Algèbre Linéaire

Author: CatMono

Date: July, 2025

Version: 0.1

Contents

Chapter 7　Jordan Forms　　　　　　　　　　　　　　　　　　　　　　　　　　**20**

Chapter 8　Quadratic Forms　　　　　　　　　　　　　　　　　　　　　　　　　**21**

Chapter 9　Inner Product Spaces　　　　　　　　　　　　　　　　　　　　　　**22**

Preface

This is the preface of the book...

Chapter 1 Determinants

1.1 Permutations

1.2 Determinant and Its Properties

1.3 Expanding by Rows (Columns)

¶ Expanding by One Row
¶ Cramer's Rule
¶ Expanding by k Rows

1.4 Special Determinants

Definition 1.1 (Vandermonde Determinant)

The Vandermonde determinant is defined as

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}$$

where $x_1, x_2, \ldots, x_n$ are distinct variables. ♣

The value of the Vandermonde determinant is given by

$$V_n = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Definition 1.2 (Arrow Determinant)

The Arrow determinant ($\searecho$) is defined as

$$A_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

The value of the Arrow determinant is given by

$$A_n = \left(a_{11} - \sum_{k=2}^{n} \frac{a_{1k} a_{k1}}{a_{kk}} \right) \prod_{k=2}^{n} a_{kk}.$$

♣

From the first column sequentially, subtract $\frac{a_{21}}{a_{22}}$ times the second column, $\cdots$, $\frac{a_{n1}}{a_{nn}}$ times the n-th column, so that the first column becomes:

$$\begin{bmatrix} a_{11} - \sum_{k=2}^{n} \frac{a_{1k} a_{k1}}{a_{kk}} & 0 & 0 & \vdots & 0 \end{bmatrix}^{\mathrm{T}}.$$

Then expand along the first column.

Definition 1.3 (Two-Triangular Determinant)

If the determinant satisfies

$$a_{ij} = \begin{cases} a, & i < j, \\ x_i, & i = j, \\ b, & i > j, \end{cases}$$

, then D_n is called a two-triangular determinant. ♣

The value of the two-triangular determinant is given by

$$\begin{vmatrix} x_1 & a & a & \ldots & a \\ b & x_2 & a & \ldots & a \\ b & b & x_3 & \ldots & a \\ \vdots & \vdots & \vdots & & \vdots \\ b & b & b & \ldots & x_n \end{vmatrix} = \begin{cases} \left[x_1 + a \sum\limits_{k=2}^{n} \frac{x_1 - a}{x_k - a} \right] \cdot \prod\limits_{k=2}^{n} (x_k - a), & a = b \\ (x_n - b)D_{n-1} + \prod\limits_{k=1}^{n-1} (x_k - a), & a \neq b \end{cases}$$

Chapter 2 System of Linear Equations

> **Definition 2.1 (System of Linear Equations)**
>
> A system of linear equations is a collection of one or more linear equations involving the same set of variables. For example, a system of m linear equations in n variables can be written as:
>
> $$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$
>
> where $x_1, x_2, \ldots, x_n$ are the variables, a_{ij} are the coefficients, and b_i are the constants.
>
> A solution to the system is an ordered set of values for the variables that satisfies all equations simultaneously. If two systems have the same solution set, they are called equivalent systems, whose relationship are equivalence. ♣

2.1 Elimination Method

2.2 Linear Space

> **Definition 2.2 (Linear Space over Field F)**
>
> Let V be a non-empty set, and F be a field. If the following two operations (binary mapping) are defined on V:
>
> **Vector addition** For any $\alpha, \beta \in V$, there exists $\gamma \in V$ such that $\gamma = \alpha + \beta$.
>
> **Scalar multiplication** For any $\alpha \in V$ and $c \in F$, there exists $\gamma \in V$ such that $\gamma = c \cdot \alpha$.
>
> and the following axioms are satisfied (for any $\alpha, \beta, \gamma \in V$ and $k, l \in F$):
>
> **A1. Commutativity of Addition** $\alpha + \beta = \beta + \alpha$.
>
> **A2. Associativity of Addition** $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
>
> **A3. Existence of Additive Identity** There exists an element $0 \in V$ such that $\alpha + 0 = 0 + \alpha = \alpha$.
>
> **A4. Existence of Additive Inverse** There exists an element $-\alpha \in V$ such that $\alpha + (-\alpha) = (-\alpha) + \alpha = 0$.
>
> **M1. Compatibility of Scalar Multiplication with Field Multiplication** $k(l\alpha) = (kl)\alpha$.
>
> **M2. Identity Element of Scalar Multiplication** $1 \cdot \alpha = \alpha$, where 1 is the multiplicative identity in F.
>
> **D1. Distributivity of Scalar Multiplication with respect to Vector Addition** $k(\alpha + \beta) = k\alpha + k\beta$.
>
> **D2. Distributivity of Scalar Multiplication with respect to Field Addition** $(k + l)\alpha = k\alpha + l\alpha$.
>
> Then V is called a linear space (or vector space) over the field F. ♣

⚡**Remark** In the definition of linear space, **A1. Commutativity of Addition** can be derived from the other axioms, so it is not strictly necessary to include it as an axiom.

