MATHS FOR ECONOMISTS

A COMPREHENSIVE HANDBOOK

© Zhenran Guo[†]
[†] SUFE, Shanghai, China

Maths for Economists: A Comprehensive Handbook

Zhenran Guo

5th February 2026

Preface

This book is written as a handbook for economists rather than a note for courses like Econ Math, where materials are selected from various sources including course handouts from SUFE, lecture notes from open scources and many great books. To make it a comprehensive one, I'll try to cover almost all maths needed for studying economics, including logic, algebra, analysis, topology, probability theory and the theory of optimization. And I'll add something more, basically based on the syllabus of the course I've taken. In this way can those(especially from SUFE) who're taking courses like Mathematical analysis refer to it.

Though served as a handbook, the book will be written as intuitive as possible so that all techniques and definitions are expressed in a natural way and readers can better understand them.

The book may be updated in a monthly basis.

Acknowledgement

I would like to thank Copilot and Deepseek for helping me writing and editing this book.

Contents

I Foundations 1

1 Logic 3
1.1 Negation . . . 3
1.2 Conjunction and Disjunction . . . 4
1.3 Quantifiers . . . 4
1.3.1 Property . . . 4
1.3.2 Forall and Exists . . . 4
1.4 Implications . . . 5
1.4.1 Implication . . . 5
1.4.2 Equivalence . . . 5
1.4.3 Prove by Contrapositive . . . 5
1.4.4 Prove by Contradiction . . . 6

2 Set, Relation and Operation 7
2.1 Operations of Sets . . . 7
2.1.1 Complement, Intersection and Union . . . 7
2.1.2 Power Set . . . 8
2.1.3 Cartesian Product . . . 8
2.1.4 Family of Sets . . . 9
2.2 Relation and Operation . . . 10
2.2.1 Order . . . 11
2.2.2 Mapping . . . 12
2.2.3 Correspondence . . . 13

2.2.4 Operation . . . 14
2.3 Countability . . . 14
2.3.1 The Axiom of Choice . . . 14
2.3.2 Cardinality . . . 15
2.3.3 Countable Set . . . 16

3 Group, Ring and Field 19
3.1 Group . . . 19
3.1.1 the Definition . . . 19
3.1.2 Homomorphisms . . . 20
3.1.3 Isomorphisms . . . 21
3.2 Ring and Field . . . 22
3.2.1 The Definition . . . 22
3.2.2 Homomorphism and Isomorphism . . . 22
3.2.3 Ordered Field . . . 23
3.3 Polynomials . . . 23

II Vector Space and Linear Algebra 25

4 Vector Space 27
4.1 the Definition of Vector Space and Subspaces . . . 27
4.2 Operation of Subspaces . . . 28
4.2.1 Sum . . . 28
4.2.2 Direct Sum . . . 28
4.3 Subspace Spanned by Vectors . . . 29
4.3.1 Span . . . 29
4.3.2 Linear Independence . . . 29
4.4 Finite Dimensional Vector Space . . . 31
4.4.1 Basis . . . 31
4.4.2 Dimension . . . 32

5 Linear Mapping and Matrix 33
5.1 Linear Mapping . . . 33
5.1.1 the Definitions . . . 33
5.1.2 Isomorphic Linear Mapping . . . 34

5.2 Matrix 35
  5.2.1 From Vector to Coordinate 37
  5.2.2 From Linear Mapping to Matrix 37
  5.2.3 Operations of Matrix 38
  5.2.4 Change of Basis 41

6 Normed Vector Space **43**

7 Multilinear Form and Determinant **45**
7.1 Multilinear Form 47
  7.1.1 Bilinear Form 47
    7.1.1.1 Symmetric and Alternating Bilinear Form 48
    7.1.1.2 Matrix of a Bilinear Form 49
  7.1.2 Multilinear Form 49
    7.1.2.1 Alternating Multilinear Form 49
    7.1.2.2 Swapping the Entries 50

8 Eigenvalue and Eigenvector **53**

III Metric Space and Topology Space **55**

9 Metric Space and Convergence **57**
9.1 Metric Space 57
  9.1.1 the Definition 57
  9.1.2 Neighborhood and Cluster Point 57
  9.1.3 Induced Metric 58
9.2 Convergence and Infinite Limit 58
  9.2.1 Convergence of a Sequence 58
  9.2.2 Infinite Limit 58
  9.2.3 Limit Superior and Limit Inferior 58
  9.2.4 Some Special Sequences 58
    9.2.4.1 Monotone Sequence 58
    9.2.4.2 Subsequence 58
  9.2.5 Banach Space 58

IV Differentiation in One Variable **59**

V More in Convergence: Uniform Convergence **61**

A Notation **63**

B Supplementary Scripts **65**

Part I

Foundations

Chapter 1

Logic

To make complex ideas clear, we often separate them into statements and check whether they're right or wrong. Here we use symbolic logic. Symbolic logic is basically about statements which can be claimed to be either true or false and "operations" of statements.

We start with what is a statement. For sure, most of the sentences can be seen as a statement. But to notice, not all sentences are statements. For example, "This sentence is false." is not a statement, that is because when it is true, it is false and vice versa.

In maths, we call statements as propositions and we can categorize them into theorems, axioms, lemmas, corollaries, etc. Axioms are statements that we accept to be true without proof. Theorems are statements that can be proven to be true based on axioms and previously proven theorems. Lemmas are "helper" theorems used to prove larger theorems and corollaries are statements that follow directly from a theorem.

1.1 Negation

We use the symbol ¬ to express "NOT". Suppose A is a statement, then $\neg A$ is also a statement and $\neg A$ means "A is NOT true". We can use a truth table to show the relationship between A and $\neg A$.

Statement	A	$\neg A$
Truth	T	F
Value	F	T

1.2 Conjunction and Disjunction

Conjunction and Disjunction are ways to combine statements together. We use the symbol $\land$ to express "AND". Suppose A and B are statements, then $A \land B$ is also a statement and $A \land B$ is true only when both A and B are true. Similarly, we use the symbol $\lor$ to express "OR". Suppose A and B are statements, then $A \lor B$ is also a statement and $A \lor B$ is false only when both A and B are false.

We can use a truth table to show the relationship between A, B, $A \land B$ and $A \lor B$.

Statement	A	B	$A \land B$	$A \lor B$
	T	T	T	T
Truth	T	F	F	T
Value	F	T	F	T
	F	F	F	F

1.3 Quantifiers

1.3.1 Property

We first define the concept of property. We say $P(x)$ is a property of x (x is from a particular class) if when x is replaced with a certain object, $P(x)$ becomes a statement. For example, let $P(x)$ be "x is even", then $P(2)$ is true and $P(3)$ is false. The set $\{x : P(x)\}$ consists of all values of x such that $P(x)$ is true.

1.3.2 Forall and Exists

Now we can create another kind of statement using quantifiers. There are two quantifiers: "for all" and "there exists". The expression $\forall x \in X : P(x)$ means "for all components x in the set X, $P(x)$ is true". And the expression $\exists x \in X : P(x)$ means "there exists at least one component x in the set X, such that $P(x)$ is true".

A very important proposition is that

$$\neg[\exists x \in X : P(x)] = \forall x \in X : \neg P(x) \tag{1.1}$$

And to understand all statements, we also need $\neg(\neg A) = A, \neg(A \wedge B) = \neg A \vee \neg B, \neg(A \vee B) = \neg A \wedge \neg B$. This becomes trivial to deal with quantifiers that we only have to interchange $\vee$ and $\wedge$ and change $\forall$ to $\exists$ and vice versa when we negate a statement.

1.4 Implications

1.4.1 Implication

The implication $(A \implies B) := (\neg A) \vee B$ is false if and only if A is true and B is false. The definition is simple, but note that $A \implies B$ not necessarily mean that A causes B to be true. When A and B are false, the implication is also true, which makes it different from the familiar meaning of "imply".

We also say that A is a sufficient condition for B, and B is a necessary condition for A when we write the statement $A \implies B$.

1.4.2 Equivalence

When A is both necessary and sufficient for B, we say that A is equivalent to B, and we write $A \iff B$. Note that $A \iff B$ is equivalent to $(A \implies B) \wedge (B \implies A)$.

1.4.3 Prove by Contrapositive

With equivalence, we know that the equation in (1.1) actually means equivalent and we can make clear of the above relationships. Moreover, using the interchanging technique, we can find the counterpositive statement

$$(A \implies B) \iff (\neg B \implies \neg A)$$

This inspires us to prove a statement, we can turn to its contrapositive if difficulties arise when proving the original statement.

1.4.4 Prove by Contradiction

Consider statements $(A \implies C), (C \implies B)$, then

$$(A \implies B) \iff (A \implies C) \land (C \implies B)$$

Suppose B is false and assume A is true, then $(C \implies B)$ and $(A \implies C)$ are false, that is to say, we can find a statement C with the false truth value.

This leads us to the proof by contradiction method. To prove $A \implies B$, we can assume A is true and B is false, then try to find a statement with a false truth value and we prove the original statement.

Chapter 2

Set, Relation and Operation

Since all readers should be familiar with the high-school-level set theory, we only give a brief review of the basic concepts and operations of sets.

We can't define what is a set but we can give it a description. A set is all objects that satisfy a certain property. The following lists some of the basics.

Definition 2.0.1 (Subset)**.** Let A, B be sets, when we say $A \subset B$, we mean $\forall x \in A, x \in B$. Notice

$$A = B \iff (A \subset B) \land (B \subset A)$$

2.1 Operations of Sets

2.1.1 Complement, Intersection and Union

Definition 2.1.1 (Complement, Intersection and Union)**.** Let A, B be sets, we define the following operations:

Complement The (relative) complement of set B in set A is defined as $A \setminus B = \{x : (x \in A) \land (x \notin B)\}$.

Intersection The intersection of sets A and B is defined as $A \cap B = \{x : (x \in A) \land (x \in B)\}$.

Union The union of sets A and B is defined as $A \cup B = \{x : (x \in A) \vee (x \in B)\}$.

By definition, one can easily check:

Proposition 2.1.1. *Let A, B, C be sets, then*

Commutativity $A \cap B = B \cap A$ *and* $A \cup B = B \cup A$.

