Honors in Linear and Abstract Algebra I

Lecture Notes for
MATH 2131

Department of Mathematics
Hong Kong University of Science and Technology

September 21, 2025

PUBLISHED BY HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

First printing, September 2025

Contents

Prefaces

This lecture notes was written by a student in the course MATH 2131 – Honors in Linear and Abstract Algebra by Professor Meng Guowu in HKUST in 2025-26 Fall.
All diagrams in this lecture notes are written in LaTeX TikZ code.
The notes is with reference to textbooks *'Linear Algebra'* by Friedberg, Insel and Spence, *'Abstract Algebra'* by Artin and *'A First Course in Abstract Algebra'* by Fraleigh.
Also, this notes is with reference to the other the notes of two other professors teaching the same course before, Professor Ivan Ip and Professor Min Yan.

List of Symbols

Symbols	Meaning
$\mathbb{F}$	a field
U, V, W	vector spaces
α, β	elements in $\mathbb{F}$
$\mathbb{F}^n$	the set of all column matrices with n entries in $\mathbb{F}$
$(\mathbb{F}^n)^*$	the set of all row matrices with n entries in $\mathbb{F}$
$\mathbb{F}[X]$	the polynomial ring
$\mathbb{F}[[X]]$	the formal power series ring
0_V	additive identity of vector space V
1_V	multiplicative identity of vector space V
e_V	an identity of vector space V
$\subset$	proper subset
$\subseteq$	subset, i.e. can be equal
ι	Inclusion map
$\hookrightarrow$	Injective arrow
π	Projection map
$\twoheadrightarrow$	Surjective arrow
S, T	Linear maps
A, B	Matrices
$\mathrm{Hom}(V, W)$	Hom-set of V to W
$\mathrm{End}(A)$	Endomorphism ring of A
$\mathrm{M}_{m\times n}(\mathbb{F})$	the set of all $m \times n$ matrices over $\mathbb{F}$
$\vec{x}$	column vector with entries x_i
$\hat{x}$	row vector with entries x_i
$\vec{e}_i$	column vector with only 1 at the i-th row and 0 at other places
$\hat{e}_i$	row vector with only 1 at the i-th column and 0 at other places
$\alpha\cdot$	a map that performs scalar multiplication
$A\cdot$	a map that performs matrix multiplication

1. Abstract Linear Spaces

"I assume you have learnt linear algebra."

Guowu Meng

1.1 Binary Operation

We start with the definition of a binary operation.

Definition 1.1 — Binary Operation. A *binary operation* on a set S is a mapping of the elements of the Cartesian product $S \times S$ to S.

$$\begin{aligned} f : S \times S &\to S \\ (x,y) &\mapsto f(x,y) \end{aligned}$$

For easier understanding, binary operation is combining two objects into one. Hence, there is something called unary and ternary operations, corresponding to the action of combining one and three objects into one respectively.

■ **Example 1.1** A common example of a binary operation is addition on the set of natural numbers $\mathbb{N}$.

$$\begin{aligned} + : \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (x,y) &\mapsto x+y \end{aligned} \tag{1.1}$$

■

Definition 1.2 — Associative. A binary operation $f : S \times S \to S$ is said to be *associative* if, for all $x,y,z \in S$,

$$f(x, f(y,z)) = f(f(x,y),z)$$

■ **Example 1.2** A common example of an associative (binary) operation is addition on the set of natural numbers $\mathbb{N}$. For all $x,y,z \in \mathbb{N}$, we have $x+(y+z) = (x+y)+z$. ■

Definition 1.3 — Identifiable. A binary operation $f : S \times S \to S$ is said to be *identifiable*, or *unital*, if there exists an element $e \in S$, the *identity* or *unit element*, such that, for all $x \in S$

$$f(e,x) = x = f(x,e)$$

■ **Example 1.3** A common example of an identifiable (binary) operation is multiplication on the set of natural numbers $\mathbb{N}$. The identity element is 1, and for all $x \in \mathbb{N}$, we have $x \cdot 1 = x = 1 \cdot x$. ■

Proposition 1.1 The identity element of an identifiable operation is unique.

Proof. Let e_1 and e_2 be two identity elements for the operation f. Then, for any element $x \in S$, we have:

$$f(x,e_1) = x = f(e_1,x)$$

$$f(x,e_2) = x = f(e_2,x)$$

Now, consider the element e_1: $f(e_1,e_2) = e_1$. But since e_2 is an identity element, we also have: $f(e_1,e_2) = e_2$. Therefore, we conclude that $e_1 = e_2$, proving the uniqueness of the identity element. ■

Note that the two-sided identity must be unique, but one-sided identities need not be. The following is an example of it.

■ **Example 1.4** Consider a set $X = \left\{ \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix} \middle| \, a \in \mathbb{R} \right\}$ with the binary operation defined as matrix multiplication. This set has many left identity elements, but no two-sided identity element. ■

Definition 1.4 — Invertible. A binary operation $f : S \times S \to S$ is said to be *invertible* if, for every element $x \in S$, there exists an element $y \in S$, called the two-sided *inverse* of x, denoted as x^{-1}, such that

$$f(x,y) = e = f(y,x)$$

where e is the identity element of the operation.

R Invertible operation exists if inverse operation exists, as there exists an identity element.

■ **Example 1.5** A common example of an invertible (binary) operation is addition on the set of integers $\mathbb{Z}$. For every integer $x \in \mathbb{Z}$, there exists an integer $y = -x$ such that:

$$x + (-x) = 0 = (-x) + x \tag{1.2}$$

where 0 is the identity element for addition. ■

Proposition 1.2 The inverse element of an invertible operation is unique.

Proof. Let y_1 and y_2 be two inverses of an element $x \in S$. Then, by definition of inverse, we have:

$$f(x,y_1) = e = f(y_1,x)$$

$$f(x,y_2) = e = f(y_2,x)$$

Now, consider the element y_1: $f(y_1,x) = e$. But since y_2 is also an inverse of x, we can substitute e with $f(x,y_2)$: $f(y_1,x) = f(x,y_2) = e$. By the associativity of the operation, we can rearrange this to:

$$y_1 = f(y_1,e) = f(y_1,f(x,y_2)) = f(f(y_1,x),y_2) = f(e,y_2) = y_2$$

Thus, the inverse element is unique. ■

Same for the inverse, one-sided need not be unique. The example is left as an exercise.