Give some examples of linear spaces:

1. The ring of univariate polynomials $P[x]$ over the number field P, with the usual polynomial addition and scalar multiplication by elements from P, forms a linear space over the number field P. If we consider only polynomials of degree less than n, along with the zero polynomial, this also forms a linear space over P, denoted as $P[x]_n$.

2. The set of all positive real numbers $\mathbb{R}^+$, with addition and scalar multiplication defined as:
$$a \oplus b = ab, \quad k \circ a = a^k,$$
forms a linear space over the number field $\mathbb{R}$.

3. The set of all real-valued functions, under function addition and scalar multiplication by elements from $\mathbb{R}$, forms a linear space over the field $\mathbb{R}$.

4. The field P itself, under its own addition and multiplication, forms a linear space over itself.

5. The set of all n-dimensional vectors (ordered n-tuples of elements from P) over the number field P, with the usual vector addition and scalar multiplication, forms a linear space over the number field P.

In this section, we primarily examine linear spaces consisting of n-dimensional vectors over the real number field $\mathbb{R}$. As a matter of fact, all elements of a linear space can be called vectors, and the definitions and properties of vectors in $\mathbb{R}^n$ can be generalized to any linear space.

¶ **Linear Independence**

> **Definition 2.3 (Linear Combination)**
>
> Let V be a linear space over the field F, and let $\alpha_1, \alpha_2, \ldots, \alpha_n \in V$. For any scalars $k_1, k_2, \ldots, k_n \in F$, the vector
> $$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$
> is called a **linear combination** of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ (we also say that β is linearly expressed by $\alpha_1, \alpha_2, \ldots, \alpha_n$), where k_i are called the coefficients of the linear combination. ♣

Linear Representation and Equivalence of Vector Groups

If every vector α_i ($i = 1, 2, \ldots, t$) in the vector group $\alpha_1, \alpha_2, \ldots, \alpha_t$ can be linearly expressed by the vector group $\beta_1, \beta_2, \ldots, \beta_s$, then the vector group $\alpha_1, \alpha_2, \ldots, \alpha_t$ is said to be linearly expressed by the vector group $\beta_1, \beta_2, \ldots, \beta_s$.

If two vector groups can be linearly expressed by each other, they are said to be **equivalent**.

⚡**Remark** Linear representation can be used to identify redundant equations in a system of linear equations. In a system of linear equations, if one equation can be linearly expressed by other equations, then through elementary transformations, the corresponding row in the matrix can be reduced to a row of zeros.

If two vector groups are equivalent, then their corresponding systems of linear equations are equivalent.

> **Definition 2.4 (Linear Dependence and Independence of Vector Groups)**
>
> A vector group $\alpha_1, \alpha_2, \ldots, \alpha_t (t \geq 2)$ is said to be **linearly dependent** if there exists a vector in the group that can be linearly expressed by the other vectors.
>
> Another equivalent definition is: A vector group $\alpha_1, \alpha_2, \ldots, \alpha_t (t \geq 1)$ is linearly dependent if there exist <u>not all zero</u> scalars $k_1, k_2, \ldots, k_t$ in the field F such that:
> $$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_t\alpha_t = 0.$$
> Conversely, a vector group $\alpha_1, \alpha_2, \ldots, \alpha_t (t \geq 1)$ is **linearly independent** if: the fact that
> $$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_t\alpha_t = 0$$
> implies:
> $$k_1 = k_2 = \cdots = k_t = 0.$$ ♣

☝ **Note** *If a subset of a vector group is linearly dependent, then the entire vector group is linearly dependent. If a vector group is linearly independent, then any non-empty subset of the group is also linearly independent. (**Partial dependence implies overall dependence; overall independence implies partial independence.**)*

Particularly, since two proportional vectors are linearly dependent, a linearly independent vector group cannot contain two proportional vectors.

A few propositions about linear dependence and independence are given below:

Proposition 2.1

1. Generally, for a vector group $\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, s$, the necessary and sufficient condition for linear dependence is that the equation:

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_s\alpha_s = 0$$

or the homogeneous system of linear equations:

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \cdots + a_{s1}x_s = 0 \\ a_{12}x_1 + a_{22}x_2 + \cdots + a_{s2}x_s = 0 \\ \vdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{sn}x_s = 0 \end{cases}$$

has a non-zero solution.

2. If a vector group $\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, s$ is linearly independent, then adding one component to each vector to form an $n+1$-dimensional vector group $\beta_i = (a_{i1}, a_{i2}, \ldots, a_{in}, a_{i,n+1})$, $i = 1, 2, \ldots, s$, will also be linearly independent.

3. For two vector groups $\alpha_1, \alpha_2, \ldots, \alpha_r$ and $\beta_1, \beta_2, \ldots, \beta_s$, if the former can be linearly expressed by the latter:

    i. and $r > s$, then the former is linearly dependent.
    ii. and $\alpha_1, \alpha_2, \ldots, \alpha_r$ is linearly independent, then $r \leqslant s$.