Associativity $(A \cap B) \cap C = A \cap (B \cap C)$ *and* $(A \cup B) \cup C = A \cup (B \cup C)$.

Distributivity $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *and* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

de Morgan's Laws $(A \cap B)^c = A^c \cup B^c$ *and* $(A \cup B)^c = A^c \cap B^c$

Subset $A \subset B \iff A \cap B = A \iff A \cup B = B$.

2.1.2 Power Set

Definition 2.1.2 (Power Set)**.** Let A be a set, define:

$$2^A = \{X : X \subset A\}$$

2.1.3 Cartesian Product

We define an ordered pair or a n-tuple $x = (x_1, x_2, \cdots, x_n)$ and the equity means that all components are equal. Denote $x_j := pr_j(x)$ and we call it the jth projection of x.

The Cartesian product is to describe the set of ordered pairs.

Definition 2.1.3 (Cartesian Product)**.** Let A, B be sets, define:

$$A \times B = \{(a, b) : a \in A, b \in B\}$$

and by deduction,

$$\prod_{i=1}^{n} A_i = (A_1 \times \cdots \times A_{n-1}) \times A_n$$

Specially, when $A_i = A$ for all i, we denote:

$$A^n = \prod_{i=1}^{n} A$$

To practice logic, we show how to prove the following proposition.

Proposition 2.1.2. *Let A, B be sets, then $A \times B = \emptyset \iff (A = \emptyset) \vee (B = \emptyset)$*

Proof. ($\Rightarrow$) We prove by contradiction. Suppose $A \times B$ is non-empty, which implies there exists $(a, b) \in A \times B$, then by definition of Cartesian product, we have $a \in A$ and $b \in B$, contradicting the assumption.

($\Leftarrow$) This side can be similarly proved by contradiction. □

2.1.4 Family of Sets

We now extend the operation of intersection and union a little bit. Consider we're intersecting a series of sets $A, B, C, \cdots$, we first rename the sets as $\{A_i\}_{i \in I}$, where I is called an index set, and intersect them according to the index i, that is

$$\bigcap_{i \in I} A_i = \{x : \forall i \in I, x \in A_i\}.$$

Similarly, for union, we have

$$\bigcup_{i \in I} A_i = \{x : \exists i \in I, x \in A_i\}.$$

Proposition 2.1.3. *Let $\{A_\alpha : \alpha \in I_A\}$ and $\{B_\beta : \beta \in I_B\}$ be families of subsets of a set X, then*

(1) Associativity:

$$\left(\bigcap_{\alpha \in I_A} A_\alpha\right) \cap \left(\bigcap_{\beta \in I_B} B_\beta\right) = \bigcap_{(\alpha, \beta) \in I_A \times I_B} (A_\alpha \cap B_\beta)$$

and

$$\left(\bigcup_{\alpha\in I_A} A_\alpha\right) \cup \left(\bigcup_{\beta\in I_B} B_\beta\right) = \bigcup_{(\alpha,\beta)\in I_A\times I_B} (A_\alpha \cup B_\beta)$$

(2) Distributivity:

$$A \cap \left(\bigcup_{\alpha\in I_A} A_\alpha\right) = \bigcup_{\alpha\in I_A}(A \cap A_\alpha)$$

and

$$A \cup \left(\bigcap_{\alpha\in I_A} A_\alpha\right) = \bigcap_{\alpha\in I_A}(A \cup A_\alpha)$$

(3) de Morgan's Laws:

$$\left(\bigcap_{\alpha\in I_A} A_\alpha\right)^c = \bigcup_{\alpha\in I_A} A_\alpha^c$$

and

$$\left(\bigcup_{\alpha\in I_A} A_\alpha\right)^c = \bigcap_{\alpha\in I_A} A_\alpha^c$$

Proof. We show how to prove (3). A common trick we use when proving the equality of 2 sets is to show that each set is the subset of the other set.

For $\subset$, $\forall a \in \left(\bigcap_{\alpha\in I_A} A_\alpha\right)^c$, since $a \in \bigcap_{\alpha\in I_A} A_\alpha \iff \forall\alpha \in I_A, a \in A_\alpha$, we have $a \in \left(\bigcap_{\alpha\in I_A} A_\alpha\right)^c \iff \exists\alpha_0 \in I_A, a \in A_{\alpha_0}^c$ by negation. Thus $a \in A_{\alpha_0}^c \subset \bigcup_{\alpha\in I_A} A_\alpha^c$.

For the other side, there exists an $\alpha_0 \in I_A$ such that $a \in A_{\alpha_0}^c$. Notice that $\bigcap_{\alpha\in I_A} A_\alpha \subset A_{\alpha_0}$, we have $A_{\alpha_0}^c \subset \left(\bigcap_{\alpha\in I_A} A_\alpha\right)^c$. □

2.2 Relation and Operation

This subsection is to formally introduce relationship between elements in a set X. A binary relation R on a set X is a subset of the Cartesian product $X \times X$. For $(x, y) \in R$, we denote it as xRy.

There're several important types of relations.

Relation	Meaning
Reflexive	$\forall x \in X, xRx$
Symmetric	$xRy \implies yRx$
Transitive	$(xRy) \wedge (yRz) \implies (xRz)$
Complete	$\forall x, y \in X, (xRy) \vee (yRx)$

We now turn to some special relations.

2.2.1 Order

A relation $\leq$ on a set X is a partial order if it's reflexive, transitive, and anti-symmetric*. We call the pair $(X, \leq)$ a partially ordered set. If, in addition to reflexiveness, transitivity, and anti-symmetric, completeness also holds for $\leq$, we say it's a total order and $(X, \leq)$ is a total ordered set.

We now turn to the elements in a partially ordered set $(X, \leq)$.

Definition 2.2.1 (Monotonicity)**.** Given a partially ordered set $(X, \leq)$, for $\{x_n\}_{n=1}^{\infty} \subset X$, we say it's **monotonically increasing** if $x_n \leq x_{n+1}, \forall n \in \mathbb{N}$; similarly, it's **monotonically decreasing** if $x_{n+1} \leq x_n, \forall n \in \mathbb{N}$.

Definition 2.2.2 (Boundedness)**.** Given a partially ordered set $(X, \leq)$ and $A \subseteq X$, an element x is an **upper bound** of A if $a \leq x, \forall a \in A$. We say A is bounded above. And similarly, an element x is a **lower bound** of A if $x \leq a, \forall a \in A$ and it's bounded below.

Definition 2.2.3 (Supremum and Infimum)**.** Given a partially ordered set $(X, \leq)$ and $A \subseteq X$, if A is bounded above, then we define

$$\sup A = \min\{x \in X : x \text{ is an upper bound of } A\}$$

Similarly, if A is bounded below,

$$\inf A = \max\{x \in X : x \text{ is a lower bound of } A\}$$

---

* Anti-symmetry means $(x \leq y) \wedge (y \leq x) \implies x = y$

Note 1. Note that, if $\sup A$ and $\max A$ exists, $\sup A \in A \iff \sup A = \max A$. Similarly, if $\inf A$ and $\min A$ exists, $\inf A \in A \iff \inf A = \min A$.

A very important example of partially order is the subset relation. We introduce the concept of lim inf and lim sup of a sequence of sets as an example, which is of great importance in measure theory.

Example 2.2.1. Let $\{A_i\}_{i=1}^{\infty}$ be a family of subsets of a set X. Let $\mathcal{A}^{(n)}$ and $\mathcal{A}_{(n)}$ denote

$$\mathcal{A}^{(n)} = \bigcap_{i=n}^{\infty} A_i, \quad \mathcal{A}_{(n)} = \bigcup_{i=n}^{\infty} A_i$$

and notice that

$$\mathcal{A}^{(n+1)} \subset \mathcal{A}^{(n)}, \mathcal{A}_{(n)} \subset \mathcal{A}_{(n+1)}, \quad \forall n \in \mathbb{N}$$

Since $\left(\left\{\mathcal{A}^{(n)}\right\}_{n=1}^{\infty}, \subset\right)$ and $\left(\left\{\mathcal{A}_{(n)}\right\}_{n=1}^{\infty}, \subset\right)$ are partially ordered sets, we have

$$\inf \left\{\mathcal{A}^{(n)}\right\}_{n=1}^{\infty} = \bigcup_{n=1}^{\infty} \mathcal{A}^{(n)} = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i$$
$$\sup \left\{\mathcal{A}_{(n)}\right\}_{n=1}^{\infty} = \bigcap_{n=1}^{\infty} \mathcal{A}_{(n)} = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$$

denoted as $\liminf_{n\to\infty} A_n$ and $\limsup_{n\to\infty} A_n$ respectively.

The concept of partial order is important in economics as preference is actually a partial order. We can define certain kinds of preferences with ease.

Definition 2.2.4 (Rational Preference)**.** A preference is rational if and only if it's complete and transitive.

2.2.2 Mapping

We've encountered mappings before, where we say mapping is a rule that assigns each element of a set, the domain, to exactly one element of the image set. If treated carefully, we may find the word "rule" vague. We now try to give it a formal definition.

Consider the set $G = \{(x, f(x)) : x \in X\}$. To make $f : X \to Y$ a mapping, given any $x \in X$, there should exist only one $y \in Y$ such that $(x, y) \in G$. This leads to the following definition.

Definition 2.2.5 (Mapping)**.** A mapping $f : X \to Y$ is a binary relation

$$G = \{(x, f(x)) : x \in X \wedge (\exists ! y \in Y \text{s.t.} f(x) = y)\}$$

and we denote $f(A) = \operatorname{Im} f|_A$, where $A \subset X$.

We now list some important concepts.

Definition 2.2.6. Let $f : X \to Y$ be a mapping.

Surjective f is surjective $\iff$ $\operatorname{Im} f = Y$.

Injective f is injective $\iff$ $(f(x) = f(y) \implies x = y)$.

Bijective f is bijective $\iff$ f is both injective and surjective.