Definition 1.5 — Commutative. A binary operation $f : S \times S \to S$ is said to be *commutative* if, for all $x, y \in S$, the following holds:

$$f(x,y) = f(y,x)$$

■ **Example 1.6** A common example of a commutative operation is addition on the set of integers $\mathbb{Z}$. For all $x, y \in \mathbb{Z}$, we have: $x + y = y + x$ ■

Definition 1.6 — Distributive (Harmonic). A binary operation $g : S \times S \to S$ is said to be *distributive* with respect to another binary operation $f : S \times S \to S$ if, for all $x, y, z \in S$, the following holds:

$$g(x, f(y,z)) = f(g(x,y), g(x,z))$$
$$g(f(y,z), x) = f(g(y,x), g(z,x))$$

The professor preferred to use the word "harmonic" instead of "distributive". Note that it is important to show that "*which binary operation* is distributive to *which binary operation*". (The two binary operation in this sentence is not commutative.)

■ **Example 1.7** A common example of a distributive operation is multiplication over addition on the set of integers $\mathbb{Z}$. For all $x, y, z \in \mathbb{Z}$, we have:

$$x \cdot (y+z) = x \cdot y + x \cdot z$$
$$(y+z) \cdot x = y \cdot x + z \cdot x$$

■

1.2 Groups, Rings, Fields

With those five properties, we can construct monoid and groups.

Definition 1.7 — Monoid. A *monoid* is a set M equipped with a binary operation $f : M \times M \to M$ such that the following properties hold:

1. *Closure Property:* For all $x, y \in M$, $f(x, y) \in M$.
2. *Associative Property*
3. *Identifiable Property*

We say (M, f) is a monoid, and f is the *monoid operation* on the set M. A set M with a monoid operation f is the *monoid structure*.

Definition 1.8 — Group. A *group* is a set G equipped with a monoid operation $f : G \times G \to G$ with the additional property that every element has an inverse.

■ **Example 1.8** $(\mathbb{R} \setminus \{0\}, \times)$ is a group, but $(\mathbb{R}, \times)$ is not a group since 0 does not have a multiplicative inverse. ■

Definition 1.9 — Abelian Monoid / Group. A monoid / group (G, f) is said to be an *abelian* if the operation f is commutative.

Definition 1.10 — Unital Ring. A *unital ring* is a set R equipped with two binary operations $f : R \times R \to R$ (addition) and $g : R \times R \to R$ (multiplication) such that the following properties hold:

1. *Additive Group:* (R, f) is an abelian group.
2. *Multiplicative Monoid:* (R, g) is a monoid.
3. *Distributive Property:* g with respect to f.

Definition 1.11 — Commutative Ring. A *commutative ring* is a unital ring R such that the multiplication operation $g : R \times R \to R$ is commutative.

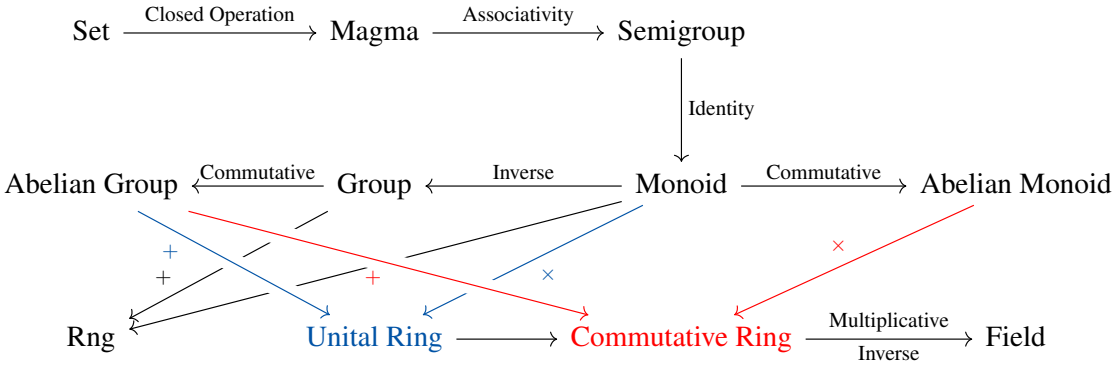
■ **Example 1.9** $(\mathbb{Z}, +, \times)$ is a commutative ring. ■

Definition 1.12 — Field. A *field* is a commutative ring $\mathbb{F}$ such that every non-zero element has a multiplicative inverse.

■ **Example 1.10** $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ and $(\mathbb{C}, +, \times)$ are fields. ■

■ **Example 1.11 — Finite Field.** $(\mathbb{Z}/2\mathbb{Z}, +, \times)$ is a field, where $\mathbb{Z}/2\mathbb{Z} = \{[0], [1]\}$, $[0]$ is the set of even integers and $[1]$ is the set of odd integers. Note that any $\mathbb{Z}/p\mathbb{Z}$ is a finite field, where p is a prime number. ■

We may draw a diagram for the relationship between the algebraic structures.



1.3 Morphisms

Normally, when we have two sets we can have a set map. Then if the two are algebraic structures? Moreover, if they are in the same structure?

Definition 1.13 — Morphisms. A *morphism* is a structure-preserving map between two algebraic structures (e.g., groups, rings, fields). Formally, let $(A, \cdot_A)$ and $(B, \cdot_B)$ be two algebraic structures. A morphism $f : A \to B$ is a set map / function such that:

$$f(x \cdot_A y) = f(x) \cdot_B f(y) \quad \forall x, y \in A$$

Homomorphisms are morphisms between two same algebraic structures.

Definition 1.14 — Monoid Homomorphism. A *monoid homomorphism* is a morphism between two monoids that preserves the monoid structure. Formally, let $(M_1, \cdot_1)$ and $(M_2, \cdot_2)$ be two monoids with identity elements e_1 and e_2, respectively. A function $f : M_1 \to M_2$ is a monoid homomorphism if:

1. $f(x \cdot_1 y) = f(x) \cdot_2 f(y) \quad \forall x, y \in M_1$
2. $f(e_1) = e_2$

Definition 1.15 — Group Homomorphism. A *group homomorphism* is a morphism between two groups that preserves the group structure. Formally, let $(G_1, \cdot_1)$ and $(G_2, \cdot_2)$ be two groups with identity elements e_1 and e_2, respectively. A function $f : G_1 \to G_2$ is a group homomorphism if:

1. $f(x \cdot_1 y) = f(x) \cdot_2 f(y) \quad \forall x, y \in G_1$
2. $f(e_1) = e_2$
3. $f(x^{-1}) = (f(x))^{-1} \quad \forall x \in G_1$

Proposition 1.3 The second and third properties of a group homomorphism are consequences of the first property.

Proof. Let $f : G_1 \to G_2$ be a group homomorphism satisfying the first property.

Second Property: For any element $x \in G_1$, we have:

$$f(x) = f(x \cdot_1 e_1) = f(x) \cdot_2 f(e_1)$$

So for any $f(x) \in G_2$, this implies that $f(e_1)$ must be the identity element in G_2, i.e., $f(e_1) = e_2$.

