

$$C = \begin{pmatrix} C_{11} & \ldots & C_{1\ell}\\ \vdots & \ldots & \vdots\\ C_{h1} & \cdots & C_{hl}\end{pmatrix},$$

where $C_{uv} = \sum_{t=1}^{k} A_{ut}B_{tv}$, $1 \leqslant u \leqslant h$, and $1 \leqslant v \leqslant \ell$.

Proof. Note that $m_1 + \cdots + m_h = m$ and $p_1 + \cdots + p_\ell = p$. For a pair (i, j) such that $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$ let u be the least number such that $i \leqslant m_1 + \cdots + m_u$ and let v be the least number such that $j \leqslant p_1 + \cdots + p_v$. The definition of u and v implies $m_1 + \cdots + m_{u-1} + 1 \leqslant i \leqslant m_1 + \cdots + m_u$ and $p_1 + \cdots + p_{v-1} + 1 \leqslant j \leqslant p_1 + \cdots + p_v$. This implies that the c_{ij} element of the product is located in the submatrix $C_{uv} = \sum_{t=1}^{k} A_{ut}B_{tv}$ of C. By the definition of the matrix product we have

$$\begin{aligned}c_{ij} &= \sum_{g=1}^{n} a_{ig}b_{gj}\\ &= \sum_{g=1}^{n_1} a_{ig}b_{gj} + \sum_{g=n_1+1}^{n_1+n_2} a_{ig}b_{gj} + \cdots + \sum_{g=n_1+\cdots+n_{k-1}+1}^{n_1+\cdots+n_s} a_{ig}b_{gj}\end{aligned}$$

Observe that the vectors $(a_{i1}, \ldots, a_{in_1})$ and $(b_{1j}, \ldots, b_{n_1 j})'$ represent the line number $i-(m_1+\cdots+m_{u-1}+1)$ and the column number $j-(p_1+\cdots+p_{v-1}+1)$ of the matrix A_{u1} and B_{1v}, etc. Similarly,

$$(a_{i,n_1+\cdots+n_{k-1}+1}, \ldots, a_{i,n_1+\cdots+n_s})$$

and

$$(b_{n_1+\cdots+n_{k-1}+1,j}, \ldots, b_{n_1+\cdots+n_s,j})'$$

represent the line number $i-(m_1+\cdots+m_{u-1}+1)$ and the column number $j-(p_1+\cdots+p_{v-1}+1)$ of the matrix A_{uk} and B_{kv}, which shows that c_{ij} is computed correctly as an element of the block C_{uv}.

Next, we explore the relationship between linear mappings and matrices. Let $h \in \mathsf{Hom}(\mathbb{C}^m, \mathbb{C}^n)$ be a linear transformation between the linear spaces $\mathbb{C}^m$ and $\mathbb{C}^n$, let $R = \{\mathbf{r}_1, \ldots, \mathbf{r}_m\}$ be a basis in $\mathbb{C}^m$, and let $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ be a basis in $\mathbb{C}^n$.

Since $h(\mathbf{r}_j) \in \mathbb{C}^n$ we can write:

$$h(\mathbf{r}_j) = a_{1j}\mathbf{s}_1 + a_{2j}\mathbf{s}_2 + \cdots + a_{nj}\mathbf{s}_n.$$

Definition 5.64. *The* matrix $A_h \in \mathbb{C}^{n\times m}$ associated to the linear mapping $h : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ *is the matrix that has*

$$h(\boldsymbol{r}_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

as its j^{th} *column for* $1 \leqslant j \leqslant m$.

Let $\mathbf{v} \in \mathbb{C}^m$ be a vector such that $\mathbf{v} = v_1\mathbf{r}_1 + \cdots + v_m\mathbf{r}_m$. Then, the image of $\mathbf{v}$ under h is

$$\begin{aligned} h(\mathbf{v}) &= h\left(\sum_{j=1}^m v_j\mathbf{r}_j\right) = \sum_{j=1}^m v_j h(\mathbf{r}_j) \\ &= \sum_{j=1}^m v_j \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}v_j \\ \sum_{j=1}^m a_{2j}v_j \\ \vdots \\ \sum_{j=1}^m a_{nj}v_j \end{pmatrix}, \end{aligned}$$

which is easily seen to equal $A_h\mathbf{v}$.

As we saw above, the matrix A_h attached to $h : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ depends on the bases chosen for the linear spaces $\mathbb{C}^m$ and $\mathbb{C}^n$.

Let A_h^R and A_h^S be the matrices associated to h that correspond to the bases R and S, respectively. We have $A_h^R = (h(\mathbf{r}_1) \ \cdots \ h(\mathbf{r}_n))$ and $A_h^S = (h(\mathbf{s}_1) \ \cdots \ h(\mathbf{s}_n))$. The Equalities (5.3) can be written succinctly as

$$A_h^R = P A_h^S, \tag{5.1}$$

where P is the matrix

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \cdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix},$$

whose entries have been introduced in Equalities (5.2).

Let $h : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ be an endomorphism of $\mathbb{C}^n$ and let $R = \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ and $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ be two bases of $\mathbb{C}^n$. The vectors $\mathbf{s}_i$ can be expressed as linear combinations of the vectors $\mathbf{r}_1, \ldots, \mathbf{r}_n$:

$$\mathbf{s}_i = p_{i1}\mathbf{r}_1 + \cdots + p_{in}\mathbf{r}_n, \tag{5.2}$$

for $1 \leqslant i \leqslant n$, which implies

$$h(\mathbf{s}_i) = p_{i1}h(\mathbf{r}_1) + \cdots + p_{in}h(\mathbf{r}_n). \tag{5.3}$$

for $1 \leqslant i \leqslant n$.

Matrix multiplication corresponds to the composition of linear mappings, as we show next.

Theorem 5.65. *Let $h \in \mathsf{Hom}(\mathbb{C}^m, \mathbb{C}^n)$ and $g \in \mathsf{Hom}(\mathbb{C}^n, \mathbb{C}^p)$. Then,*

$$A_{gh} = A_g A_h.$$

Proof. If $\mathbf{p}_1, \ldots, \mathbf{p}_m$ is a basis for $\mathbb{C}^m$, then $A_{gh}(\mathbf{p}_i) = gh(\mathbf{p}_i) = g(h(\mathbf{p}_i)) = g(A_h\mathbf{p}_i) = A_g(A_h(\mathbf{p}_i))$ for every i, where $1 \leqslant i \leqslant n$. This proves that $A_{gh} = A_g A_h$.

Thus, if h is an idempotent endomorphism of a linear space the matrix A_h is idempotent.

Starting from a matrix $A \in \mathbb{C}^{n\times m}$ we can define a *linear operator associated to* A, $h_A : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ as $h_A(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x} \in \mathbb{C}^m$. If $\mathbf{p}_1, \ldots, \mathbf{p}_m$ is a basis for $\mathbb{C}^m$, then $h_A(\mathbf{p}_i)$ is the i^{th} column of the matrix A.

It is immediate that $A_{h_A} = A$ and $h_{A_h} = h$.

Attributes of a matrix A are transferred to the linear operator h_A. For example, if A is Hermitian we say that h_A is Hermitian.

The association of matrices in $\mathbb{C}^{n\times m}$ with linear operators, described in Definition 5.64, suggests the association of certain subspaces of the linear spaces $\mathbb{C}^n$ and $\mathbb{C}^m$ to A.

Definition 5.66. *Let $A \in \mathbb{C}^{n\times m}$ be a matrix. The* range of A *is the subspace $\mathsf{Img}(h_A)$ of $\mathbb{C}^n$. The* null space of A *is the subspace $\mathsf{Ker}(h_A)$.*

The range of A and the null space of A are denoted by $\mathsf{Ran}(A)$ and $\mathsf{NullSp}(A)$, respectively.

Clearly, $C_{A,n} = \mathsf{Ran}(A)$. The null space of $A \in \mathbb{C}^{m\times n}$ consists of those $\mathbf{x} \in \mathbb{C}^n$ such that $A\mathbf{x} = \mathbf{0}$.

Let $\{\mathbf{p}_1, \ldots, \mathbf{p}_m\}$ be a basis of $\mathbb{C}^m$. Since $\mathsf{Ran}(A) = \mathsf{Img}(h_A)$ it follows that this subspace is generated by the set $\{h_A(\mathbf{p}_1), \ldots, h_A(\mathbf{p}_m)\}$, that is, by the columns of the matrix A. For this reason the subspace $\mathsf{Ran}(A)$ is also known as the *column subspace* of A.

Several important facts concerning idempotent endomorphisms that were previously presented can now be formulated in terms of matrices. For example, Theorem 5.32 applied to $\mathbb{C}^n$ states that if A is an idempotent matrix, then $\mathbb{C}^n = \mathsf{NullSp}(A) \boxplus \mathsf{Ran}(A)$. Conversely, by Theorem 5.33 if U and W are two subspaces of $\mathbb{C}^n$ such that $\mathbb{C}^n = U \boxplus W$, then there exists an idempotent matrix $A \in \mathbb{C}^{n\times n}$ such that $U = \mathsf{NullSp}(A)$ and $W = \mathsf{Ran}(A)$.

Let $A \in \mathbb{C}^{n\times n}$ be a square matrix. Suppose that there exist two matrices U and V such that $AU = I_n$ and $VA = I_n$. This implies

$$V = VI_n = V(AU) = (VA)U = I_nU = U.$$

Thus, if $AU = VA = I_n$, the two matrices involved, U and V, must be equal.

Definition 5.67. *A matrix $A \in \mathbb{C}^{n\times n}$ is invertible if there exists a matrix $B \in \mathbb{C}^{n\times n}$ such that $AB = BA = I_n$.*

Suppose that C is another matrix such that $AC = CA = I_n$. By the associativity of the matrix product we have $C = CI_n = C(AB) = (CA)B = I_nB = B$. Therefore, if A is invertible there is exactly one matrix B such that $AB = BA = I_n$. We denote the matrix B by A^{-1} and we refer to it as the *inverse of the matrix A*.

Note that $A \in \mathbb{C}^{n\times n}$ is a unitary matrix if and only if $A^{-1} = A^{\mathsf{H}}$.

Theorem 5.68. *If $A, B \in \mathbb{C}^{n\times n}$ are two invertible matrices, then the product AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. Applying the definition of the inverse of a matrix we obtain

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n,$$

which implies $(AB)^{-1} = B^{-1}A^{-1}$.

Theorem 5.69. *If $A \in \mathbb{C}^{n\times n}$ is invertible, then A^{H} is invertible and $(A^{\mathsf{H}})^{-1} = (A^{-1})^{\mathsf{H}}$.*

Proof. Since $AA^{-1} = I_n$, we have $(A^{-1})^{\mathsf{H}}A^{\mathsf{H}} = I_n$, which shows that $(A^{-1})^{\mathsf{H}}$ is the inverse of A^{H} and $(A^{\mathsf{H}})^{-1} = (A^{-1})^{\mathsf{H}}$.

Example 5.70. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

be a matrix in $\mathbb{R}^{2\times 2}$. We seek to determine conditions under which A is invertible. Suppose that

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

is a matrix in $\mathbb{R}^{2\times 2}$ such that $AX = I_2$. This matrix equality amounts to four scalar equalities:

$$\begin{aligned} a_{11}x_{11} + a_{12}x_{21} &= 1, \; a_{11}x_{12} + a_{12}x_{22} = 0, \\ a_{21}x_{11} + a_{22}x_{21} &= 1, \; a_{21}x_{12} + a_{22}x_{22} = 0, \end{aligned}$$

which, under certain conditions, can be solved with respect to $x_{11}, x_{12}, x_{21}, x_{22}$.

By multiplying the first equality by a_{22} and the third by $-a_{12}$ and adding the resulting equalities we obtain $(a_{11}a_{22} - a_{12}a_{21})x_{11} = -a_{22}$. Thus, if $a_{11}a_{22} - a_{12}a_{21} \neq 0$, we have $x_{11} = -\frac{a_{22}}{a_{11}a_{22}-a_{12}a_{21}}$. The same condition, $a_{11}a_{22} - a_{12}a_{21} \neq 0$, suffices to allow us to obtain the value of the remaining components of X, as the reader can easily verify. Thus, A is an invertible matrix if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Definition 5.71. *A* stochastic matrix *is a matrix $A \in \mathbb{R}^{n\times n}$ such that $a_{ij} \geqslant 0$ for $1 \leqslant i, j \leqslant n$ and $\sum_{j=1}^{n} a_{ij} = 1$ for every i, $1 \leqslant i \leqslant n$.*

A doubly stochastic matrix *is a matrix $A \in \mathbb{R}^{n\times n}$ such that both A and A' are stochastic.*

The rows of a stochastic matrix can be regarded as discrete probability distributions.

Example 5.72. The matrix $A \in \mathbb{R}^{3\times 3}$ defined by

$$A = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

is a stochastic matrix.

Example 5.73. Let

$$\phi : \begin{pmatrix} 1 & \cdots & k & \cdots & n \\ a_1 & \cdots & a_k & \cdots & a_n \end{pmatrix},$$

be a permutation of the set $\{1, \ldots, n\}$, where $a_k = \phi(k)$ for $1 \leqslant k \leqslant n$.

The matrix of this permutation is the square matrix $P_\phi = (p_{ij}) \in \{0,1\}^{n\times n}$, where

$$p_{ij} = \begin{cases} 1 & \text{if } j = \phi(i), \\ 0 & \text{otherwise,} \end{cases} \tag{5.4}$$

for $1 \leqslant i, j \leqslant n$.

Note that the matrix of the permutation $1_{1,\ldots,n}$ is the matrix I_n.

Also, if ϕ, ψ are two permutations of the set $\{1, \ldots, n\}$, then $P_{\psi\phi} = P_\phi P_\psi$. Indeed, since $(P_\phi P_\psi)_{ij} = \sum_{k=1}^{n} (P_\phi)_{ik} (P_\psi)_{kj}$, observe that only the term $(P_\phi)_{ik}(P_\psi)_{kj}$ in which $k = \phi(i)$ and $j = \psi(k)$ is different from 0. Thus, $(P_\phi P_\psi)_{ij} \neq 0$ if and only if $j = \psi(\phi(i))$, which means that $P_{\psi\phi} = P_\phi P_\psi$.

Thus, if ϕ and ϕ^{-1} are two inverse permutations in $PERM_n$, we have $P_\phi P_{\phi^{-1}} = I_n$, so P_ϕ is invertible and $P_\phi^{-1} = P_{\phi^{-1}}$.

For instance, if $\phi \in PERM_4$ is

$$\phi : \begin{pmatrix} 1\,2\,3\,4 \\ 3\,1\,4\,2 \end{pmatrix},$$

then

$$P_\phi = \begin{pmatrix} 0\,0\,1\,0 \\ 1\,0\,0\,0 \\ 0\,0\,0\,1 \\ 0\,1\,0\,0 \end{pmatrix}$$

Its inverse is

$$P_\phi^{-1} = P_{\phi^{-1}} = \begin{pmatrix} 0\,1\,0\,0 \\ 0\,0\,0\,1 \\ 1\,0\,0\,0 \\ 0\,0\,1\,0 \end{pmatrix}.$$

It is easy to verify that the inverse of a permutation matrix P_ϕ coincides with its transpose $(P_\phi)'$.

Observe that if $A \in \mathbb{R}^{n \times n}$ having the rows $\mathbf{r}_1, \ldots, \mathbf{r}_n$ and P_ϕ is a permutation matrix, then $P_\phi A$ is the matrix whose rows are $\mathbf{r}_{\phi(1)}, \mathbf{r}_{\phi(2)}, \ldots, \mathbf{r}_{\phi(n)}$. Similarly, if the columns of A are $\mathbf{c}_1, \ldots, \mathbf{c}_n$, the columns of the matrix AP_ϕ are $\mathbf{c}_{\phi(1)}, \ldots, \mathbf{c}_{\phi(n)}$. In other words, $P_\phi A$ is obtained from A be permuting its rows according to the permutation ϕ and AP_ϕ is obtained from A by permuting the columns according to the same permutation.

Since every column and row of a permutation matrix contains exactly one 1, it follows that each such matrix is also a doubly-stochastic matrix.

Theorem 5.74. *Let $A \in \mathbb{R}^{n \times n}$ be a lower (upper) triangular matrix such that $a_{ii} \neq 0$ for $1 \leqslant i \leqslant n$. The matrix A is invertible and its inverse is a lower (upper) triangular matrix having diagonal elements equal to the reciprocal of the diagonal elements of A.*

Proof. Let A be a lower triangular matrix

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix},$$

where $a_{ii} \neq 0$ for $1 \leqslant i \leqslant n$. The proof is by induction on $n \geqslant 1$.

The base case, $n = 1$ is immediate, since the inverse of the matrix (a_{11}) is $\left(\frac{1}{a_{11}}\right)$.

Suppose that the statement holds for matrices in $\mathbb{R}^{(n-1)\times(n-1)}$. Then A can be written as

$$A = \left(\begin{array}{c|c} B & \mathbf{0}_{n-1} \\ \hline a_{n1}\ a_{n2}\ \cdots\ a_{n\ n-1} & a_{nn} \end{array}\right),$$

where $B \in \mathbb{R}^{(n-1)\times(n-1)}$ is a lower triangular matrix. By the inductive hypothesis, this matrix is invertible, its inverse B^{-1} is also lower triangular and the diagonal elements of B^{-1} are the reciprocal elements of the corresponding diagonal elements of B. The matrix

$$\left(\begin{array}{c|c} B^{-1} & \mathbf{0}_{n-1} \\ \hline \mathbf{v} & \frac{1}{a_{nn}} \end{array}\right)$$

is the inverse of A, where $\mathbf{v} = -\frac{1}{a_{nn}}\mathbf{a}'B^{-1}$, and $\mathbf{a}' = (a_{n1}, a_{n2}, \ldots, a_{n\ n-1})$, as the reader can easily verify.

A similar argument can be used for upper triangular matrices.

Theorem 5.75. *Let $A \in \mathbb{R}^{n\times n}$ be an invertible matrix. Then, its transpose A' is invertible and $(A')^{-1} = (A^{-1})'$.*

Proof. Observe that $A'(A^{-1})' = (A^{-1}A)' = I_n' = I_n$. Therefore, $(A')^{-1} = (A^{-1})'$.

If $A \in \mathbb{R}^{n\times n}$ is invertible we have $AA^{-1} = I_n$, so $\mathit{trace}(A)\mathit{trace}(A^{-1}) = \mathit{trace}(I_n) = n$. This implies

$$\mathit{trace}(A^{-1}) = \frac{n}{\mathit{trace}(A)}. \tag{5.5}$$

Theorem 5.76. *Let $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ be a basis in $\mathbb{C}^n$. A matrix $A \in \mathbb{C}^{n\times n}$ is invertible if and only if the set of vectors $\{A\boldsymbol{r}_1, \ldots, A\boldsymbol{r}_n\}$ is a basis in $\mathbb{C}^n$.*

Proof. Suppose that A is an invertible matrix. Note that $A\mathbf{x}_i = A\mathbf{x}_j$ implies $\mathbf{x}_i = \mathbf{x}_j$, so $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ consists on n distinct vectors. We claim that the set $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ is linearly independent. Indeed, suppose that $c_1A\mathbf{r}_1 + \cdots + c_nA\mathbf{r}_n = \mathbf{0}_n$ such that not all coefficients c_i equal 0. Then, by multiplying by A^{-1} to the left we obtain $c_1\mathbf{r}_1 + \cdots + c_n\mathbf{r}_n = \mathbf{0}_n$, which contradicts the fact that $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ is a basis. Thus, $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ is a linearly independent that consists of n vectors, which means that this set is a basis in $\mathbb{C}^n$.

Conversely, suppose that for any basis $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ of $\mathbb{C}^n$, $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ is a basis in $\mathbb{C}^n$. Each of the vectors $\mathbf{r}_i$ can be uniquely expressed as a linear combination of $A\mathbf{r}_1, \ldots, A\mathbf{r}_n$. In particular, for the *standard basis* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, each of the vectors $\mathbf{e}_i$ can be uniquely expressed as a linear combination of the vectors $A\mathbf{e}_1 = \mathbf{a}_1, \ldots, A\mathbf{e}_n = \mathbf{a}_n$, where $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the columns of the matrix A. In other words, we have the equalities

$$\mathbf{e}_i = b_{i1}\mathbf{a}_1 + \cdots + b_{in}\mathbf{a}_n$$

for $1 \leqslant i \leqslant n$. In a succinct form, these equalities can be written as $I_n = BA$, where B is the matrix of the coefficient b_{ij}, which shows that A is an invertible matrix.

Theorem 5.77. *Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ be two bases of $\mathbb{C}^n$. There exists an invertible matrix $P \in \mathbb{C}^{n \times n}$ such that*

$$(\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n) = (\boldsymbol{w}_1 \ \cdots \ \boldsymbol{w}_n)P.$$

Proof. Since $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is a basis of $\mathbb{C}^n$ each vector $\mathbf{u}_i$ is a unique linear combination of the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$, that is

$$\mathbf{u}_i = p_{1i}\mathbf{w}_1 + \cdots + p_{ni}\mathbf{w}_n = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_n) \begin{pmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{pmatrix},$$

for $1 \leqslant i \leqslant n$, so the equality of the theorem holds for the matrix $P = (p_{ij})$. We have to show that P is an invertible matrix.

Assume that $P\mathbf{t} = \mathbf{0}_n$. The equality of the theorem implies

$$(\mathbf{u}_1 \ \cdots \ \mathbf{u}_n) \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_n)P\mathbf{t} = \mathbf{0}_n.$$

which implies $t_1\mathbf{u}_1 + \cdots + t_n\mathbf{u}_n = \mathbf{0}_n$. Since $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis we obtain $t_1 = \cdots = t_n = 0$, so $\mathbf{t} = \mathbf{0}_n$, which implies that P is an invertible matrix.

Corollary 5.78. *Let $\boldsymbol{z} \in \mathbb{C}^n$ and assume that $\boldsymbol{z} \in \mathbb{C}^n$ can be expressed relatively to the bases $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ as $\boldsymbol{z} = \sum_{i=1}^n x_i\boldsymbol{u}_i$ and as $\boldsymbol{z} = \sum_{i=1}^n y_i\boldsymbol{w}_i$, respectively. If $(\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n) = (\boldsymbol{w}_1 \ \cdots \ \boldsymbol{w}_n)P$, then*

$$P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Proof. We have

$$\mathbf{z} = (\mathbf{u}_1 \ \cdots \ \mathbf{u}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_n)P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

The linear independence of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ implies the desired equality.

5.4 Rank

The subspace $\mathsf{Ran}(A)$ of a matrix A is generated by the columns of this matrix. An analogous space is $\mathsf{rows}(A)$, the linear space spanned by the rows of A.