Inverse Suppose f is bijective, then there exists an inverse function $f^{-1} : Y \to X$ such that $f^{-1} \circ f = \operatorname{Id}_X$.[†]

Proposition 2.2.1 (Set Valued Mapping)**.** *Suppose we have $f : X \to Y$ and families of subsets of X $\{A_i\}_{i \in I_A}$ and $\{B_j\}_{j \in I_B}$. Then we have*

(1) $f\left(\bigcup_{i\in I_A} A_i\right) = \bigcup_{i \in I_A} f(A_i)$

(2) $f\left(\bigcap_{i\in I_A} A_i\right) \subset \bigcap_{i \in I_A} f(A_i)$

(3) $f^{-1}\left(\bigcup_{j\in I_B} B_j\right) = \bigcup_{j \in I_B} f^{-1}(B_j)$

(4) $f^{-1}\left(\bigcap_{j\in I_B} B_j\right) = \bigcap_{j \in I_B} f^{-1}(B_j)$

2.2.3 Correspondence

For mappings, we assign only one element in the image set to each element in the domain. We now consider assigning multiple elements in the image set to each element in the domain, which is called correspondence.

Definition 2.2.7 (Correspondence)**.** A correspondence $\varphi : X \twoheadrightarrow Y$ is a binary relation

$$G = \{(x, \varphi(x)) : x \in X \wedge (\varphi(x) \subset Y)\}.$$

[†] f^{-1} is well-defined as we'll show later that an operation has at most one identity element and then we can construct the inverse function with the bijective property. An important proposition is $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

and for the subset A of X, we call $\varphi^u(A) := \{x \in X : \varphi(x) \subset A\}$ the upper inverse image of A; similarly, we call $\varphi^l(A) := \{x \in X : \varphi(x) \cap A \neq \emptyset\}$ the lower inverse image of A.

2.2.4 Operation

An operation is a mapping $\circledast : X \times X \to Y$ and we denote $\circledast(x, y)$ as $x \circledast y$ in convention. We say the subset A of X is closed under the operation $\circledast$ if $A \circledast A \subset A$.

Two important kinds of operation we often meet are associative and commutative operations.

> **Associative** An operation $\circledast$ is associative if $(x \circledast y) \circledast z = x \circledast (y \circledast z)$ for all $x, y, z \in X$.
>
> **Commutative** An operation $\circledast$ is commutative if $x \circledast y = y \circledast x$ for all $x, y \in X$.

Proposition 2.2.2 (Uniqueness of Identity Element). *We call an element e an identity element if and only if $x \circledast e = e \circledast x = x$ for all $x \in X$. For a certain operation, there exsists at most one identity element.*

Proof. We prove by contradiction. Suppose there're two identity elements e and e'. Then $e = e \circledast e' = e'$, contradicting with $e \neq e'$. □

2.3 Countability

For finite sets, we can easily count the element and say it's countable. But what if the set have infinite many elements? Mathematicians believe that natural numbers can be counted and is the smallest infinite set. In this section we formally treat infinity, before which we introduce the axiom of choice.

2.3.1 The Axiom of Choice

Axiom 2.3.1 (the Axiom of Choice). *For a family of non-empty sets $\{A_i\}_{i \in I}$, the Cartesian product $\prod_{i \in I} A_i$ is non-empty.*

That is equivalent to say there exists a choice function f such that $\forall i \in I, f(i) \in A_i$.

2.3.2 Cardinality

The cardinality of a finite set is defined as the number of elements, that is $|A| = \# A$. For infinite sets, the definition is complicated and we turn to compare the cardinality of 2 sets.

Definition 2.3.1 (Equality of Cardinality)**.** Two sets A and B are said to have the same cardinality if there exists a bijection from A to B, denoted as $|A| = |B|$. If there exists a subsection of B and a bijection from A to that subsection, we say $|A| \leq |B|$.

Proposition 2.3.1. *Let X, Y be sets, the following statements are equivalent:*

(1) $|X| \leq |Y|$
(2) there exists an injection from X to Y.
(3) there exists a surjection from Y to X.

Proof. (1) $\implies$ (2) is done by definition.

To show (2) $\implies$ (3), let g be a mapping from $f(X)$ to X such that $g \circ f|_{f(X)} = \mathrm{Id}_{f(X)}$, which is a bijection. For a fixed $x \in X$, define

$$h(y) = \begin{cases} g(y), & y \in f(X) \\ x, & y \notin f(X) \end{cases}$$

which is a surjection from Y to X.

(3) $\implies$ (1) needs the axiom of choice. Let $A_x := \{y \in Y : f(y) = x\}$ and A_x is non-empty as f is surjective. By **axiom 2.3.1, the Axiom of Choice**, there exists a choice function h such that $h(x) \in A_x$. Notice that

$$x \neq x' \implies (A_x \cap A_{x'}) = \emptyset \implies h(x) \neq h(x')$$

Then h is an injection from X to Y. ☐

Actually, $(|\cdot|, \leq)$ is a total ordered set. See the following theorems.

Theorem 2.3.2 (Anti-symmetry and Completeness of Cardinality)**.** *Let X, Y be sets.*

Bernstein-Schroeder $(|Y| \leq |X| \wedge |X| \leq |Y|) \implies |X| = |Y|$

Completeness $|X| \leq |Y| \vee |Y| \leq |X|$

Note 2. Note that the completeness is equivalent to the axiom of choice.

2.3.3 Countable Set

Definition 2.3.2 (Countability)**.** A set X is countable if $|X| = |\mathbb{N}|$.

Theorem 2.3.3 (Existence of Countable Subsets)**.** *There exists an infinite countable subset for any infinite set.*

Proof. Let X be an infinite set. Denote $\mathcal{A}_n = \{A \subset X : |A| = 2^n\}, \forall n \in \mathbb{N}$. By **axiom 2.3.1, the Axiom of Choice**, there exists a family of non-empty sets $\{B_n\}_{n=1}^{\infty}$ such that $B_n \in \mathcal{A}_n$.

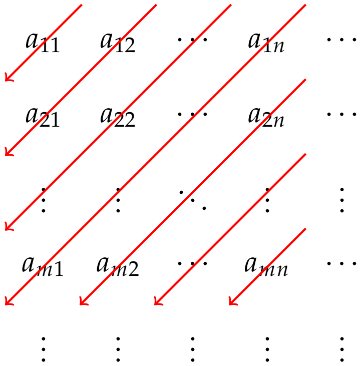
We now use **axiom 2.3.1, the Axiom of Choice**, again to construct a series $\{c_n\}_{n=1}^{\infty}$ such that $c_n \neq c_{n'}$. To do that, we construct a family of pairwise disjoint sets $\{C_n\}$ from B_n. An intuitive way to do so is to extract c_n from the new part of B_n, that is to construct $C_n = B_n \setminus \bigcup_{i=1}^{n-1} B_i$.

By construction, $f(n) : \mathbb{N} \mapsto c_n$ is an injection and by **proposition 2.3.1**, $|\mathbb{N}| \leq |\{c_n\}|$. Thus $\{c_n\}$ is a countable subset of X. □

Theorem 2.3.4. *A countable union of countable sets is countable.*

Proof. Let $\{A_n\}_{n=1}^{\infty}$ be a countable family of countable sets and denote $A_i = \{a_{ij}\}_{j=1}^{\infty}$. We try to find an injection from $\bigcup_{n=1}^{\infty} A_n$ to $\mathbb{N}$.

Suppose $\{A_n\}_{n=1}^{\infty}$ is pairwise disjoint, define $f : a_{mn} \mapsto m + \frac{(m+n-1)(m+n-2)}{2}$, the construction is illustrated below:


a_{11} a_{12} $\cdots$ a_{1n} $\cdots$
a_{21} a_{22} $\cdots$ a_{2n} $\cdots$
$\vdots$ $\vdots$ $\ddots$ $\vdots$ $\vdots$
a_{m1} a_{m2} $\cdots$ a_{mn} $\cdots$
$\vdots$ $\vdots$ $\vdots$ $\vdots$ $\vdots$

Otherwise, denote $C_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$. Then C_n is a countable set and is pairwise disjoint. It's done with the result of the previous case. □

Proposition 2.3.5. *A finite product of countable sets is countable.*

Proof. Notice that $\mathcal{A}_1 \times \mathcal{A}_2 = \bigcup_{a \in \mathcal{A}_2} \{(a, a') : a' \in \mathcal{A}_1\}$, which is a countable union of countable sets. Therefore, $\prod_{i=1}^{n} \mathcal{A}_i$ is countable by the deductive relation $\prod_{i=1}^{n} \mathcal{A}_i = \prod_{i=1}^{n-1} \mathcal{A}_i \times \mathcal{A}_n$. □

Corollary 2.3.6. *$\mathbb{N}^n$ and $\mathbb{Q}$ is countable.*

Proof. Only have to show $\mathbb{Q}$ is countable. We can construct the surjection $f(m, n) = m/n$. Notice it maps $\mathbb{N}^2$ to $\mathbb{Q}$, thus $|\mathbb{Q}| \leq |\mathbb{N}^2| = |\mathbb{N}|$. Therefore, $\mathbb{Q}$ is countable by **theorem 2.3.2, Bernstein-Schroeder theorem**. □

Chapter 3

Group, Ring and Field

With the concept of relation, we've added some structures to a set. For example, we've encountered $(\text{Funct}(X), \circ)$ and we know that it's associative, it has an identity element Id_X and for bijective ones, each element has an inverse.

In this chapter, we'll go deeper into understanding the abstract properties of a set and the added relation. We introduce some basic algebraic concepts and important propositions.

3.1 Group

3.1.1 the Definition

Definition 3.1.1 (Group)**.** If, for the 2-tuple $(G, \odot)$, the relation $\odot$ is associative, then we call a semi-group. Additionally, if there exists an identity element, we call it a monoid.

A group is a monoid such that for each element $g \in G$, there exists an inverse element $g^{-1} \in G$ and we call it an Abelian group if the relation $\odot$ is commutative.

Note 3. Note that the identity element is unique by **proposition 2.2.2, the uniqueness of identity element**. The following 2 propositions are also useful:

- $\forall g \in G, g \odot h = g \iff h \odot g = g$

- $(g \odot h)^{-1} = h^{-1} \odot g^{-1}$

Example 3.1.1. Drirect sum of $\mathcal{F}^n$

Definition 3.1.2 (Subgroup)**.** For a nonempty subset H of G, if it's closed under the operation $\odot$ [*] and $\forall h \in H, h^{-1} \in H$, we say $(H, \odot)$ is a subgroup of G[†]. We call $g \odot H$ and $H \odot g$ the left and right coset of $g \in G$ with respect to H and if $g \odot H = H \odot g$, we say H is a normal subgroup.