Third Property: We have:

$$e_2 = f(e_1) = f(x \cdot_1 x^{-1}) = f(x) \cdot_2 f(x^{-1})$$

This shows that $f(x^{-1})$ is the inverse of $f(x)$ in G_2, i.e., $f(x^{-1}) = (f(x))^{-1}$. ■

For monoid homomorphisms, the second property cannot be derived from the first property. Consider the identity element e_1 in M_1. If we apply the first property, we get $f(e_1 \cdot_1 e_1) = f(e_1) \cdot_2 f(e_1)$. This simplifies to $f(e_1) = f(e_1) \cdot_2 f(e_1)$, which does not necessarily imply that $f(e_1)$ is the identity element in M_2, i.e., $f(e_1) \neq e_2$, but $f(e_1)$ is the idempotent element in M_2. Therefore, the second property must be explicitly stated for monoid homomorphisms.

However in the case of group homomorphisms, the existence of inverses ensures that there is only one element that can be idempotent under the group operation, which is the identity element. Thus, for group homomorphisms, the second property can be derived from the first property.

Definition 1.16 — Idempotent Elements. An element a is said to be *idempotent* if $a = a^2$.

To introduce the vector space, the following two morphisms are required.

Definition 1.17 — Ring Homomorphism. A *ring homomorphism* is a morphism between two rings that preserves both the additive and multiplicative structures. Formally, let $(R_1, +_1, \cdot_1)$ and $(R_2, +_2, \cdot_2)$ be two rings with identity elements 0_1, 1_1 and 0_2, 1_2, respectively. A function $f : R_1 \to R_2$ is a ring homomorphism if:

1. $f(x +_1 y) = f(x) +_2 f(y) \quad \forall x, y \in R_1$
2. $f(x \cdot_1 y) = f(x) \cdot_2 f(y) \quad \forall x, y \in R_1$
3. $f(1_1) = 1_2$

Definition 1.18 — Endomorphism. An *endomorphism* is a morphism from an algebraic structure to itself. Formally, let $(A, \cdot)$ be an algebraic structure. An endomorphism $f : A \to A$ is a set map such that:

$$f(x \cdot y) = f(x) \cdot f(y) \quad \forall x, y \in A$$

The following two sets are the sets of all structure-preserving maps.

Definition 1.19 — Hom-set. The set of all morphisms from an algebraic structure A to another algebraic structure B is called the *hom-set*, denoted by $\mathrm{Hom}(A, B)$.

Definition 1.20 — Endomorphism Ring. The set of all endomorphisms of an abelian group $(A, +)$, denoted by $\mathrm{End}(A)$, forms a (non-commutative) ring under pointwise addition and composition of set maps. The addition and multiplication operations are defined as follows:

$$\begin{aligned} + : \mathrm{End}(A) \times \mathrm{End}(A) &\to \mathrm{End}(A) \\ (f, g) &\mapsto (f + g : x \mapsto f(x) + g(x)) \qquad f + g : A \to A \end{aligned}$$

$$\begin{aligned} \circ : \mathrm{End}(A) \times \mathrm{End}(A) &\to \mathrm{End}(A) \\ (f, g) &\mapsto (f \circ g : x \mapsto f(g(x))) \qquad f \circ g : A \to A \end{aligned}$$

The identity element for addition is the zero endomorphism, which maps every element to the identity element of the group.

$$\begin{aligned} 0 : A &\to A \\ x &\mapsto 0 \end{aligned}$$

The identity element for multiplication is the identity endomorphism, which maps every element to itself.

$$\begin{aligned} 1 : A &\to A \\ x &\mapsto x \end{aligned}$$

Note that all endomorphisms in $\mathrm{End}(A)$ are group homomorphisms and $\mathrm{End}(A) = \mathrm{Hom}(A, A)$.

1.4 Vector Spaces

Then we can define what a linear structure is.

Definition 1.21 — Linear Structure. A *linear structure* over a field $\mathbb{F}$ on a set V is a pair $(+,\cdot)$ where $(V,+)$ is an abelian group with a ring homomorphism $\mathbb{F} \to \mathrm{End}(V)$, where $\mathrm{End}(V)$ is the endomorphism ring of the abelian group $(V,+)$.

$$\begin{aligned} \cdot &: \mathbb{F} \to \mathrm{End}(V) \\ \alpha &\mapsto (\alpha\cdot : \vec{x} \mapsto \alpha\vec{x}) \qquad \alpha\cdot : V \to V \end{aligned}$$

The ring homomorphism is a (ring) action of the field $\mathbb{F}$ on the abelian group $(V,+)$, called *scalar multiplication*. The ring action can be written as a binary operation:

$$\begin{aligned} \cdot &: \mathbb{F} \times V \to V \\ (\alpha,\vec{x}) &\mapsto \alpha\vec{x} \end{aligned}$$

A linear space / vector space is a set with a linear structure over a field on the set. In normal textbook, a linear space will be defined as follows:

Corollary 1.1 — Linear Spaces. A linear space over a field $\mathbb{F}$ is a set V equipped with two operations: vector addition $+ : V \times V \to V$ and scalar multiplication $\cdot : \mathbb{F} \times V \to V$, satisfying the following axioms for all $\vec{u},\vec{v},\vec{w} \in V$ and $\alpha,\beta \in \mathbb{F}$:

Axiom	Statement
1. Associativity of addition	$(\vec{u}+\vec{v})+\vec{w} = \vec{u}+(\vec{v}+\vec{w})$
2. Existence of additive identity	$\exists\vec{0} \in V$ such that $\forall\vec{u} \in V$, $\vec{u}+\vec{0} = \vec{u}$
3. Existence of additive inverses	$\forall\vec{u} \in V$, $\exists -\vec{u} \in V$ such that $\vec{u}+(-\vec{u}) = \vec{0}$
4. Commutativity of addition	$\vec{u}+\vec{v} = \vec{v}+\vec{u}$
5. Distributivity of scalar multiplication with respect to vector addition	$\alpha(\vec{u}+\vec{v}) = \alpha\vec{u}+\alpha\vec{v}$
6. Distributivity of scalar multiplication with respect to field addition	$(\alpha+\beta)\cdot = \alpha\cdot+\beta\cdot$
7. Compatibility of scalar multiplication with field multiplication	$(\alpha\beta)\cdot = (\alpha\cdot)\circ(\beta\cdot)$
8. Identity element of scalar multiplication	$\mathbb{F} \ni 1 \mapsto (1\cdot : x \mapsto x) \in \mathrm{End}(V)$

R The first four axioms ensure that $(V,+)$ is an abelian group, while the fifth axiom describes the distributivity inside $\mathrm{End}(A)$ and the last three axioms describe the ring homomorphism.