Definition 5.79. *Let $A \in \mathbb{C}^{m\times n}$ be a matrix. Its* column rank *is the number*

$$\textit{c-rank}(A) = \dim(\mathsf{Ran}(A)).$$

The row rank *of A is the number r-rank(A) equal to the dimension of the row space of A.*

Theorem 5.80. *Let $A \in \mathbb{C}^{m\times n}$ be a matrix. We have r-rank(A) = c-rank(A).*

Proof. Let $\mathbf{r}_1, \ldots, \mathbf{r}_m$ be the rows of A and let $\mathbf{c}_1, \ldots, \mathbf{c}_n$ be the columns of the same, so

$$A = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{pmatrix} = (\mathbf{c}_1 \ \cdots \ \mathbf{c}_n).$$

Suppose that *r-rank*$(A) = r$. There exists a basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$ of the subspace $\mathsf{rows}(A)$ such that every row $\mathbf{r}_i$ of A can be written as a linear combination: $\mathbf{r}_i = \mathbf{u}_i B$, where $\mathbf{u}_i = (u_{i1} \ \cdots u_{ir})$ for $1 \leqslant i \leqslant m$ and

$$B = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_r \end{pmatrix} \in \mathbb{C}^{r\times n}.$$

Since $a_{ij} = \mathbf{r}_i \mathbf{e}_j$, we can write $a_{ij} = \mathbf{r}_i \mathbf{e}_j = \mathbf{u}_i B \mathbf{e}_j$.

Let $U \in \mathbb{C}^{m\times r}$ be the matrix whose rows are $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and let $\mathbf{d}_1, \ldots, \mathbf{d}_r$ be the columns of this matrix. The j^{th} column of A can be written as

$$\begin{aligned} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} &= \begin{pmatrix} \mathbf{u}_1 B \mathbf{e}_j \\ \vdots \\ \mathbf{u}_m B \mathbf{e}_j \end{pmatrix} = UB\mathbf{e}_j = (\mathbf{d}_1 \ \cdots \ \mathbf{d}_r) B \mathbf{e}_j \\ &= (\mathbf{d}_1 \ \cdots \ \mathbf{d}_r) \begin{pmatrix} b_{1j} \\ \vdots \\ b_{rj} \end{pmatrix} = b_{1j}\mathbf{d}_1 + \cdots + b_{rj}\mathbf{d}_r. \end{aligned}$$

This shows that the column space of A is generated by the set $\{\mathbf{d}_1, \ldots, \mathbf{d}_r\}$, so *c-rank*$(A) \leqslant r$. The same argument applied to A' implies that the $r =$ *c-rank*$(A') \leqslant$ *r-rank*$(A') =$ *c-rank*(A), so *c-rank*$(A) = r$, which concludes the argument.

Definition 5.81. *The* rank *of a matrix A is the number denoted by rank(A) given by rank$(A) = \dim(\mathsf{Ran}(A)) = \dim(\mathsf{Img}(h_A))$.*

In other words, the rank of A is the maximal size of a set of linearly independent columns of A. By Theorem 5.80, the rank of A equals the maximal size of a set of linearly independent rows of A.

Consider the matrices $T_n^{i\leftrightarrow j}, T_n^{i\overset{+}{\leftarrow}j}$ and $T_{n,i}^{(a)}$ in $\mathbb{R}^{n\times n}$ defined by:

$$\begin{aligned} T_n^{i\leftrightarrow j} &= (\mathbf{e}_1 \ \cdots \ \mathbf{e}_{i-1} \ \mathbf{e}_j \ \mathbf{e}_{i+1} \ \cdots \ \mathbf{e}_{j-1} \ \mathbf{e}_i \ \mathbf{e}_{j+1} \ \cdots \ \mathbf{e}_n) \\ T_n^{i\overset{+}{\leftarrow}j} &= (\mathbf{e}_1 \ \cdots \ \mathbf{e}_{i-1} \ \mathbf{e}_i + \mathbf{e}_j \ \mathbf{e}_i \ \cdots \ \mathbf{e}_n), \end{aligned}$$

and

$$T_{n,i}^{(a)} = \mathsf{diag}(1, 1, \ldots, 1, a, 1, \ldots, 1),$$

where a occupies the i^{th} diagonal position.

By multiplying a matrix $A \in \mathbb{C}^{m\times n}$ at the right with any of these matrices, certain transformations on the set of columns of A take place. Namely, the matrix $AT_n^{i\leftrightarrow j}$ is obtained from A by permuting the i^{th} and the j^{th} column. The matrix $AT_n^{i\overset{+}{\leftarrow}j}$ results by adding the j^{th} column to the i^{th} column. Finally, $AT_{n,i}^{(a)}$ is obtained from A by multiplying the i^{th} column by a, where $a \neq 0$.

Similar effects are obtained on the rows of A by multiplying A at the left with $T_m^{i\leftrightarrow j}, (T_m^{i\overset{+}{\leftarrow}j})'$ and $T_{m,i}^{(a)}$.

Example 5.82. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

We have

$$T_3^{2\leftrightarrow 3} A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix},$$

and

$$AT_4^{2\leftrightarrow 3} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} & a_{12} & a_{14} \\ a_{21} & a_{23} & a_{22} & a_{24} \\ a_{31} & a_{33} & a_{32} & a_{34} \end{pmatrix}.$$

For $A \in \mathbb{C}^{m\times n}$ the column space of the matrices $AT_n^{i\leftrightarrow j}, AT_n^{i\overset{+}{\leftarrow}j}$ and $AT_{n,i}^{(a)}$ is the same as the column space of A and the row space of $T_m^{i\leftrightarrow j}A, (T_m^{i\overset{+}{\leftarrow}j})'A$ and $T_{m,i}^{(a)}A$ is the same as the row space of A. Therefore, any of the matrices

$$AT_n^{i\leftrightarrow j}, AT_n^{i\overset{+}{\leftarrow}j}, AT_{n,i}^{(a)}, T_{m,i}^{i\leftrightarrow j}A, (T_m^{i\overset{+}{\leftarrow}j})'A, T_{m,i}^{(a)}A$$

has the same rank as A.

Theorem 5.83. *If* $A \in \mathbb{C}^{m\times n}$ *we have*

$$\dim(\mathit{NullSp}(A)) + \mathit{rank}(A) = n. \tag{5.6}$$

Proof. Suppose that $\dim(\mathsf{NullSp}(A)) = q$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$ is a basis of $\mathsf{NullSp}(A)$, and that $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_q, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_n\}$ is its extension to a basis B of $\mathbb{C}^m$.

If $\mathbf{y} \in \mathsf{Ran}(A)$, then $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{C}^n$. Since $\mathbf{x}$ is a linear combination of the basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ we can write

$$\mathbf{x} = a_1\mathbf{v}_1 + \cdots + a_q\mathbf{v}_q + a_{q+1}\mathbf{v}_{q+1} + \cdots + a_n\mathbf{v}_n,$$

so

$$\begin{aligned}\mathbf{y} &= a_1A\mathbf{v}_1 + \cdots + a_qA\mathbf{v}_q + a_{q+1}A\mathbf{v}_{q+1} + \cdots + a_nA\mathbf{v}_n\\ &= a_{q+1}A\mathbf{v}_{q+1} + \cdots + a_nA\mathbf{v}_n.\end{aligned}$$

Therefore, $\{A\mathbf{v}_{q+1}, \ldots, A\mathbf{v}_n\}$ spans $\mathsf{Ran}(A)$. We claim that this set of vectors is linearly independent. Suppose this is not the case. Then, there exist $n-q$ numbers $d_{q+1}, \ldots, d_n$ such that

$$d_{q+1}A\mathbf{v}_{q+1} + \cdots + d_nA\mathbf{v}_n = \mathbf{0},$$

so $A(d_{q+1}\mathbf{v}_{q+1} + \cdots + d_n\mathbf{v}_n) = \mathbf{0}$, which means that $\mathbf{w} = d_{q+1}\mathbf{v}_{q+1} + \cdots + d_n\mathbf{v}_n \in \mathsf{NullSp}(A)$. Therefore, $\mathbf{w}$ is a linear combination of $\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$, which implies the existence of $d_1, \ldots, d_q$ such that

$$d_{q+1}\mathbf{v}_{q+1} + \cdots + d_n\mathbf{v}_n = d_1\mathbf{v}_1 + \cdots + d_q\mathbf{v}_q.$$

This contradicts the fact that B is a basis, so the set $\{A\mathbf{v}_{q+1}, \ldots, A\mathbf{v}_n\}$ is linearly independent and, therefore, is a basis of $\mathsf{Ran}(A)$. This implies that $\mathit{rank}(A) = n - q$, which is the desired equality.

Example 5.84. For the matrix

$$A = \begin{pmatrix} 1 & 0 & 2\\ 1 & -1 & 1\\ 2 & 1 & 5\\ 1 & 2 & 4\end{pmatrix}$$

we have $\mathit{rank}(A) = 2$. Indeed, if $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ are its columns, then it is easy to see that $\{\mathbf{c}_1, \mathbf{c}_2\}$ is a linearly independent set, and $\mathbf{c}_3 = 2\mathbf{c}_1 + \mathbf{c}_2$. Thus, the maximal size of a set of linearly independent columns of A is 2.

Example 5.85. Let $A \in \mathbb{C}^{n\times m}$ and $B \in \mathbb{C}^{p\times q}$. For the matrix $C \in \mathbb{C}^{(n+p)\times(m+q)}$ defined by

$$C = \begin{pmatrix} A & O_{n,q}\\ O_{p,m} & B\end{pmatrix}$$

we have $rank(C) = rank(A) + rank(B)$.

Suppose that $rank(C) = \ell$ and let $\mathbf{c}_1, \ldots, \mathbf{c}_\ell$ be a maximal set of linearly independent columns of C. Without loss of generality we may assume that the first k columns are among the first m columns of A and the remaining $\ell - k$ columns are among the last q columns of C. The first k columns of C correspond to k linearly independent columns of A, while the last $\ell - k$ columns correspond to $\ell - k$ linearly independent columns of B. Thus, $rank(C) = k \leqslant rank(A) + rank(B)$.

Conversely, suppose that $rank(A) = s$ and $rank(B) = t$ and let $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_s}$ be a maximal set of linearly independent columns of A and let $\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_t}$ be a maximal set of linearly independent columns of B. Then, it is easy to see that the vectors

$$\begin{pmatrix} \mathbf{a}_{i_1} \\ \mathbf{0}_n \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{a}_{i_s} \\ \mathbf{0}_n \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{0}_n \\ \mathbf{b}_{j_1} \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{0}_n \\ \mathbf{b}_{j_t} \end{pmatrix}$$

constitute a linearly independent set of columns of C, so $rank(A) + rank(B) \leqslant rank(C)$. Thus, $rank(C) = rank(A) + rank(B)$.

Example 5.86. Let $\mathbf{x}$ and $\mathbf{y}$ be two vectors in $\mathbb{C}^n - \{\mathbf{0}\}$. The matrix $\mathbf{x}\mathbf{y}^{\mathsf{H}}$ has rank 1. Indeed, if $\mathbf{y}^{\mathsf{H}} = (y_1, y_2, \ldots, y_n)$, then $\mathbf{x}\mathbf{y}^{\mathsf{H}} = (y_1\mathbf{x}\ y_2\mathbf{x}\ \cdots\ y_n\mathbf{x})$, which implies that the maximum number of linearly independent columns of $\mathbf{x}\mathbf{y}^{\mathsf{H}}$ is 1.

The above discussion also shows that if $A \in \mathbb{C}^{n \times m}$, then $rank(A) \leqslant \min\{m, n\}$.

Theorem 5.87. *Let* $A \in \mathbb{C}^{m \times n}$ *be a matrix. We have* $rank(A) = rank(\overline{A})$.

Proof. Suppose that $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ and that the set $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_p}\}$ is a set of linearly independent columns of A. Then, the set $\{\overline{\mathbf{a}_{i_1}}, \ldots, \overline{\mathbf{a}_{i_p}}\}$ is a set of linearly independent columns of $\overline{A}$. This implies $rank(\overline{A}) = rank(A)$.

Corollary 5.88. *We have* $rank(A) = rank(A^{\mathsf{H}})$ *for every matrix* $A \in \mathbb{C}^{m \times n}$.

Proof. Since $A^{\mathsf{H}} = \overline{A'}$, the statement follows immediately.

Definition 5.89. *A matrix* $A \in \mathbb{C}^{n \times m}$ *is a* full-rank *matrix if* $rank(A) = \min\{m, n\}$.

If $A \in \mathbb{C}^{m \times n}$ is a full-rank matrix and $m \geqslant n$, then the n columns of the matrix are linearly independent; similarly, if $n \geqslant m$, the m rows of the matrix are linearly independent.

A matrix that is not a full-rank is said to be *degenerate*. A degenerate square matrix is said to be *singular*. A *non-singular matrix* $A \in \mathbb{C}^{n \times n}$ is a matrix that is not singular and, therefore has $rank(A) = n$.

Theorem 5.90. *A matrix* $A \in \mathbb{C}^{n\times n}$ *is non-singular if and only if it is invertible.*

Proof. Suppose that A is non-singular, that is, $rank(A) = n$. In other words the set of columns $\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ of A is linearly independent, and therefore, is a basis of $\mathbb{C}^n$. Then, each of the vectors $\mathbf{e}_i$ can be expressed as a unique combination of the columns of A, that is

$$\mathbf{e}_i = b_{1i}\mathbf{c}_1 + b_{2i}\mathbf{c}_2 + \cdots + b_{ni}\mathbf{c}_n,$$

for $1 \leqslant i \leqslant n$. These equalities can be written as

$$(\mathbf{c}_1 \ \cdots \ \mathbf{c}_n)\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = I_n.$$

Consequently, the matrix A is invertible and

$$A^{-1} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}$$

Suppose now that A is invertible and that

$$d_1\mathbf{c}_1 + \cdots + d_n\mathbf{c}_n = \mathbf{0}.$$

This is equivalent to

$$A\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \mathbf{0}.$$

Multiplying both sides by A^{-1} implies

$$\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \mathbf{0},$$

so $d_1 = \cdots = d_n = 0$, which means that the set of columns of A is linearly independent, so $rank(A) = n$.

Corollary 5.91. *A matrix* $A \in \mathbb{C}^{n\times n}$ *is non-singular if and only if* $A\boldsymbol{x} = \boldsymbol{0}$ *implies* $\boldsymbol{x} = \boldsymbol{0}$ *for* $\boldsymbol{x} \in \mathbb{C}^n$.

Proof. If A is non-singular then, by Theorem 5.90, A is invertible. Therefore, $A\mathbf{x} = \mathbf{0}$ implies $A^{-1}(Ax) = A^{-1}\mathbf{0}$, so $\mathbf{x} = \mathbf{0}$.

Conversely, suppose that $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$. If $A = (\mathbf{c}_1 \ \cdots \ \mathbf{c}_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)'$, the previous implication means that $x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \mathbf{0}$ implies $x_1 = \cdots = x_n = 0$, so $\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ is linearly independent. Therefore, $\mathit{rank}(A) = n$, so A is non-singular.

Let $A \in \mathbb{C}^{n\times m}$ be a matrix. It is easy to see that the square matrix $B = A^{\mathsf{H}}A \in \mathbb{C}^{m\times m}$ is Hermitian.

Theorem 5.92. *Let $A \in \mathbb{C}^{n\times m}$ be a matrix and let $B = A^{\mathsf{H}}A$. The matrices A and B have the same rank.*

Proof. We prove that $\mathsf{NullSp}(A) = \mathsf{NullSp}(B)$. If $A\mathbf{u} = \mathbf{0}$, then $B\mathbf{u} = A^{\mathsf{H}}(A\mathbf{u}) = \mathbf{0}$, so $\mathsf{NullSp}(A) \subseteq \mathsf{NullSp}(B)$. If $\mathbf{v} \in \mathsf{NullSp}(B)$, then $A^{\mathsf{H}}A\mathbf{v} = 0$, which implies that $\mathbf{v}^{\mathsf{H}}A^{\mathsf{H}}A\mathbf{v} = 0$. This, in turn can be written as $(A\mathbf{v})^{\mathsf{H}}(A\mathbf{v}) = 0$, so, by a previous observation, we have $A\mathbf{v} = \mathbf{0}$, which means that $\mathbf{v} \in \mathsf{NullSp}(A)$. We conclude that $\mathsf{NullSp}(A) = \mathsf{NullSp}(A^{\mathsf{H}}A)$. The equalities

$$\begin{aligned}\dim(\mathsf{NullSp}(A)) + \mathit{rank}(A) &= m,\\ \dim(\mathsf{NullSp}(A)) + \mathit{rank}(A^{\mathsf{H}}A) &= m,\end{aligned}$$

imply that $\mathit{rank}(A^{\mathsf{H}}A) = m$.

Corollary 5.93. *Let $A \in \mathbb{C}^{n\times m}$ be a matrix of full-rank. If $m \geqslant n$, then the matrix $A^{\mathsf{H}}A$ is non-singular; if $n \leqslant m$, then AA^{H} is non-singular.*

Proof. Suppose that $m \geqslant n$. Then, $\mathit{rank}(A^{\mathsf{H}}A) = \mathit{rank}(A) = m$ because A is a full-rank matrix. Thus, $A^{\mathsf{H}}A \in \mathbb{C}^{m\times m}$ is non-singular. The argument for the second part of the corollary is similar.

Example 5.94. Let $A = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_m) \in \mathbb{C}^{n\times m}$. Since $AA^{\mathsf{H}} = \mathbf{a}_1\mathbf{a}_1^{\mathsf{H}} + \cdots + \mathbf{a}_m\mathbf{a}_m^{\mathsf{H}}$ it follows that the rank of the matrix $\mathbf{a}_1\mathbf{a}_1^{\mathsf{H}} + \cdots + \mathbf{a}_m\mathbf{a}_m^{\mathsf{H}}$ equals the rank of the matrix A and, therefore, it cannot exceed m.

Theorem 5.95. (Sylvester's Rank Theorem) *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times p}$ be two matrices. We have*

$$\mathit{rank}(AB) = \mathit{rank}(B) - \dim(\mathit{NullSp}(A) \cap \mathit{Ran}(B)).$$

Proof. Both $\mathsf{NullSp}(A)$ and $\mathsf{Ran}(B)$ are subspaces of $\mathbb{C}^n$, so $\mathsf{NullSp}(A) \cap \mathsf{Ran}(B)$ is a subspace of $\mathbb{C}^n$. If $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is a basis of the subspace $\mathsf{NullSp}(A) \cap \mathsf{Ran}(B)$, then exists a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \cdots, \mathbf{u}_l$ of the subspace $\mathsf{Ran}(B)$.

The set $\{A\mathbf{u}_{k+1}, \ldots, A\mathbf{u}_l\}$ is linearly independent. Indeed, suppose that there exists a linear combination

$$a_1A\mathbf{u}_{k+1} + \cdots + a_{l-k}A\mathbf{u}_l = \mathbf{0}.$$

Then, $A(a_1\mathbf{u}_{k+1}+\cdots+a_{l-k}\mathbf{u}_l)=\mathbf{0}$, so $a_1\mathbf{u}_{k+1}+\cdots+a_{l-k}\mathbf{u}_l \in \mathsf{NullSp}(A)$. Since $\mathbf{u}_{k+1},\ldots,\mathbf{u}_l \in \mathsf{Ran}(B)$, it follows that $a_1\mathbf{u}_{k+1}+\cdots+a_{l-k}\mathbf{u}_l \in \mathsf{NullSp}(A)\cap \mathsf{Ran}(B)$. Since $\mathbf{u}_1,\ldots,\mathbf{u}_k$ is a basis of the subspace $\mathsf{NullSp}(A)\cap\mathsf{Ran}(B)$, we have

$$a_1\mathbf{u}_{k+1}+\cdots+a_{l-k}\mathbf{u}_l = d_1\mathbf{u}_1+\cdots+d_k\mathbf{u}_k$$

for some $d_1,\ldots,d_k \in \mathbb{C}$, which implies

$$a_1\mathbf{u}_{k+1}+\cdots+a_{l-k}\mathbf{u}_l - d_1\mathbf{u}_1-\cdots-d_k\mathbf{u}_k=\mathbf{0}.$$

Since $\mathbf{u}_1,\ldots,\mathbf{u}_k,\mathbf{u}_{k+1},\cdots,\mathbf{u}_l$ is a basis of $\mathsf{Ran}(B)$, it follows that $a_1=\cdots=a_{l-k}=d_1=\cdots=d_k=0$, so $A\mathbf{u}_{k+1},\ldots,A\mathbf{u}_l$ is indeed linear independent.

Next, we show that $A\mathbf{u}_{k+1},\ldots,A\mathbf{u}_l$ spans the subspace $\mathsf{Ran}(AB)$. Since $\mathbf{u}_j \in \mathsf{Ran}(B)$ it is clear that $A\mathbf{u}_j\in\mathsf{Ran}(AB)$ for $k+1 \leqslant j \leqslant l$. If $\mathbf{w}\in\mathsf{Ran}(AB)$, then $\mathbf{w}=AB\mathbf{x}$ for some $\mathbf{x}\in\mathbb{C}^p$. Since $B\mathbf{x}\in\mathsf{Ran}(B)$ we can write $B\mathbf{x}=b_1\mathbf{u}_1+\cdots+b_k\mathbf{u}_k+b_{k+1}\mathbf{u}_{k+1}+\cdots+b_l\mathbf{u}_l$, which implies

$$\mathbf{w}=AB\mathbf{x}=b_{k+1}A\mathbf{u}_{k+1}+\cdots+b_lA\mathbf{u}_l,$$

because $A\mathbf{u}_1=\cdots=A\mathbf{u}_k=\mathbf{0}$, as $\mathbf{u}_1,\ldots,\mathbf{u}_k$ belong to $\mathsf{NullSp}(A)$. Thus, $A\mathbf{u}_{k+1},\ldots,A\mathbf{u}_l$ spans the subspace $\mathsf{Ran}(AB)$, which allows us to conclude that this linearly independent set is a basis for this subspace that contains $l-k$ elements. This allows us to conclude that $\mathit{rank}(AB)=\dim(\mathsf{Ran}(AB))=\mathit{rank}(B)-\dim(\mathsf{NullSp}(A)\cap\mathsf{Ran}(B))$.

Corollary 5.96. *Let $A\in\mathbb{C}^{m\times n}$. If $R\in\mathbb{C}^{m\times m}$ and $Q\in\mathbb{C}^{n\times n}$ are invertible matrices then*

$$\mathit{rank}(A)=\mathit{rank}(RA)=\mathit{rank}(AQ)=\mathit{rank}(RAQ).$$

Proof. Note that $\mathit{rank}(R)=m$ and $\mathit{rank}(Q)=n$. Thus, $\mathsf{NullSp}(R)=\{\mathbf{0}_m\}$ and $\mathsf{NullSp}(Q)=\{\mathbf{0}_n\}$. By Sylvester's Rank Theorem we have

$$\begin{aligned}\mathit{rank}(RA)&=\mathit{rank}(A)-\dim(\mathsf{NullSp}(R)\cap\mathsf{Ran}(A))\\&=\mathit{rank}(A)-\dim(\{\mathbf{0}\})=\mathit{rank}(A).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\mathit{rank}(AQ)&=\mathit{rank}(Q)-\dim(\mathsf{NullSp}(A)\cap\mathsf{Ran}(Q))\\&=n-\dim(\mathsf{NullSp}(A))=\mathit{rank}(A),\end{aligned}$$

because $\mathsf{Ran}(Q)=\mathbb{C}^n$.