3.1.2 Homomorphisms

Now we turn to mappping one group to the other. A very special kind is the one that preserves the algebraic structure and we call it a homomorphism.

Definition 3.1.3 (Homomorphism)**.** A homomorphism φ is a mapping that maps $(G, \odot)$ to $(H, \circledast)$ such that $\forall g_1, g_2 \in G : \varphi(g_1 \odot g_2) = \varphi(g_1) \circledast \varphi(g_2)$. A homomorphism that maps G to itslef is called an endomorphism.

We list some important definitions and propositions.

Proposition 3.1.1. *A homomorphism maps the identity element to the identity element.*

Proof. Since $e_H \circledast \varphi(e_G) = \varphi(e_G \odot e_G) = \varphi(e_G) \circledast \varphi(e_G)$, we have $e_H = \varphi(e_G)$. □

Proposition 3.1.2. *A homomorphism maps the inverse element to the inverse of element.*

Proof. Notice that

$$\varphi(g) \circledast \varphi(g^{-1}) = \varphi(g \odot g^{-1}) = e_H \implies \varphi(g)^{-1} = \varphi(g^{-1}).$$

□

Definition 3.1.4 (Kernel)**.** The kernel of a homomorphism $\varphi : G \to H$ is defined by

$$\operatorname{Ker}(\varphi) := \varphi^{-1}(e_H) = \{g \in G : \varphi(g) = e_H\}$$

---

[*] Meaning

$$H \odot H \subset H :\iff \forall h_1, h_2 \in H : (h_1 \odot h_2) \in H.$$

[†] We can simply denote it as H if the operation is clear from the context.

and it's a normal subgroup of G[‡].

Proposition 3.1.3. *φ is injective $\iff$ $\mathrm{Ker}(\varphi)$ is trivial*[§].

Proof. ($\Rightarrow$) is trivial as injectivity makes $\varphi^{-1}(e_H)$ single-valued and by **proposition 3.1.1** we have $e_G \in \mathrm{Ker}(\varphi)$.

For ($\Leftarrow$), suppose there exist $g_1 \neq g_2$ such that $\varphi(g_1) = \varphi(g_2)$. Then $\varphi(e_G) = \varphi(g_2) \circledast \varphi(g_2^{-1}) = \varphi(g_1) \circledast \varphi(g_2^{-1}) = \varphi(g_1 \odot g_2^{-1}) = e_H$. This contradicts with $\mathrm{Ker}(\varphi) = \{e_G\}$ as $g_1 \odot g_2^{-1} \neq e_G$. ☐

3.1.3 Isomorphisms

Definition 3.1.5 (Isomorphism)**.** An isomorphism is a bijective homomorphism, denoted by $\cong$. If it maps a group to itself, we call it an automorphism.

Example 3.1.2 (The Set of All Automorphisms as a Group)**.** Let $\mathrm{Aut}(G)$ be the set of all automorphisms of G, then $(\mathrm{Aut}(G), \circ)$ is a group, which we call it an automorphism group.

Proposition 3.1.4 (Induced Operation)**.** *Let $(G, \odot)$ be a group and H is a nonempty set. Given a bijection $\varphi : G \longrightarrow H$, we define the operation $\circledast$*

$$\mu \circledast \nu := \varphi(\varphi^{-1}(\mu) \odot \varphi^{-1}(\nu)), \mu, \nu \in H$$

then $(H, \circledast)$ is a group and φ is an isomorphism.

Proof. Showing that $(H, \circledast)$ is a group is trivial. To show that φ is a homomorphism, given $h_1, h_2 \in H$, there exist $g_1, g_2 \in G$ such that $\varphi(g_1) = h_1$ and $\varphi(g_2) = h_2$. Then by definition, $h_1 \circledast h_2 = \varphi(g_1 \odot g_2)$. ☐

A very important result to know is that isomorphism hold exactly the same algebraic structure, which provides some convenience in calculation. For example, when it's hard to deal with $g_1 \odot g_2$, we may turn to $h_1 \circledast h_2$ and use φ^{-1} to get $g_1 \odot g_2$.

[‡] To show this, first verify that $\mathrm{Ker}(\varphi)$ is indeed a subgroup. And it's normal as we notice that $\forall k \in g \odot \mathrm{Ker}(\varphi), \exists \mu \in \mathrm{Ker}(\varphi)$ such that $k = g \odot \mu$. Let $\nu = g \odot \mu \odot g^{-1}$, we have $\nu \in \mathrm{Ker}(\varphi)$ and since $\nu \odot g = g \odot \mu = k$, we have $k \in \mathrm{Ker}(\varphi) \odot g$. The other side can be proved similarly.

[§] That is $\mathrm{Ker}(\varphi) = \{e_G\}$.

3.2 Ring and Field

In this section, we add another operations to a group. We require that the set forms an Abelian group with respect to one of the operations and we want a certain kind of distributivity. This leads us to ring. And we can further add more requirements to rings and make it a field.

3.2.1 The Definition

Definition 3.2.1 (Ring)**.** For a 3-tuple $(R, +, \cdot)$, if

(1) $(R, +)$ is an Abelian group[¶],
(2) $(R, \cdot)$ is a semi-group,
(3) Distributivity:

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c, \qquad \forall a, b, c \in R$$

With more specific requirements, we have different kinds of rings:

> **Commutative Ring** If $\cdot$ is commutative, then $(R, +, \cdot)$ is a commutative ring.
>
> **Ring with Unity** If there exists an identity element $1_R \in R$ for multiplication, then $(R, +, \cdot)$ is a ring with unity.

The definition of subring is similar to that of subgroup.

Definition 3.2.2 (Subring)**.** Let S be a nonempty subset of a ring $(R, +, \cdot)$. If S is a subgroup of $(R, +)$ and is closed under multiplication, we say it's a subring.

Remark 1. Given the existence of additive and multiplicative identity elements in a ring, we know that they're unique. Also, there may exists $a \neq 0_R$ such that $a \cdot b = 0_R$ or $b \cdot a = 0_R$ for some $b \neq 0_R$. Such elements are called **zero divisors**.

3.2.2 Homomorphism and Isomorphism

The definition here is similar to that of group homomorphism and isomorphism.

[¶] A conventional notation for the identity element is 0_R.

Definition 3.2.3 (Ring Homomorphism and Isomorphism)**.** Let R, S be rings. A mapping $\varphi : R \to S$ is a ring homomorphism if it preserve the multiplication and addition operations. If φ is bijective, then it is a ring isomorphism.

Note that if φ is a ring homomorphism, then it's a group homomorphism for $(R, +)$ and $(S, +)$ and the kernel is defined by $\mathrm{Ker}(\varphi) = \{r \in R : \varphi(r) = 0_S\}$.

3.2.3 Ordered Field

In this subsubsection, we add an order relation to a field and make it an ordered field.

Definition 3.2.4 (Ordered Field)**.** An ordered field is a field $(F, +, \cdot)$ equipped with a total order relation $\leq$, if the following statements hold for all $a, b, c \in F$:

(1) If $a \leq b$, then $a + c \leq b + c$.
(2) If $0 \leq a$ and $0 \leq b$, then $0 \leq a \cdot b$.

3.3 Polynomials

Part II

Vector Space and Linear Algebra

Chapter 4

Vector Space

From this chapter we start to formally introduce linear algebra. We'll see that linear algebra is similar to a ring but with some fine properties so that we can achieve more powerful results. To study linear algebra, we start with the vector space and gradually add more structures to it.

4.1 the Definition of Vector Space and Subspaces

Definition 4.1.1 (Vector Space)**.** A vector space over a field $\mathcal{F}$ is a triple $(V, +, \cdot)$ consisting of a nonempty set V whose elements are called **vectors**, and two operations called **vector addition** and **scalar multiplication** respectively, defined as

$$+ : (u, v) \in V^2 \mapsto u + v, \quad \cdot : (a, v) \in \mathcal{F} \times V \mapsto a \cdot v$$

which satisfy the following axioms:

(1) $(V, +)$ is an Abelian group.
(2) Identity element: $\exists 1 \in \mathcal{F}$ such that $\forall v \in V, 1 \cdot v = v$.
(3) Associativity: $\forall a, b \in \mathcal{F}, \forall v \in V, (ab) \cdot v = a \cdot (b \cdot v)$.
(4) Distributivity: $\forall a \in \mathcal{F}, \forall u, v \in V, a \cdot (u + v) = a \cdot u + a \cdot v$ and $\forall a, b \in \mathcal{F}, \forall v \in V, (a + b) \cdot v = a \cdot v + b \cdot v$.

We can define the concept of subspace similar to that of subgroup and subring.

Definition 4.1.2 (Subspace of a Vector Space)**.** Let U be a nonempty subset of a $\mathcal{F}$-vector space $(V,+,\cdot)$. If U is closed under vector addition and scalar multiplication, then we say U is a subspace of V.

4.2 Operation of Subspaces

4.2.1 Sum

Definition 4.2.1 (Sum)**.** Let $V_1,\cdots,V_m$ be vector spaces over the same field $\mathcal{F}$. We define their sum by

$$V_1+V_2+\cdots+V_m := \{v_1+v_2+\cdots+v_m : v_i \in V_i, i=1,2,\ldots,m\}.$$

Remark 2. Note that $V_1+V_2+\cdots+V_m$ is also a subspace of V and is the smallest subspace containing all $V_1,\cdots,V_m$.

4.2.2 Direct Sum

For an element v in the sum of subspaces, suppose that

$$v = v_1+v_2+\cdots+v_m = u_1+u_2+\cdots+u_m.$$

Then

$$(v_1-u_1)+\cdots+(v_m-u_m)=0.$$

The most special case is when $v_i = u_i$ ($\forall i \in \{1,2,\ldots,m\}$). This special case is defined as a direct sum.

Definition 4.2.2 (Direct Sum of Subspaces)**.** If every element in the sum of subspaces can be represented uniquely, then the sum is called the **direct sum** of the subspaces, denoted by $V_1 \oplus V_2 \oplus \cdots \oplus V_m$.