■ **Example 1.12** $\mathbb{F}$ is a linear space over itself with the usual addition and multiplication operations.

$$\begin{aligned} \cdot &: \mathbb{F} \times \mathbb{F} \to \mathbb{F} \\ (\alpha,\beta) &\mapsto \alpha\beta \end{aligned}$$

The first $\mathbb{F}$ is the field acting on the second $\mathbb{F}$, which is the abelian group. ■

■ **Example 1.13** Let X be a set and $\mathbb{F}$ be a field. (f is a set map)

$$\begin{aligned} \mathbb{F}[[X]] = \mathrm{Map}(X,\mathbb{F}) &\overset{\text{def}}{=\!=} \text{the set of all } \mathbb{F}\text{-valued functions on } X \\ &=\!= \{f : X \to \mathbb{F}\} \end{aligned}$$

$\mathbb{F}[[X]]$ is a linear space over $\mathbb{F}$ with the following operations defined pointwisely:

$$\begin{aligned} +&: \mathbb{F}[[X]] \times \mathbb{F}[[X]] \to \mathbb{F}[[X]] \\ &\qquad (f,g) \mapsto (f+g : x \mapsto f(x)+g(x)) \qquad f+g : X \to \mathbb{F} \end{aligned}$$

$$\begin{aligned} \cdot&: \mathbb{F} \times \mathbb{F}[[X]] \to \mathbb{F}[[X]] \\ &\qquad (\alpha,f) \mapsto (\alpha f : x \mapsto \alpha f(x)) \qquad \alpha f : X \to \mathbb{F} \end{aligned}$$

■

■ **Example 1.14** Let X be a set and $\mathbb{F}$ be a field.

$$\begin{aligned} \mathbb{F}[X] = \mathrm{Map}_{\mathrm{fin}}(X,\mathbb{F}) &\stackrel{\mathrm{def}}{=\!=} \text{the set of all finitely supported } \mathbb{F}\text{-valued functions on } X \\ &=\!= \{f : X \to \mathbb{F} \mid f \text{ is finitely supported}\} \end{aligned}$$

$\mathbb{F}[X]$ is a linear space over $\mathbb{F}$ as $\mathbb{F}[X] \subseteq \mathbb{F}[[X]]$ and the operations are defined pointwisely as in the previous example.

$f : X \to \mathbb{F}$ is finitely supported if the set $\{x \in X \mid f(x) \neq 0\}$ is finite or $f(x) \neq 0$ for only finitely many $x \in X$. ■

■ **Example 1.15** Let t be a formal variable. Then $\mathbb{F}[[t]] \stackrel{\mathrm{def}}{=\!=} \mathbb{F}[[\{1,t,t^2,\cdots\}]] = \sum_{n=0}^{\infty} a_n t^n$ is the set of all formal power series in t with coefficients in $\mathbb{F}$ and $\mathbb{F}[t] \stackrel{\mathrm{def}}{=\!=} \mathbb{F}[\{1,t,t^2,\cdots\}] = \sum_{n=0}^{N} a_n t^n$ is the set of all polynomials in t with coefficients in $\mathbb{F}$. Both $\mathbb{F}[[t]]$ and $\mathbb{F}[t]$ are linear spaces over $\mathbb{F}$. ■

There are another names for $\mathbb{F}[X]$ and $\mathbb{F}[[X]]$: Polynomial ring and Formal Power Series ring, respectively.

■ **Example 1.16** Let n be a positive integer and $\mathbb{F}$ be a field. Then

$$\mathbb{F}^n \stackrel{\mathrm{def}}{=\!=} \left\{ \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \;\middle|\; c_i \in \mathbb{F} \right\}$$

is the set of all *column matrices* with n entries in $\mathbb{F}$. Elements in $\mathbb{F}^n$ are written as $\vec{x}$ and are called *column vectors*. $\mathbb{F}^n$ is a linear space over $\mathbb{F}$ with the following operations defined entrywisely:

$$\begin{aligned} +&: \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n \\ &\qquad (\vec{a},\vec{b}) \mapsto \vec{a}+\vec{b} = \begin{bmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \cdot&: \mathbb{F} \times \mathbb{F}^n \to \mathbb{F}^n \\ &\qquad (\alpha,\vec{a}) \mapsto \alpha\vec{a} = \begin{bmatrix} \alpha a_1 \\ \vdots \\ \alpha a_n \end{bmatrix} \end{aligned}$$

$\mathbb{F}^n$ is a linear space over $\mathbb{F}$ automatically as $\mathbb{F}$ is a linear space over itself. ■

2. Linear Maps and Matrices

"Linear algebra is the easiest in Mathematics"

Guowu Meng

2.1 Linear Maps

Linear map, sometimes linear transformation, is a homomorphism preserving linear structure.

Definition 2.1 — Linear Maps. Let V and W be two linear spaces over a field $\mathbb{F}$. A *linear map* is a set map $T : V \to W$ such that for all $\vec{u}, \vec{v} \in V$ and $\alpha \in \mathbb{F}$, the following holds:

$$T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$$
$$T(\alpha\vec{u}) = \alpha T(\vec{u})$$

The set of all linear maps from V to W is denoted by $\mathrm{Hom}(V,W)$. Some may write $\mathscr{L}(V,W)$.

Definition 2.2 — Linear Combinations. Let V be a linear space over a field $\mathbb{F}$. A *linear combination* of vectors $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n \in V$ is a vector of the form:

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{F}$ are scalars.

The following is the definition of linear map using linear combination. Note that this definition is equivalence to the definition above. The following is also equivalence to linear combination with n vectors.

Corollary 2.1 — Linear Maps and Linear Combinations. A set map $f : V \to W$ between two linear spaces over a field $\mathbb{F}$ is a linear map if and only if T respects linear combinations, i.e., for all $\vec{v}_1, \vec{v}_2 \in V$ and all scalars $\alpha_1, \alpha_2 \in \mathbb{F}$, the following holds:

$$T(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2) = \alpha_1 T(\vec{v}_1) + \alpha_2 T(\vec{v}_2)$$

■ **Example 2.1** Let A be an $m \times n$ matrix with entries in a field $\mathbb{F}$. The map $T : \mathbb{F}^n \to \mathbb{F}^m$ defined by

$$T\vec{x} = T(\vec{x}) = A\vec{x}$$

where right-hand side is the usual matrix multiplication, is a linear map over $\mathbb{F}$. ■

Proposition 2.1 A linear map $T : \mathbb{F}^n \to \mathbb{F}^m$ is a matrix multiplication by a unique $m \times n$ matrix A with entries in $\mathbb{F}$. The matrix A is called the *standard matrix* of the linear map T.

$$\begin{array}{ccc} \mathrm{Hom}(\mathbb{F}^n, \mathbb{F}^m) & \overset{\text{natural identification}}{=\!=\!=} & \mathrm{M}_{m\times n}(\mathbb{F}) \\ T & \longmapsto & A \\ A\cdot & \longleftarrow\!\shortmid & A \end{array}$$

where $A\cdot : \vec{x} \mapsto A\vec{x}$ and A can be expressed as follows:

$$A = \begin{bmatrix} | & | & & | \\ T\vec{e}_1 & T\vec{e}_2 & \cdots & T\vec{e}_n \\ | & | & & | \end{bmatrix}$$

The vector $\vec{e}_i$ is the column vectors where only has the value 1 at the i-th place and 0 at other places.