The last equality of the theorem follows from the first two.

Corollary 5.97. *Let $A\in\mathbb{C}^{m\times n}$ and $B\in\mathbb{C}^{n\times p}$ be two matrices. We have*

$$\dim(\mathit{NullSp}(AB))=\dim(\mathit{NullSp}(B))+\dim(\mathit{NullSp}(A)\cap\mathit{Ran}(B)).$$

Proof. By Equality (5.6) we have:

$$\dim(\mathsf{NullSp}(AB)) + \mathit{rank}(AB) = p,$$
$$\dim(\mathsf{NullSp}(B)) + \mathit{rank}(B) = p.$$

An application of Sylvester's Rank Theorem implies

$$\dim(\mathsf{NullSp}(AB)) = \dim(\mathsf{NullSp}(B)) + \dim(\mathsf{NullSp}(A) \cap \mathsf{Ran}(B)).$$

Corollary 5.98. *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times p}$ be two matrices. We have*

$$\mathit{rank}(A) + \mathit{rank}(B) - n \leqslant \mathit{rank}(AB) \leqslant \min\{\mathit{rank}(A), \mathit{rank}(B)\},$$

and

$$\begin{aligned}\max\{\dim(\mathsf{NullSp}(A)), \dim(\mathsf{NullSp}(B))\} &\leqslant \dim(\mathsf{NullSp}(AB)) \\ &\leqslant \dim(\mathsf{NullSp}(A)) + \dim(\mathsf{NullSp}(B)).\end{aligned}$$

Proof. Since $\dim(\mathsf{NullSp}(A) \cap \mathit{rank}(B)) \leqslant \dim(\mathsf{NullSp}(A)) = n - \mathit{rank}(A)$ it follows that $\mathit{rank}(AB) \geqslant \mathit{rank}(B) - (n - \mathit{rank}(A)) = \mathit{rank}(A) + \mathit{rank}(B) - n$.

For the second inequality, observe that Sylvester's Rank Theorem implies immediately $\mathit{rank}(AB) \leqslant \mathit{rank}(B)$. Also, $\mathit{rank}(AB) = \mathit{rank}((AB)') = \mathit{rank}(B'A') \leqslant \mathit{rank}(A') = \mathit{rank}(A)$, so $\mathit{rank}(AB) \leqslant \min\{\mathit{rank}(A), \mathit{rank}(B)\}$.

The second part of the Corollary follows from the first part.

Corollary 5.99. *If $A \in \mathbb{C}^{m\times n}$ is a full-rank matrix with $m \geqslant n$, then $\mathit{rank}(AB) = \mathit{rank}(B)$ for any $B \in \mathbb{C}^{n\times p}$.*

Proof. Since $m \geqslant n$, we have $\mathit{rank}(A) = n$; therefore, the n columns of A are linearly independent so $\mathsf{NullSp}(A) = \{\mathbf{0}\}$. By Sylvester's Rank Theorem we have $\mathit{rank}(AB) = \mathit{rank}(B)$.

Theorem 5.100. (The Full-Rank Factorization Theorem) *Let $A \in \mathbb{C}^{m\times n}$ be a matrix with $\mathit{rank}(A) = r > 0$. There exists $B \in \mathbb{C}^{m\times r}$ and $C \in \mathbb{C}^{r\times n}$ such that $A = BC$.*

Furthermore, if $A = DE$, where $D \in \mathbb{C}^{m\times r}$, $E \in \mathbb{C}^{r\times n}$, then both D and E are full-rank matrices, that is, we have $\mathit{rank}(D) = \mathit{rank}(E) = r$.

Proof. Let $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\} \subseteq \mathbb{C}^m$ be a basis for the $\mathsf{Ran}(A)$. Define $B = (\mathbf{b}_1 \ \cdots \ \mathbf{b}_r) \in \mathbb{C}^{m\times r}$. The columns of A, $\mathbf{a}_1, \ldots, \mathbf{a}_n$ can be written as $\mathbf{a}_i = c_{1i}\mathbf{b}_1 + \cdots c_{ri}\mathbf{b}_r$ for $1 \leqslant i \leqslant n$, which amounts to

$$A = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_n) = (\mathbf{b}_1 \ \cdots \ \mathbf{b}_r)\begin{pmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & \cdots & \vdots \\ c_{r1} & \cdots & c_r \end{pmatrix}.$$

Thus, $A = BC$, where

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & \cdots & \vdots \\ c_{r1} & \cdots & c_r \end{pmatrix}.$$

Suppose now that $A = DE$, where $D \in \mathbb{C}^{m \times r}$, $E \in \mathbb{C}^{r \times n}$. It is clear that we have both $rank(D) \leqslant r$ and $rank(E) \leqslant r$. On another hand, by Corollary 5.98, $r = rank(A) = rank(DE) \leqslant \min\{rank(D), rank(E)\}$ implies $r \leqslant rank(D)$ and $r \leqslant rank(E)$, so $rank(D) = rank(E) = r$.

Corollary 5.101. *Let $A \in \mathbb{C}^{m \times n}$ be a matrix such that $rank(A) = r > 0$, and let $A = BC$ be a full-rank factorization of A.*

If the columns of B constitute a basis of the column space of A then C is uniquely determined. Furthermore, if the rows of C constitute a basis of the row space of A and, then B is uniquely determined.

Proof. This statement is an immediate consequence of the full-rank factorization theorem.

Corollary 5.102. *If $A \in \mathbb{C}^{m \times n}$ is a matrix with $rank(A) = r > 0$, then A can be written as*

$$A = \boldsymbol{b}_1 \boldsymbol{c}_1' + \cdots + \boldsymbol{b}_r \boldsymbol{c}_r',$$

where $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r\} \subseteq \mathbb{C}^m$ and $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_r\} \subseteq \mathbb{C}^n$ are linearly independent sets.

Proof. The corollary follows from Theorem 5.100 by adopting the set of columns of B as $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$ and the transposed rows of C as $\{\mathbf{c}_1, \ldots, \mathbf{c}_r\}$.

Theorem 5.103. *Let $A \in \mathbb{C}^{m \times n}$ be a full-rank matrix. If $m \geqslant n$, then there exists a matrix $D \in \mathbb{C}^{n \times m}$ such that $DA = I_n$. If $n \geqslant m$, then there exists a matrix $E \in \mathbb{C}^{n \times m}$ such that $AE = I_m$.*

Proof. Suppose that $A = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_n) \in \mathbb{C}^{m \times n}$ is a full-rank matrix and $m \geqslant n$. Then, the n columns of A are linearly independent and we can extend the set of columns to a basis of $\mathbb{C}^m$, $\{\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{d}_1, \ldots, \mathbf{d}_{m-n}\}$. The matrix $T = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_n \ \mathbf{d}_1 \ \cdots \ \mathbf{d}_{m-n})$ is invertible, so there exists

$$T^{-1} = \begin{pmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_n \\ \mathbf{t}_{n+1} \vdots \\ \mathbf{t}_m \end{pmatrix}$$

such that $T^{-1}T = I_m$. If we define

$$D = \begin{pmatrix} \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_n \end{pmatrix}$$

it is immediate that $DA = I_n$.

The argument for the second part is similar.

Definition 5.104. *Let $A \in \mathbb{C}^{m\times n}$. A* left inverse *of A is a matrix $D \in \mathbb{C}^{n\times m}$ such that $DA = I_n$. A* right inverse *of A is a matrix $E \in \mathbb{C}^{n\times m}$ such that $AE = I_m$.*

Theorem 5.103 can now be restated as follows. Let $A \in \mathbb{C}^{m\times n}$ be a full-rank matrix. If $m \geqslant n$, then A has a left inverse; if $n \geqslant m$, then A has a right inverse.

Corollary 5.105. *Let $A \in \mathbb{C}^{n\times n}$ be a square matrix. The following statements are equivalent.*

(i) *A has a left inverse;*
(ii) *A has a right inverse;*
(iii) *A has an inverse.*

Proof. It is clear that (iii) implies both (i) and (ii). Suppose now that A has a left inverse, so $DA = I_n$. Then, the columns of A, $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are linearly independent, for if $a_1\mathbf{c}_1 + \cdots a_n\mathbf{c}_n = \mathbf{0}$, we have $a_1D\mathbf{c}_1 + \cdots + a_nD\mathbf{c}_n = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{c}_n = \mathbf{0}$, which implies $a_1 = \cdots = a_n = 0$. Thus, *rank*$(A) = n$, so A has an inverse.

In a similar manner (using the rows of A) we can show that (ii) implies (iii).

Theorem 5.106. *Let $A \in \mathbb{C}^{m\times n}$ be a matrix with rank$(A) = r > 0$. There exists a non-singular matrix $G \in \mathbb{C}^{m\times m}$ and a non-singular matrix $H \in \mathbb{C}^{n\times n}$ such that*

$$A = G\begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r}\end{pmatrix} H.$$

Proof. By the Full-Rank Factorization Theorem (Theorem 5.100) there are two full-rank matrices $B \in \mathbb{C}^{m\times r}$ and $C \in \mathbb{C}^{r\times n}$ such that $A = BC$. Let $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$ be the columns of B and let $\mathbf{c}'_1, \ldots, \mathbf{c}'_r$ be the rows of C. It is clear that both sets of vectors are linearly independent and, therefore, for the first set there exist $\mathbf{b}_{r+1}, \ldots, \mathbf{b}_m$ such that $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ is a basis of $\mathbb{C}^m$; for the second set we have the vectors $\mathbf{c}'_{r+1}, \ldots, \mathbf{c}'_n$ such that $\{\mathbf{c}'_1, \ldots, \mathbf{c}'_n\}$ is a basis for $\mathbb{R}^n$. Define $G = (\mathbf{b}_1 \ \cdots, \mathbf{b}_m)$ and

$$H = \begin{pmatrix} \mathbf{c}'_1 \\ \vdots \\ \mathbf{c}'_n \end{pmatrix}.$$

Clearly, both G and H are non-singular and

$$A = G\begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r}\end{pmatrix} H.$$

Lemma 5.107. *If $A \in \mathbb{C}^{m\times n}$ is a matrix and $\boldsymbol{x} \in \mathbb{C}^m$, $\boldsymbol{y} \in \mathbb{C}^n$ are two vectors such that $\boldsymbol{x}^{\mathsf{H}} A\boldsymbol{y} \neq 0$, then $rank(A\boldsymbol{y}\boldsymbol{x}^{\mathsf{H}}A) = 1$.*

Proof. By the associative property of matrix product we have $A\mathbf{y}\mathbf{x}^{\mathsf{H}}A = A(\mathbf{y}\mathbf{x}^{\mathsf{H}})A$, so $rank(A\mathbf{y}\mathbf{x}^{\mathsf{H}}A) \leqslant \min\{rank(\mathbf{y}\mathbf{x}^{\mathsf{H}}, rank(A)\} = 1$, by Corollary 5.98.

We claim that $A\mathbf{y}\mathbf{x}^{\mathsf{H}}A \neq O_{m,n}$. Suppose that $A\mathbf{y}\mathbf{x}^{\mathsf{H}}A = O_{m,n}$. This implies $\mathbf{x}^{\mathsf{H}}A\mathbf{y}\mathbf{x}^{\mathsf{H}}A\mathbf{y} = 0$. If $z = \mathbf{x}^{\mathsf{H}}A\mathbf{y}$, the previous equality amounts to $z^2 = 0$, which yields $z = \mathbf{x}^{\mathsf{H}}A\mathbf{y} = 0$. This contradicts the hypothesis of the lemma, so $A\mathbf{y}\mathbf{x}^{\mathsf{H}}A \neq O_{m,n}$, which implies $rank(A\mathbf{y}\mathbf{x}^{\mathsf{H}}A) \geqslant 1$. This allows us to conclude that $rank(A\mathbf{y}\mathbf{x}^{\mathsf{H}}A) = 1$.

The rank-1 matrix $A\mathbf{y}\mathbf{x}^{\mathsf{H}}A$ discussed in Lemma 5.107 plays a central role in the next statement.

Theorem 5.108. (Wedderburn's Theorem) *Let $A \in \mathbb{C}^{m\times n}$ be a matrix. If $\boldsymbol{x} \in \mathbb{C}^m$ and $\boldsymbol{y} \in \mathbb{C}^n$ are two vectors such that $\boldsymbol{x}^{\mathsf{H}}A\boldsymbol{y} \neq 0$ and B is the matrix*

$$B = A - \frac{1}{\boldsymbol{x}^{\mathsf{H}}A\boldsymbol{y}}A\boldsymbol{y}\boldsymbol{x}^{\mathsf{H}}A,$$

then $rank(B) = rank(A) - 1$.

Proof. Observe that if $\mathbf{z} \in \mathsf{NullSp}(A)$, then $A\mathbf{z} = \mathbf{0}$. Therefore, we have

$$B\mathbf{z} = -\frac{1}{\mathbf{x}^{\mathsf{H}}A\mathbf{y}}A\mathbf{y}\mathbf{x}^{\mathsf{H}}A\mathbf{z} = \mathbf{0},$$

so $\mathsf{NullSp}(A) \subseteq \mathsf{NullSp}(B)$. Conversely, if $\mathbf{z} \in \mathsf{NullSp}(B)$, we have

$$A\mathbf{z} - \frac{1}{\mathbf{x}^{\mathsf{H}}A\mathbf{y}}A\mathbf{y}\mathbf{x}^{\mathsf{H}}A\mathbf{z} = \mathbf{0},$$

which can be written as

$$A\mathbf{z} = \frac{1}{\mathbf{x}^{\mathsf{H}}A\mathbf{y}}A\mathbf{y}(\mathbf{x}^{\mathsf{H}}A\mathbf{z}) = \frac{\mathbf{x}^{\mathsf{H}}A\mathbf{z}}{\mathbf{x}^{\mathsf{H}}A\mathbf{y}}A\mathbf{y}.$$

Thus, we obtain $A(\mathbf{z} - k\mathbf{y}) = \mathbf{0}$, where $k = \frac{\mathbf{x}^{\mathsf{H}}A\mathbf{z}}{\mathbf{x}^{\mathsf{H}}A\mathbf{y}}$. Since $A\mathbf{y} \neq \mathbf{0}$, this shows that a basis of $\mathsf{NullSp}(B)$ can be obtained by adding $\mathbf{y}$ to a basis of $\mathsf{NullSp}(A)$. Therefore, $\dim(\mathsf{NullSp}(B)) = \dim(\mathsf{NullSp}(A)) + 1$, so $rank(B) = rank(A) - 1$.

Theorem 5.109. *A square matrix $A \in \mathbb{C}^{n\times n}$ generates an increasing sequence of null spaces*

$$\{\mathbf{0}\} = \mathsf{NullSp}(A^0) \subseteq \mathsf{NullSp}(A^1) \subseteq \cdots \subseteq \mathsf{NullSp}(A^k) \subseteq \cdots$$

and a decreasing sequence of subspaces

$$\mathbb{C}^n = \mathsf{Ran}(A^0) \supseteq \mathsf{Ran}(A^1) \supseteq \cdots \supseteq \mathsf{Ran}(A^k) \supseteq \cdots$$

Furthermore, there exists a number ℓ such that

$$\mathsf{NullSp}(A^0) \subset \mathsf{NullSp}(A^1) \subset \cdots \subset \mathsf{NullSp}(A^\ell) = \mathsf{NullSp}(A^{\ell+1}) = \cdots$$

and

$$\mathsf{Ran}(A^0) \supset \mathsf{Ran}(A^1) \supset \cdots \supset \mathsf{Ran}(A^\ell) = \mathsf{Ran}(A^{\ell+1}) = \cdots$$

Proof. The proof of the existence of the increasing sequence of null subspaces and the decreasing sequence of ranges is immediate. Since $\mathsf{NullSp}(A^k) \subseteq \mathbb{C}^n$ for every k there exists a least number p such that $\mathsf{Ran}(A^p) = \mathsf{Ran}(A^{p+1})$. Therefore, $\mathsf{Ran}(A^{p+i}) = A^i\mathsf{Ran}(A^p) = A^i\mathsf{Ran}(A^{p+1}) = \mathsf{Ran}(A^{p+i+1})$ for every $i \in \mathbb{N}$. Thus, once two consecutive subspaces $\mathsf{Ran}(A^\ell)$ and $\mathsf{Ran}(A^{\ell+1})$ are equal the sequence of range subspaces stops growing.

By Equality (5.6), we have $\dim(\mathsf{Ran}(A^k)) + \dim(\mathsf{NullSp}(A^k)) = n$, so the sequence of null spaces stabilizes at the same number ℓ.

Definition 5.110. *The* index *of a square matrix $A \in \mathbb{C}^{n\times n}$ is the number ℓ defined in Theorem 5.109.*

We denote the index of a matrix $A \in \mathbb{C}^{n\times n}$ by $\mathsf{index}(A)$.

Observe that if $A \in \mathbb{C}^{n\times n}$ is a non-singular matrix, then $\mathsf{index}(A) = 0$ because in this case $\mathbb{C}^n = \mathsf{Ran}(A^0) = \mathsf{Ran}(A)$.

Theorem 5.111. *Let $A \in \mathbb{C}^{n\times n}$ be a square matrix. The following statements are equivalent:*

(i) $\mathsf{Ran}(A^k) \cap \mathsf{NullSp}(A^k) = \{\mathbf{0}\}$;
(ii) $\mathbb{C}^n = \mathsf{Ran}(A^k) \boxplus \mathsf{NullSp}(A^k)$;
(iii) $k \geqslant \mathsf{index}(A)$.

Proof. We prove this theorem by showing that (i) and (ii) are equivalent, (i) implies (iii), and (iii) implies (ii).

Suppose that the first statement holds. The set

$$T = \{\mathbf{t} \in V \mid \mathbf{t} = \mathbf{u} + \mathbf{v}, \mathbf{u} \in \mathsf{Ran}(A^k), \mathbf{v} \in \mathsf{NullSp}(A^k)\}$$

is a subspace of $\mathbb{C}^n$ and $\dim(T) = \dim(\mathsf{Ran}(A^k)) + \dim(\mathsf{NullSp}(A^k)) = n$. Therefore, $T = \mathbb{C}^n$, so $\mathbb{C}^n = \mathsf{Ran}(A^k) \boxplus \mathsf{NullSp}(A^k)$. The second statement clearly implies the first.

Suppose now that $\mathbb{C}^n = \mathsf{Ran}(A^k) \boxplus \mathsf{NullSp}(A^k)$. Then,

$$\mathsf{Ran}(A^k) = A^k\mathbb{C}^n = A\mathsf{Ran}(A^k) = \mathsf{Ran}(A^{k+1})$$

so $k \geqslant \mathsf{index}(A)$.

Conversely, if $k \geqslant \mathsf{index}(A)$ and $\mathbf{x} \in \mathsf{Ran}(A^k) \cap \mathsf{NullSp}(A^k)$, then $\mathbf{x} = A^k\mathbf{y}$ and $A^k\mathbf{x} = \mathbf{0}$, so $A^{2k}\mathbf{y} = \mathbf{0}$. Thus, $\mathbf{y} \in \mathsf{NullSp}(A^{2k}) = \mathsf{NullSp}(A^k)$, which means that $\mathbf{x} = A^k\mathbf{y} = \mathbf{0}$. Thus, the first statement holds.

5.5 Multilinear Forms

The notion of linear mapping can be extended as follows.

Definition 5.112. *Let $\mathcal{F}$ be a field and let $\{M_1, \ldots, M_n\}$ be a family of n F-linear spaces.*

*An F-*multilinear *mapping is a mapping $f : M_1 \times \cdots \times M_n \longrightarrow M$, where M is an F-linear space that is linear in each of its arguments. In other words, f satisfies the conditions*

$$\begin{aligned} & f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \sum_{j=1}^{k} a_j \boldsymbol{x}_i^j, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n) \\ &= \sum_{j=1}^{k} a_j f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_i^j, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n), \end{aligned}$$

for every $\boldsymbol{x}_i, \boldsymbol{x}_i^j \in M_i$ and $a_1, \ldots, a_k \in F$.

*If M is the field F itself, then we refer to f as an n-*linear form*. For the special case $n = 2$ we use the terms* bilinear mapping *or* bilinear form*.*

We introduce next a class of multilinear forms that plays a central role in this chapter.

Definition 5.113. *Let $\mathcal{F} = (F, \{0, 1, +, -, \cdot\})$ be a field and let M be an F-linear space. An F-multilinear form $f : M^n \longrightarrow F$ is* skew-symmetric *if $\boldsymbol{x}_i = \boldsymbol{x}_j$ for $1 \leqslant i \neq j \leqslant n$ implies $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_n) = 0$.*

The next statement shows that when two arguments of f are interchanged, then the value of f is multiplied by -1.

Theorem 5.114. *Let L be an F-linear space and let $f : L^n \longrightarrow F$ be a skew-symmetric F-multilinear form. We have*

$$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_n) = -f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_n),$$

for $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in L$.

Proof. Since f is a multilinear form we have:

$$\begin{aligned} & f(\mathbf{x}_1, \ldots, \mathbf{x}_i + \mathbf{x}_j, \ldots, \mathbf{x}_i + \mathbf{x}_j, \ldots, \mathbf{x}_n) \\ &= f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n) + f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) \\ &\quad + f(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n) + f(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) \\ &= f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) + f(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n). \end{aligned}$$

By the defining property of skew-symmetry we have the equalities

$$\begin{aligned} f(\mathbf{x}_1, \ldots, \mathbf{x}_i + \mathbf{x}_j, \ldots, \mathbf{x}_i + \mathbf{x}_j, \ldots, \mathbf{x}_n) &= 0, \\ f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n) &= 0, \\ f(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) &= 0, \end{aligned}$$

which yield

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) = -f(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n),$$

for $\mathbf{x}_1, \ldots, \mathbf{x}_n \in L$.

Corollary 5.115. *Let $\mathcal{F}$ be a field, L be an F-linear space and let $f : L^n \longrightarrow F$ be a skew-symmetric F-multilinear form.*

If $\boldsymbol{x}_i = \boldsymbol{x}_j$ for $i \neq j$, then $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_n) = 0$.

Proof. This follows immediately from Theorem 5.114.

Theorem 5.114 has the following useful extension.

Theorem 5.116. *Let L be an F-linear space and let $f : L^n \longrightarrow F$ be a skew-symmetric F-multilinear form.*

If $\phi \in PERM_n$ is a permutation given by

$$\phi : \begin{pmatrix} 1 & \cdots & i & \cdots & n \\ j_1 & \cdots & j_i & \cdots & j_n \end{pmatrix},$$

then $f(\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_n}) = (-1)^{inv(\phi)} f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ for $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in M$.