4. Any $n+1$ n-dimensional vectors must be linearly dependent.

5. Two equivalent linearly independent vector groups must contain the same number of vectors.

6. If a vector group $\alpha_1, \alpha_2, \ldots, \alpha_r$ is linearly independent, and adding a vector β to form the vector group $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta$ makes it linearly dependent, then β can be uniquely expressed as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_r$.

7. If a vector β can be linearly expressed by a vector group $\alpha_1, \alpha_2, \ldots, \alpha_r$, then $\alpha_1, \alpha_2, \ldots, \alpha_r$ is linearly independent if and only if the expression for β is unique.

♠

¶ **Dimension, Basis, and Coordinates**

Definition 2.5 (Dimension)

If a linear space V over field F contains n vectors that are linearly independent, but any more than n vectors are linearly dependent, then V is called an n-dimensional linear space, noted as $\dim V_F = n$, simplified as $\dim V = n$ when the field is clear from context.

If a linear space does contain as many vectors that are linearly independent as you wish, it is called an infinite-dimensional linear space.

♣

> **Definition 2.6 (Basis and Coordinates)**
>
> In an n-dimensional linear space V, if $\alpha_1, \alpha_2, \ldots, \alpha_n$ are n linearly independent vectors in V, then they form a **basis** of V.
>
> Any vector $\alpha \in V$ can be uniquely expressed as a linear combination of the basis vectors:
> $$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n,$$
> where $k_1, k_2, \ldots, k_n \in F$ are called the **coordinates** of α with respect to the basis $\alpha_1, \alpha_2, \ldots, \alpha_n$. ♣

The definition leads to a theorem that is easy to see: If there are n linearly independent vectors in a linear space V, and any vector in V can be expressed as a linear combination of these n vectors, then these n vectors form a basis of V, and V is an n-dimensional linear space.

¶ **Linear Subspaces**

> **Definition 2.7 (Linear Subspace)**
>
> Let V be a linear space over field F, and let $W \subseteq V$ be a non-empty subset of V. If W itself is a linear space under the same addition and scalar multiplication as in V, then W is called a **linear subspace** of V.
>
> Zero Subspace and whole space V itself are called trivial subspaces of V. ♣

It can be easily derived that: Let W be a non-empty subset of a linear space V. Then W is a linear subspace of V if and only if:

1. For any $\alpha, \beta \in W$, $\alpha + \beta \in W$.
2. For any $\alpha \in W$ and $k \in F$, $k\alpha \in W$.

2.3 Rank of Vector Groups and Matrices

¶ **Maximal Linearly Independent Group and Rank of Vector Groups**

> **Definition 2.8 (Maximal Linearly Independent Group)**
>
> A subset of a vector group is called a **maximal linearly independent group** if:
>
> 1. The subset itself is linearly independent.
> 2. Adding any vector (if available) from the vector group to this subset makes the new subset linearly dependent. ♣

🖊 **Property**

1. *The maximal linearly independent group is not unique; any maximal linearly independent group is equivalent to the original vector group.*
2. *All maximal linearly independent groups of a vector group contain the same number of vectors.*

> **Definition 2.9 (Rank of Vector Groups)**
>
> The number of vectors contained in the maximal linearly independent group of a vector group is called the **rank** of the vector group. ♣

🖊 **Property**

1. *A vector group is linearly independent if and only if its rank is equal to the number of vectors it contains.*
2. *Equivalent vector groups have the same rank, but vector groups with the same rank are not necessarily equivalent.*

3. *If rank of a vector group is r, then any r linearly independent vectors from the group form a maximal linearly independent group.*

4. *A vector group containing non-zero vectors always has a maximal linearly independent group, and any linearly independent subset of vectors can be expanded into a maximal linearly independent group.*

5. *If vector group (I) can be linearly expressed by vector group (II), then the rank of (I) does not exceed the rank of (II).*

¶ `Rank of Matrices`

2.4 Solution to Linear Systems

¶ `Necessary and Sufficient Conditions for the Existence of Solutions`

¶ `The Structure of Solutions to Homogeneous Systems of Linear Equations`

For a homogeneous system of linear equations, it is evident that the linear combination of solutions is still a solution to the system. Thus, the concept of a **fundamental system of solutions** is introduced.

Definition 2.10 (Fundamental System of Solutions)

The homogeneous system of linear equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0 \end{cases}$$

(or in matrix form $AX = O$) has a set of solutions $\eta_1, \eta_2, \ldots, \eta_t$, which is called a **fundamental system of solutions**, if:

1. Any solution of the system can be expressed as a linear combination of $\eta_1, \eta_2, \ldots, \eta_t$.
2. $\eta_1, \eta_2, \ldots, \eta_t$ are linearly independent[a].