Theorem 4.2.1 (Equivalent definition of direct sum)**.** *$V_1+\cdots+V_m$ is a direct sum $\iff$ the equation $v_1+v_2+\cdots+v_m=0$ holds only when $v_i=0$ for all i.*

Proof. It follows directly from the above discussion. □

For the case of 2 subspaces, we have a more specific result.

Theorem 4.2.2 (Direct Sum of 2 Subspaces). *Let U and W be subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.*

Proof. ($\Rightarrow$) For any $u \in U \cap W$, there exists $(-u) \in W$ such that $0 = u + (-u)$. Since the representation is unique, we must have $u = 0$.

($\Leftarrow$) Suppose $u \in U, w \in W$ satisfy $u + w = 0$, i.e., $u = -w \in W$. Then $u \in U \cap W$, which implies $u = w = 0$. By the equivalent definition, $U + W$ is a direct sum. ☐

4.3 Subspace Spanned by Vectors

4.3.1 Span

We now turn to the elements of the vector space. For any $v \in V$, we can find a list of $a_1, \cdots, a_n \in \mathcal{F}$ and $v_1, \cdots, v_n \in V$ such that $v = a_1 v_1 + \cdots + a_n v_n$. In this way, given a list of vectors, we can construct a vector space by generating vectors with all $(a_1, \cdots, a_n) \in \mathcal{F}^n$. We call such a space a **span**, defined by

$$\operatorname{span}(v_1, \cdots, v_n) := \{a_1 v_1 + \cdots + a_n v_n : a_i \in \mathcal{F}, v_i \in V\}.$$

If $\operatorname{span}(v_1, \cdots, v_n) = V$, we say that $v_1, \cdots, v_n$ spans V.

Notice that $x^0, \cdots, x^n$ can span $\mathcal{P}_n(\mathcal{F})$ and $x^0, \cdots, x^n, 2x^n$ can do it as well. This inspires us to find out which is the smallest one that can span a vector space.

4.3.2 Linear Independence

In this section, we introduce the concept of linear independence and we'll find out that the smallest spanning list is linearly independent. Consider the elements in the spanned subspace. Let

$$v = a_1 v_1 + \cdots + a_m v_m = b_1 v_1 + \cdots + b_m v_m,$$

then

$$(a_1 - b_1) v_1 + \cdots + (a_m - b_m) v_m = 0.$$

The most special case is when $a_i = b_i$ (for all $i \in \{1, \ldots, m\}$). To characterize this situation, we introduce linear independence.

Definition 4.3.1 (Linear Independence)**.** If $a_1v_1+\cdots+a_mv_m = 0$ holds only when $a_i = 0, \forall i \in \{1, \cdots, m\}$, then the vectors $v_1, \ldots, v_m$ are said to be **linearly independent**.

This following propositions directly follow from the above discussion.

Proposition 4.3.1. *Removing vectors from a linearly independent list does not change the linear independence of the list.*

Proposition 4.3.2. *The vectors $v_1, \ldots, v_m$ are linearly independent if and only if every element in* $\operatorname{span}(v_1, \ldots, v_m)$ *can be represented uniquely.*

We know proceed to study the length of a spanning list and a linearly independent list and we first do some preperations. Intuitively, a spanning list should be longer than a linearly independent list. The following lemma formalizes this idea.

Lemma 4.3.3. *Suppose $v_1, \ldots, v_m$ are linearly dependent. Then there exists $k \in \{1, 2, \ldots, m\}$ such that $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$. Furthermore, removing v_k does not change the spanned subspace, i.e.,*

$$\operatorname{span}(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m) = \operatorname{span}(v_1, \ldots, v_m).$$

Proof. By definition, we have $a_1v_1 + \cdots + a_mv_m = 0$ where $a_1, \cdots, a_m$ are not all zeros. Suppose that $a_{k+1} = \cdots = a_m = 0$ and $a_1a_2\cdots a_k \neq 0$. We have $v_k = -\frac{a_1}{a_k}v_1 - \cdots - \frac{a_{k-1}}{a_k}v_k$, implying that $v_k \in \operatorname{span}(v_1, \cdots, v_{k-1})$.

For any vector in $\operatorname{span}(v_1, \ldots, v_m)$, it can be expressed as the linear combination of $v_1, \ldots, v_m$. By substituting the expression of v_k into this linear combination, we can express it solely in terms of $v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m$. Thus, the spanned subspace remains unchanged. □

Theorem 4.3.4. *Let $v_1, \ldots, v_m$ be a spanning list of a vector space V. Then any linearly independent list of vectors in V has length at most m.*

Proof. Let $u_1, \cdots, u_m$ be a linearly independent list. Consider adding u_1 to $v_1, \cdots, v_n$ and by **lemma** 4.3.3 we know $u_1, v_1, \cdots, v_n$ is linearly dependent and we can remove one of the vectors without changing the spanned space. Notice that we can choose one vector from $v_1, \cdots, v_n$ since u_1 is none-zero.

We now proceed to add u_2 to the previous list. Since $u_2 \notin \text{span}(u_1)$ by linear independence, we can choose a vector from the v's and remove it. We continue doing so until no vectors are left in the linearly independent list. We now get a list with n vectors which contains $u_1, \cdots, u_m$, implying $m \leq n$. $\square$

4.4 Finite Dimensional Vector Space

In the previous section, we've discovered that the smallest spanning list is linearly independent. This inspire us to present a vector space with the help of a linearly independent spanning list. In this section, we move on to discover the properties of the subspace spanned by a linearly independent list with finite length.

4.4.1 Basis

Definition 4.4.1 (Basis)**.** A **basis** of a vector space V is a linearly independent list of vectors that spans V.

The most important result is the following, which shows how to construct a basis.

Lemma 4.4.1 (Steinitz)**.** *Suppose $v_1, \ldots, v_m$ is a spanning list of a vector space V, then there exists a subset of $\{v_1, \ldots, v_m\}$ that forms a basis of V. Similarly, if $u_1, \ldots, u_n$ is a linearly independent list in V, then there exists a list $u_{n+1}, \ldots, u_m$ in V such that $u_1, \ldots, u_m$ forms a basis of V.*

Proof. For the first part, if $v_1, \ldots, v_m$ is linearly independent, then it is a basis. Otherwise, by **lemma 4.3.3**, we can remove one vector without changing the spanned space. Continuing this process, we will eventually obtain a linearly independent list that spans V, which is a basis.

For the second part, if $u_1, \ldots, u_n$ spans V, then it is a basis. Otherwise, there exists $v \in V$ such that $v \notin \text{span}(u_1, \ldots, u_n)$. We can add v to the list to form $u_1, \ldots, u_n, v$. Repeating this process, we will eventually obtain a spanning list that contains $u_1, \ldots, u_n$. By the first part, we can extract a basis from this spanning list. $\square$

4.4.2 Dimension

It's natural to define the length of the basis as the dimension of the vector space. Before we present the definition, we need to show that the length of the basis is unique.

Theorem 4.4.2. *Any two basis of a vector space V have the same length.*

Proof. Suppose B_1 and B_2 are two basis of V. We can view B_1 as a spanning list and B_2 as a linearly independent list. By **theorem 4.3.4**, we know that the length of B_2 is at most that of B_1. By symmetry, the length of B_1 is at most that of B_2. Thus, they have the same length. □

Definition 4.4.2 (Dimension)**.** The **dimension** of a vector space V, denoted by $\dim V$, is defined as the length of any basis of V.

Theorem 4.4.3 (Criterion for Basis of a Finite-dimensional Vector Space)**.** *A spanning or linearly independent list of a vector space V is a basis if and only if its length equals* $\dim V$.

Proof. It's done with **lemma 4.4.1, the Steinitz exchange lemma**. □

We close this subsection with the following results concerning operations of vector spaces and their dimension.

Theorem 4.4.4 (the Dimension Inequality)**.** *For vector space V and its subspace U, suppose $U \subset V$. Then $0 \leq \dim U \leq \dim V$ and it becomes an equation only when U is trivial or $U = V$.*

Theorem 4.4.5. *For vector space V and U,*

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

Theorem 4.4.6. *Let $V_1, \cdots, V_m$ be vector spaces, then $V_1 + \cdots + V_m$ is a direct sum* $\iff \dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + V_m$.

Chapter 5

Linear Mapping and Matrix

By definition, for a vector sapce V, we know $(V,+)$ is an Abelian group. It's natural to consider the group homomorphisms and isomorphisms on vector spaces. We'll see linear mapping as a group homomorphism and discover some more fine properties in this chapter.

5.1 Linear Mapping

5.1.1 the Definitions

Definition 5.1.1 (Linear Mapping)**.** For a function φ from the vector space V to the vector space W, if the following statements hold, we call it a linear mapping[*].

- $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2), \forall v_1, v_2 \in V$.
- $\varphi(\lambda v) = \lambda\varphi(v), \forall \lambda \in \mathcal{F}$.

We denote the set of all linear mappings from V to W as $\mathcal{L}(V, W)$.

It's clear that a linear mapping is indeed a homomorphism, so we can easily translate some of the definitions and propositions from group homomorphisms to linear mappings.

---

[*] One may ask if such a mapping exists. The answer is yes as there exists a unique linear mapping that takes the basis of V to the basis of W, see (Axler, 2024, p. 54) for proof of this proposition.

Definition 5.1.2 (Kernel)**.** The kernel of a linear mapping $\varphi \in \mathcal{L}(V, W)$ is defined as

$$\operatorname{Ker}\varphi = \{v \in V : \varphi(v) = 0\}$$

Proposition 5.1.1. *Suppose $\varphi \in \mathcal{L}(V, W)$, then the following holds:*

- $\varphi(0) = 0$
- *φ is injective $\iff$ $\operatorname{Ker}(\varphi)$ is trivial.*

We can also define the operations on $\mathcal{L}(V, W)$.

Definition 5.1.3 (Operations on $\mathcal{L}(V, W)$)**.** Let $\mathcal{L}(V, W)$ be the set of all linear mappings from V to W, we have 3 operations on it:

Addition $\forall \varphi, \varpi \in \mathcal{L}(V, W), (\varphi + \varpi)(v) = \varphi(v) + \varpi(v)$

Scalar Multiplication $\forall \lambda \in \mathcal{F}, (\lambda\varphi)(v) = \lambda(\varphi(v))$

Product (Combination) $\forall \varphi, \varpi \in \mathcal{L}(V, W), (\varphi\varpi)(v) = \varphi \circ \varpi(v)$

Then $\mathcal{L}(V, W)$ is a vector space.