Proof. The argument is by induction on $p = inv(\phi)$. The basis case, $p = 0$ is immediate because in this case, ϕ is the identity mapping.

Suppose that the argument holds for permutations that have no more than p inversions and let ϕ be a permutation that has $p+1$ inversions. Then, as we saw in the proof of Theorem 3.8, there exists a standard transposition ψ such that for the permutation ϕ' defined as $\phi' = \psi\phi$ we have $inv(\phi') = inv(\phi) - 1$. Suppose that ϕ' is the permutation

$$\phi' : \begin{pmatrix} 1 & 2 & \cdots & \ell & \ell+1 & \cdots & n \\ j_1 & j_2 & \cdots & j_\ell & j_{\ell+1} & \cdots & j_n \end{pmatrix}$$

and ψ is the standard transposition that exchanges j_ℓ and $j_{\ell+1}$, so

$$\phi : \begin{pmatrix} 1 & 2 & \cdots & \ell & \ell+1 & \cdots & n \\ j_1 & j_2 & \cdots & j_{\ell+1} & j_\ell & \cdots & j_n \end{pmatrix}$$

By the inductive hypothesis,

$$f(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_\ell}, \mathbf{x}_{j_{\ell+1}}, \ldots, \mathbf{x}_{j_n}) = (-1)^{inv(\phi')} f(\mathbf{x}_1, \ldots, \mathbf{x}_n)$$

and

$$\begin{aligned}
&f(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_{\ell+1}}, \mathbf{x}_{j_\ell}, \ldots, \mathbf{x}_{j_n}) \\
&= -f(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_\ell}, \mathbf{x}_{j_{\ell+1}}, \ldots, \mathbf{x}_{j_n}) \\
&= -(-1)^{inv(\phi')} f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = (-1)^{inv(\phi)} f(\mathbf{x}_1, \ldots, \mathbf{x}_n),
\end{aligned}$$

which concludes the argument.

Theorem 5.117. *Let $\mathcal{F}$ be a field, L be an F-linear space, $f : L^n \longrightarrow F$ be a skew-symmetric F-multilinear form, and let $a \in F$.*

If $i \neq j$ and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in M^n$, then

$$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i + a\boldsymbol{x}_j, \ldots, \boldsymbol{x}_n).$$

Proof. Suppose that $i < j$. Then, by the linearity of f we have

$$\begin{aligned}
&f(\mathbf{x}_1, \ldots, \mathbf{x}_i + a\mathbf{x}_j, \ldots, \mathbf{x}_n) \\
&= f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n) + af(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_n) \\
&= f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_n),
\end{aligned}$$

by Corollary 5.115.

Theorem 5.118. *Let L be an F-linear space and let $f : L^n \longrightarrow \mathbb{R}$ be a skew-symmetric linear form on L. If $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is a linearly dependent subset of L, then $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = 0$.*

Proof. Suppose that $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is linearly dependent set, that is, one of the vectors can be expressed as a linear combination of the remaining vectors, say $\mathbf{x}_n = a_1\mathbf{x}_1 + \cdots + a_{n-1}\mathbf{x}_{n-1}$. Then,

$$\begin{aligned}
f(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_n) &= f(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}, a_1\mathbf{x}_1 + \cdots + a_{n-1}\mathbf{x}_{n-1}) \\
&= \sum_{i=1}^{n-1} a_i f(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_i) = 0,
\end{aligned}$$

by Corollary 5.115.

Theorem 5.119. *Let L be an n-dimensional linear space and let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ be a basis of L. There exists a unique, skew-symmetric multilinear form $d_n : L^n \longrightarrow \mathbb{R}$ such that $d_n(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = 1$.*

Proof. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be n vectors such that

$$\mathbf{x}_i = a_{i1}\mathbf{u}_1 + a_{i2}\mathbf{u}_2 + \cdots + a_{in}\mathbf{u}_n$$

for $1 \leqslant i \leqslant n$. If d_n is a skew symmetric multilinear form, $d_n : L^n \longrightarrow \mathbb{R}$, then

$$\begin{aligned}
&d_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\\
&= d_n\left(\sum_{j_1=1}^{n} a_{1j_1}\mathbf{u}_{j_1}, \sum_{j_2=1}^{n} a_{2j_2}\mathbf{u}_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{u}_{j_n}\right)\\
&= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_n=1}^{n} a_{1j_1}a_{2j_2}\cdots a_{nj_n} d_n(\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, \ldots, \mathbf{u}_{j_n})
\end{aligned}$$

We need to retain only the terms of this sum in which the arguments of $d_n(\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_n})$ are pairwise distinct (because term where $j_p = j_q$ for $p \neq q$ is zero, by Corollary 5.115). In other words, only the terms in which the list $(j_1, \ldots, j_n)$ is a permutation of $(1, \ldots, n)$ have a non-zero contribution to the sum. By Theorem 5.116, we can write

$$\begin{aligned}
&d_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\\
&= d_n(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n) \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}.
\end{aligned}$$

where the sum extends to all $n!$ permutations $(j_1, \ldots, j_n)$ of $(1, \ldots, n)$. Since $d_n(\mathbf{u}_1, \ldots, \mathbf{u}_n) = 1$, it follows that

$$d_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

Note that $d_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ is expressed using the elements of the matrix A, where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

5.6 Linear Systems

Consider the following set of linear equalities

$$\begin{aligned}
a_{11}x_1 + \ldots + a_{1n}x_n &= b_1,\\
a_{21}x_1 + \ldots + a_{2n}x_n &= b_2,\\
\vdots\quad &\ \ \vdots\\
a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m,
\end{aligned}$$

where a_{ij} and b_i belong to a field F. This set constitutes a *system of linear equations* and solving it means finding $x_1, \ldots, x_n$ that satisfy all equalities.

The system can be written succinctly in a matrix form as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \text{ and } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

In terms of linear transformations, solving this linear system amounts to determining those vectors $\mathbf{x}$ such that $h_A(\mathbf{x}) = \mathbf{b}$.

If the set of solutions of a system $A\mathbf{x} = \mathbf{b}$ is not empty we say that the system is *consistent*. Note that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b} \in \mathsf{Ran}(A)$.

Let $A\mathbf{x} = \mathbf{b}$ be a linear system in matrix form, where $A \in \mathbb{C}^{m\times n}$. The matrix $(A \ \ \mathbf{b}) \in \mathbb{C}^{m\times(n+1)}$ is the *augmented matrix* of the system $A\mathbf{x} = \mathbf{b}$.

Theorem 5.120. *Let $A \in \mathbb{C}^{m\times n}$ be a matrix and let $\boldsymbol{b} \in \mathbb{C}^{n\times 1}$. The linear system $A\boldsymbol{x} = \boldsymbol{b}$ is consistent if and only if rank$(A \ \boldsymbol{b})$ = rank(A).*

Proof. If $A\mathbf{x} = \mathbf{b}$ is consistent and $\mathbf{x}$ is a solution of this system, then $\mathbf{b} = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n$, where $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are the columns of A. This implies *rank*$(A \ \ \mathbf{b})$ = *rank*(A).

Conversely, if *rank*$(A \ \mathbf{b})$ = *rank*(A), the vector $\mathbf{b}$ is a linear combination of the columns of A, which means that $A\mathbf{x} = \mathbf{b}$ is a consistent system.

Definition 5.121. *An* homogeneous linear system *is a linear system of the form $A\boldsymbol{x} = \boldsymbol{0}_m$, where $A \in \mathbb{C}^{m\times n}$, $\boldsymbol{x} \in \mathbb{C}^{n,1}$ and $\boldsymbol{0} \in \mathbb{C}^{m\times 1}$.*

Clearly, any homogeneous system $A\mathbf{x} = \mathbf{0}_m$ is consistent and has the solution $\mathbf{x} = \mathbf{0}_n$. This solution is referred to as the *trivial solution*. The set of solutions of such a system is $\mathsf{NullSp}(A)$, the null space of the matrix A.

Let $\mathbf{u}$ and $\mathbf{v}$ be two solutions of the system $A\mathbf{x} = \mathbf{b}$. Then $A(\mathbf{u}-\mathbf{v}) = \mathbf{0}_m$, so $\mathbf{z} = \mathbf{u}-\mathbf{v}$ is a solution of the homogeneous system $A\mathbf{x} = \mathbf{0}_m$, so $\mathbf{z} \in \mathsf{NullSp}(A)$. Thus, the set of solutions of $A\mathbf{x} = \mathbf{b}$ can be obtained as a translation of the null space of A by any particular solution of $A\mathbf{x} = \mathbf{b}$. In other words the set of solution of $A\mathbf{x} = \mathbf{b}$ is $\{\mathbf{x} + \mathbf{z} \mid \mathbf{z} \in \mathsf{NullSp}(A)\}$.

Thus, for $A \in \mathbb{C}^{m\times n}$, the system $A\mathbf{x} = \mathbf{b}$ has a unique solution if and only if $\mathsf{NullSp}(A) = \{\mathbf{0}_n\}$, that is, according to Equality (5.6), if *rank*$(A) = n$.

Theorem 5.122. *Let $A \in \mathbb{C}^{n\times n}$. Then, A is invertible (which is to say that rank$(A) = n$) if and only if the system $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for every $\boldsymbol{b} \in \mathbb{C}^n$.*

Proof. If A is invertible, then $\mathbf{x} = A^{-1}\mathbf{b}$, so the system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

Conversely, if the system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{C}^n$, let $\mathbf{c}_1, \ldots, \mathbf{c}_n$ be the solution of the systems $A\mathbf{x} = \mathbf{e}_1, \ldots, A\mathbf{x} = \mathbf{e}_n$, respectively. Then, we have

$$A(\mathbf{c}_1 \ \ \cdots \ \ \mathbf{c}_n) = I_n,$$

which shows that A is invertible and $A^{-1} = (\mathbf{c}_1 | \cdots | \mathbf{c}_n)$.

Corollary 5.123. *An homogeneous linear system* $A\boldsymbol{x} = \boldsymbol{0}$*, where* $A \in \mathbb{C}^{n\times n}$ *has a non-trivial solution if and only if* A *is a singular matrix.*

Proof. This statement follows from Theorem 5.122.

Thus, by calculating the inverse of A we can solve any linear system of the form $A\mathbf{x} = \mathbf{b}$.

Definition 5.124. *A matrix* $A \in \mathbb{C}^{n\times n}$ *is* diagonally dominant *if* $|a_{ii}| > \sum\{|a_{ik}| \mid 1 \leqslant k \leqslant n \text{ and } k \neq i\}$.

Theorem 5.125. *A diagonally dominant matrix is non-singular.*

Proof. Suppose that $A \in \mathbb{C}^{n\times n}$ is a diagonally dominant matrix that is singular. By Corollary 5.123, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution $\mathbf{x} \neq \mathbf{0}$. Let x_k be a component of $\mathbf{x}$ that has the largest absolute value. Since $\mathbf{x} \neq \mathbf{0}$, we have $|x_k| > 0$. We can write

$$a_{kk}x_k = -\sum\{a_{kj}x_j \mid 1 \leqslant j \leqslant n \text{ and } j \neq k\},$$

which implies

$$\begin{aligned}
|a_{kk}|\,|x_k| &= \Big|\sum\{a_{kj}x_j \mid 1 \leqslant j \leqslant n \text{ and } j \neq k\}\Big| \\
&\leqslant \sum\{|a_{kj}|\,|x_j| \mid 1 \leqslant j \leqslant n \text{ and } j \neq k\}\Big| \\
&\leqslant |x_k| \sum\{|a_{kj}| \mid 1 \leqslant j \leqslant n \text{ and } j \neq k\}.
\end{aligned}$$

Thus, we obtain

$$|a_{kk}| \leqslant \sum\{|a_{kj}| \mid 1 \leqslant j \leqslant n \text{ and } j \neq k\},$$

which contradicts the fact that A is diagonally dominant.

5.7 Determinants

Determinants are a class of numerical multilinear functions defined on the set of square matrices. They play an important role in theoretical considerations of linear algebra and are useful for symbolic computations. As we shall see, determinants can be used to solve certain small and well-behaved linear system; however, they are of limited use for large or numerically difficult linear systems. [1]

---

[1] Historically, determinants appeared long before matrices related to solving linear systems. In modern times, determinants were introduced by Leibniz at the end of the 17th century and Cramer formula appeared in 1750. The term "determinant" was introduced by Gauss in 1801. The term "matrix", the Latin word for womb, was introduced in 1848 by James Joseph Sylvester (1814-1897), a British mathematician whose name is linked to many fundamental results in linear algebra. The term was suggested by the role of matrices as generators of determinants.

Definition 5.126. *Let* $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ *be a square matrix. The* determinant of A *is the number* $\sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$.

The determinant of A *is denoted either by* $\det(A)$ *or by:*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Example 5.127. We have $\det(I_n) = 1$ since $(I_n)_{ij} = 1$ if $j = i$ and $(I_n)_{ij} = 0$ otherwise. Thus, there exists only one non-zero term in the sam

$$\det(I_n) = \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} (I_n)_{1j_1} (I_n)_{2j_2} \cdots (I_n)_{nj_n},$$

which is obtained when $j_i = i$ for $1 \leqslant i \leqslant n$, and this unique term is 1.

Example 5.128. Let $A \in \mathbb{R}^{3 \times 3}$ be the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

The number $\det(A)$ is the sum of six terms corresponding to the six permutations of the set $\{1, 2, 3\}$, as shown below.

Permutation ϕ	$inv(\phi)$	Term
$(1,2,3)$	0	$a_{11}a_{22}a_{33}$
$(3,1,2)$	2	$a_{13}a_{21}a_{32}$
$(2,3,1)$	2	$a_{12}a_{23}a_{31}$
$(2,1,3)$	1	$-a_{12}a_{21}a_{33}$
$(3,2,1)$	3	$-a_{13}a_{22}a_{31}$
$(1,3,2)$	1	$-a_{11}a_{23}a_{32}$

Thus, we have

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} \\ &= -a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}. \end{aligned}$$

The number of terms of a determinant of order n is $n!$; this number grows very fast with n. For instance, for $n = 10$, we have $10! = 3,682,800$ terms. Thus, direct computations of determinants are very expensive.

Theorem 5.129. *Let* $A \in \mathbb{C}^{n \times n}$ *be a matrix. We have* $\det(A') = \det(A)$.

Proof. The definition of A' allows us to write

$$\det(A') = \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{j_1 1} a_{j_2 2} \cdots a_{j_n n},$$

where the sum extends to all permutations of $(1, \ldots, n)$. Due to the commutativity of numeric multiplication we can rearrange the term $a_{j_1 1} a_{j_2 2} \cdots a_{j_n n}$ as $a_{1k_1} a_{2k_2} \cdots a_{nk_n}$, where

$$\phi : \begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} \text{ and } \psi : \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$$

are inverse permutations. Since both ϕ and ψ have the same parity, it follows that

$$(-1)^{inv(j_1,\ldots,j_n)} a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} = (-1)^{inv(k_1,\ldots,k_n)} a_{1k_1} a_{2k_2} \cdots a_{nj_n},$$

which implies $\det(A') = \det(A)$.

Corollary 5.130. *If* $A \in \mathbb{C}^{n\times n}$*, then* $\det(A^{\mathsf{H}}) = \overline{\det(A)}$*. Furthermore, if* A *is a Hermitian matrix,* $\det(A)$ *is a real number.*

Proof. Let $\bar{A}$ be the matrix obtained from A by replacing each a_{ij} by its conjugate. Since conjugation of complex numbers permutes with both the sum and product of complex numbers it follows that $\det(\bar{A}) = \overline{\det(A)}$. Thus, $\det(A^{\mathsf{H}}) = \det(\bar{A})' = \det(\bar{A}) = \overline{\det(A)}$.

The second part of the corollary follows from the equality $\det(A) = \overline{\det(A)}$.

Corollary 5.131. *If* $A \in \mathbb{C}^{n\times n}$ *is a unitary matrix, then* $|\det(A)| = 1$.

Proof. Since A is unitary we have $A^{\mathsf{H}}A = AA^{\mathsf{H}} = I_n$. By Theorem 5.133, $\det(AA^{\mathsf{H}}) = \det(A)\det(A^{\mathsf{H}}) = \det(A)\overline{\det(A)} = |\det(A)|^2 = 1$. Thus, $|\det(A)| = 1$.

Theorem 5.132. *The following properties of* $\det(A)$ *hold for any* $A \in \mathbb{C}^{n\times n}$*:*

(i) $\det(A)$ *is a linear function of the rows of* A *(of the columns of* A*);*
(ii) *if two rows (columns) are permuted, then* $\det(A)$ *is changing signs;*
(iii) *if* A *has two equal rows (columns), then* $\det(A) = 0$*;*
(iv) *if a row of of a matrix, multiplied by a constant, is added to another row, then* $\det(A)$ *remains unchanged; the same holds if instead of rows we consider columns;*
(v) *if a row (column) equals* $\mathbf{0}_n$*, then* $\det(A) = 0$.

Proof. We begin with the above statements that involve columns of A.

To prove Part (i) let

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \beta\mathbf{b}_k + \gamma\mathbf{c}_k \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}, B = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \mathbf{b}_k \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \text{ and } C = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \mathbf{c}_k \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix},$$

where $\mathbf{b}_j, \mathbf{c}_j$ are row vectors and $\beta, \gamma \in \mathbb{R}$. By the definition of $\det(A)$ we have

$$\begin{aligned} \det(A) &= \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1} \cdots (\beta b_{kj_k} + \gamma c_{kj_k}) \cdots a_{nj_n} \\ &= \beta \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1} \cdots b_{kj_k} \cdots a_{nj_n} \\ &\quad + \gamma \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1} \cdots c_{kj_k} \cdots a_{nj_n} \\ &= \beta \det(B) + \gamma \det(C), \end{aligned}$$

which proves that $\det(\cdot)$ is linear.

Let now

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_p \\ \vdots \\ \mathbf{a}_q \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \text{ and let } \tilde{A} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_q \\ \vdots \\ \mathbf{a}_p \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

be the matrix obtained by swaping the p^{th} and the q^{th} row of A.

By the definition of determinants,

$$\det(\tilde{A}) = \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_p,\ldots,j_q,\ldots,j_n)} a_{1j_1} \cdots a_{qj_q} \cdots a_{pj_p} \cdots a_{nj_n}.$$

Note that the permutation $(j_1, \ldots, j_p, \ldots, j_q, \ldots, j_n)$ is obtained by the composition of $(j_1, \ldots, j_q, \ldots, j_p, \ldots, j_n)$ with the transposition that swaps j_p with j_q. Therefore, $\det(A) = -\det(\widetilde{(A)})$, which proves Part (ii).

If two rows of A are equal, then by swapping these rows we get $\det(A) = -\det(A)$, so $\det(A) = 0$, which proves Part (iii).

Part (iv) follows from the first three parts; the last part is a direct consequence of the definition of $\det(A)$.

The corresponding statements concerning rows of A follow from Theorem 5.129 because the rows of A are the transposed columns of A'.

Theorem 5.133. *Let* $A, B \in \mathbb{C}^{n\times n}$ *be two matrices. We have* $\det(AB) = \det(A)\det(B)$.

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be the rows of the matrices A and B respectively, where $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$ for $1 \leqslant i \leqslant n$. Then, the rows $\mathbf{c}_1, \ldots, \mathbf{c}_n$ of the matrix $C = AB$ are given by $\mathbf{c}_i = a_{i1}\mathbf{b}_1 + \cdots + a_{in}\mathbf{b}_n$, as it can be easily seen.

If $d_n : (\mathbb{C}^n)^n \longrightarrow \mathbb{C}$ is the skew-symmetric multilinear that defines the determinant whose existence and uniqueness was shown in Theorem 5.119, then we have

$$\begin{aligned}\det(AB) &= d_n(\mathbf{c}_1, \ldots, \mathbf{c}_i, \ldots, \mathbf{c}_n)\\ &= d_n\left(\sum_{j_1=1}^{n} a_{1j_1}\mathbf{b}_{j_1}, \ldots, \sum_{j=1}^{n} a_{ij_i}\mathbf{b}_{j_i}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}\mathbf{b}_{j_n}\right)\\ &= \sum_{j_1=1}^{n}\cdots\sum_{j_1=1}^{n}\cdots\sum_{j_1=1}^{n} a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n} d_n(\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_i}, \ldots, \mathbf{b}_{j_n}),\end{aligned}$$

due to the linearity of d_n. Observe now that only the sequences $(j_1, \ldots, j_n)$ that represent permutations of the set $\{1, \ldots, n\}$ contribute to the sum because d_n is skew-symmetric. Furthermore, if $(j_1, \ldots, j_n)$ represents a permutation ϕ, then $d_n(\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_i}, \ldots, \mathbf{b}_{j_n}) = (-1)^{inv(\phi)} d_n(\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Thus, we can write

$$\begin{aligned}&\det(AB)\\ &= \sum_{j_1=1}^{n}\cdots\sum_{j_1=1}^{n}\cdots\sum_{j_1=1}^{n} a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n} d_n(\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_i}, \ldots, \mathbf{b}_{j_n})\\ &= \left(\sum_{j_1=1}^{n}\cdots\sum_{j_1=1}^{n}\cdots\sum_{j_1=1}^{n} (-1)^{inv(j_1,\ldots,j_n)} a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n}\right) d_n(\mathbf{b}_1, \ldots, \mathbf{b}_n)\\ &= \det(A)\det(B).\end{aligned}$$

Corollary 5.134. *Let* $A \in \mathbb{R}^{n\times n}$. *We have*

$$\begin{aligned}&\det(AT_n^{i\leftrightarrow j}) = -\det(A), \det(AT_n^{i\overset{+}{\leftarrow}j}) = \det(A), \det(AT_{n,i}^{(a)}) = a\det(A).\\ &\det(T_n^{i\leftrightarrow j}A) = -\det(A), \det((T_n^{i\overset{+}{\leftarrow}j})'A) = \det(A), \det(T_{n,i}^{(a)}A) = a\det(A).\end{aligned}$$

Proof. Note that $\det(T_n^{i\leftrightarrow j} = -1$, $\det(T_n^{i\overset{+}{\leftarrow}j}) = 1$ and $\det(T_{n,i}^{(a)} = a$. The statement follows immediately from Theorem 5.133.