---

[a] ensuring there are no redundant solutions in the fundamental system of solutions

---

⚡**Remark** This definition is similar to the maximal linearly independent group of a vector group. In fact, the fundamental solution set is a maximal linearly independent group of the solution vector group. Therefore, the number of solutions it contains equals its rank.

All solutions of a homogeneous system of linear equations ($AX = O$) form a linear space (solution space), whose basis is the fundamental system of solutions. Dimension of solution space is $n - r(A)$ (where n is the number of variables, and $r(A)$ is the rank of the coefficient matrix A). Thus, a fundamental system of solutions contains exactly $n - r(A)$ solutions.

¶ `The Structure of Solutions to Nonhomogeneous Linear Systems`

The homogeneous system of linear equations $AX = O$ is called the **derived system** of the nonhomogeneous system $AX = B$.

It can be easily proven that:

- The difference between two solutions of $AX = B$ is a solution of its derived system $AX = 0$.
- The sum of a solution of $AX = B$ and a solution of $AX = 0$ is still a solution of $AX = B$.

If γ_0 is a particular solution of $AX = B$, then any solution γ of $AX = B$ can be expressed as:

$$\gamma = \gamma_0 + \eta,$$

where η is a solution of $AX = 0$.

Thus, for any particular solution γ_0 of $AX = B$, when η takes all the solutions of its derived system $AX = 0$, the above formula provides all the solutions of $AX = B$.

Chapter 3 Matrices

3.1 Basic Operations

¶ **Addition**

¶ **Scalar Multiplication**

¶ **Transpose**

¶ **Matrix Multiplication**

> **Theorem 3.1 (Determinant of a Product)**
>
> For square matrices A and B of the same order, the determinant of their product is equal to the product of their determinants:
> $$|AB| = |A| \cdot |B|.$$
> It can be easily extended to multiple matrices:
> $$|A_1 A_2 \cdots A_k| = |A_1| \cdot |A_2| \cdots |A_k|.$$

> **Theorem 3.2 (Cauchy-Binet Formula)**
>
> Let $A = (a_{ij})_{m \times n}$ and $B = (b_{ij})_{n \times m}$:
> 1. If $m > n$, then $|AB| = 0$;
> 2. If $m \leq n$, then $|AB|$ is equal to the sum of products of all m-step minors of A and the corresponding m-step minors of B, that is:
> $$|AB| = \sum_{1 \leq v_1 < v_2 < \cdots < v_m \leq n} \left| A \begin{pmatrix} 1, 2, \ldots, m \\ v_1, v_2, \ldots, v_m \end{pmatrix} \right| \cdot \left| B \begin{pmatrix} v_1, v_2, \ldots, v_m \\ 1, 2, \ldots, m \end{pmatrix} \right|.$$

¶ **Matrix Equivalence**

> **Definition 3.1 (Matrix Equivalence)**
>
> Two matrices A and B are said to be **equivalent**, denoted as $A \cong B$, if they can be transformed into one another by a combination of elementary row and column operations.

Note *Obviously, two matrices are equivalent if and only if they have the same rank.*

> **Definition 3.2 (Canonical Form of Matrix Equivalence)**
>
> For any matrix A with rank r, it can be equivalent to the matrix
> $$J_r = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}_{m \times n},$$
> which is called the **canonical form** of matrix equivalence.

3.2 Special Matrices

¶ **Diagonal Matrix**

¶ **Fundamental Matrix**

> **Definition 3.3 (Fundamental Matrix)**
>
> E_{ij} is called a **fundamental matrix** if the element in the i-th row and j-th column is 1, and all other elements are 0. ♣

🖉 Property

¶ `Triangular Matrix`

¶ `Elementary Matrix`

> **Definition 3.4 (Elementary Matrix)**
>
> An **elementary matrix** is obtained by performing a single elementary row operation on an identity matrix E. There are three types of elementary matrices (also hold for column operations):
>
> **Row swapping** Interchanges rows i and j of E: $P(i, j)$.
>
> **Row scaling** Multiplies row i of E by a non-zero scalar c: $P(i(c))$.
>
> **Row addition** Adds k times row j to row i of E: $P(i, j(k))$. ♣

> **🖋 Note** *It can be proven that $P(i, j)$ can be implemented by a sequence of $P(i(c))$ and $P(i, j(k))$ operations.*

¶ `Symmetric Matrix`

¶ `Circulant-Shift Permutation Matrix`

¶ `Nilpotent Matrix`

¶ `Row Full Rank Matrix`

3.3 Inverse Matrix

¶ `Inverse Matrix and Its Operations`

¶ `Equivalent Propositions and Method of Inversion`

¶ `Generalized Inverse`

3.4 Block Matrix

> **Theorem 3.3 (Determinant Reduction Formula)**
>
> Let $A_{m \times m}, B_{m \times n}, C_{n \times m}, D_{n \times n}$ be matrices. Then:
>
> 1. If A is invertible, then:
> $$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B|.$$
>
> 2. If D is invertible, then:
> $$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \cdot |A - BD^{-1}C|.$$
>
> 3. If both A and D are invertible, then:
> $$|D| \cdot |A - BD^{-1}C| = |A| \cdot |D - CA^{-1}B|.$$ ♡

⚡Remark The mnemonic is: For $\begin{vmatrix} A & B \\ C & D \end{vmatrix}$, for example, if A is invertible, one factor is $|A|$, and the other factor is D (the diagonal element of A) minus the product of the other three terms arranged clockwise, where the

middle one is the inverse matrix.