5.1.2 Isomorphic Linear Mapping

Before we start to deal with isomorphisms, we first dig deeper into the linear mappings of finite-dimensional vector spaces.

Theorem 5.1.2 (Fundamental Theorem of Linear Mappings)**.** *Suppose $\varphi \in \mathcal{L}(V, W)$, then*

$$\dim V = \dim \operatorname{Im}\varphi + \dim \operatorname{Ker}\varphi$$

Proof. Let $u_1, \cdots, u_n$ be a basis of $\operatorname{Ker}\varphi$. By **lemma 4.4.1, the Steinitz exchange lemma**, we can extend them as $u_1, \cdots, u_n, v_1, \cdots, v_m$ so that it becomes a basis of V. We close the proof by showing $\varphi(v_1), \cdots, \varphi(v_n)$ is a basis of $\operatorname{Im}\varphi$.

We first show that $\varphi(v_1), \cdots, \varphi(v_n)$ spans $\operatorname{Im}\varphi$. Notice that $\forall v \in V$, we can rewrite it as $a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_nv_n = v$. Thus

$$\varphi(v) = b_1\varphi(v_1) + \cdots + b_n\varphi(v_n)$$

implying $\varphi(v) \in \operatorname{span}(\varphi(v_1), \cdots, \varphi(v_n))$.

Now we show the linear independence[†]. Suppose $c_1v_1 + \cdots + c_nv_n = 0$, then $\varphi(c_1v_1 + \cdots + c_nv_n) = 0$, meaning $c_1v_1 + \cdots + c_nv_n \in \operatorname{Ker}\varphi$. We can rewrite it with the linear combination of $u_1, \cdots, u_m$, say $c_1v_1 + \cdots + c_nv_n = d_1u_1 + \cdots + d_mu_m$. Since $u_1, \cdots, u_n, v_1, \cdots, v_m$ is linearly independent, all c's and d's are zero. □

Corollary 5.1.3. *Suppose V and W are vector spaces. If $\dim V < \dim W$, $\varphi \in \mathcal{L}(V, W)$ is NOT surjective. If $\dim V > \dim W$, $\varphi \in \mathcal{L}(V, W)$ is NOT injective*

With these results, we can depict an isomorphism with ease.

Definition 5.1.4 (Isomorphic Linear Mapping)**.** If a linear mapping is bijective, we say it's isomorphic.

We know by **corollary 5.1.3** that a necessary condition for a linear mapping to be isomorphic is that the finite-dimensional vector spaces have the same dimension.

Corollary 5.1.4. *Suppose $\varphi \in \mathcal{L}(V, W)$ with $\dim V = \dim W$, Then*

$$\varphi \textit{ is isomorphic} \iff \operatorname{Ker}\varphi \textit{ is trivial} \iff \operatorname{Im}\varphi = W \iff \varphi \textit{ is invertible}$$

Proof. Use **theorem 5.1.2, the fundamental theorem of linear mappings**, we have

$$\varphi \text{ is surjective} \iff \operatorname{Im}\varphi = W \iff \dim\operatorname{Ker}\varphi = 0 \iff \varphi \text{ is injective}$$

□

If there exists an isomorphism from V to W, we say V and W are isomorphic.

Theorem 5.1.5. *V and W are isomorphic $\iff \dim V = \dim W$.*

5.2 Matrix

We have already explored the abstract concept of linear mappings and we know that isomorphic linear mappings are of great importance as they maintain the structure of a certain vector space. We now turn to a more concrete representation of linear mappings, that is matrix, before which we consider a specific vector space, $\mathcal{F}^n$.

[†] From this part we can also show that an isomorphic linear mapping maps a linearly independent list to a linearly independent list.

Example 5.2.1 ($\mathcal{F}^n$ as a Vector Space)**.** For a field $\mathcal{F}$, you may verify that the direct product $\mathcal{F}^{n\ddagger}$ is a $\mathcal{F}$-vector space with the group addition and scalar multiplication defined by $\cdot : (\lambda, (f_1, \cdots, f_n)) \in \mathbb{F} \times \mathcal{F}^n \mapsto (\lambda f_1, \cdots, \lambda f_n)$. We have $\dim \mathcal{F}^n = n$ and its basis is given by

$$\underbrace{(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,0,\cdots,1)}_{n \text{ vectors}}.$$

Proof. We show the basis part. It's clear that these n vectors span $\mathcal{F}^n$. Now we show the linear independence. Suppose $a_1(1,0,\cdots,0) + a_2(0,1,\cdots,0) + \cdots + a_n(0,0,\cdots,1) = 0$, then we have $(a_1, a_2, \cdots, a_n) = 0$, implying all a_i's are zero. □

Note that we can also write $(f_1, \cdots, f_n)$ in a column, that is

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

or simply denote it by $(f_1, \cdots, f_n)^t$. Wether to write a vector in $\mathcal{F}^n$ in a row or in a column is basically for simplicity of the notation. We'll see the benefits when presenting the m-direct product of $\mathcal{F}^n$. Suppose $\mathbf{f}_1, \cdots \mathbf{f}_m \in \mathcal{F}^n$, elements of the direct product $\mathcal{F}^{n,m}$ can be written as $(\mathbf{f}_1, \cdots \mathbf{f}_m)$ and we drop the brackets of $\mathbf{f}$'s, that is

$$\begin{pmatrix} f_{11} & \cdots & f_{1m} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nm} \end{pmatrix}.$$

A similar result of the previous example is that $\mathcal{F}^{n,m}$ is a vector space with dimension nm.

---

‡ What we mean here is to see the field $\mathcal{F}$ as a group $(\mathcal{F}, +)$.

5.2.1 From Vector to Coordinate

From **theorem** 5.1.5, we know that V, whose dimension is n, and $\mathcal{F}^n$ is isomorphic. That is to say there exists a bijection $\mathcal{M} : V \longrightarrow \mathcal{F}^n$. For a basis of V, denoted by $v_1, \cdots, v_n$, we construct the isomorphism as follows:

$$v_i \mapsto (\underbrace{0, \cdots, 1, \cdots, 0}_{\text{the } i\text{th slot equals to 1}}).$$

Then, for $v \in V$, we can write it as $v = a_1 v_1 + \cdots + a_n v_n$ and

$$\mathcal{M}(v) = a_1 \mathcal{M}(v_1) + \cdots + a_n \mathcal{M}(v_n)$$

this leads us to the definition of coordinate.

Definition 5.2.1 (Coordinate)**.** Suppose for $v \in V$, we have $v = a_1 v_1 + \cdots + a_n v_n$, then its coordinate is defined as

$$\mathcal{M}(v) = (a_1, a_2, \cdots, a_n).$$

5.2.2 From Linear Mapping to Matrix

Notice that $\dim \mathcal{L}(V, W) = nm$ if $\dim V = m$ and $\dim W = n$. So it's natural to construct an isomorphism to give a specific representation of linear mappings. Suppose $\varphi \in \mathcal{L}(V, W)$ with $\dim V = n$ and $\dim W = m$. Let $v_1, \cdots, v_m$ be a basis of V and $w_1, \cdots, w_n$ be a basis of W. We try to calculate $\mathcal{M}(\varphi(v_i))$ so that we can know the coordinate of any vectors in $\varphi(V)$ by linear combination. Suppose $\varphi(v_i) = a_{1i} w_1 + a_{2i} w_2 + \cdots + a_{ni} w_n$, then

$$\mathcal{M}(\varphi(v_i)) = (a_{1i}, a_{2i}, \cdots, a_{ni})^t.$$

Thus, we can present φ by

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}.$$

Definition 5.2.2 (Matrix)**.** Suppose $\varphi \in \mathcal{L}(V, W)$ and the basis of V and W are

$(v_1, \cdots, v_m)$ and $(w_1, \cdots, w_n)$ respectively, then the matrix of φ with respect to the basis $(v_1, \cdots, v_m)$ and $(w_1, \cdots, w_n)$ is defined as

$$\mathcal{M}(\varphi, (v_1, \cdots, v_m), (w_1, \cdots, w_n))^{\S} = \begin{array}{cc} & \begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_m \end{array} \\ \begin{array}{c} w_1 \\ \vdots \\ w_n \end{array} & \begin{pmatrix} & & a_{1k} & & \\ & & \vdots & & \\ & & a_{nk} & & \end{pmatrix} \end{array}$$

where $\{a_{ij}\}_{(i,j)=(1,1)}^{(n,m)}$ satisfy $\varphi(v_k) = a_{1k}w_1 + a_{2k}w_2 + \cdots + a_{nk}w_n$ for $k = 1, \ldots, m$. We denote the set of matrix with n rows and m columns as $\mathcal{F}^{n,m}$.

5.2.3 Operations of Matrix

As mentioned before, a matrix is basically a linear mapping. In this way we define the operations of matrixes analogous to the operations of linear mappings.

Addition Suppose $\varphi, \varpi \in \mathcal{L}(V, W)$ and we have $\varphi(v_k) = a_{1k}w_1 + \cdots + a_{nk}w_n$ and $\varpi(v_k) = b_{1k}w_1 + \cdots + b_{nk}w_n$.

$$(\varphi + \varpi)(v_k) = \varphi(v_k) + \varpi(v_k) = (a_{1k} + b_{1k})w_1 + \cdots + (a_{nk} + b_{nk})w_n$$

Scalar Multiplication Suppose $\varphi \in \mathcal{L}(V, W)$ and $\varphi(v_k) = a_{1k}w_1 + \cdots + a_{nk}w_n$. For $\lambda \in \mathbb{F}$,

$$(\lambda\varphi)(v_k) = \lambda\left(\varphi(v_k)\right) = \lambda a_{1k}w_1 + \cdots + \lambda a_{nk}w_n$$

Product (Composition) Suppose $\varphi \in \mathcal{L}(V, W)$ and $\psi \in \mathcal{L}(W, U)$, with $\varphi(v_k) = a_{1k}w_1 + \cdots + a_{nk}w_n$ and $\psi(w_k) = b_{1k}u_1 + \cdots + b_{pk}u_p$. Then the

---

§ Simply use the notation $\mathcal{M}(\varphi)$ when the basis is clear from the context.

composition $\psi \circ \varphi$ satisfies:

$$\begin{aligned}(\psi \circ \varphi)(v_k) = \psi\left(\varphi(v_k)\right) &= a_{1k}\psi(w_1) + \cdots + a_{nk}\psi(w_n) \\ &= a_{1k}\sum_{i=1}^{p} b_{i1}u_i + \cdots + a_{nk}\sum_{i=1}^{p} b_{in}u_i \\ &= \left(\sum_{j=1}^{n} b_{1j}a_{jk}\right)u_1 + \cdots + \left(\sum_{j=1}^{n} b_{pj}a_{jk}\right)u_p\end{aligned}$$

This leads us to the following definitions.