Lemma 5.135. *Let* $B \in \mathbb{R}^{(n+1)\times(n+1)}$ *be*

$$B = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

We have $\det(B) = \det(A)$, *where*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Proof. If $B = (b_{ij})$, then

$$b_{1j} = \begin{cases} 1 & \text{if } j = 1 \\ 0 & \text{otherwise,} \end{cases} \text{ and } b_{i1} = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Also, if $i > 1$ and $j > 1$, then $b_{ij} = a_{i-1,j-1}$ for $2 \leqslant i, j \leqslant n + 1$. By the definition of the determinant, each term of the sum that defines $\det(B)$ must include an element of the first row. However, only the first element of this row is non-zero, so

$$\begin{aligned} \det(B) &= \sum_{(j_1,j_2,\ldots,j_{n+1})} (-1)^{inv(j_1,j_2,\ldots,j_{n+1})} b_{1j_1} b_{2j_2} \ldots b_{n+1,j_{n+1}}, \\ &= \sum_{(j_2,\ldots,j_{n+1})} (-1)^{inv(1,j_2,\ldots,j_{n+1})} a_{1,j_2-1} \cdots a_{n-1,j_n-1} a_{n,j_{n+1}-1}, \end{aligned}$$

where $(j_2, \ldots, j_{n+1})$ is a permutation of the set $\{2, \ldots, n + 1\}$. Since

$$inv(1, j_2, \ldots, j_{n+1}) = inv(j_2, \ldots, j_{n+1}),$$

it follows that

$$\det(B) = \sum_{(j_2,\ldots,j_{n+1})} (-1)^{inv(j_2,\ldots,j_{n+1})} a_{1,j_2-1} a_{2,j_3-1} \ldots a_{n,j_{n+1}-1}.$$

Observe now that if $(j_2, \ldots, j_{n+1})$ is a permutation of the set $\{2, \ldots, n + 1\}$, then $(k_1, \ldots, k_n)$, where $k_i = j_{i+1} - 1$ for $1 \leqslant i \leqslant n$ is a permutation of $(1, \ldots, n)$ that has the same number of inversions as $(j_2, \ldots, j_{n+1})$. Therefore,

$$\det(B) = \sum_{(k_1,\ldots,k_n)} (-1)^{inv(k_1,\ldots,k_n)} a_{1k_1} a_{2k_2} \ldots a_{nk_n} = \det(A).$$

Lemma 5.136. *Let* $A \in \mathbb{R}^{n \times n}$ *be a matrix partitioned as:*

$$A = \left(\begin{array}{ccc|ccc}
a_{11} & \cdots & a_{1q} & a_{1,q+1} & \cdots & a_{1n} \\
\vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
a_{p1} & \cdots & a_{pq} & a_{p,q+1} & \cdots & a_{pn} \\
\hline
a_{p+1,1} & \cdots & a_{p+1,q} & a_{p+1,q+1} & \cdots & a_{p+1,n} \\
\vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
a_{n1} & \cdots & a_{nq} & a_{n,q+1} & \cdots & a_{nn}
\end{array}\right)$$

and let $B \in \mathbb{R}^{(n+1)\times(n+1)}$ be defined by

$$B = \left(\begin{array}{ccc|c|ccc}
a_{11} & \cdots & a_{1q} & 0 & a_{1,q+1} & \cdots & a_{1n} \\
\vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
a_{p1} & \cdots & a_{pq} & 0 & a_{p,q+1} & \cdots & a_{pn} \\
\hline
0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\hline
a_{p+1,1} & \cdots & a_{p+1,q} & 0 & a_{p+1,q+1} & \cdots & a_{p+1,n} \\
\vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
a_{n1} & \cdots & a_{nq} & 0 & a_{n,q+1} & \cdots & a_{nn}
\end{array}\right).$$

Then, $\det(B) = (-1)^{p+q}\det(A)$.

Proof. By permuting the $(p+1)^{\text{st}}$ row of B with each of the p rows preceding it in the matrix B and, then, by permuting the $(q+1)^{\text{st}}$ column with each of the q columns preceding it we obtain the matrix C given by

$$C = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & a_{11} & \cdots & a_{1q} & a_{1,q+1} & \cdots & a_{1n} \\
\vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
0 & a_{p1} & \cdots & a_{pq} & a_{p,q+1} & \cdots & a_{pn} \\
0 & a_{p+1,1} & \cdots & a_{p+1,q} & a_{p+1,q+1} & \cdots & a_{p+1,n} \\
\vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
0 & a_{n1} & \cdots & a_{nq} & a_{n,q+1} & \cdots & a_{nn}
\end{pmatrix}.$$

By the third part of Theorem 5.132, each of these row or column permutations multiplies $\det(B)$ by -1, so $\det(C) = (-1)^{p+q}\det(B)$ By Lemma 5.135 we have $\det(C) = \det(A)$, so $\det(B) = (-1)^{p+q}\det(A)$.

Definition 5.137. *Let $A \in \mathbb{C}^{m\times n}$. A* minor *of* order *$k$ of A is a determinant of the form*

$$\mathit{det}\left(A\begin{bmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{bmatrix}\right).$$

A principal minor of order *k of A is a determinant of the form:*

$$\det\left(A\begin{bmatrix} i_1 & \cdots & i_k \\ i_1 & \cdots & i_k \end{bmatrix}\right)$$

The leading principal minor *of order k is the determinant*

$$\det\left(A\begin{bmatrix}1 & \cdots & k\\ 1 & \cdots & k\end{bmatrix}\right).$$

For $A \in \mathbb{C}^{n\times n}$, $\det(A)$ is the unique principal minor of order n, and that the principal minors of order 1 of A are just the diagonal entries of A: $a_{11}, \ldots, a_{nn}$.

Theorem 5.138. *Let $A \in \mathbb{C}^{n\times n}$. Define the matrix $A_{ij} \in \mathbb{C}^{(n-1)\times(n-1)}$ as*

$$A_{ij} = A\begin{bmatrix}1 & \cdots & i-1 & i+1 & \cdots & n\\ 1 & \cdots & j-1 & j+1 & \cdots & h\end{bmatrix},$$

that is, the matrix obtained from A by removing the i^{th} row and the j^{th} column. Then, we have

$$\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{\ell j}) = \begin{cases}\det(A) & \textit{if } i=\ell,\\ 0 & \textit{otherwise},\end{cases}$$

for every i, ℓ, $1 \leqslant i, \ell \leqslant n$.

Proof. Let $\mathbf{x}_i$ be the i^{th} row of A, which can be expressed as

$$\mathbf{x}_i = \sum_{j=1}^{n} a_{ij}\mathbf{e}_j,$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$ such that $d_n(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$.

By the linearity of d_n we have

$$\begin{aligned}d_n(A) &= d_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)\\ &= d_n(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \sum_{j=1}^{n} a_{ij}\mathbf{e}_j, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n)\\ &= \sum_{j=1}^{n} a_{ij} d_n(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{e}_j, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n)\end{aligned}$$

The determinant $d_n(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{e}_j, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n)$ corresponds to a matrix $D^{(i,j)}$ obtained from A by replacing the i^{th} row by the sequence

$$(0, \ldots, 0, 1, 0, \ldots, 0),$$

whose unique non-zero component is on the j^{th} position. Next, by multiplying the i^{th} row by $-a_{kj}$ and adding the result to the k^{th} row for $1 \leqslant k \leqslant i-1$ and $i+1 \leqslant k \leqslant n$ yields a matrix $E^{(i,j)}$ that coincides with the matrix A with the following exceptions:

(i) the elements of row i are 0 with the exception of the j^{th} element of this row that equals 1, and
(ii) the elements of column j are 0 with the exception of the element mentioned above.

Clearly, $\det(D^{(i,j)}) = \det(E^{(i,j)})$. By applying Lemma 5.136 we obtain $\det(E^{(i,j)}) = (-1)^{i+j}\det(A_{ij})$, so

$$d_n(A) = \sum_{j=1}^{n} a_{ij} d_n(E^{(i,j)}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

which is the first case of the desired formula.

Suppose now that $i \neq \ell$. The same determinant could be computed by using an expansion on the ℓ^{th} row:

$$d_n(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{\ell j} \det(A_{\ell j}).$$

Then, $\sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{\ell j})$ is the determinant of a matrix obtained from A by replacing the ℓ^{th} row by the i^{th} row and such a determinant is 0 because the new matrix has two identical rows. This proves the second case of the equality of the theorem.

The equality of the theorem is known as the *Laplace expansion of* $\det(A)$ *by row* i.

Since the determinant of a matrix A equals the determinant of A', $\det(A)$ can be expanded by the j^{th} row as

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

for every $1 \leqslant j \leqslant n$. Thus, we have

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{\ell j}) = \begin{cases} \det(A) & \text{if } i = \ell, \\ 0 & \text{if } i \neq \ell. \end{cases}$$

This formula is *Laplace expansion of* $\det(A)$ *by column* j.

The number $\mathit{cof}(a_{ij}) = (-1)^{i+j}\det(A_{ij})$ is the *cofactor* of a_{ij} in either kind of Laplace expansion. Thus, the both types of Laplace expansions can be succinctly expressed by the equalities

$$\det(A) = \sum_{j=1}^{n} a_{ij}\, \mathit{cof}(a_{ij}) = \sum_{i=1}^{n} a_{ij}\, \mathit{cof}(a_{ij}), \tag{5.7}$$

for all $i, j \in \{1, \ldots, n\}$.

Cofactors of the form $\mathit{cof}(a_{ii})$ are known as *principal cofactors* of A.

Example 5.139. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be a sequence of n real numbers. The *Vandermonde determinant* $V_{\mathbf{a}}$ is defined by

$$V_{\mathbf{a}} = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & & \cdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

By subtracting the first line from the remaining lines we have

$$V_{\mathbf{a}} = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & & \cdots & \vdots \\ 0 & a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix} = \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

Factoring now $a_{i+1} - a_1$ from the i^{th} line of the new determinant for $1 \leqslant i \leqslant n$ yields

$$V_{\mathbf{a}} = (a_2 - a_1) \cdots (a_n - a_1) \begin{vmatrix} 1 & a_2 + a_1 & \cdots & \sum_{i=0}^{n-2} a_2^{n-2-i} a_1^i \\ \vdots & \vdots & \cdots & \vdots \\ 1 & a_n + a_1 & \cdots & \sum_{i=0}^{n-2} a_n^{n-2-i} a_1^i \end{vmatrix}$$

Consider two successive columns of this determinant:

$$\mathbf{c}_k = \begin{pmatrix} \sum_{i=0}^{k-1} a_2^{k-1-i} a_1^i \\ \vdots \\ \sum_{i=0}^{k-1} a_n^{k-1-i} a_1^i \end{pmatrix} \text{ and } \mathbf{c}_{k+1} = \begin{pmatrix} \sum_{i=0}^{k} a_2^{k-i} a_1^i \\ \vdots \\ \sum_{i=0}^{k} a_n^{k-i} a_1^i \end{pmatrix}$$

Observe that

$$\mathbf{c}_{k+1} = \begin{pmatrix} a_2^k \\ \vdots \\ a_n^k \end{pmatrix} + a_1 \mathbf{c}_k,$$

it follows that my subtracting from each column $\mathbf{c}_{k+1}$ be the previous column multiplied by a_1 (from right to left) we obtain

$$\begin{aligned} V_{\mathbf{a}} &= (a_2 - a_1) \cdots (a_n - a_1) \begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{vmatrix} \\ &= (a_2 - a_1) \cdots (a_n - a_1) V_{(a_2, \ldots, a_n)}. \end{aligned}$$

By applying repeatedly this formula, it follows that $V_{\mathbf{a}} = \prod_{p>q} (a_p - a_q)$, where $1 \leqslant p, q \leqslant n$.

Theorem 5.133 can be extended to products of rectangular matrices.

Theorem 5.140. *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$ be two matrices, where $m \leqslant n$. We have*

$$\det(AB) = \sum \Big\{ \det\left(A \begin{bmatrix} 1 & \cdots & m \\ k_1 & \cdots & k_m \end{bmatrix}\right) \det\left(B \begin{bmatrix} k_1 & \cdots & k_m \\ 1 & \cdots & m \end{bmatrix}\right) \Big| 1 \leqslant k_1 < k_2 < \cdots < k_m \leqslant n \Big\}.$$

This equality is known as the Cauchy-Binet formula*.*

Proof. Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the rows of the matrix A and let $C = AB$. The first column of the matrix AB equals $\sum_{k_1=1}^{n} \mathbf{a}_{k_1} b_{k_1 1}$ Since $\det(C)$ is linear we can write

$$\det(C) = \sum_{k_1=1}^{n} b_{k_1 1} \begin{vmatrix} a_{k_1 1} & c_{12} & \cdots & c_{1n} \\ a_{k_1 2} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{k_1 m} & c_{m2} & \cdots & c_{mn} \end{vmatrix}$$

Similarly, the second row of C equals $\sum_{k_2=1}^{n} \mathbf{a}_{k_2} b_{k_2 1}$. A further decomposition yields the sum

$$\det(C) = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} b_{k_1 1} b_{k_2 2} \begin{vmatrix} a_{k_1 1} & a_{k_2 1} & \cdots & c_{1n} \\ a_{k_1 2} & a_{k_2 2} & \cdots & c_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{k_1 m} & a_{k_2 2} & \cdots & c_{mn} \end{vmatrix},$$

and so on. Eventually, we can write

$$\det(C) = \sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} b_{k_1 1} \cdots b_{k_m m} \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & \vdots & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix}.$$

due to the multilinearity of the determinants. Only terms involving distinct numbers $k_1, \ldots, k_m$ can be retained in this sum because any such term with $k_p = k_q$ equals 0. Suppose that $\{k_1, \ldots, k_m\} = \{h_1, \ldots, h_m\}$, where $h_1 < \cdots < h_m$ and ϕ is the bijection defined by $k_i = \phi(h_i)$ for $1 \leqslant i \leqslant m$. Then,

$$\begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & \vdots & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix} = (-1)^{inv(k_1,\ldots,k_m)} \begin{vmatrix} a_{1h_1} & \cdots & a_{1h_m} \\ \vdots & \vdots & \vdots \\ a_{mh_1} & \cdots & a_{mh_m} \end{vmatrix},$$

which allows us to write

$$\det(C) = \sum_{h_1 < \cdots < h_m} \begin{vmatrix} a_{1h_1} & \cdots & a_{1h_m} \\ \vdots & \vdots & \vdots \\ a_{mh_1} & \cdots & a_{mh_m} \end{vmatrix} \sum_{\phi} (-1)^{inv(\phi)} b_{\phi(h_1)1} \cdots b_{\phi(h_m)m},$$

where ϕ is a permutation of the set $\{h_1, \ldots, h_m\}$. The last equality is equivalent to the Cauchy-Binet formula.

Example 5.141. Let $A \in \mathbb{C}^{2\times n}$ and $B \in \mathbb{C}^{n\times 2}$ be the matrices

$$A = \begin{pmatrix} a_1 \cdots a_n \\ b_1 \cdots b_n \end{pmatrix} \text{ and } B = \begin{pmatrix} c_1 & d_1 \\ \vdots & \vdots \\ c_n & d_n \end{pmatrix}.$$

Note that

$$AB = \begin{pmatrix} \sum_{i=1}^n a_i c_i & \sum_{i=1}^n a_i d_i \\ \sum_{i=1}^n b_i c_i & \sum_{i=1}^n b_i d_i \end{pmatrix}.$$

By applying Binet-Cauchy formula we obtain:

$$\begin{aligned} &\left(\sum_{i=1}^n a_i c_i\right)\left(\sum_{i=1}^n b_i d_i\right) - \left(\sum_{i=1}^n a_i d_i\right)\left(\sum_{i=1}^n b_i c_i\right) \\ &= \sum_{1 \leqslant i < j \leqslant n} (a_i b_j - a_j b_i)(c_i d_j - d_i c_j). \end{aligned}$$

This equality is known as *Lagrange's Identity.*

Let $C \in \mathbb{C}^{n\times n}$ be a square matrix and let $\mathbf{b}$ be a vector in $\mathbb{C}^n$. Denote by $(C \overset{i}{\leftarrow} \mathbf{b})$ the matrix obtained from C by replacing the i^{th} column by $\mathbf{b}$.

Example 5.142. Let $A \in \mathbb{C}^{n\times n}$ be a matrix. Then, $(A \overset{q}{\leftarrow} \mathbf{e}_p)$ is the (p,q)-minor of A and $(-1)^{p+q}\det(A \overset{q}{\leftarrow} \mathbf{e}_p)$ is the cofactor of a_{pq}.

Theorem 5.143. *Let $\{g_{ij} : \mathbb{R} \longrightarrow \mathbb{R} \mid 1 \leqslant i,j \leqslant n\}$ be a collection of n^2 differentiable functions and let $G(x)$ the matrix defined by*

$$G(x) = \begin{pmatrix} g_{11}(x) & \cdots & g_{1n}(x) \\ \vdots & \cdots & \vdots \\ g_{n1}(x) & \cdots & g_{nn}(x) \end{pmatrix}$$

for $x \in \mathbb{R}$. The derivative of the function $\det(G(x))$ is given by

$$(\det(G(x))' = \sum_{i=1}^n \det(G(x) \overset{i}{\leftarrow} \boldsymbol{g}_i(x)'),$$

where

$$\boldsymbol{g}_i(x)' = \begin{pmatrix} g'_{1i} \\ \vdots \\ g'_{ni} \end{pmatrix}$$

is the column of the derivatives of the functions positioned in column i of the matrix $G(x)$, for $1 \leqslant i \leqslant n$.

Proof. By the definition of determinants we have:

$$\det(G(x)) = \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} g_{1j_1}(x) g_{2j_2}(x) \cdots g_{nj_n}(x)$$

Therefore, we can write

$$\begin{aligned}\det(G(x))' &= \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} g'_{1j_1}(x) g_{2j_2}(x) \cdots g_{nj_n}(x) \\ &\quad + \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} g_{1j_1}(x) g_{2j_2}(x)' \cdots g_{nj_n}(x) \\ &\quad + \sum_{j_1,\ldots,j_n} (-1)^{inv(j_1,\ldots,j_n)} g_{1j_1}(x) g_{2j_2}(x) \cdots g_{nj_n}(x)' \\ &= \sum_{i=1}^{n} \det(G(x) \overset{i}{\leftarrow} \mathbf{g}_i(x)'),\end{aligned}$$

which concludes the argument.

Example 5.144. Let $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ and let $G(x) = \det(A - xI_3)$. We have

$$\begin{aligned}&(\det(G(x))' \\ &= \det(G(x) \overset{1}{\leftarrow} (-\mathbf{e}_1)) + \det(G(x) \overset{2}{\leftarrow} (-\mathbf{e}_2) + \det(G(x) \overset{3}{\leftarrow} (-\mathbf{e}_3)) \\ &= \begin{vmatrix} -1 & a_{12} & a_{13} \\ 0 & a_{22}-x & a_{23} \\ 0 & a_{32} & a_{33}-x \end{vmatrix} + \begin{vmatrix} a_{11}-x & 0 & a_{13} \\ a_{21} & -1 & a_{23} \\ a_{31} & 0 & a_{33}-x \end{vmatrix} + \begin{vmatrix} a_{11}-x & a_{12} & 0 \\ a_{21} & a_{22}-x & 0 \\ a_{31} & a_{32} & -1 \end{vmatrix} \\ &= -\begin{vmatrix} a_{22}-x & a_{23} \\ a_{32} & a_{33}-x \end{vmatrix} - \begin{vmatrix} a_{11}-x & a_{13} \\ a_{31} & a_{33}-x \end{vmatrix} - \begin{vmatrix} a_{11}-x & a_{12} \\ a_{21} & a_{22}-x \end{vmatrix}.\end{aligned}$$

The same technique is applied to compute the second derivative

$$\begin{aligned}(\det(G(x))'' &= -\begin{vmatrix} -1 & a_{23} \\ 0 & a_{33}-x \end{vmatrix} - \begin{vmatrix} a_{22}-x & 0 \\ a_{32} & -1 \end{vmatrix} \\ &\quad -\begin{vmatrix} -1 & a_{13} \\ 0 & a_{33}-x \end{vmatrix} - \begin{vmatrix} a_{11}-x & 0 \\ a_{31} & -1 \end{vmatrix} \\ &\quad -\begin{vmatrix} -1 & a_{12} \\ 0 & a_{22}-x \end{vmatrix} - \begin{vmatrix} a_{11}-x & 0 \\ a_{21} & -1 \end{vmatrix} \\ &= 2(a_{11} + a_{22} + a_{33} - 3x).\end{aligned}$$

Note that $G(0) = \det(A)$, $-G'(0)$ equals the sum of order 2 principal minors of A, while $G''(0)$ is twice the sum of order 1 principal minors of A.

This observation can be generalized to square matrices of any size: if $A \in \mathbb{R}^{n\times n}$, then for the k^{th} derivative of the function $G(x) = \det(A - xI_n)$ we have

$$G^{(k)}(0) = (-1)^k k! S_{n-k}(A),$$

where $S_p(A)$ is the sum of all order-p principal minors of A (see Exercise 52).

Example 5.145. Let $Q_n(a,b)$ be the determinant of order n:

$$Q_n(a,b) = \begin{vmatrix} a & 1 & 1 & 1 & \cdots & 1 \\ 1 & b & 1 & 0 & \cdots & 0 \\ 1 & 0 & b & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & b \end{vmatrix}$$

We have $Q_2(a,b) = ab - 1$. To compute the value of $Q_n(a,b)$ note that, by expanding this determinant by its last column we have:

$$Q_n(a,b) = (-1)^{n+1} \begin{vmatrix} 1 & b & 0 & 0 & \cdots & 0 \\ 1 & 0 & b & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & b \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{vmatrix} + bQ_{n-1}(a,b) = -b^{n-2} + bQ_{n-1}(a,b).$$

It is easy to verify that $Q_n(a,b) = b^{n-1}a - (n-1)b^{n-2}$ for $n \geqslant 2$.