3.5 Operations of Rank

> *Proposition 3.1*
>
> The matrices A and B in the following operations do not need to be square matrices; they only need to be compatible for multiplication or addition.
>
> **1. Addition**
> $$\mathrm{rank}(A + B) \leqslant \mathrm{rank}(A) + \mathrm{rank}(B).$$
>
> **2. Multiplication**
> $$\mathrm{rank}(AB) \leqslant \mathrm{rank}(A), \quad \mathrm{rank}(AB) \leqslant \mathrm{rank}(B).$$
>
> **2.1. Sylvester's Inequality**
> $$\mathrm{rank}(AB) \geqslant \mathrm{rank}(A) + \mathrm{rank}(B) - n \quad (A_{s\times n}, B_{n\times m}).$$
>
> Specially, if $AB = O$, then:
> $$\mathrm{rank}(A) + \mathrm{rank}(B) \leqslant n.$$
>
> **2.2. Frobenius Inequality**
> $$\mathrm{rank}(ABC) \geqslant \mathrm{rank}(AB) + \mathrm{rank}(BC) - \mathrm{rank}(B).$$
>
> **3. Transpose**
> $$\mathrm{rank}(AA^T) = \mathrm{rank}(A^T A) = \mathrm{rank}(A) = \mathrm{rank}(A^T).$$
>
> **4. Inverse**
> $$\mathrm{rank}(A) = \mathrm{rank}(A^{-1}) = n.$$
>
> **5. Block Matrix**
> $$\mathrm{rank}\begin{pmatrix} A & O \\ O & D \end{pmatrix} = \mathrm{rank}(A) + \mathrm{rank}(D).$$

3.6 Low–Rank Update

Due to all the row and column vectors of a rank-1 matrix are linearly dependent, it can be expressed as the outer product of two non-zero vectors; in other words, a rank-1 matrix can be expressed as $\alpha\beta^T$, where α and β are non-zero column vectors.

Based on the decomposition $A = \alpha\beta^T$, the matrix of rank-1 has simplified calculation rules:

🔗 **Property**

Exponentiation *For any positive integer $k \geq 1$,*
$$A^k = (\beta^T \alpha)^{k-1} \cdot A,$$

where $\beta^T \alpha$ is a constant (the inner product of vectors).

Rank Transmission *If B is any matrix, then:*
$$\mathrm{rank}(AB) \leq 1 \quad and \quad \mathrm{rank}(BA) \leq 1,$$

(rank 1 matrices multiplied by arbitrary matrices result in ranks not exceeding 1).

> **Theorem 3.4 (Sherman-Morrison Formula)**
>
> If $A \in \mathbb{R}^{n \times n}$ is an invertible matrix, and $\alpha, \beta \in \mathbb{R}^n$ are column vectors, then $A + \alpha\beta^T$ is invertible if and only if $1 + \beta^T A^{-1} \alpha \neq 0$. In this case, the inverse of $A + \alpha\beta^T$ is given by:
>
> $$\left(A + \alpha\beta^T\right)^{-1} = A^{-1} - \frac{A^{-1}\alpha\beta^T A^{-1}}{1 + \beta^T A^{-1}\alpha},$$
>
> where $\alpha\beta^{\mathrm{T}}$ is the outer product of α and β.

🖋 **Note** *Combining the properties of determinants, we can derive the determinant version of the Sherman-Morrison formula:*

$$\left|A + \alpha\beta^T\right| = |A| \cdot \left(1 + \beta^T A^{-1}\alpha\right),$$

*which is known as the **matrix determinant lemma**.*

The theorem can also be stated in terms of the adjugate matrix of A:

$$\det(A + uv^T) = \det(A) + v^T \operatorname{adj}(A)u,$$

in which case it applies whether or not the matrix A is invertible.

Chapter 4 Linear Spaces

4.1 Linear Spaces and Their Bases

¶ `Basis Transformation and Coordinate Transformation`

4.2 Subspaces

¶ `Intersection and Sum of Subspaces`

> **Definition 4.1 (Intersection and Sum of Subspaces)**
>
> Let V be a linear space over field F, and let $V_1, V_2 \subseteq V$ be two subspaces of V. The **intersection** of V_1 and V_2 is defined as:
> $$V_1 \cap V_2 = \{\alpha | \alpha \in V_1 \text{ and } \alpha \in V_2\}.$$
>
> The **sum** of V_1 and V_2 is defined as:
> $$V_1 + V_2 = \{\alpha | \alpha = \alpha_1 + \alpha_2, \alpha_1 \in V_1, \alpha_2 \in V_2\}.$$
> ♣

If V_1 and V_2 are two subspaces of V, then both their intersection $V_1 \cap V_2$ and their sum $V_1 + V_2$ are also subspaces of V. Simultaneously, $V_1 + V_2$ is the smallest subspace of V that contains both V_1 and V_2 ($V_1 \cup V_2$).