Definition 5.2.3 (Operations of Matrix)**.** In the space of matrices, we can define the following operations:

- **Addition:** For matrices $A = \mathcal{M}(\varphi)$ and $B = \mathcal{M}(\varpi)$ in $\mathcal{F}^{n,m}$ (with respect to the same bases),

  $$\mathcal{M}(\varphi + \varpi) = \mathcal{M}(\varphi) + \mathcal{M}(\varpi),$$

  that is,

  $$\begin{array}{cc} & \begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_m \end{array} \\ \begin{array}{c} w_1 \\ \vdots \\ w_n \end{array} & \begin{pmatrix} & & a_{1k} & & \\ & & \vdots & & \\ & & a_{nk} & & \end{pmatrix} \end{array} + \begin{array}{cc} & \begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_m \end{array} \\ \begin{array}{c} w_1 \\ \vdots \\ w_n \end{array} & \begin{pmatrix} & & b_{1k} & & \\ & & \vdots & & \\ & & b_{nk} & & \end{pmatrix} \end{array}$$

  $$= \begin{array}{cc} & \begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_m \end{array} \\ \begin{array}{c} w_1 \\ \vdots \\ w_n \end{array} & \begin{pmatrix} & & a_{1k} + b_{1k} & & \\ & & \vdots & & \\ & & a_{nk} + b_{nk} & & \end{pmatrix} \end{array}$$

- **Scalar Multiplication:** For a matrix $A = \mathcal{M}(\varphi)$ in $\mathcal{F}^{n,m}$ and a scalar $\lambda \in \mathbb{F}$,

  $$\mathcal{M}(\lambda\varphi) = \lambda\mathcal{M}(\varphi),$$

that is,

$$\lambda\ \begin{array}{c|ccccc|} & v_1 & \cdots & v_k & \cdots & v_m \\ \hline w_1 & & & a_{1k} & & \\ \vdots & & & \vdots & & \\ w_n & & & a_{nk} & & \\ \hline \end{array} = \begin{array}{c|ccccc|} & v_1 & \cdots & v_k & \cdots & v_m \\ \hline w_1 & & & \lambda a_{1k} & & \\ \vdots & & & \vdots & & \\ w_n & & & \lambda a_{nk} & & \\ \hline \end{array}$$

- **Multiplication (Composition):** For matrices $A = \mathcal{M}(\varphi) \in \mathcal{F}^{n,m}$ (with respect to bases of V and W) and $B = \mathcal{M}(\psi) \in \mathcal{F}^{p,n}$ (with respect to bases of W and U),

$$\mathcal{M}(\psi \circ \varphi) = \mathcal{M}(\psi)\mathcal{M}(\varphi) = BA,$$

that is,

$$\begin{array}{c|ccccc|} & w_1 & \cdots & w_k & \cdots & w_n \\ \hline u_1 & & & b_{1k} & & \\ \vdots & & & \vdots & & \\ u_p & & & b_{pk} & & \\ \hline \end{array} \cdot \begin{array}{|ccccc|c} v_1 & \cdots & v_k & \cdots & v_m & \\ \hline & & a_{1k} & & & w_1 \\ & & \vdots & & & \vdots \\ & & a_{nk} & & & w_n \\ \hline \end{array}$$

$$= \begin{array}{c|ccccc|} & v_1 & \cdots & v_k & \cdots & v_m \\ \hline u_1 & & & \sum_{j=1}^{n} b_{1j}a_{jk} & & \\ \vdots & & & \vdots & & \\ u_p & & & \sum_{j=1}^{n} b_{pj}a_{jk} & & \\ \hline \end{array}$$

Note that the product matrix BA is in $\mathcal{F}^{p,m}$, corresponding to the linear mapping $\psi \circ \varphi : V \longrightarrow U$.

Remark 3. There're several ways to understand product oa matrix. Let Col_k and Row_k denote the kth column and row respectively. Consider the matrix multiplication $C = AB$:

- entry of matrix product equals row times column: $c_{ij} = \mathrm{Row}_i(A) \cdot \mathrm{Col}_j(B)$.
- column(row) of matrix product equals matrix times column(row): $\mathrm{Col}_k(C) = A\,\mathrm{Col}_k(B)$.
- matrix multiplication as linear combinations of columns(rows): Notice that, suppose b is a vector presented by $(b_1, \cdots, b_n)$, $Ab = \sum_{i=1}^{n} b_i \,\mathrm{Col}_i(A)$. Thus, if we

see columns of B as vectors,

$$\mathrm{Col}_k(C) = \sum_{i=1}^{n} b_{ik}\,\mathrm{Col}_i(A).$$

The third one leads to the CR factorization, which plays a crucial role in the definition of rank.

5.2.4 Change of Basis

The definition implies that matrix may change if we choose a diiferent basis. In this section, we're going to find a formula for changing the basis, which can help us find the simplist matrix to present a certain linear mapping.

Before the matrix version, we consider changing the basis of a vector. Let $e_1, \dots, e_n$ be a basis of V. For any $\alpha \in V$, we write $\alpha = \lambda_1 e_1 + \cdots + \lambda_n e_n$, which can be represented as an $n \times 1$ matrix (column vector):$\alpha = (\lambda_1, \cdots, \lambda_n)'$

Let $f_1, \dots, f_n$ be another basis of V. We investigate how to express α with respect to the new basis f, starting from its representation in the basis e. Write each f_i in terms of e:

$$f_i = \sum_{k=1}^{n} a_{ik} e_k, \quad i = 1, 2, \dots, n.$$

Now suppose α is expressed in the f-basis as $\alpha = \mu_1 f_1 + \cdots + \mu_n f_n$. Then

$$\begin{aligned}
\alpha &= \mu_1 \left(\sum_{k=1}^{n} a_{1k} e_k \right) + \cdots + \mu_n \left(\sum_{k=1}^{n} a_{nk} e_k \right) \\
&= \left(\sum_{k=1}^{n} \mu_k a_{k1} \right) e_1 + \cdots + \left(\sum_{k=1}^{n} \mu_k a_{kn} \right) e_n \\
&= \left(a_{11}, \dots, a_{n1} \right) \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} e_1 + \cdots + \left(a_{1n}, \dots, a_{nn} \right) \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} e_n \\
&= \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.
\end{aligned}$$

Thus, the coordinates of α in the e-basis are obtained by multiplying the transition matrix by its coordinate vector in the f-basis.

Chapter 6

Normed Vector Space

hi

Chapter 7

Multilinear Form and Determinant

One good way to introduce determinant is through solving equation systems and define it by deduction. Consider solving

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

then we have

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$$

and denote

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

We can rewrite the solution as

$$x_1 = \frac{\det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}, x_2 = \frac{\det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

Now we turn to the 3-dimensional case:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

Similarly, by elimination method, we have

$$x_1 = \frac{b_1 \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det\begin{pmatrix} b_2 & a_{23} \\ b_3 & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} b_2 & a_{22} \\ b_3 & a_{32} \end{pmatrix}}{a_{11} \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}$$

$$x_2 = \frac{a_{11} \det\begin{pmatrix} b_2 & a_{23} \\ b_3 & a_{33} \end{pmatrix} - b_1 \det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} a_{21} & b_2 \\ a_{31} & b_3 \end{pmatrix}}{a_{11} \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}$$

$$x_3 = \frac{a_{11} \det\begin{pmatrix} a_{22} & b_2 \\ a_{32} & b_3 \end{pmatrix} - a_{12} \det\begin{pmatrix} a_{21} & b_2 \\ a_{31} & b_3 \end{pmatrix} + b_1 \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}{a_{11} \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}$$

This leads us to denote

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

and rewrite the solution as

$$x_1 = \frac{\det\begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}, x_2 = \frac{\det\begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}, x_3 = \frac{\det\begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}$$

In this way can we define determinant in a deductive way and achieve the important Cramer's Law easily. However, this definition arises only as a notation, lacking the geometric understanding of determinants. To do this, we start from the concept of **multilinear form**.

7.1 Multilinear Form

7.1.1 Bilinear Form

A bilinear form on V is a function on $V \times V$ and is linear in each slot while holding the other one fixed.

Definition 7.1.1 (Bilinear Form)**.** Let V be a vector space over field $\mathbb{F}$. A function $\beta : V \times V \to \mathbb{F}$ is called a bilinear form on V if

$$v \mapsto \beta(u, v), \quad u \mapsto \beta(u, v)$$

are both linear mapping on V. We denote $V^{(2)}$ as the set of all bilinear forms in V.

An example of bilinear form is the inner product.

Example 7.1.1 (Inner Product)**.** Let β be the inner product $\langle \cdot, \cdot \rangle$. Recall the linearity of inner product in each slot and we can verify by definition that inner product is indeed a bilinear form.

7.1.1.1 Symmetric and Alternating Bilinear Form

However, though the inner product is symmetric, that is $\beta(u,v) = \beta(v,u)$, symmetry does not always hold for bilinear forms. There is a special kind of bilinear form called **symmetric bilinear form**.

Definition 7.1.2 (Symmetric Bilinear Form)**.** A bilinear form β from $V^{(2)}$ is called symmetric if

$$\beta(u,v) = \beta(v,u), \quad \forall u,v \in V.$$

We denote $V^{(2)}_{sym}$ as the set of all symmetric bilinear forms in $V^{(2)}$. It's trivial to show that $V^{(2)}_{sym}$ is a subspace of $V^{(2)}$.