Theorem 5.146. *Let $A, B \in \mathbb{C}^{n\times n}$. Then $\det(A+B)$ is equal to the sum of the determinants of the 2^n matrices obtained by replacing each subset of the columns of A by the corresponding subset of columns of B.*

Proof. Let $A = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_n)$ and $B = (\mathbf{b}_1 \ \cdots \ \mathbf{b}_n)$. Since $\det(A+B) = \det(\mathbf{a}_1 + \mathbf{b}_1 \ \cdots \ \mathbf{a}_n + \mathbf{b}_n)$, by the linearity of determinants, we can write

$$\begin{aligned}
&\det(\mathbf{a}_1 + \mathbf{b}_1 \ \cdots \ \mathbf{a}_n + \mathbf{b}_n) \\
&= \det(\mathbf{a}_1 \ \mathbf{a}_2 + \mathbf{b}_2 \ \cdots, \mathbf{a}_n + \mathbf{b}_n) + \det(\mathbf{b}_1 \ \mathbf{a}_2 + \mathbf{b}_2 \ \cdots, \mathbf{a}_n + \mathbf{b}_n) \\
&= \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots, \mathbf{a}_n + \mathbf{b}_n) + \det(\mathbf{a}_1 \ \mathbf{b}_2 \ \ldots, \mathbf{a}_n + \mathbf{b}_n) \\
&\quad + \det(\mathbf{b}_1 \ \mathbf{a}_2 \ \cdots, \mathbf{a}_n + \mathbf{b}_n) + \det(\mathbf{b}_1 \ \mathbf{b}_2 \ \ldots, \mathbf{a}_n + \mathbf{b}_n) \\
&= \cdots \\
&= \det(\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n) + \det(\mathbf{a}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{a}_n) + \cdots + \det(\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n).
\end{aligned}$$

In principle, determinants can be used for solving linear systems of equation and we discuss a formula that allows us to do just that. Let $A \in \mathbb{C}^{n\times n}$, $\mathbf{b} \in \mathbb{C}^n$ and consider the linear system $A\mathbf{x} = \mathbf{b}$. The columns of the matrix A are denoted by $\mathbf{a}_1, \ldots, \mathbf{a}_n$, that is, $A = (\mathbf{a}_1 \ \cdots \ \mathbf{a}_n)$. Note that

$$\begin{aligned}
A(I_n \overset{i}{\leftarrow} \mathbf{x}) &= A(\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_{i-1} \ \mathbf{x} \ \mathbf{e}_{i+1} \ \cdots \ \mathbf{e}_n) \\
&= (\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_{i-1} \ \mathbf{b} \ \mathbf{a}_{i+1} \ \cdots \ \mathbf{a}_n) \\
&= (A \overset{i}{\leftarrow} \mathbf{b}). \qquad (5.8)
\end{aligned}$$

By expanding the determinant

$$\det(I_n \overset{i}{\leftarrow} \mathbf{x}) = \begin{vmatrix} 1 & 0 & \cdots & x_1 & \cdots & 0 \\ 0 & 1 & \cdots & x_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & x_i & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & x_n & \cdots & 1 \end{vmatrix}$$

by its i^{th} row we obtain $\det(I_n \overset{i}{\leftarrow} \mathbf{x}) = x_i$. Thus, computing the determinants on both sides of Equality (5.8) we have

$$\det(A)x_i = \det(A \overset{i}{\leftarrow} \mathbf{b}),$$

for $1 \leqslant i \leqslant n$. This method for computing the components of the solution of the system $A\mathbf{x} = \mathbf{b}$ is known as *Cramer's formula.*

Definition 5.147. *Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. The* adjoint matrix *of A is the matrix*

$$\mathit{adj}(A) = \begin{pmatrix} \mathit{cof}(a_{11}) & \cdots & \mathit{cof}(a_{n1}) \\ \vdots & \cdots & \vdots \\ \mathit{cof}(a_{1n}) & \cdots & \mathit{cof}(a_{nn}) \end{pmatrix},$$

that is, the transposed matrix of the matrix whose entries are the cofactors of the elements of A.

Example 5.148. For the matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix}$$

the matrix of cofactors is

$$C = \begin{pmatrix} -3 & 1 & 1 \\ 4 & -1 & -2 \\ 2 & -1 & -1 \end{pmatrix},$$

so the adjoint matrix is

$$\mathsf{adj}(A) = \begin{pmatrix} -3 & 4 & 2 \\ 1 & -1 & -1 \\ 1 & -2 & -1 \end{pmatrix}.$$

Theorem 5.149. *If $A \in \mathbb{R}^{n\times n}$, then $A\ \mathit{adj}(A) = \det(A)I_n$.*

Proof. Equalities (5.7) allow us to write:

$$(A\ \mathsf{adj}(A))_{ij} = \sum_{k=1}^{n} a_{ik}(\mathsf{adj}(A))_{kj} = \sum_{k=1}^{n} a_{ik}\,\mathit{cof}(a_{jk}) = \begin{cases} \det(A) & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $A\ \mathsf{adj}(A) = \det(A)I_n$.

This allows us to give an explicit formula for computing the inverse of a non-singular matrix A as

$$A^{-1} = \frac{1}{\det(A)}\mathsf{adj}(A). \tag{5.9}$$

By Theorem 5.133, if A is a non-singular matrix, we have $\det(A)\det(A^{-1}) = 1$.

Corollary 5.150. *If* $A \in \mathbb{R}^{n\times n}$ *then* $\det(A) \neq 0$ *if and only rank*$(A) = n$*.*

Proof. By Theorem 5.122 *rank*$(A) = n$ if and only if A is invertible; A is invertible if and only if $\det(A) \neq 0$.

Example 5.151. For the matrix $A \in \mathbb{R}^{3\times 3}$ introduced in Example 5.148 we have $det(A) = -1$ and

$$A^{-1} = \frac{1}{\det(A)}\mathsf{adj}(A) = \begin{pmatrix} 3 & -4 & -2 \\ -1 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix}.$$

5.8 Partitioned Matrices and Determinants

Lemma 5.152. *Let* $A \in \mathbb{C}^{n\times n}$*,* $B \in \mathbb{C}^{m\times m}$ *be two square matrices and let* $D \in \mathbb{C}^{(m+n)\times(m+n)}$ *be the matrix*

$$D = \begin{pmatrix} A & O_{m,n} \\ C & B \end{pmatrix},$$

where $C \in \mathbb{C}^{m\times n}$*. Then, we have* $\det(D) = \det(A)\det(B)$*.*

If the matrix $F \in \mathbb{C}^{(m+n)\times(m+n)}$ *is*

$$F = \begin{pmatrix} A & E \\ O_{m,n} & B \end{pmatrix},$$

where $E \in \mathbb{C}^{n\times m}$*, then* $\det(F) = \det(A)\det(B)$*.*

Proof. Suppose that $D = (d_{ij})$. The definition of $\det(D)$ implies that

$$\det(D) = \sum_{j_1,\ldots,j_n,j_{n+1},\ldots,j_{n+m}} (-1)^{inv(j_1,\ldots,j_{n+m})} d_{1j_1} \cdots d_{n+m\, j_{n+m}}.$$

Each term of this sum involves factors chosen from each row (specified by the first subscript of d_{ij}). Note that any term

$$d_{1j_1} \cdots d_{nj_n} d_{n+1\, j_{n+1}} \cdots d_{n+m\, j_{n+m}}$$

in which any of the first n subscripts $j_1, \ldots, j_n$ is at least equal to $n+1$ equals 0. Therefore, non-zero terms are those in which $(j_1, \ldots, j_n)$ is a permutation of $(1, \ldots, n)$. In such terms $(j_{n+1}, \ldots, j_{n+m})$ is a permutation of $(n+1, \ldots, n+m)$. By the definition of the matrix D the product $d_{1j_1} \cdots d_{nj_n}$ is actually $a_{1j_1} \cdots a_{nj_n}$; the product $d_{n+1\, j_{n+1}} \cdots d_{n+m\, j_{n+m}}$ equals $b_{1k_1} \cdots b_{mk_m}$, where $k_1 = j_{n+1} - n, \ldots, k_m = j_{n+m} - n$ and

$$inv(j_1, \ldots, j_n, j_{n+1}, \ldots, j_{n+m}) = inv(j_1, \ldots, j_n) + inv(k_1, \ldots, k_m).$$

This allows us to write

$$\det(D) = \left(\sum_{j_1 \cdots j_n} a_{1j_1} \cdots a_{nj_n} \right) \cdot \left(\sum_{k_1 \cdots k_m} b_{1k_1} \cdots b_{mk_m} \right) = \det(A)\det(B).$$

The proof of the second part of the lemma is similar.

Theorem 5.153. *Let A be an block upper (or lower) triangular partitioned matrix given by*

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ O & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ O & O & \cdots & A_{mm} \end{pmatrix},$$

where $A_{ii} \in \mathbb{R}^{p_i \times p_i}$ for $1 \leqslant i \leqslant m$. Then,

$$\det(A) = \det(A_{11})\det(A_{22}) \cdots \det(A_{mm}).$$

If A is a block lower triangular matrix

$$A = \begin{pmatrix} A_{11} & O & \cdots & O \\ A_{21} & A_{22} & \cdots & O \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{pmatrix},$$

the same equality holds.

Proof. The argument is by induction on m, where $m \geqslant 2$. The base case, $m = 2$ was shown in Lemma 5.152. Suppose that the statement holds for partitioned upper diagonal matrices having $m-1$ diagonal blocks. Note that if

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ O & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ O & O & \cdots & A_{mm} \end{pmatrix},$$

we can also regard A as a partitioned matrix

$$A = \begin{pmatrix} B & C \\ O & A_{mm} \end{pmatrix},$$

where B is a partitioned upper diagonal matrices having $m-1$ diagonal blocks. Then, by the base case and by the inductive hypothesis we have

$$\det(A) = \det(B)\det(A_{mm}) = \det(A_1)\cdots\det(A_{m-1})\det(A_{mm}).$$

Theorem 5.154. *Let* $A \in \mathbb{C}^{m\times m}$, $D \in \mathbb{C}^{n\times n}$ *be two square matrices and let* $E \in \mathbb{C}^{(m+n)\times(m+n)}$ *be the matrix*

$$E = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where $B \in \mathbb{C}^{m\times n}$ *and* $C \in \mathbb{C}^{n\times m}$*. If the matrix* A *is invertible, then* $\det(E) = \det(A)\det(D - CA^{-1}B)$.

Proof. If A is invertible, then

$$\begin{pmatrix} I_m & O \\ -CA^{-1} & I_n \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix},$$

which implies

$$\det(E) = \det\begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}.$$

Theorem 5.153 implies the desired equality.

Theorem 5.155. *Let* $A, B, C, D \in \mathbb{C}^{m\times m}$ *be four square matrices such that* $AC = CA$ *and let* $E \in \mathbb{C}^{2m\times 2m}$ *be the matrix*

$$E = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

We have $\det(E) = det(AD - CB)$.

Proof. Suppose initially that A is invertible, so $\det(E) = \det(A^{-1})\det(D - CA^{-1}B)$ by Theorem 5.154. Then,

$$\begin{aligned}\det(E) &= \det(A)\det(D - CA^{-1}B) = \det(AD - ACA^{-1}B) \\ &= \det(AD - CAA^{-1}B) = \det(AD - CB).\end{aligned}$$

If A is not invertible, that is, if A is singular consider the continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = \det(A + xI)$ for $x \in \mathbb{R}$. There exists $\delta > 0$ such that if $x \in (0, \delta)$, then $f(x) \neq 0$, which means that $\det(A+xI) \neq 0$, which implies that $A + xI$ is an invertible matrix. Note that if $AC = CA$, then $(A+xI)C = C(A+xI)$, so the first part of the argument can be applied to the matrix

$$E_x = \begin{pmatrix} A + xI & B \\ C & D \end{pmatrix}.$$

This implies $\det(E_x) = \det((A + xI)D - CB)$ if $x \in (0, \delta)$. If x tends towards 0, by the continuity of $f(x)$ it follows that

$$\det(E) = \lim_{x \to 0} \det((A + xI)D - CB) = \det(AD - BC),$$

which concludes our argument.

The argument presented in the second part of the proof of Theorem 5.155 is known as a *continuity argument.*

Theorem 5.156. *Let $A \in \mathbb{C}^{n \times n}$ be a square matrix. The rank of A equals the largest size of a non-zero minor of A.*

Proof. Let $r = \mathit{rank}(A)$ and let s be the largest size of a non-zero minor of A that is the determinant of the submatrix $S = A \begin{bmatrix} i_1 & \cdots & i_s \\ j_1 & \cdots & j_s \end{bmatrix}$. By permuting the rows and the columns of A we obtain a matrix B of the same rank as A such that $B \begin{bmatrix} 1 & \cdots & s \\ 1 & \cdots & s \end{bmatrix} = S$. Since any permutation of rows or columns preserves the non-nullity of a determinant we have

$$\det \left(B \begin{bmatrix} 1 & \cdots & s \\ 1 & \cdots & s \end{bmatrix} \right) \neq 0.$$

Thus, the rows of S are linearly independent and, therefore, the first s rows of B are linearly independent, so $s \leqslant r$.

Since $\mathit{rank}(A) = r$, A has r linearly independent rows. By permuting the rows of A, these rows can be brought in the first r position in a matrix B which has the same rank r as A. Then, r linearly independent columns of B are brought on the first r position to result into a matrix C. Let $P = C \begin{bmatrix} 1, \ldots, r \\ 1, \ldots, r \end{bmatrix}$. Since P is of rank r it follows that $\det(P) \neq 0$, so $r \leqslant s$.

5.9 The Kronecker and Hadamard products

Definition 5.157. *Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times q}$ be two matrices. The* Kronecker product *of these matrices is the matrix $A \otimes B \in \mathbb{C}^{mp \times nq}$ defined by*

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

Example 5.158. Consider the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

Their Kronecker product is

$$A \otimes B = \left(\begin{array}{ccc|ccc} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\ a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} \\ a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & a_{12}b_{31} & a_{12}b_{32} & a_{12}b_{33} \\ \hline a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \\ a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} \\ a_{21}b_{31} & a_{21}b_{32} & a_{21}b_{33} & a_{22}b_{31} & a_{22}b_{32} & a_{22}b_{33} \end{array}\right).$$

The next theorem contains a few elementary properties of Kronecker's product.

Theorem 5.159. *For any matrices A, B, C, D we have:*
(i) $(A \otimes B)' = A' \otimes B'$,
(ii) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$,
(iii) $(A \otimes B)(C \otimes D) = (AC \otimes BD)$,
(iv) $A \otimes B + A \otimes C = A \otimes (B + C)$,
(v) $A \otimes D + B \otimes D = (A + B) \otimes D$,
(vi) $(A \otimes B)' = A' \otimes B'$,
(vii) $(A \otimes B)^{\mathsf{H}} = A^{\mathsf{H}} \otimes B^{\mathsf{H}}$,
when the usual matrix sum and multiplication are well-defined in each of the above equalities.

Proof. The proof is straightforward and is left to the reader.

Example 5.160. Let $\mathbf{x} \in \mathbb{C}^n$ and $\mathbf{y} \in \mathbb{C}^m$. We have

$$\mathbf{x} \otimes \mathbf{y} = \begin{pmatrix} x_1\mathbf{y} \\ \vdots \\ x_n\mathbf{y} \end{pmatrix} = \begin{pmatrix} y_1\mathbf{x} \\ \vdots \\ y_m\mathbf{x} \end{pmatrix} \in \mathbb{C}^{mn}.$$

Theorem 5.161. *If $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{m \times m}$ are two invertible matrices, then $A \otimes B$ is invertible and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.*

Proof. Since

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1} \otimes BB^{-1}) = I_n \otimes I_m,$$

the theorem follows by noting that $I_n \otimes I_m = I_{nm}$.

Theorem 5.162. *The Kronecker product $A \otimes B$ of two normal (unitary) matrices A, B is a normal (a unitary) matrix.*

Proof. By Theorem 5.159 we can write

$$\begin{aligned}(A \otimes B)'(A \otimes B) &= (A' \otimes B')(A \otimes B)\\ &= (A'A \otimes B'B)\\ &= (AA' \otimes BB')\\ &\quad \text{(because both } A \text{ and } B \text{ are normal)}\\ &= (A \otimes B)(A \otimes B)',\end{aligned}$$

which implies that $A \otimes B$ is normal.

Definition 5.163. *Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be two square matrices. Their* Kronecker sum *is the matrix $A \oplus B \in \mathbb{C}^{mn\times mn}$ defined by*

$$A \oplus B = (A \otimes I_n) + (I_m \otimes B).$$

The Kronecker difference *is the matrix $A \ominus B \in \mathbb{C}^{mn\times mn}$ defined by*

$$A \ominus B = (A \otimes I_n) - (I_m \otimes B).$$

Definition 5.164. *Let $A, B \in \mathbb{C}^{m\times n}$. The* Hadamard product *of A and B is the matrix $A \odot B \in \mathbb{C}^{m\times n}$ defined by*

$$A \odot B = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn} \end{pmatrix}.$$

The Hadamard quotient *$A \oslash B$ is defined only if $b_{ij} \neq 0$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. In this case*

$$A \oslash B = \begin{pmatrix} \frac{a_{11}}{b_{11}} & \frac{a_{12}}{b_{12}} & \cdots & \frac{a_{1n}}{b_{1n}} \\ \frac{a_{21}}{b_{21}} & \frac{a_{22}}{b_{22}} & \cdots & \frac{a_{2n}}{b_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{m1}}{b_{m1}} & \frac{a_{m2}}{b_{m2}} & \cdots & \frac{a_{mn}}{b_{mn}} \end{pmatrix}.$$

Theorem 5.165. *If $A, B, C \in \mathbb{C}^{m\times n}$ and $c \in \mathbb{C}$ we have*
(i) *$A \odot B = B \odot A$;*
(ii) *$A \odot J_{m,n} = J_{m,n} \odot A = A$;*
(iii) *$A \odot (B + C) = A \odot B + A \odot C$;*
(iv) *$A \odot (cB) = c(A \odot B)$.*

Proof. The proof is straightforward and is left to the reader.

Note that the Hadamard product of two matrices $A, B \in \mathbb{C}^{m\times n}$ is a submatrix of the Kronecker product $A \otimes B$.

Example 5.166. Let $A, B \in \mathbb{C}^{2\times 3}$ be the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}.$$

The Kronecker product of these matrices is $A \otimes B \in \mathbb{C}^{4\times 9}$ given by:

$$A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} & a_{13}b_{11} & a_{13}b_{12} & a_{13}b_{13} \\ a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} & a_{13}b_{21} & a_{13}b_{22} & a_{13}b_{23} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} & a_{23}b_{11} & a_{23}b_{12} & a_{23}b_{13} \\ a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} & a_{23}b_{21} & a_{23}b_{22} & a_{23}b_{23} \end{pmatrix}.$$

The Hadamard product of the same matrices is

$$A \odot B = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & a_{13}b_{13} \\ a_{21}b_{21} & a_{22}b_{22} & a_{23}b_{23} \end{pmatrix},$$

and we can regard the Hadamard product as a submatrix of the Kronecker product $A \otimes B$,

$$A \odot B = (A \otimes B) \begin{bmatrix} 1, 5, 9 \\ 4, 4, 4 \end{bmatrix}.$$

5.10 Topological Linear Spaces

We are examining now the interaction between the algebraic structure of linear spaces and topologies that can be defined on linear spaces that are compatible in a certain sense with the algebraic structure. Compatibility, in this case, is defined as the continuity of addition and scalar multiplication.

Definition 5.167. *Let F be the real field $\mathbb{R}$ or complex field $\mathbb{C}$. An F-*topological linear space *is a topological space $(V, \mathcal{O})$ such that*
(i) *V is an F-linear space;*
(ii) *the vector addition is a continuous function between V^2 and V;*
(iii) *the scalar multiplication is a continuous function between $F \times V$ and V.*

Unless stated otherwise, we assume that the field F is either the real or the complex field.

Theorem 5.168. *Let $(V, \mathcal{O})$ be an F-*topological linear space *and let $\boldsymbol{z} \in V$. The translation mapping $\mathtt{t}_{\mathbf{z}} : V \longrightarrow V$ is a homeomorphism.*

Proof. It is immediate that $\mathtt{t}_{\mathbf{z}}$ is a bijection whose inverse is $\mathtt{t}_{-\mathbf{z}}$. The continuity of both $\mathtt{t}_{\mathbf{z}}$ and $\mathtt{t}_{-\mathbf{z}}$ follows from the continuity of the vector addition of V.

Example 5.169. If $a \neq 0$, then each homotety h_a of a topological linear space $(V, \mathcal{O})$ is a homeomorphism. Indeed, the inverse of h_a is $\mathsf{h}_{a^{-1}}$. The continuity of both h_a and $\mathsf{h}_{a^{-1}}$ follows from the continuity of the scalar multiplication of V.

Theorem 5.170. *Let $(V, \mathcal{O})$ be a topological linear space. If W is a neighborhood of $\mathbf{0}$, then $\mathsf{t}_{\boldsymbol{x}}(W)$ is a neighborhood of $\boldsymbol{x}$. Moreover, every neighborhood of x can be obtained by a translation of a neighborhood of $\mathbf{0}$.*

Proof. Since W is a neighborhood of the origin, there exists an open subset L of V such that $\mathbf{0} \in L \subseteq W$. This implies $\mathbf{x} = \mathsf{t}_{\mathbf{x}}(\mathbf{0}) \in \mathsf{t}_{\mathbf{x}}(L) \subseteq \mathsf{t}_{\mathbf{x}}(W)$. Since every translation is a homeomorphism of $(V, \mathcal{O})$ it follows that $\mathsf{t}_{\mathbf{x}}(L)$ is an open set and this, in turn, implies that $\mathsf{t}_{\mathbf{x}}(W)$ is a neighborhood of $\mathbf{x}$.

Conversely, let U be a neighborhood of $\mathbf{x}$ and let K be an open set such that $\mathbf{x} \in K \subseteq U$. Then, we have $\mathbf{0} = \mathsf{t}_{-\mathbf{x}}(\mathbf{x}) \in \mathsf{t}_{-\mathbf{x}}(K) \subseteq \mathsf{t}_{-\mathbf{x}}(U)$. Since $\mathsf{t}_{-\mathbf{x}}(K)$ is an open set, it follows that $\mathsf{t}_{-\mathbf{x}}(K)$ is a neighborhood of $\mathbf{0}$ and the desired conclusion follows from the fact that $U = \mathsf{t}_{\mathbf{x}}(\mathsf{t}_{-\mathbf{x}}(U))$.

Theorem 5.170 shows that in a topological linear space the neighborhoods of any point are obtained by translating the neighborhoods of $\mathbf{0}$.

Corollary 5.171. *If $\mathcal{F}_x$ is a fundamental system of neighborhoods of $\boldsymbol{x}$ in the topological linear space $(V, \mathcal{O})$, then $\mathcal{F}_x$ can be obtained by a translation of a fundamental system of neighborhoods $\mathcal{F}_0$ of $\mathbf{0}$.*

Proof. This statement follows immediately from Theorem 5.170.

The next theorem shows that a linear function between two topological linear spaces is continuous if and only if it is continuous in the zero element of the first space.

Theorem 5.172. *Let $(V_1, \mathcal{O}_1)$ and $(V_2, \mathcal{O}_2)$ be two topological F-linear spaces having $\mathbf{0}_1$ and $\mathbf{0}_2$ as zero elements, respectively. A linear operator $f \in \mathsf{Hom}(V_1, V_2)$ is continuous in $\boldsymbol{x} \in V_1$ if and only if it is continuous in $\mathbf{0}_1 \in V_1$.*

Proof. Let f be a function that is continuous in a point $\mathbf{x} \in V_1$. If $U \in neigh_{\mathbf{0}_2}(\mathcal{O}_2)$, then $f(\mathbf{x}) + U$ is a neighborhood of $f(\mathbf{x})$. Since f is continuous, there exists a neighborhood W of $\mathbf{x}$ such that $f(W) \subseteq f(\mathbf{x}) + U$.

Observe that the set $-\mathbf{x} + W$ is a neighborhood of $\mathbf{0}_1$. Moreover, any neighborhood of $\mathbf{0}_1$ has this form. If $\mathbf{t} \in -\mathbf{x} + W$, then $\mathbf{t} + \mathbf{x} \in W$ and, therefore, $f(\mathbf{t}) + f(\mathbf{x}) = f(\mathbf{t} + \mathbf{x}) \in f(\mathbf{x}) + U$. This shows that $f(\mathbf{t}) \in U$, which proves that f is continuous in $\mathbf{0}_1$.