¶ `Dimension Formula`

> **Proposition 4.1**
>
> In a finite-dimensional linear space V,
> $$L(\alpha_1, \alpha_2, \ldots, \alpha_s) + L(\beta_1, \beta_2, \ldots, \beta_t) = L(\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_1, \beta_2, \ldots, \beta_t).$$
> ♠

> **Theorem 4.1 (Dimension Formula)**
>
> Let V be a finite-dimensional linear space over field F, and let $V_1, V_2 \subseteq V$ be two subspaces of V. Then:
> $$\dim(V_1) + \dim(V_2) = \dim(V_1 + V_2) + \dim(V_1 \cap V_2).$$
> ♡

¶ `Direct Sum of Subspaces`

> **Definition 4.2 (Direct Sum of Subspaces)**
>
> Let V be a linear space over field F, and let $V_1, V_2 \subseteq V$ be two subspaces of V. If any vector $\alpha \in V_1 + V_2$ can be uniquely expressed as:
> $$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in V_1, \alpha_2 \in V_2,$$
>
> then $V_1 + V_2$ is called the **direct sum** of V_1 and V_2, denoted as $V_1 \oplus V_2$.
> ♣

> **Proposition 4.2**
>
> The necessary and sufficient condition for $V_1 + V_2$ to be a direct sum is:
> 1. $0 = 0 + 0$ (the zero vector can only be expressed as the sum of two zero vectors);
> 2. $V_1 \cap V_2 = \{0\}$ (the intersection of the two subspaces is only the zero vector);
> 3. $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2)$, or equivalently, $\dim(V_1 \cap V_2) = 0$.

4. Any bases of V_1 and V_2 together form a basis of $V_1 + V_2$.

4.3 Isomorphisms

4.4 Quotient Spaces

Chapter 5 Linear Mappings

5.1 Linear Mappings and Their Computation

¶ **Definition of Linear Mappings**

> *Definition 5.1 (Linear Mapping)*
>
> Let V and V' be linear spaces over field F. A mapping $\mathcal{A} : V \to V'$ is called a **linear mapping** if:
> $$\mathcal{A}(\alpha + \beta) = \mathcal{A}\alpha + \mathcal{A}\beta, \quad \forall \alpha, \beta \in V,$$
> and
> $$\mathcal{A}(k\alpha) = k\mathcal{A}\alpha, \quad \forall \alpha \in V, k \in F.$$
> If $V = V'$, then $\mathcal{A}$ is called a **linear transformation** of V. The linear mapping from V over field F to F is called a **linear function** on V. ♣

✗**Remark** The field naturally constitutes a one-dimensional linear space over itself, so a linear function is a special case of a linear mapping.

✗**Remark** The set of all linear mappings from V to V' over the field F is denoted by $\mathrm{Hom}(V, V')$ or $\mathcal{L}(V, V')$.

The set of all linear transformations is denoted by $\mathrm{Hom}(V)$, $\mathcal{L}(V)$, or $\mathrm{End}(V)$.

The set of all $m \times n$ matrices over the number field P is denoted by $M_{m \times n}(P)$ or $P^{m \times n}$.

¶ **Existence and Uniqueness of Linear Mappings**

¶ **Operations of Linear Mappings**

In the following, all linear spaces are over the field F.

Addition Let $\mathcal{A}, \mathcal{B} \in \mathrm{Hom}(V, V')$. The sum of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A} + \mathcal{B}$, is defined by:
$$(\mathcal{A} + \mathcal{B})(\alpha) = \mathcal{A}(\alpha) + \mathcal{B}(\alpha), \quad \forall \alpha \in V.$$

The sum of linear maps is still a linear map, i.e., $\mathcal{A} + \mathcal{B} \in \mathrm{Hom}(V, V')$.

1. **Commutativity:** $\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$,
2. **Associativity:** $(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$,
3. **Additive Identity:** The zero map $\prime$ satisfies $\mathcal{A} + \prime = \mathcal{A}$,
4. **Additive Inverse:** For every $\mathcal{A}$, the additive inverse $-\mathcal{A}$ satisfies $\mathcal{A} + (-\mathcal{A}) = \prime$.

Scalar Multiplication Let $k \in F$ and $\mathcal{A} \in \mathrm{Hom}(V, V')$. The scalar product $k\mathcal{A}$ is defined by:
$$(k\mathcal{A})(\alpha) = k\mathcal{A}(\alpha), \quad \forall \alpha \in V.$$

Scalar multiplication of a linear map is still a linear map, i.e., $k\mathcal{A} \in \mathrm{Hom}(V, V')$.