Proof. Consider a column matrix $\vec{x} \in \mathbb{F}^n$ with entries $x_1, x_2, \cdots, x_n \in \mathbb{F}$. Then $\vec{x}$ can be expressed as a linear combination of the vectors $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$:

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n = \sum_{i=1}^{n} x_i\vec{e}_i$$

Since T is a linear map, it respects linear combinations. Therefore, we have:

$$T\vec{x} = T\left(\sum_{i=1}^{n} x_i\vec{e}_i\right) = \sum_{i=1}^{n} x_i T(\vec{e}_i) = \sum_{i=1}^{n} x_i\vec{a}_i = A\vec{x}$$

where $\vec{a}_i = T(\vec{e}_i)$ is the i-th column of the matrix $A = \begin{bmatrix} | & | & & | \\ T\vec{e}_1 & T\vec{e}_2 & \cdots & T\vec{e}_n \\ | & | & & | \end{bmatrix}$. Thus, we have $T\vec{x} = A\vec{x}$ for all $\vec{x} \in \mathbb{F}^n$. This shows that T can be represented as a matrix multiplication by the matrix A. ■

There is a simpler way to write $\sum_{i=1}^{n} x_i\vec{e}_i$: The Einstein Summation Convention. When an index variable appears twice in a single term and is not otherwise defined, it implies summation of that term over all the values of the index. Therefore, we can write:

$$\vec{x} = x_i\vec{e}_i$$

where i is summed from 1 to n.

Definition 2.3 — Homogeneous Linear Functions. A linear map $f : \mathbb{F}^n \to \mathbb{F}$ is called a *homogeneous linear function* or a *linear functional* if it satisfies the property:

$$f(\alpha\vec{x}) = \alpha f(\vec{x}) \quad \text{for all } \alpha \in \mathbb{F} \text{ and } \vec{x} \in \mathbb{F}^n.$$

Corollary 2.2 — Standard Matrix of a Linear Map. The standard matrix of a linear map $T : \mathbb{F}^n \to \mathbb{F}^m$ can be written as:

$$A = \begin{bmatrix} f_1(\vec{e}_1) & f_1(\vec{e}_2) & \cdots & f_1(\vec{e}_n) \\ f_2(\vec{e}_1) & f_2(\vec{e}_2) & \cdots & f_2(\vec{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ f_m(\vec{e}_1) & f_m(\vec{e}_2) & \cdots & f_m(\vec{e}_n) \end{bmatrix}$$

where $f_i : \mathbb{F}^n \to \mathbb{F}$ is the i-th component function of T, i.e., $T\vec{x} = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$ for all $\vec{x} \in \mathbb{F}^n$.

R Each component function f_i is a homogeneous linear function, and the standard matrix A is constructed by evaluating these functions at the standard basis vectors $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$ of $\mathbb{F}^n$.

■ **Example 2.2** Let $D : \mathbb{F}[t] \to \mathbb{F}[t]$ be the differentiation operator defined by:

$$D\left(\sum_{n=0}^{N} a_n t^n\right) = \sum_{n=1}^{N} n a_n t^{n-1}$$

for all polynomials $\sum_{n=0}^{N} a_n t^n \in \mathbb{F}[t]$. The differentiation operator D is a linear map over $\mathbb{F}$. The standard matrix of D with respect to the standard basis $\{1, t, t^2, \cdots, t^N\}$ of $\mathbb{F}[t]$ is given by:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & N \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

■

Proposition 2.2 Let X be a set and W be a linear space over a field $\mathbb{F}$. Then the set of all set maps from X to W, denoted by $\mathrm{Map}(X, W)$, is a linear space over $\mathbb{F}$ with the following operations defined pointwisely:

$$\begin{aligned} + : \mathrm{Map}(X,W) \times \mathrm{Map}(X,W) &\to \mathrm{Map}(X,W) \\ (f,g) &\mapsto (f+g : x \mapsto f(x)+g(x)) \qquad f+g : X \to W \end{aligned}$$

$$\begin{aligned} \cdot : \mathbb{F} \times \mathrm{Map}(X,W) &\to \mathrm{Map}(X,W) \\ (\alpha,f) &\mapsto (\alpha f : x \mapsto \alpha f(x)) \qquad \alpha f : X \to W \end{aligned}$$

Proof. The $\mathrm{Map}(X, W)$ is defined pointwisely by $\mathbb{F}$, hence it is trivial to be a linear map. ■

Proposition 2.3 Let V and W be two linear spaces over a field $\mathbb{F}$. Then $\mathrm{Hom}(V, W)$ is a linear space over $\mathbb{F}$ with the following operations defined pointwisely:

$$\begin{aligned} + : \mathrm{Hom}(V,W) \times \mathrm{Hom}(V,W) &\to \mathrm{Hom}(V,W) \\ (f,g) &\mapsto (f+g : \vec{v} \mapsto f(\vec{v})+g(\vec{v})) \qquad f+g : V \to W \end{aligned}$$

$$\begin{aligned} \cdot : \mathbb{F} \times \mathrm{Hom}(V,W) &\to \mathrm{Hom}(V,W) \\ (\alpha,f) &\mapsto (\alpha f : \vec{v} \mapsto \alpha f(\vec{v})) \qquad \alpha f : V \to W \end{aligned}$$

Proof. Note that $\mathrm{Hom}(V,W) \subseteq \mathrm{Map}(V,W)$. We need to show that the operations defined above are closed in $\mathrm{Hom}(V,W)$, i.e., for all $f,g \in \mathrm{Hom}(V,W)$ and $\alpha \in \mathbb{F}$, $f+g \in \mathrm{Hom}(V,W)$ and $\alpha f \in \mathrm{Hom}(V,W)$ or equivalently, f respects linear combinations.