Conversely, suppose that f is continuous in $\mathbf{0}_1$. Let $\mathbf{x} \in V_1$ and let $Z \in neigh_{f(\mathbf{x})}(\mathcal{O}_2)$. The set $-f(\mathbf{x}) + Z$ is a neighborhood of $\mathbf{0}_2$ in V_2. The continuity of f in $\mathbf{0}_1$ implies the existence of a neighborhood T of $\mathbf{0}_1$ such that $f(T) \subseteq -f(\mathbf{x}) + Z$. Note that $\mathbf{x} + T$ is a neighborhood of $\mathbf{x}$ in V_1 and every neighborhood of $\mathbf{x}$ in V_1 has this form. Since $f(\mathbf{x} + T) \subseteq Z$, it follows that f is continuous in $\mathbf{x}$.

Corollary 5.173. *Let $(V_1, \mathcal{O}_1)$ and $(V_2, \mathcal{O}_2)$ be two topological F-linear spaces. A linear operator $f \in \mathit{Hom}(V_1, V_2)$ is either continuous on V_1 or is discontinuous in every point of V_1.*

Proof. This statement is a direct consequence of Theorem 5.172.

Theorem 5.174. *Let C and D be two subsets of $\mathbb{R}^n$ such that C is compact and D is closed. Then the set $C + D = \{\boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} \in C, \boldsymbol{y} \in D\}$ is closed.*

Proof. Let $\mathbf{x} \in \mathbf{K}(C + D)$. There exists a sequence $(\mathbf{x}_0, \mathbf{x}_1, \ldots)$ such that $\mathbf{x}_i \in C + D$ and $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}$. The definition of $C + D$ means that there is a sequence $(\mathbf{u}_0, \mathbf{u}_1, \ldots) \in \mathbf{Seq}_\infty(C)$ and a sequence $(\mathbf{v}_0, \mathbf{v}_1, \ldots) \in \mathbf{Seq}_\infty(D)$ such that $\mathbf{x}_i = \mathbf{u}_i + \mathbf{v}_i$ for $i \in \mathbb{N}$.

Since C is compact, the sequence $(\mathbf{v}_0, \mathbf{v}_1, \ldots)$ contains a convergent subsequence $(\mathbf{u}_{i_0}, \mathbf{u}_{i_1}, \ldots)$. Let $\mathbf{u} = \lim_{m\to\infty} \mathbf{u}_{i_m}$. Clearly, $\lim_{m\to\infty} \mathbf{x}_{i_m} = \mathbf{x}$. Since D is a closed set, $\lim_{m\to\infty} \mathbf{v}_{i_m} = \mathbf{x} - \mathbf{u} \in D$. Therefore, $\mathbf{x} = \mathbf{u} + \mathbf{v} \in C + D$, so $\mathbf{K}(C + D) = C + D$, which means that $C + D$ is closed.

Exercises and Supplements

1. Let L be an F-linear space. Prove that $0\mathbf{x} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ for every $a \in F$ and $\mathbf{x} \in L$.
2. Prove that a subset K of a linear space is linearly dependent if and only of there is $\mathbf{x} \in K$ that can be expressed as a linear combination of $K - \{\mathbf{x}\}$.
3. Let L, M be two F-linear spaces, $h : L \longrightarrow M$ be a linear mapping, and let $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ be a subset of L. Prove that if $\{h(\mathbf{x}_1), \ldots, h(\mathbf{x}_m)\}$ is a linearly independent set in M, then X is linearly independent in L.
4. Let U, V be two subspaces of the finite-dimensional F-linear space L. Prove that there exists a vector $\mathbf{w} \neq \mathbf{0}$ in both U and V only if $\dim(U) + \dim(V) \geqslant \dim(L)$.
5. Let L be a finite-dimensional F-linear space and let U, V be two subspaces of L. Show that $U + V = U \boxplus V$ if and only if $\dim(U + V) = \dim(U) + \dim(V)$.
6. Let U, V, W be subspaces of a finite-dimensional F-linear space L. Prove that
$$\dim(U \cap V \cap W) \geqslant \dim(U) + \dim(V) - 2\dim(L).$$
7. Let $\mathbf{w}$ be a vector in $\mathbb{R}^n$. Prove that the set
$$P_{\mathbf{w}} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{w}'\mathbf{x} = 0\}$$
is a subspace of $\mathbb{R}^n$.
8. Let POL be the real linear space of all polynomials in the variable x and let POL_n be the set of all polynomials of degree at most n. Prove that
   a) any sequence of polynomials of degrees $0, 1, 2, \cdots$ is a basis for POL;
   b) POL_n is a subspace of dimension $n + 1$ of POL.
9. Prove that there exists a unique linear function $L : \text{POL} \longrightarrow \mathbb{R}$ such that $L(1) = 1$ and $L([x]_k) = 1$ for $k \geqslant 1$, where the polynomial $[x]_k = x(x-1)\cdots(x-k+1)$ was introduced in Exercise 27 of Chapter 3.

10. Prove that
$$B_{n+1} = \frac{1}{e}\left(1^n + \frac{2^n}{1!} + \frac{3^n}{2!} + \cdots\right)$$
(Dobinski's Formula), where B_n is the n^{th} Bell number introduced in Exercise 33 of Chapter 3.

    Solution: As observed in Exercise 28 of Chapter 3, we have $e = \sum_{k=0}^{\infty} \frac{[k]_n}{k!}$. By applying L (defined in Exercise 9) to both members of the equality of Exercise 30 of Chapter 3 we have
$$\sum\{L([u]_k) \mid k = |\pi|, \pi \in PART(A)\} = L(u^m),$$
which implies $B_n = L(u^m)$.

    If $p(u)$ is a polynomial, taking into account that L is linear and that the polynomials of the form $[u]_n$ constitute a basis for POL, we can write $p(u) = \sum_n a_n [u]_n$. Therefore,
$$\begin{aligned} L(p(u)) &= \sum_n a_n L([u]_n) = \sum_n a_n \frac{1}{e} \sum_k \frac{[k]_n}{k!} \\ &= \frac{1}{e} \sum_n \sum_k a_n \frac{[k]_n}{k!} = \frac{1}{e} \sum_k \sum_n a_n \frac{[k]_n}{k!} \\ &= \frac{1}{e} \sum_k \frac{p(k)}{k!} \end{aligned}$$
for any polynomial p. Choosing $p(u) = u^n$ yields Dobinski's formula.

An *affine subspace* of a F-linear space $(L, +, \cdot)$ is a nonempty subset U of L such that there exists $\mathbf{u} \in L$ such that the set $U - \{\mathbf{u}\} = \{\mathbf{x} - \mathbf{u} \mid \mathbf{x} \in U\}$ is a linear subspace.

11. Let $\mathbf{x}_0, \mathbf{a} \in \mathbb{R}^n$. The *line that passes through $\boldsymbol{x}_0$ and has the direction $\boldsymbol{a}$* is the set
$$L_{\mathbf{x}_0,\mathbf{a}} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_0 + t\mathbf{a} \text{ for some } t \in \mathbb{R}\}.$$
Prove that $L_{\mathbf{x}_0,\mathbf{a}}$ is an affine subspace of $\mathbb{R}^n$.
12. Let $\mathbf{x}_0$ and $\mathbf{w}$ be two vectors in $\mathbb{R}^n$. Prove that the set
$$H_{\mathbf{x}_0,\mathbf{w}} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{w}'(\mathbf{x} - \mathbf{x}_0) = 0\}$$
is an affine subset of $\mathbb{R}^n$.
13. Let $\mathbf{e}_1, \ldots, \mathbf{e}_m$ be the vectors of the standard basis of $\mathbb{C}^m$. Prove that $\sum_{k=1}^{m} \mathbf{e}_k \mathbf{e}_k' = I_m$.
14. Let A, B, C be three matrices such that $A = BC$. Prove that
    a) $\mathsf{Ran}(A) \subseteq \mathsf{Ran}(B)$; also, if C is a square invertible matrix, show that $\mathsf{Ran}(A) = \mathsf{Ran}(B)$;
    b) $\mathsf{NullSp}(C) \subseteq \mathsf{NullSp}(A)$; also, if B is a square invertible matrix, $\mathsf{NullSp}(C) = \mathsf{NullSp}(A)$.
15. Let $A \in \mathbb{R}^{n \times n}$. Prove that:

a) the matrices

$$B = \frac{1}{2}(A + A') \text{ and } C = \frac{1}{2}(A - A')$$

are symmetric and skew-symmetric, respectively;

b) any square real matrix $A \in \mathbb{R}^{n \times n}$ can be uniquelly written as the sum of a symmetric and a skew-symmetric matrix.

Conclude that $\mathbb{R}^{n \times n}$ is the direct sum of the subspace of symmetric matrices and the subspace of skew-symmetric matrices.

16. Prove that there exist 2^{n^2} matrices in $\{0,1\}^{n \times n}$.
17. Let $A, B \in \mathbb{R}^{n \times n}$ such that there exists a non-singular matrix $X \in \mathbb{C}^{n \times n}$ such that $AX = XB$. Prove that there exists a non-singular matrix $Y \in \mathbb{R}^{n \times n}$ such that $AY = YB$.
18. Let $A \in \mathbb{C}^{n \times n}$. Prove that:

    a) the matrices

    $$B = \frac{1}{2}(A + A') \text{ and } C = \frac{1}{2}(A - A')$$

    are Hermitian and skew-Hermitian, respectively;

    b) any square complex matrix $A \in \mathbb{C}^{n \times n}$ can be uniquelly written as the sum of a Hermitian and a skew-Hermitian matrix.

    Conclude that $\mathbb{C}^{n \times n}$ is the direct sum of the subspace of Hermitian matrices and the subspace of skew-Hermitian matrices.
19. Let K be a finite set that spans the F-linear space $(L, +, \cdot)$ and let H be a subset of L that is linearly independent. There exists a basis B such that $H \subseteq B \subseteq K$.
20. Let $C, D \in \mathbb{C}^{n \times m}$ and let $\{\mathbf{t}_1, \ldots, \mathbf{t}_m\}$ be a basis in $\mathbb{C}^m$. Prove that if $C\mathbf{t}_i = D\mathbf{t}_i$ for $1 \leqslant i \leqslant m$, then $C = D$.
21. Let $A \in \mathbb{C}^{n \times n}$ be the matrix

    $$A = \begin{pmatrix} & \mathbf{0}'_{n-1} & & & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

    Prove that

    a) $A\mathbf{e}_k = \mathbf{e}_{k+1}$ for $1 \leqslant k \leqslant n-1$ and $A\mathbf{e}_n = \mathbf{a}$, where $\mathbf{a}$ is the last column of A.

    b) if $p(t) = t^{n-1} + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ we have $p(A)\mathbf{e}_i = \mathbf{0}_n$ for $1 \leqslant i \leqslant n$, and therefore, $p(A) = O_{n,n}$;

    c) there is no polynomial q of degree less than n such that $q(A) = O_{n,n}$.

    The matrix A is referred as the *companion matrix* of the polynomial p.
22. Let S be a finite set $S = \{x_1, \ldots, x_n\}$ and let $*$ be a binary operation on S. For t in S, define the matrices $\mathbf{L}_t, \mathbf{M}_t \in S^{n \times n}$ as $(\mathbf{L}_t)_{ij} = u$ if $(x_i * t) * x_j = u$ and $(\mathbf{R}_t)_{ij} = v$ if $x_i * (t * x_j) = u$.

    Prove that "$*$" is an associative operation on S if and only if for every $t \in S$ we have $\mathbf{L}_t = \mathbf{R}_t$.
23. Let $\mathbf{A} = (a_{ij})$ be an $(m \times n)$-matrix of real numbers. Prove that

    $$\max_j \min_i a_{ij} \leqslant \min_i \max_j a_{ij}$$

    (the *minimax inequality*).

Solution: Note that $a_{ij_0} \leqslant \max_j a_{ij}$ for every i and j_0, so $\min_i a_{ij_0} \leqslant \min_i \max_j a_{ij}$, again for every j_0. Thus, $\max_j \min_i a_{ij} \leqslant \min_i \max_j a_{ij}$.

24. Let $A \in \mathbb{C}^{n\times n}$ be a matrix and let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ be two vectors such that $A\mathbf{x} = a\mathbf{x}$ and $A'\mathbf{y} = a\mathbf{y}$ for some $a \in \mathbb{C}$, and $\mathbf{x}'\mathbf{y} = 1$. Let L be the rank-1 matrix $L = \mathbf{x}\mathbf{y}'$. Prove that:
    a) $L\mathbf{x} = \mathbf{x}$ and $\mathbf{y}'L = \mathbf{y}'$;
    b) L is an idemponent matrix;
    c) $A^m L = LA^m = a^m L$ for $m \geqslant 1$;
    d) $L(A - aL) = O_{n,n}$;
    e) $(A - aL)^m = A^m - a^m L$ for $m \geqslant 1$.
25. Let $A \in \mathbb{R}^{n\times n}$ be a matrix such that $A > O_{n,n}$ and let $\mathbf{x} \in \mathbb{R}^n$ be a non-negative vector such that $\mathbf{x} \neq \mathbf{0}_n$. Prove that $A\mathbf{x} > 0$.

    Solution: It is clear that $A\mathbf{x} \geqslant \mathbf{0}_n$. Suppose that there exists i, $1 \leqslant i \leqslant n$ such that $(A\mathbf{x})_i = \sum_{j=1}^n a_{ij}x_j = 0$. Since A is positive and $\mathbf{x}$ is non-negative, it follows that $a_{ij}x_j = 0$ for $1 \leqslant j \leqslant n$; this is possible only if $x_j = 0$ for $1 \leqslant j \leqslant n$, that is, if $\mathbf{x} = \mathbf{0}_n$. This contradiction shows that $A\mathbf{x} > 0$.
26. Prove that the product of two doubly-stochastic matrices is a doubly-stochastic matrix.
27. Let $A \in \mathbb{R}^{n\times n}$ be a matrix such that for every doubly-stochastic matrix $B \in \mathbb{R}^{n\times n}$ we have $AB = BA$. Prove that there exist $a, b \in \mathbb{R}$ such that $A = aI_n + bJ_n$.
28. Let $S = \{x_1, \ldots, x_n\}$ be a finite set. If $\pi = \{B_1, \ldots, B_m\}$ and $\sigma = \{C_1, \ldots, C_p\}$ are two partitions on S, prove that for the matrix $Q = M'_\pi M_\sigma \in \mathbb{N}^{m\times p}$ we have $q_{hk} = |B_h \cap C_k|$ for $1 \leqslant h \leqslant m$ and $1 \leqslant k \leqslant p$.
29. Let $\phi \in PERM_n$ be
$$\phi : \begin{pmatrix} 1 & \cdots & i & \cdots & n \\ a_1 & \cdots & a_i & \cdots & a_n \end{pmatrix},$$
    and let $v_p(\phi) = |\{(i_k, i_l) \mid i_l = p, k < l, i_k > i_l\}$ be the number of inversions of ϕ that have p as their second component, for $1 \leqslant p \leqslant n$. Prove that
    a) $v_p \leqslant n - p$ for $1 \leqslant p \leqslant n$;
    b) for every sequence of numbers $(v_1, \ldots, v_n) \in \mathbb{N}^n$ such that $v_p \leqslant n - p$ for $1 \leqslant p \leqslant n$ there exists a unique permutation ϕ that has $(v_1, \ldots, v_n) \in \mathbb{N}^n$ as its sequence of inversions.
30. Let p be the polynomial $p(x_1, \ldots, x_n) = \prod_{i<j}(x_i - x_j)$. For a permutation $\phi \in PERM_n$,
$$\phi : \begin{pmatrix} 1 & \cdots & i & \cdots & n \\ a_1 & \cdots & a_i & \cdots & a_n \end{pmatrix},$$
    define the number p_ϕ as $p_\phi = p(a_1, \ldots, a_n)$. Prove that
    a) $(-1)^{inv}(\phi) = \frac{p_\phi}{p_{\iota_n}}$ for any permutation ϕ;
    b) $(-1)^{inv(\psi\phi)} = (-1)^{inv(\psi)}(-1)^{inv(\phi)}$.
31. exer:nov1413a Let $A \in \mathbb{C}^{n\times n}$ be a matrix. Prove that for every permutation matrix P_ϕ, A is a symmetric matrix if and only if $P_\phi^{-1}AP_\phi$ is a symmetric matrix.
32. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Prove that if $trace(A) = 0$ and the sum of principal minors of order 2 equals 0, then $A = O_{n,n}$.

    Solution: A principal minor of order 2 has the form $m_{ij} = a_{ii}a_{jj} - a_{ij}^2$ for $1 \leqslant i < j \leqslant n$ (because A is a symmetric matrix) and there are $\binom{n}{2}$ such minors. Therefore, the sum of these minors is $M = \sum_{i<j} a_{ii}a_{jj} - \sum_{i<j} a_{ij}^2 = 0$. Since

trace$(A) = 0$ we have *trace*$(A)^2 = \sum_{i=1}^n a_{ii}^2 + 2\sum_{i<j} a_{ii}a_{jj} = 0$. These equalities imply

$$0 = \sum_{i=1}^n a_{ii}^2 + 2\sum_{i<j} a_{ii}a_{jj} = \sum_{i=1}^n a_{ii}^2 + 2\sum_{i<j} a_{ij}^2$$

which yields $a_{11} = \cdots = a_{nn} = a_{12} = \cdots = a_{n-1\,n} = 0$. Thus, $A = O_{n,n}$.

33. Let D_n be the determinant having n rows, defined by

$$D_n = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ a & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & a & 1 \end{vmatrix},$$

where $n \geqslant 2$. Prove that $D_n = 1$ for every $n \geqslant 2$.

34. Using elementary properties of determinants prove that

$$\begin{vmatrix} a^2 + x^2 & ab + xy & ac + xz \\ ab + xy & b^2 + y^2 & bc + yz \\ ac + xz & bc + yz & c^2 + z^2 \end{vmatrix} = 0.$$

35. Let $T_n(a)$ be the tri-diagonal determinant

$$T_n(a) = \begin{vmatrix} a & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & a & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & a & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & a \end{vmatrix}$$

having n rows, where $n \geqslant 2$.

Prove that $T_2(a) = a^2 - 1$ and that $T_{n+2}(a) = aT_{n+1}(a) - T_n(a)$ for $n \geqslant 2$.

36. Let $A \in \mathbb{R}^{n \times n}$ and $c \in \mathbb{R}$. Prove that $\det(cA) = c^n \det(A)$.

A matrix $A \in \mathbb{C}^{n \times n}$ is *unimodular* if $|\det(A)| = 1$.

37. Let $A \in \mathbb{R}^{n \times n}$ be a matrix such that $a_{ij} \in \mathbb{Z}$. Prove that A is nonsingular and the matrix A^{-1} has integer entries if and only if A is unimodular.

Solution: Since $A^{-1} = \frac{1}{\det(A)}\mathsf{adj}(A)$, it is clear that for a unimodular matrix whose elements are integers, the entries of A^{-1} are integers.

Conversely, if A^{-1} exists and its entries are integers, $A^{-1}A = I_n$ implies $\det(A)\det(A^{-1}) = 1$, which, in turn implies $\det(A) \in \{-1, 1\}$.

38. Let $A \in \mathbb{Z}^{m \times n}$ be a matrix having integer entries. Prove that there exist unimodular matrices $C \in \mathbb{R}^{m \times m}, D \in \mathbb{R}^{n \times n}$ such that $CAD = \mathsf{diag}(z_1, \ldots, z_r, 0, \ldots, 0)$, where $r \leqslant \min\{m, n\}$ and $z_1, \ldots, z_r$ are positive integers such that $z_i | z_{i+1}$ for $1 \leqslant i \leqslant r - 1$. This decomposition of A is known as the *Smith normal form of* A.

Solution: Let $a = \min_{i,j}\{|a_{ij}| \mid a_{ij} \neq 0, 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n\}$. Using row and column transpositions (which can be achieved by multiplications at left and at right by matrices of the form $T_n^{i \leftrightarrow j}$ place a in the top leftmost position of the matrix $T_1 A T_2$, where T_1 and T_2 are products of matrices of the form $T_n^{i \leftrightarrow j}$.

Suppose initially that a divides all entries of the matrix A. Then, by multiplications at left and right with matrices of the form $T_n^{i \overset{+}{\leftarrow} j}$ and $T_{n,i}^{(a)}$ we can place 0s everywhere in the first row and the first column except in the position $(1, 1)$ yielding a matrix of the form

$$\begin{pmatrix} 1 & \mathbf{0}'_{n-1} \\ \mathbf{0}_{n-1} & \tilde{A} \end{pmatrix}.$$

Applying the same argument to $\tilde{A}$, we reach the desired conclusion.

Suppose that a_{1j} is not divisible by a_{11}. Then, we have $a_{1j} = a_{11}q + r$ with $0 < |r| < |a_{11}|$. After subtracting the first column multiplied by q from the j^{th} column, we obtain r in position $(1, j)$. Then, we apply the argument used in the previous case. Since $|r| < |a_{11}|$ the process must end. The case, when a_{j1} is not divisible by a_{11} can be treated similarly. Thus, we have shown that there exist invertible matrices $T_1, \ldots, T_p, S_1, \ldots, S_q$ such that $T_p \cdots T_1 A S_1 \cdots S_q = \mathsf{diag}(z_1, \ldots, z_r, 0, \ldots, 0)$. By choosing $C = (T_p \cdots T_1)^{-1}$ and $D = (S_1 \cdots S_q)^{-1}$ we reach the desired conclusion.

39. Prove that if a matrix $A \in \mathbb{C}^{n \times n}$ is skew-Hermitian and n is an odd number, then $\Re(\det(A)) = 0$.
40. Prove that $\mathit{trace}(J_{n,n} A J_{n,n} B) = \mathit{trace}(J_{n,n} A) \mathit{trace}(J_{n,n} B)$, where $A, B \in \mathbb{C}^{n \times n}$.
41. Prove that $\det(aI_n + bJ_{n,n}) = a^n + na^{n-1}b$ for $n \geqslant 1$.
42. Let $A \in \mathbb{C}^{n \times n}$ and let $x, y \in \mathbb{C}$. Prove that

$$(xI_n - A)^{-1} - (yI_n - A)^{-1} = (y - x)(xI_n - A)^{-1}(yI_n - A)^{-1}.$$

43. Let $A \in \mathbb{R}^{n \times n}$ be a skew-symmetric matrix and let $B = A + bJ_{n,n}$, where $b \in \mathbb{R}$. Prove that if n is even, then $\det(A) = \mathit{det}(B)$.
44. Let $U, V \in \mathbb{C}^{n \times m}$ be two matrices. Prove that $\det(I_n + UV^{\mathsf{H}}) = \det(I_m + V^{\mathsf{H}}U)$ (Sylvester's Identity).