1. **Unit Identity:** $1 \cdot \mathcal{A} = \mathcal{A}$,
2. **Associativity of Scalars:** $(kl)\mathcal{A} = k(l\mathcal{A})$,
3. **Distributive Properties:**
$$k(\mathcal{A} + \mathcal{B}) = k\mathcal{A} + k\mathcal{B},$$
$$(k + l)\mathcal{A} = k\mathcal{A} + l\mathcal{A}.$$

Multiplication (Composition of Linear Maps) Let $\mathcal{A} \in \mathrm{Hom}(V', W)$ and $\mathcal{B} \in \mathrm{Hom}(U, V')$. The product (composition) of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A}\mathcal{B}$, is defined by:
$$(\mathcal{A}\mathcal{B})(\alpha) = \mathcal{A}(\mathcal{B}(\alpha)), \quad \forall \alpha \in U.$$

The composition of linear maps is still a linear map, i.e., $\mathcal{AB} \in \mathrm{Hom}(U, W)$.

1. **Associativity of Composition:** $(\mathcal{AB})\mathcal{C} = \mathcal{A}(\mathcal{BC})$,
2. **Distributive Properties:**

$$\mathcal{A}(\mathcal{B} + \mathcal{C}) = \mathcal{AB} + \mathcal{AC},$$
$$(\mathcal{A} + \mathcal{B})\mathcal{C} = \mathcal{AC} + \mathcal{BC},$$

3. **Non-Commutativity:** In general, $\mathcal{AB} \neq \mathcal{BA}$,
4. **Identity Transformation:** The identity transformation $\mathcal{E}$ satisfies:

$$\mathcal{EA} = \mathcal{AE} = \mathcal{A}, \quad \forall \mathcal{A} \in \mathrm{Hom}(V).$$

Let V and V' be vector spaces over the field F.

1. The collection of all linear mappings from V to V', denoted by $\mathrm{Hom}(V, V')$, under the addition and scalar multiplication defined above, forms a linear space over the field F.
2. The collection of all linear transformations on V, denoted by $\mathrm{Hom}(V)$, under the addition and composition of maps defined above, forms a unital ring. In this case, it is referred to as the **endomorphism ring** and is denoted by $\mathrm{End}(V)$.

> **Definition 5.2 (Algebra over a Field)**
>
> Let V be a linear space over the field F equipped with an additional binary operation from $V \times V$ to V, denoted by $\cdot$. Then V is called an **algebra over the field** F if (forall $\alpha, \beta, \gamma \in V$ and $k, l \in F$):
>
> **Right Distributivity** $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$;
>
> **Left Distributivity** $\gamma \cdot (\alpha + \beta) = \gamma \cdot \alpha + \gamma \cdot \beta$;
>
> **Compatibility with Scalar Multiplication** $(k\alpha) \cdot (l\beta) = kl(\alpha \cdot \beta)$. ♣

⚡**Remark** The left and right distributivity conditions can be combined into a single condition: V forms a unital ring under the addition and scalar multiplication.

It is evident that $\mathrm{Hom}(V)$ is an algebra over the field F.

5.2 Kernel and Image of Linear Mappings

> **Definition 5.3 (Kernel and Image of Linear Mappings)**
>
> Let V and V' be linear spaces over field F, and let $\mathcal{A} : V \to V'$ be a linear mapping. The **kernel** of $\mathcal{A}$ is defined as:
>
> $$\mathrm{Ker}(\mathcal{A}) = \{\alpha | \alpha \in V, \mathcal{A}(\alpha) = 0\}.$$
>
> The **image** of $\mathcal{A}$ is defined as:
>
> $$\mathrm{Im}(\mathcal{A}) = \{\beta | \beta = \mathcal{A}(\alpha), \alpha \in V, \beta \in V'\}.$$
>
> They can be also denoted as $\mathcal{A}^{-1}(0)$ and $\mathcal{A}(V)$ respectively.
>
> The rank of $\mathrm{Im}(\mathcal{A}$ is called the **rank** of $\mathcal{A}$; the dimension of $\mathrm{Ker}\ \mathcal{A}$ is called the **nullity** of $\mathcal{A}$. ♣

🔗 **Property**

> **Theorem 5.1 (Rank-Nullity Theorem)**
>
> Let V and V' be finite-dimensional linear spaces over field F, and let $\mathcal{A} : V \to V'$ be a linear mapping. Then:
> $$\dim(\operatorname{Ker} \mathcal{A}) + \dim(\operatorname{Im} \mathcal{A}) = \dim(V).$$

> **Definition 5.4 (Cokernel)**
>
> Let V and V' be linear spaces over field F, and let $\mathcal{A} : V \to V'$ be a linear mapping. The **cokernel** of $\mathcal{A}$ is defined as the quotient space:
> $$\operatorname{Coker}(\mathcal{A}) = V'/\operatorname{Im}(\mathcal{A}),$$
>
> denoted as $\operatorname{Coker}(\mathcal{A})$.