Let $\vec{u},\vec{v} \in V$ and $\alpha,\beta \in \mathbb{F}$. Since $f,g \in \mathrm{Hom}(V,W)$, we have:

$$\begin{aligned}(f+g)(\alpha\vec{u}+\beta\vec{v}) &\overset{\text{def}}{=\!=} f(\alpha\vec{u}+\beta\vec{v})+g(\alpha\vec{u}+\beta\vec{v}) \\ &\overset{\text{lin}}{=\!=} \alpha f(\vec{u})+\beta f(\vec{v})+\alpha g(\vec{u})+\beta g(\vec{v}) \\ &=\!= \alpha(f(\vec{u})+g(\vec{u}))+\beta(f(\vec{v})+g(\vec{v})) \\ &\overset{\text{def}}{=\!=} \alpha(f+g)(\vec{u})+\beta(f+g)(\vec{v})\end{aligned}$$

where "lin" denotes the linearity of f and g. Thus, $f+g \in \mathrm{Hom}(V,W)$ and $\alpha f \in \mathrm{Hom}(V,W)$. ■

R Note that $\mathrm{End}(V) = \mathrm{Hom}(V,V)$ is a linear space over $\mathbb{F}$ and also a ring with the addition and multiplication operations defined in the previous section. The addition operation is commutative, but the multiplication operation is not necessarily commutative.

Then we can say that

$$\mathrm{Map}(\mathbb{F}^n,\mathbb{F}^m) \supseteq \mathrm{Hom}(\mathbb{F}^n,\mathbb{F}^m) \cong \mathrm{M}_{m\times n}(\mathbb{F})$$

2.2 Injections, Surjections and Isomorphisms

Similar to normal maps, there are injective, surjective and bijective linear maps.

Definition 2.4 — Injective Lienar Maps. A linear map $f : V \to W$ between two linear spaces over a field $\mathbb{F}$ is said to be *injective* (or one-to-one) if for all $\vec{u}, \vec{v} \in V$, the following holds:

$$f(\vec{u}) = f(\vec{v}) \implies \vec{u} = \vec{v}$$

Equivalently, f is injective if the only vector in V that maps to the zero vector in W is the zero vector itself:

$$f(\vec{u}) = 0 \implies \vec{u} = 0$$

Definition 2.5 — Surjective Lienar Maps. A linear map $f : V \to W$ is said to be *surjective* (or onto) if for every $\vec{w} \in W$, there exists at least one $\vec{v} \in V$ such that:

$$f(\vec{v}) = \vec{w}$$

Definition 2.6 — Invertible Linear Maps / Linear Equivalences. A linear map $T : V \to W$ is said to be *invertible* if T has a unique two-sided inverse S, denoted by T^{-1}, i.e., there exists a linear map $S : W \to V$ such that:

$$TS = e_W \quad \text{and} \quad ST = e_V$$

where $e_V : V \to V$ and $e_W : W \to W$ are the identity maps on V and W, respectively. In this case, we say that the linear spaces V and W are *isomorphic* or *linear equivalent*, denoted by $V \cong W$.

Corollary 2.3 — Invertible Linear Maps. A linear map $T : V \to W$ is invertible if and only if T is both injective and surjective, i.e., bijective / one-to-one correspondence.

Proof. ($\Rightarrow$) Assume $T : V \to W$ is invertible. By definition, there exists a linear map $S : W \to V$ such that $TS = e_W$ and $ST = e_V$.

To show that T is injective, suppose $T(\vec{u}) = T(\vec{v})$ for some $\vec{u}, \vec{v} \in V$. We have:

$$S(T(\vec{u})) = S(T(\vec{v})) \implies (ST)(\vec{u}) = (ST)(\vec{v}) \implies e_V(\vec{u}) = e_V(\vec{v}) \implies \vec{u} = \vec{v}$$

Thus, T is injective. Then, to show that T is surjective, let $\vec{w} \in W$. Since $TS = e_W$, we have:

$$T(S(\vec{w})) = e_W(\vec{w}) = \vec{w}$$

Then for every $\vec{w} \in W$, there exists a $\vec{v} = S(\vec{w}) \in V$ such that $T(\vec{v}) = \vec{w}$. Thus, T is surjective.

($\Leftarrow$) Now assume that $T : V \to W$ is both injective and surjective. We need to show that there exists a linear map $S : W \to V$ such that $TS = e_W$ and $ST = e_V$.

Since T is surjective, for each $\vec{w} \in W$, there exists at least one $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$. Define the map $S : W \to V$ by choosing one such preimage for each $\vec{w}$:

$$S(\vec{w}) = \text{a chosen } \vec{v} \text{ such that } T(\vec{v}) = \vec{w}$$

To show that S is well-defined, we need to ensure that if $T(\vec{v}_1) = T(\vec{v}_2)$, then $\vec{v}_1 = \vec{v}_2$. This follows from the injectivity of T.

Now we verify that $TS = e_W$: $(TS)(\vec{w}) = T(S(\vec{w})) = \vec{w}$ for all $\vec{w} \in W$. Thus, $TS = e_W$. Next, we verify that $ST = e_V$: $(ST)(\vec{v}) = S(T(\vec{v})) = \vec{v}$ for all $\vec{v} \in V$. Thus, $ST = e_V$.

Therefore, T has a two-sided inverse S, and hence T is invertible. ■

Definition 2.7 — Characteristic of a Field. The *characteristic* of a field $\mathbb{F}$ is the smallest positive integer n such that:

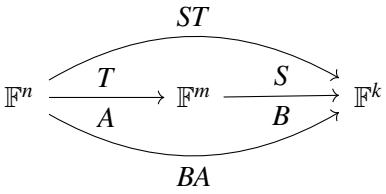
$$\underbrace{1+1+\cdots+1}_{n \text{ times}} = 0$$

If no such positive integer exists, the characteristic of $\mathbb{F}$ is defined to be 0.

■ **Example 2.3** The differentiation operator $D : \mathbb{F}[t] \to \mathbb{F}[t]$ is not an injective linear map as $D(1) = 0 = D(2)$ but is a surjective linear map if $\mathbb{F}$ is a field of characteristic 0. ■

2.3 Matrix Multiplications and Compositions of Linear Maps

We consider two linear maps $T : \mathbb{F}^n \to \mathbb{F}^m$ and $S : \mathbb{F}^m \to \mathbb{F}^k$ with standard matrices A and B, respectively. We want to find the standard matrix of the composition $ST : \mathbb{F}^n \to \mathbb{F}^k$.