Solution: Starting from the matrix equalities

$$\begin{aligned}
\begin{pmatrix} I_n & -U \\ V^{\mathsf{H}} & I_m \end{pmatrix} &= \begin{pmatrix} I_n & O_{n,m} \\ V^{\mathsf{H}} & I_m \end{pmatrix} \cdot \begin{pmatrix} I_n & -U \\ O_{m,n} & I_m + V^{\mathsf{H}}U \end{pmatrix} \\
&= \begin{pmatrix} I_n & -U \\ O_{m,n} & I_m \end{pmatrix} \cdot \begin{pmatrix} I_n + UV^{\mathsf{H}} & O_{n,m} \\ V^{\mathsf{H}} & I_m \end{pmatrix},
\end{aligned}$$

which are immediate, and taking the determinants of both sides the Sylvester's Identity follows immediately taking into account that

$$\det \begin{pmatrix} I_n & U \\ O_{m,n} & I_m + V^{\mathsf{H}}U \end{pmatrix} = \det(I_m + V^{\mathsf{H}}U)$$

and

$$\det \begin{pmatrix} I_n + UV^{\mathsf{H}} & O_{n,m} \\ V^{\mathsf{H}} & I_m \end{pmatrix} = \det(I_n + UV^{\mathsf{H}}).$$

We used the fact that the determinant of a block upper triangular matrix equals the product of the determinants of matrices situated on the diagonal.

45. Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Prove that:

a) $\det(I_n + \mathbf{x}\mathbf{y}^{\mathsf{H}}) = 1 + \mathbf{y}^{\mathsf{H}}\mathbf{x}$ and $\det(I_n - \mathbf{x}\mathbf{y}^{\mathsf{H}}) = 1 - \mathbf{y}^{\mathsf{H}}\mathbf{x}$;
b) if $A \in \mathbb{C}^{n\times n}$ is an invertible matrix, then $\det(A + \mathbf{x}\mathbf{y}^{\mathsf{H}}) = \det(A)(1 + \mathbf{y}^{\mathsf{H}}A^{-1}\mathbf{x})$.

Solution: The first part is a direct consequence of Supplement 44 by taking $m = 1$.

For the second part we have $A + \mathbf{x}\mathbf{y}^{\mathsf{H}} = A(I_n + A^{-1}\mathbf{x}\mathbf{y}^{\mathsf{H}})$, which yields the desired inequality.

46. Let $A \in \mathbb{R}^{m\times n}$, where $m \leqslant n$. Prove that

$$\det(AA') = \sum \left\{ \det\left(A \begin{bmatrix} 1 & \cdots & m \\ k_1 & \cdots & k_m \end{bmatrix} \right)^2 \middle| 1 \leqslant k_1 < k_2 < \cdots < k_m \leqslant n \right\}.$$

Solution: This equality follows from Cauchy-Binet formula.

47. Let $v_1, \ldots, v_n$ be n complex numbers such that $\sum_{k=1}^{n} |v_k|^2 = 2$. Prove that

$$\begin{vmatrix} 1 - v_1\bar{v}_1 & v_1\bar{v}_2 & v_1\bar{v}_3 & \cdots & v_1\bar{v}_n \\ v_2\bar{v}_1 & 1 - v_2\bar{v}_2 & v_2\bar{v}_3 & \cdots & v_2\bar{v}_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ v_n\bar{v}_1 & v_n\bar{v}_2 & v_n\bar{v}_3 & \cdots & 1 - v_n\bar{v}_n \end{vmatrix} = -1.$$

Solution: Note that the matrix whose determinant is to be computed equals $I_n - \mathbf{v}\mathbf{v}^{\mathsf{H}}$. Thus, by Supplement 45, we have $\det(I_n - \mathbf{v}\mathbf{v}^{\mathsf{H}}) = 1 + \mathbf{v}^{\mathsf{H}}\mathbf{v} = -1$.

48. Let $D = \mathsf{diag}(a_1, \ldots, a_n)$ and let $A = D + \mathbf{1}_n\mathbf{1}_n'$. Prove that

$$\det(A) = \prod_{i=1}^{n} a_i \left(1 + \sum_{i=1}^{n} \frac{1}{a_i} \right).$$

Solution: By Part (b) of Supplement 45, we have

$$\det(A) = \det(D + \mathbf{1}_n\mathbf{1}_n^{\mathsf{H}}) = \det(D)(1 + \mathbf{1}_n^{\mathsf{H}}D^{-1}\mathbf{1}_n),$$

which amounts to the formula we need to prove.

49. Let $Z \in \mathbb{C}^{n\times n}$, $W \in \mathbb{C}^{m\times m}$ and let $U, V \in \mathbb{C}^{n\times m}$ be four matrices such that each of the matrices W, Z, $Z + UWV'$, and $W^{-1} + UWV'$ have an inverse. Prove that

$$(Z + UWV')^{-1} = Z^{-1} - Z^{-1}U(W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1}$$

(the *Woodbury-Sherman-Morrison identity*).

Solution: Consider the system of matrix equations:

$$ZX + UY = I_n, V'X - W^{-1}Y = O_{m,n},$$

where $X \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{m\times n}$

The second equation implies $V'X = W^{-1}Y$, so $UWV'X = UY$. Substituting UY in the first equation yields $ZX + UWV'X = I$, so $(Z + UWV')X = I$, which implies

$$X = (Z + UWV')^{-1}. \tag{5.10}$$

On the other hand, we have $X = Z^{-1}(I - UY)$ from the first equation. Substituting X in the second equation yields $V'Z^{-1}(I - UY) = W^{-1}Y$, which is equivalent to

$$\begin{aligned} V'Z^{-1} &= +W^{-1}Y + V'Z^{-1}UY \\ &= (W^{-1} + V'Z^{-1}U)Y. \end{aligned}$$

Thus, we have $Y = (W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1}$. Substituting the values of X and Y in the first equality implies

$$ZX + U(W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1} = I.$$

Therefore, $ZX = I - U(W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1}$, which implies

$$X = Z^{-1} - Z^{-1}U(W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1}. \tag{5.11}$$

The Woodbury-Sherman-Morrison identity follows immediately from Equalities (5.10) and (5.11).

50. Using the notations introduced in the previous Supplement prove the Woodbury-Sherman-Morrison identity for determinants:

$$\det(Z + UWV') = \det(Z)\det(W)\det(W^{-1} + V'Z^{-1}U).$$

Solution: We can write:

$$\begin{aligned} &\det(I_n - U(W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1}) \\ &= \det(I_n - V'Z^{-1}U(W^{-1} + V'Z^{-1}U)^{-1}) \\ &\quad \text{(by Sylvester's identity)} \\ &= \det((W^{-1} + V'Z^{-1}U)(W^{-1} + V'Z^{-1}U)^{-1} \\ &\quad -V'Z^{-1}U(W^{-1} + V'Z^{-1}U)^{-1}) \\ &= \det((W^{-1} + V'Z^{-1}U - V'Z^{-1}U)\det(W^{-1} + V'Z^{-1}U)^{-1} \\ &= \det(W^{-1})\det(W^{-1} + V'Z^{-1}U)^{-1}. \end{aligned}$$

This allows us to write $\det(Z)\det(X) = \det(I_n - U(W^{-1} + V'Z^{-1}U)^{-1}V'Z^{-1})$. Since $X = (Z + UWV')^{-1}$, it follows that

$$\frac{det(Z)}{\det(Z + UWV')} = \frac{1}{\det(W^{-1} + V'Z^{-1}U)\det(W)},$$

which is the desired equality.

51. Prove that if $U, V \in \mathbb{C}^{n\times m}$ are two matrices such that $I_n + UV'$ and $I_m + V'U$ are invertible matrices, then $\det(I_n + UV') = \det(I_m + V'U)$.
52. Let $A \in \mathbb{R}^{n\times n}$. Prove that for the k^{th} derivative of the function $G(x) = \det(A - xI_n)$ we have

$$G^{(k)}(0) = (-1)^k k! S_{n-k}(A),$$

where $S_p(A)$ is the sum of all order-p principal minors of A.

53. Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ be two sequences of real numbers such that $x_i + y_j \neq 0$ for $1 \leqslant i, j \leqslant n$ and $n \geqslant 2$. The *Cauchy matrix* of these sequences is the matrix $C_{\mathbf{x},\mathbf{y}}$ given by

$$(C_{\mathbf{x},\mathbf{y}})_{ij} = \frac{1}{x_i + y_j}$$

for $1 \leqslant i, j \leqslant n$. Prove that

$$\det(C_{\mathbf{x},\mathbf{y}}) = \frac{\prod_{1\leqslant j<i\leqslant n}(x_i - x_j)\prod_{1\leqslant j<i\leqslant n}(y_i - y_j)}{\prod_{1\leqslant i,j\leqslant n}(x_i + y_j)}.$$

54. The *n-Hilbert matrix* H_n is a special Cauchy matrix, where $x_i = y_i = i - \frac{1}{2}$ for $1 \leqslant i \leqslant n$. Let $h_n = 1!2! \cdots (n-1)!$ for $n \in \mathbb{N}$ and $n \geqslant 1$. Prove that:
    a) the determinant of the Hilbert matrix H_n is $\det(H_n) = \frac{h_n^4}{h_{2n}}$;
    b) the number $\frac{1}{\det(H_n)}$ is an integer.
55. Let $A \in \mathbb{R}^{n \times n}$ be defined by $a_{ij} = i + j$ for $1 \leqslant i, j \leqslant n$. Prove that if $n \geqslant 3$, $\det(A) = 0$.
56. Let $A \in \mathbb{C}^{n \times n}$. Prove that $rank(A) = n$ if and only if $rank(A^{sH}) = n$.
57. Let $A \in \mathbb{C}^{m \times n}$ be a matrix, $\mathbf{x} \in \mathbb{C}^n$ and $\mathbf{y} \in \mathbb{C}^m$. Prove that if $rank(A - aA\mathbf{x}\mathbf{y}^{\mathsf{H}}A) = rank(A) - 1$, then $\mathbf{y}^{\mathsf{H}}A\mathbf{x} \neq 0$ and $a = \frac{1}{\mathbf{y}^{\mathsf{H}}A\mathbf{x}}$.
58. Let $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times p}$, and $C \in \mathbb{C}^{p \times q}$ be three matrices. Prove the following inequality
$$rank(AB) + rank(BC) \leqslant rank(B) + rank(ABC),$$
known as Frobenius' Inequality.
59. Let $A \in \mathbb{C}^{m \times n}$ be a matrix, $\mathbf{u} \in \mathbb{C}^m$ and $\mathbf{v} \in \mathbb{C}^n$ be two vectors, $a \in \mathbb{C} - \{0\}$, and let $B = A - \frac{1}{a}\mathbf{u}\mathbf{v}^{\mathsf{H}}$. We have $rank(B) < rank(A)$, if and only if there are vectors $\mathbf{x} \in \mathbb{C}^m$ and $\mathbf{y} \in \mathbb{C}^n$ such that $\mathbf{u} = A\mathbf{y}$, $\mathbf{v} = A^{\mathsf{H}}\mathbf{x}$, and $a = \mathbf{x}^{\mathsf{H}}A\mathbf{y}$, in which case $rank(B) = rank(A) - 1$.
60. Let d_n be the determinant defined by
$$d_n = \begin{vmatrix} \cos\alpha & \cos 2\alpha & \dots \cos n\alpha \\ \cos(n+1)\alpha & \cos(n+2)\alpha & \dots \cos 2n\alpha \\ \vdots & \vdots & \dots \;\vdots \\ \cos(n^2-n+1)\alpha & \cos(n^2-n+2)\alpha & \dots \cos n^2\alpha \end{vmatrix}$$
Prove that for $n \geqslant 3$ we have $d_n = 0$.
Hint: add the third column to the first column.

Let $A \in \mathbb{C}^{m \times n}$ be a partitioned matrix given by
$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$
where A_{11} is an invertible square matrix, $A_{11} \in \mathbb{C}^{p \times p}$. Note that $A_{21} \in \mathbb{C}^{(m-p)\times p}$, $A_{12} \in \mathbb{C}^{p \times (n-p}$, and $A_{22} \in \mathbb{C}^{(m-p)\times(n-p)}$. Therefore, the matrix $B = A_{22} - A_{21}A_{11}^{-1}A_{12} \in \mathbb{C}^{(m-p)\times(n-p)}$ is well-defined. We refer to B as the *Schur's complement of* A_{11} relative to A and is denoted by A/A_{11}.

61. Let $A \in \mathbb{C}^{n \times n}$ be a square partitioned matrix given by
$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$
where A_{11} is an invertible matrix. Prove that
    a) we have $\det(A/A_{11}) = \frac{\det(A)}{\det(A_{11})}$;
    b) we have the equalities:
$$A = \begin{pmatrix} I & O \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ O & I \end{pmatrix}.$$
    and
$$\det(A) = \det(A_{11})\det(A_{22} - A_{21}A_{11}^{-1}A_{12}).$$

62. Let
$$A = \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \vdots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{pmatrix}$$
be a partitioned matrix such that each matrix
$$B_p = \begin{pmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \vdots & \vdots \\ A_{p1} & \cdots & A_{pp} \end{pmatrix}$$
is invertible for $1 \leqslant p \leqslant m-1$.
Prove that A can be uniquely written as a product $A = LDU$ such that the following conditions are satisfied:
   a) L is a lower block triangular matrix whose diagonal blocks are unit matrices;
   b) D is a block diagonal matrix, $D = \mathsf{diag}(D_1, \ldots, D_m)$ whose diagonal blocks are invertible matrices such that $D_1 = A_{11} = B_1$, and $D_j = B_j/B_{j-1}$ for $1 \leqslant p \leqslant m$;
   c) U is an upper block diagonal matrix whose diagonal blocks are unit matrices.
63. A matrix $A \in \mathbb{C}^{n \times n}$ is *strongly non-singular* if all its principal matrices are non-singular. Prove that if $A \in \mathbb{C}^{n \times n}$ is strongly non-singular, then there exists a unique factorization $A = LDU$ such that the following conditions are satisfied:
   a) L is a lower block triangular matrix whose diagonal elements are equal to 1;
   b) D is a diagonal matrix, $D = \mathsf{diag}(d_1, \ldots, d_n)$ such that $d_1 = a_{11}$ and $d_k = \frac{A_k}{A_{k-1}}$, where $A_k = A\begin{bmatrix} 1 \cdots k \\ 1 \cdots k \end{bmatrix}$;
   c) U is an upper diagonal matrix whose diagonal elements are equal to 1.
   Hint: This follows immediately from Supplement 62.
64. Prove that if $H \in \mathbb{R}^{n \times n}$ is a Hadamard matrix, then $|\det(H)| = n^{\frac{n}{2}}$.
65. Prove that if $H \in \mathbb{R}^{n \times n}$ is a Hadamard matrix and $n > 2$, then n is a multiple of 4.
66. Let $a_1, \ldots, a_n$ be n numbers and let $b_{ij} = |a_i - a_j|$ for $1 \leqslant i, j \leqslant n$. Compute the determinant of the matrix $B = (b_{ij})$.
67. Let $S = \{(x_i, y_i) \in \mathbb{C}^2 \mid 1 \leqslant i \leqslant n\}$ be a set of n pairs of complex numbers such that $i \neq j$ implies $x_i \neq x_j$ and and let $p(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$ be a polynomial such that $p(x_i) = y_i$ for $1 \leqslant i \leqslant n$. Prove that
   a) the coefficients c_i of p are given by the formula
$$c_i = \frac{V_{\mathbf{x}} \overset{i}{\leftarrow} \mathbf{y}}{V_{\mathbf{x}}},$$
   where $V_{\mathbf{x}} \overset{i}{\leftarrow} \mathbf{y}$ is the determinant obtained from $V_{\mathbf{x}}$ by replacing the i^{th} column by $\mathbf{y}$;
   b) the polynomial $p(x)$ can be written as
$$p(x) = \sum_{j=1}^{n} y_j p_j(x),$$

where $p_j(x) = \prod\{\frac{x-x_k}{x_j-x_k} \mid k \in \{1,\ldots,n\} - \{j\}\}$. The polynomial p is known as the *Lagrange interpolation polynomial* for S.

68. Let $A \in \mathbb{C}^{n\times n}$ be a matrix such that $A\mathbf{1}_n = A'\mathbf{1}_n = \mathbf{0}_n$. Prove that all cofactors $cof(a_{ij})$ of A are equal.
69. Prove that if $H \in \mathbb{R}^{n\times n}$ is a Hadamard matrix, then
$$\begin{pmatrix} H & H \\ H & -H \end{pmatrix} \in \mathbb{R}^{2n\times 2n}$$
is also a Hadamard matrix.
70. Let
$$M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$
be a partitioned matrix, where $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{m\times p}$, $C \in \mathbb{R}^{p\times m}$, and $D \in \mathbb{R}^{p\times p}$. Prove that if $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{p\times p}$ are orthogonal matrices, then
$$\begin{aligned}\det(M) &= \det\begin{pmatrix} UA & UB \\ C & D \end{pmatrix} = \det\begin{pmatrix} A & BV \\ C & DV \end{pmatrix} \\ &= \det\begin{pmatrix} AU & B \\ CU & D \end{pmatrix} = \det\begin{pmatrix} A & B \\ VC & VD \end{pmatrix}.\end{aligned}$$
71. Let $A \in \mathbb{R}^{n\times n}$ be a skew-symmetric matrix. If n is an odd number, prove that $\det(A) = 0$.
72. A *generalized inverse* (or a g-inverse) of a matrix $A \in \mathbb{C}^{m\times n}$ is a matrix $B \in \mathbb{C}^{n\times m}$ such that $ABA = A$. Prove that if B is a g-inverse and $A\mathbf{x} = \mathbf{b}$ has a solution, then $\mathbf{x} = G\mathbf{b}$ is one of these solutions.
73. Let $A \in \mathbb{C}^{m\times n}$ be a matrix such that $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, where $A_{11} \in \mathbb{C}^{r\times r}$ and $rank(A_{11}) = rank(A) = r$. Prove that:
    a) there exists a matrix $S \in \mathbb{C}^{r\times n-r}$ such that $A_{12} = A_{11}S$ and $A_{22} = A_{21}S$;
    b) the matrix $B = \begin{pmatrix} A_{11}^{-1} & O_{r,m-r} \\ O_{n-r,r} & 0_{n-r,m-r} \end{pmatrix}$ is a g-inverse of A.
74. Let $A \in \mathbb{R}^{n\times n}$ be a matrix such that $a_{ii} > 0$ for $1 \leqslant i \leqslant n$, such that $a_{ii} > \sum\{|a_{ik}| \mid 1 \leqslant k \leqslant n \text{ and } k \neq i\}$ and $1 \leqslant i \leqslant n$. Prove that $\det(A) > 0$.
75. Prove that for $A \in \mathbb{C}^{m\times n}$, there exists at most one matrix $M \in \mathbb{C}^{n\times m}$ such that the matrices MA and AM are Hermitian, $AMA = A$, and $MAM = M$.

The matrix M introduced above is a special g-inverse referred to as the *Moore-Penrose pseudoinverse of* A and is denoted by $A^\dagger$.

76. Let $A \in \mathbb{C}^{n\times n}$. Prove that if A is invertible, then $A^\dagger$ exists and equals A^{-1}.
77. Give an example of a matrix that is not invertible but has a Moore-Penrose pseudoinverse.
    Hint: Consider the matrix $O_{m,n}$.
78. Prove that if $A^\dagger$ exists then $(A^\dagger)^\dagger = A$.
79. Let $\mathsf{GL}(n,\mathbb{C})$ be the set of invertible matrices in $\mathbb{C}^{n\times n}$. Prove that
    a) the algebra $(\mathsf{GL}(n,\mathbb{C}), \{I_n, \cdot, {}^{-1}\}$, where $\cdot$ is the usual matrix multiplication is a group (this is known as the linear group);

b) the mapping $\phi : \mathsf{GL}(2,\mathbb{C}) \longrightarrow \mathsf{GL}(3,\mathbb{C})$ given by

$$\phi\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}^2 & 2a_{11}a_{12} & a_{12}^2 \\ a_{11}a_{21} & a_{11}a_{22} + a_{12}a_{22} & a_{12}a_{22} \\ a_{21}^2 & 2a_{21}a_{22} & a_{22}^2 \end{pmatrix}$$

is a group mprphism.

80. Let $A \in \mathbb{C}^{m\times n}$ be a matrix with *rank*$(A) = r > 0$ and let $A = BC$ be a full-rank factorization of A, where $B \in \mathbb{C}^{m\times r}$ and $C \in \mathbb{C}^{r\times n}$ are full-rank matrices. Prove that
   a) the matrices $B^{\mathsf{H}}B \in \mathbb{C}^{r\times r}$ and $CC^{\mathsf{H}} \in \mathbb{C}^{r\times r}$ are non-singular;
   b) the matrix $B^{\mathsf{H}}AC^{\mathsf{H}}$ is non-singular;
   c) the Moore-Penrose pseudoinverse of A is given by

$$A^{\dagger} = C^{\mathsf{H}}(CC^{\mathsf{H}})^{-1}(B^{\mathsf{H}}B)^{-1}B^{\mathsf{H}}.$$

A *data matrix* is a matrix $D \in \mathbb{R}^{m\times n}$. The columns of D, $\mathbf{v}_1,\ldots,\mathbf{v}_n$ are the *features* of the data; its rows $\mathbf{u}_1',\ldots,\mathbf{u}_m'$ are the *observations.*

The *mean* of D is the vector $\tilde{D} = \frac{1}{m}D'\mathbf{1}_m \in \mathbb{R}^n$. D is *centered* if $\tilde{D} = \mathbf{0}_n$.

81. Let $D \in \mathbb{R}^{m\times n}$ be a data matrix and let $H_m = I_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m'$. Prove that H_mD is a centered data matrix.
   Solution: We have

$$\begin{aligned}\widetilde{(H_mD)} &= \frac{1}{m}D'H_m'\mathbf{1}_m = \frac{1}{m}D'\left(I_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m'\right)'\mathbf{1}_m \\ &= \frac{1}{m}D'\left(I_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m'\right)\mathbf{1}_m = \frac{1}{m}D'\left(\mathbf{1}_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m'\mathbf{1}_m\right) = \mathbf{0}_n.\end{aligned}$$

82. Prove that the centering matrix $H_m = I_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m'$ is both symmetric and idempotent; further, prove that $H_m\mathbf{1}_m = \mathbf{0}_m$.

Bibliographical Comments

MacLane and Birkhoff [127] is a fundamental reference for algebra. Artin's book [7] contains a vast amount of material from many areas of algebra presented in a lucid manner. Anther basic referrence is [78]. Concise and important sources are [138, 132] and [73].

Exercise 9 is a result of G. Rota (see [158]). The Russian literature has produced several readable books [167, 115] which were translated into English.