5.3 Matrix Representation of Linear Mappings

Let V and V' be finite-dimensional linear spaces over field F, $\dim(V) = n$ and $\dim(V') = m$, and let $\mathcal{A} : V \to V'$ be a linear mapping. Let $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ be a basis of V, and $\{\eta_1, \eta_2, \ldots, \eta_m\}$ be a basis of V'. Then the image of each basis vector of V under $\mathcal{A}$ can be expressed as the basis vectors of V':

$$\begin{cases} \mathcal{A}(\varepsilon_1) = a_{11}\eta_1 + a_{21}\eta_2 + \cdots + a_{m1}\eta_m, \\ \mathcal{A}(\varepsilon_i) = a_{12}\eta_1 + a_{22}\eta_2 + \cdots + a_{m2}\eta_m, \\ \qquad\qquad\qquad\vdots \\ \mathcal{A}(\varepsilon_n) = a_{1n}\eta_1 + a_{2n}\eta_2 + \cdots + a_{mn}\eta_m. \end{cases}$$

It can be expressed in matrix form as:

$$\mathcal{A}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = (\mathcal{A}\varepsilon_1, \mathcal{A}\varepsilon_2, \ldots, \mathcal{A}\varepsilon_n) = (\eta_1, \eta_2, \ldots, \eta_m)A,$$

where

$$A = (a_{ij})_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

The matrix A is called the **matrix representation** of the linear mapping $\mathcal{A}$ under the bases $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ of V and $\{\eta_1, \eta_2, \ldots, \eta_m\}$ of V'.

5.4 Special Linear Transformations

Some common linear transformations include:

Identity Transformation The identity transformation $\mathcal{E}$ on V is defined by:
$$\mathcal{E}(\alpha) = \alpha, \quad \forall \alpha \in V.$$

Zero Transformation The zero transformation $\mathcal{O}$ from V to V' is defined by:
$$\mathcal{O}(\alpha) = 0, \quad \forall \alpha \in V.$$

Scalar Transformation For a fixed scalar $k \in F$, the scalar transformation $\mathcal{K}$ on V is defined by:

$$\mathcal{K}(\alpha) = k\alpha, \quad \forall \alpha \in V.$$

When $k = 1$, it is the identity transformation; when $k = 0$, it is the zero transformation.

Spin Axis Transformation The vectors on a plane form a two-dimensional linear space over real number field $\mathbb{R}$. Rotate the plane counterclockwise by an angle θ around the origin, then the transformation of any vector on the plane is a linear transformation, denoted by $\mathcal{P}_\theta$. Similarly, in three-dimensional space, rotating around a fixed axis by an angle θ also defines a linear transformation.

And some special linear transformations are given as follows:

Definition 5.5 (Idempotent Transformation)

A linear transformation $\mathcal{A}$ on V is called an **idempotent transformation** if:

$$\mathcal{A}^2 = \mathcal{A}.$$

♣

Two linear transformations $\mathcal{A}$ and $\mathcal{B}$ on V are called **orthogonal** if:

$$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} = \mathcal{O}.$$

With the definitions above, we can define the projection transformation as follows:

Definition 5.6 (Projection Transformation)

Let V be a linear space over field F, and let $U, W \subseteq V$ be two subspaces of V such that $V = U \oplus W$. For all $\alpha \in V$, there exist unique $\alpha_1 \in U$ and $\alpha_2 \in W$ such that:

$$\alpha = \alpha_1 + \alpha_2.$$

Let

$$\mathcal{P}_U : \begin{array}{c} V \to V \\ \alpha \mapsto \alpha_1 \end{array}$$

then $\mathcal{P}_U$ is called the **projection** onto U along W, which is a linear transformation on V. Similarly, the projection onto W along U can be defined.

♣

$\mathcal{P}_U$ satisfies and uniquely satisfies:

$$\mathcal{P}_U(\alpha) = \begin{cases} \alpha, & \alpha \in U, \\ 0, & \alpha \in W. \end{cases}$$

5.5 Linear Functions and Dual Spaces

Chapter 6   Diagonalization

6.1 Similarity of Matrices

6.2 Eigenvectors and Diagonalization

6.2.1 Eigenvalues and Eigenvectors

6.2.2 Necessary and Sufficient Conditions for Diagonalization

¶ Geometric Multiplicity of Eigenvectors
¶ Algebraic Multiplicity

6.3 Space Decomposition and Diagonalization

6.3.1 Invariant Subspace

6.3.2 Hamilton–Cayley Theorem

6.4 Least Squares and Diagonalization

Chapter 7 Jordan Forms

Chapter 8　Quadratic Forms

Chapter 9   Inner Product Spaces

Bibliography

[1] 丘维声, 高等代数 *(2nd edition)*, 清华大学出版社, 2019.

[2] 谢启鸿, 姚慕生, 吴泉水, 高等代数学 *(4th edition)*, 复旦大学出版社, 2022.

[3] 王萼芳, 石生明 高等代数 *(5th edition)*, 高等教育出版社, 2019.

[4] 樊启斌, 高等代数典型问题与方法 *(1st edition)*, 高等教育出版社, 2021.

[5] Wikipedia. https://en.wikipedia.org/wiki/.