Proposition 2.4 The standard matrix of the composition $ST : \mathbb{F}^n \to \mathbb{F}^k$ is the matrix multiplication BA, i.e., for all $\vec{x} \in \mathbb{F}^n$,

$$(ST)\vec{x} = B(A\vec{x}) = (BA)\vec{x}$$

Proof. Let $\vec{x} \in \mathbb{F}^n$ be a column matrix with entries $x_1, x_2, \cdots, x_n \in \mathbb{F}$. Then $\vec{x}$ can be expressed as a linear combination of the standard basis vectors $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$:

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n = \sum_{i=1}^{n} x_i\vec{e}_i$$

Consider the j-th column of BA, it is given by:

$$(ST)\vec{e}_j = S(T(\vec{e}_j)) = S(\vec{a}_j) = B\vec{a}_j = (BA)\vec{e}_j$$

for all $j = 1, 2, \ldots, n$. This shows that the standard matrix of the composition ST is indeed the matrix multiplication BA. ■

R Note that B is a $k \times m$ matrix and A is an $m \times n$ matrix, so the matrix multiplication BA is defined and results in a $k \times n$ matrix.

The matrix multiplication BA can be computed as follows:

$$BA = B\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{a}_1 & B\vec{a}_2 & \cdots & B\vec{a}_n \\ | & | & & | \end{bmatrix}$$

where $\vec{a}_i = T(\vec{e}_i)$ is the i-th column of the matrix A. Also,

$$B\vec{x} = x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_n\vec{b}_n = \sum_{i=1}^{n} x_i\vec{b}_i$$

where $\vec{b}_i = B\vec{a}_i$ is the i-th column of the matrix B. Note that B is a $k \times m$ matrix, and $\vec{x} \in \mathbb{F}^m$. Thus, the matrix multiplication $B\vec{x}$ is defined and results in a column matrix in $\mathbb{F}^k$.

2.4 Elementary Row Operations

Definition 2.8 — Elementary Row Operations. Let A be an $m \times n$ matrix over a field $\mathbb{F}$. An *elementary row operation* on A is one of the following operations:

1. Row Interchange: $R_i \leftrightarrow R_j$.
2. Row Multiplication: $R_i \to \alpha R_i$, where $\alpha \in \mathbb{F} \setminus \{0\}$.
3. Row Addition: $R_i \to R_i + \alpha R_j$, where $\alpha \in \mathbb{F}$ and $i \neq j$.

Each elementary row operation can be represented by *left multiplication* of A by an appropriate $m \times m$ matrix over $\mathbb{F}$. Note that all of them are invertible linear maps from $\mathbb{F}^{m\times n}$ to $\mathbb{F}^{m\times n}$.

For easier notations, we introduce the idea of matrix units, which is similar to the standard basis vectors $\vec{e}_i$.

Definition 2.9 — Matrix Units. Let m and n be two positive integers and $\mathbb{F}$ be a field. The *matrix unit* E_{ij} is the $m \times n$ matrix over $\mathbb{F}$ with 1 in the (i, j)-th position and 0 elsewhere, i.e.,

$$(E_{ij})_{kl} = \begin{cases} 1 & \text{if } (k,l) = (i,j) \\ 0 & \text{otherwise} \end{cases}$$

for all $1 \leq k \leq m$ and $1 \leq l \leq n$.

It can also be defined as $E_{ij} = \vec{e}_i \hat{e}_j \in M_{m\times n}(\mathbb{F})$ where $\vec{e}_i \in \mathbb{F}^m$ and $\vec{e}_j^{\,T} = \hat{e}_j \in \mathbb{F}^n$ are the i-th and j-th standard basis vectors, respectively. The $\hat{e}_j$ is the row matrix with 1 in the j-th column and 0 anywhere else.

R Note that for any $m \times n$ matrix A over a field $\mathbb{F}$, we have:

$$A\vec{e}_j = \text{the } j\text{-th column of } A \in \mathbb{F}^n$$
$$\hat{e}_i A = \text{the } i\text{-th row of } A \in (\mathbb{F}^m)^*$$

where $(\mathbb{F}^m)^*$ is the set of all row matrices with n entries in $\mathbb{F}$. $\hat{e}_i$ is an element in $(\mathbb{F}^m)^*$ for any $1 \leq i \leq m$. Then we have:

$$a_{ij} = \hat{e}_i A \vec{e}_j = \text{the } (i,j)\text{-th entry of } A$$

We can write the $E_{i,j}$ as:

$$E_{i,j} = \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \text{the } i\text{-column} \\ \downarrow \\ \end{array} \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{the } j\text{-row} \\ \\ \\ \\ \end{array}$$

Then we consider the row operations by using the matrix units.

Proposition 2.5 The row operation $R_i \leftrightarrow R_j$ is a linear map where the standard matrix is $A_{R_i \leftrightarrow R_j} = I - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$.

Proof. The linear map $T : \mathbb{F}^n \to \mathbb{F}^n$ is defined pointwisely. We can say the map is:

$$\vec{e}_k \mapsto \begin{cases} \vec{e}_j & \text{if } k = i \\ \vec{e}_i & \text{if } k = j \\ \vec{e}_k & \text{if } k \neq i, j \end{cases}$$

Then the standard matrix of T is:

$$A_{R_i \leftrightarrow R_j} = [\vec{e}_1 \ \cdots \ \vec{e}_j \ \cdots \ \vec{e}_i \ \cdots \ \vec{e}_n] = I - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$$

where I is the $n \times n$ identity matrix. ■

Proposition 2.6 The row operation $R_i \to \alpha R_i$ where $\alpha \in \mathbb{F}^x := \mathbb{F} \setminus \{0\}$ is a linear map where the standard matrix is $A_{R_i \to \alpha R_i} = I + (\alpha - 1)E_{i,i}$.

Proof. The linear map $T : \mathbb{F}^n \to \mathbb{F}^n$ is defined pointwisely. We can say the map is:

$$\vec{e}_k \mapsto \begin{cases} \alpha \vec{e}_i & \text{if } k = i \\ \vec{e}_k & \text{if } k \neq i \end{cases}$$

Then the standard matrix of T is:

$$A_{R_i \to \alpha R_i} = [\vec{e}_1 \ \cdots \ \alpha \vec{e}_i \ \cdots \ \vec{e}_n] = I + (\alpha - 1)E_{i,i}$$

where I is the $n \times n$ identity matrix. ■

Proposition 2.7 The row operation $R_i \to R_i + \alpha R_j$ where $\alpha \in \mathbb{F}$ and $i \neq j$ is a linear map where the standard matrix is $A_{R_i \to R_i + \alpha R_j} = I + \alpha E_{i,j}$.

Proof. The linear map $T : \mathbb{F}^n \to \mathbb{F}^n$ is defined pointwisely. We can say the map is:

$$\vec{e}_k \mapsto \begin{cases} \vec{e}_i + \alpha \vec{e}_j & \text{if } k = i \\ \vec{e}_k & \text{if } k \neq i \end{cases}$$

Then the standard matrix of T is:

$$A_{R_i \to R_i + \alpha R_j} = [\vec{e}_1 \ \cdots \ (\vec{e}_i + \alpha \vec{e}_j) \ \cdots \ \vec{e}_n] = I + \alpha E_{i,j}$$

where I is the $n \times n$ identity matrix. ■

2.5 Dimensions of Vector Spaces

Definition 2.10 — Finite Dimensional Vector Spaces. A linear space V over a field $\mathbb{F}$ is said to be *finite dimensional* if there exists a linear equivalence $T : V \to \mathbb{F}^n$ for some positive integer n. In this case, we say that the dimension of V is n, denoted $\dim_{\mathbb{F}} V = n$ or simply $\dim V = n$.