Another special and important kind of bilinear form is **alternating bilinear form**. We'll see its importance when we reach the final theorem that $V^{(2)}$ can be separated as the direct sum of $V^{(2)}_{sym}$ and $V^{(2)}_{alt}$.

Definition 7.1.3 (Alternating Bilinear Form)**.** A bilinear form β from $V^{(2)}$ is called alternating if

$$\beta(v,v) = 0, \quad \forall v \in V.$$

We denote $V^{(2)}_{alt}$ as the set of all alternating bilinear forms in $V^{(2)}$. It's trivial to show that $V^{(2)}_{alt}$ is a subspace of $V^{(2)}$.

Proposition 7.1.1 (Characterization of Alternating Bilinear Form)**.** *A bilinear form α is alternating if and only if*

$$\alpha(u,v) = -\alpha(v,u), \quad \forall u,v \in V.$$

Proof. Notice that, by definition of bilinear form, we have

$$\alpha(u+v,u+v) = \alpha(u,u) + \alpha(u,v) + \alpha(v,u) + \alpha(v,v)$$

Suppose α is alternating, then $\alpha(u+v,u+v) = \alpha(u,v) + \alpha(v,u) = 0$, which implies $\alpha(u,v) = -\alpha(v,u)$.

Conversely, suppose $\alpha(u,v) = -\alpha(v,u)$, then $\alpha(v,v) = -\alpha(v,v)$, which implies $\alpha(v,v) = 0$ for all v. Therefore, α is alternating. ☐

Now we turn to the final theorem of alternating and symmetric bilinear forms.

Theorem 7.1.2 (Decomposition of Bilinear Form)**.** *Given a vector space V, we have*

$$V^{(2)} = V^{(2)}_{sym} \oplus V^{(2)}_{alt}.$$

Proof. We first show that $V^{(2)} = V^{(2)}_{sym} + V^{(2)}_{alt}$. Suppose $\beta \in V^{(2)}$, Let

$$\beta_{sym} = \frac{\beta(u,v)+\beta(v,u)}{2}, \quad \beta_{alt} = \frac{\beta(u,v)-\beta(v,u)}{2}$$

It's trivial that $\beta_{sym} \in V^{(2)}_{sym}$ and $\beta_{alt} \in V^{(2)}_{alt}$, and $\beta = \beta_{sym} + \beta_{alt}$.

Then we only have to show that $V^{(2)}_{sym} \cap V^{(2)}_{alt} = \{0\}$. Suppose $\beta \in V^{(2)}_{sym} \cap V^{(2)}_{alt}$. Then β is both symmetric and alternating, so $\beta(u,v) = \beta(v,u)$ and $\beta(u,v) = -\beta(v,u)$, which implies $\beta(u,v) = 0$ for all u, v. Therefore, $V^{(2)}_{sym} \cap V^{(2)}_{alt} = \{0\}$. □

7.1.1.2 Matrix of a Bilinear Form

7.1.2 Multilinear Form

We now turn to define determinants and this subsection is for preparation. The definition of multilinear form is a natural extension of bilinear form.

Definition 7.1.4 (Multilinear Form)**.** Let V be a vector space over field $\mathbb{F}$. A k-multilinear form

$$\beta : \underbrace{V \times V \times \cdots \times V}_{k \text{ times}} \to \mathbb{F}$$

satisfies that

$$v_i \mapsto \beta(v_1, \ldots, v_i, \ldots, v_k)$$

are all linear mapping on V for each $i = 1, 2, \ldots, k$. We denote $V^{(k)}$ as the set of all k-multilinear forms in V.

7.1.2.1 Alternating Multilinear Form

For multilinear forms, we only concern about alternating ones.

Definition 7.1.5 (Alternating Multilinear Form)**.** A multilinear form β from $V^{(k)}$ is called alternating if

$$\beta(v_1, \dots, v_k) = 0$$

whenever $v_i = v_j$ for some $i \neq j$. We denote $V_{alt}^{(k)}$ as the set of all alternating multilinear forms in $V^{(k)}$. It's trivial to show that $V_{alt}^{(k)}$ is a subspace of $V^{(k)}$.

By linearity at each slot, it's trivial to show that for a linearly dependent set $\{v_1, \dots, v_k\}$, we have $\alpha(v_1, \dots, v_k) = 0$ for any alternating multilinear form $\alpha \in V_{alt}^{(k)}$. The counterpositive satement implies that if $\alpha(v_1, \dots, v_k) \neq 0$ for some alternating multilinear form $\alpha \in V_{alt}^{(k)}$, then $\{v_1, \dots, v_k\}$ is linearly independent. Actucally, with the theorem $\dim V_{alt}^{(\dim V)} = 1$, which will be shown later, the converse statement also holds.

Theorem 7.1.3. *For a nonzero* $\alpha \in V_{alt}^{(k)}$, $\alpha(v_1, \dots, v_k) \neq 0 \iff \{v_1, \dots, v_k\}$ *is linearly independent.*

7.1.2.2 Swapping the Entries

Theorem 7.1.4 (Swapping 2 Slots of an Alternating Multilinear Form)**.** *For alternating multilinear form* α *and distinct* $i, j \in \{1, 2, \dots, k\}$, *we have*

$$\alpha(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\alpha(v_1, \dots, v_j, \dots, v_i, \dots, v_k)$$

Proof. We show the case of swapping v_1 and v_2. Notice that

$$\begin{aligned}
0 &= \alpha(v_1 + v_2, v_1 + v_2, \dots, v_k) \\
&= \alpha(v_1, v_1, \dots, v_k) + \alpha(v_1, v_2, \dots, v_k) + \alpha(v_2, v_1, \dots, v_k) + \alpha(v_2, v_2, \dots, v_k) \\
&= \alpha(v_1, v_2, \dots, v_k) + \alpha(v_2, v_1, \dots, v_k)
\end{aligned}$$

We have $\alpha(v_1, v_2, \dots, v_k) = -\alpha(v_2, v_1, \dots, v_k)$. □

To generalize the swapping theorem, we consider swapping multiple slots and multiple times. The intuitive idea asks us to count how many times we just swapped and we can use the above theorem. That is to consider, given the original list $(v_1, \dots, v_k)$ and the swapped list $(v_{j_1}, \dots, v_{j_k})$ with exactly the same elements but in different orders, how many swaps are made. We don't know how many swaps are made since there're multiple ways to do it and actually, how the swaps were made wouldn't change the result. So we can reconstruct the process of swapping by doing adjacent swaps, that is only swapping v_i and v_{i+1}. In this manner, we can count the swap by comparing the original order and the final order. We introduce the concept of permutation to understand the full picture following our intuition.

Definition 7.1.6 (Permutation and its Sign)**.** A permutation of the list $(1, \dots, k)$ is an element of perm $k := \{(j_1, \dots, j_k) \in \mathbb{N}^k : j_u \neq j_v \quad \text{and} \quad 1 \leq j_i \leq k\}$.

The sign of a permutation maps a certain element of perm k to either $+1$ or -1. For a permutation $(j_1, \dots, j_k) \in$ perm k, its sign $\text{sign}(j_1, \dots, j_k)$ is defined as

$$\text{sign}(j_1, \dots, j_k) = (-1)^N$$

where N is the cardinality of $\{(m, n) : 1 \leq m < n \leq k \text{ and } j_m > j_n\}$.

The next result is straightforward.

Proposition 7.1.5 (Generalization of the Swapping)**.** *Given a permutation* $(j_1, \dots, j_k) \in$ perm k *and an alternating multilinear form* $\alpha \in V_{alt}^{(k)}$*, we have*

$$\alpha(v_{j_1}, \dots, v_{j_k}) = \text{sign}(j_1, \dots, j_k)\alpha(v_1, \dots, v_k)$$

Chapter 8

Eigenvalue and Eigenvector

hi

Part III

Metric Space and Topology Space

Chapter 9

Metric Space and Convergence

9.1 Metric Space

9.1.1 the Definition

Definition 9.1.1 (Metric Space)**.**

Example 9.1.1 (the Discrete and Natural Metric)**.**

Proposition 9.1.1 (the Reversed Triangle Inequality)**.**

$$d(x, y) \geq |d(x, z) - d(z, y)|.$$

Definition 9.1.2 (Diameter)**.**

9.1.2 Neighborhood and Cluster Point

Definition 9.1.3 (Ball)**.**

Definition 9.1.4 (Neighborhood)**.**

Definition 9.1.5 (Cluster Point)**.**

Proposition 9.1.2. *(Amann and Escher, 2005, p. 134)*

9.1.3 Induced Metric

9.2 Convergence and Infinite Limit

9.2.1 Convergence of a Sequence

Definition 9.2.1 (Convergence of a Sequence)**.**

Proposition 9.2.1 (Every Convergent Sequence is Bounded)**.**

Proposition 9.2.2 (Uniqueness of Limit)**.**

9.2.2 Infinite Limit

9.2.3 Limit Superior and Limit Inferior

9.2.4 Some Special Sequences

9.2.4.1 Monotone Sequence

Theorem 9.2.3 (Monotone and Bounded)**.**

Theorem 9.2.4 (Stolze)**.**

9.2.4.2 Subsequence

9.2.5 Banach Space

Part IV

Differentiation in One Variable

Part V

More in Convergence: Uniform Convergence

Appendix A

Notation

Appendix B

Supplementary Scripts

Bibliography

Amann, H. and Escher, J. (2005). *Analysis I*. Birkhäuser Basel.

Amann, H. and Escher, J. (2006). *Analysis II*. Birkhäuser Basel.

Armstrong, M. A. (1983). *Basic Topology*. Springer New York.

Axler, S. (2024). *Linear Algebra Done Right*. Springer International Publishing.

Mei, J. (2020). *Mathematical Analysis*. Higher Education Press.

Qiao, L. (2024). Mathematical economics. Lecture Note of Mathematical Economics.

Stein, E. M. and Shakarchi, R. (2009). *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press.

Wang, Y. (2025). A concise course in mathematical analysis. Lecture Note of Mathematical Analysis I and II.

Xie, Q. (2022). *Advanced Algebra*. Fudan University Press.

Zhang, H. (2025). Probability and statistics. Class Handout of Probability Theory and Mathematical Statistics.