Definition 2.11 — Infinite Dimensional Vector Spaces. A linear space V over a field $\mathbb{F}$ is said to be *infinite dimensional* if V is not finite dimensional.

We have to proof if the dimension of a finite dimensional vector space is well-defined.

Proposition 2.8 If there exists two linear equivalences $T : V \to \mathbb{F}^m$ and $S : V \to \mathbb{F}^n$, then $n = m$.

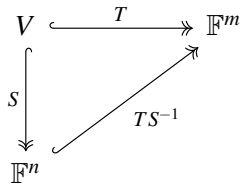
Proof. Since S is linear equivalence, it has a unique two-sided inverses $S^{-1} : \mathbb{F}^n \to V$. Consider the composition of this map:

$$TS^{-1} : \mathbb{F}^n \to \mathbb{F}^m$$

Since TS^{-1} is compositions of linear equivalences, it is also a linear equivalence. Mutantis mutandis, for the opposite direction.

Now, we know that a linear equivalence between two finite-dimensional vector spaces. Then we have $\dim \mathbb{F}^n = \dim \mathbb{F}^m$ or $n = m$. Thus, the dimension of a finite dimensional vector space is well-defined. ■

Graphically, we have the following commutative diagram:



R In drawing commutative diagram, we can use $\hookrightarrow$ to denote an injective linear map, $\twoheadrightarrow$ to denote a surjective linear map, and $\cong$ or combining the two to denote an invertible linear map.

2.6 Elementary Column Operations, Canonical Form and Rank

Definition 2.12 — Elementary Column Operations. Let A be an $m \times n$ matrix over a field $\mathbb{F}$. An *elementary column operation* on A is one of the following operations:

1. Interchange two columns of A: $C_i \leftrightarrow C_j$.
2. Multiply a column of A by a nonzero scalar in $\mathbb{F}$: $C_i \to \alpha C_i$ where $\alpha \in \mathbb{F} \setminus \{0\}$.
3. Add a scalar multiple of one column of A to another column of A: $C_i \to C_i + \alpha C_j$ where $\alpha \in \mathbb{F}$ and $i \neq j$.

Each elementary column operation can be represented by *right multiplication* of A by an appropriate $n \times n$ matrix over $\mathbb{F}$. Note that all of them are invertible linear maps from $\mathbb{F}^{m\times n}$ to $\mathbb{F}^{m\times n}$.

Proposition 2.9 Any $m \times n$ matrix A can be transformed into a matrix of the form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ by a finite sequence of elementary row and column operations on A, where r is the rank of A.

The following is the commutative diagram of the proposition above, where $B = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$:

$$\begin{array}{ccc} \mathbb{F}^n & \xrightarrow{\quad A \quad} & \mathbb{F}^m \\ \big\downarrow Q & & \big\downarrow P \\ \mathbb{F}^n & \xrightarrow{\quad B \quad} & \mathbb{F}^m \end{array}$$

Note that P is the product of a finite sequence of elementary row operation matrices and Q is the product of a finite sequence of elementary column operation matrices. Both P and Q are elementary and invertible matrices. Thus, we have:

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = PAQ^{-1}$$

Definition 2.13 — Canonical Form of a Matrix. The matrix $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ obtained from an $m \times n$ matrix A by a finite sequence of elementary row and column operations on A is called the *canonical form* of A.

R The canonical form of a matrix defined is also called the *Smith Normal Form* or *Normal Form* of a matrix.

Definition 2.14 — Rank of a Matrix. The *rank* of an $m \times n$ matrix A over a field $\mathbb{F}$, denoted by $\operatorname{rank}(A)$, is the number of leading 1's in the matrix $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ obtained from A by a finite sequence of elementary row and column operations on A.

R The value r is uniquely determined by A.

Proposition 2.10 Let A be an $m \times n$ matrix over a field $\mathbb{F}$. Then the following statements are equivalent:

$$A \text{ is invertible} \iff m\left\{\underbrace{\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}}_{n}\right. \text{ is invertible} \iff \operatorname{rank}(A) = m = n \iff \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = I_m = I_n$$

Proof. If A is invertible, then the matrix PAQ^{-1} is also invertible, as P and Q are elementary and invertible matrices, and hence the product is invertible.

If PAQ^{-1} is invertible, and note that $m = n$ is automatically true. As only square matrix is invertible. Without the loss of generality, let say PAQ^{-1} is a $m \times m$ matrix, then we have $\text{rank}(PAQ^{-1}) = m$. Also note that the rank is invarient under multiplication by invertible matrices, so $\text{rank}(A) = \text{rank}(PAQ^{-1})$. Hence, $\text{rank}(A) = m = n$.

If $\text{rank}(A) = m = n$, as the canonical matrix remains the $m \times n$ structure, we know that the canonical form is actually a square matrix, let say $m \times m$. Also $r = \text{rank}(A) = m$. Hence the whole canonical form become an identity matrix I_m.

If the canonical form is an identity matrix I, i.e., it is invertible. Then the matrix $P^{-1}IQ = A$ is also invertible for some elementary and invertible matrices P and Q. ■

Proposition 2.11 Let A be an $m \times n$ matrix over a field $\mathbb{F}$. Then the following statements are equivalent:

$$A \text{ has a left inverse} \iff A \text{ is injective} \iff \text{rank}(A) = n \iff \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_n \\ 0 \end{bmatrix}$$

Proof. If A has a left inverse, let say B, then we have $BA = I_n$. Then for $B(A(\vec{x}_1)) = B(A(\vec{x}_2))$, we have $(BA)\vec{x}_1 = (BA)\vec{x}_2$, which implies $x_1 = x_2$. Hence it is injective.

If A is injective, we can consider $A = P^{-1}CQ$, where C is the canonical form of the matrix A. Then we consider $P^{-1}CQ\vec{x} = \vec{0}$. Since P^{-1} is invertible, it won't produce non-trivial solutions. We can consider $C(Q\vec{x}) = \vec{0} = C\vec{y}$. Then we have

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where $\vec{y}_1$ and $\vec{y}_2$ are column vecotrs with size r and $n-r$ respectively. Then $I_r\vec{y}_1 = 0$, which implies $\vec{y}_1 = 0$, while $\vec{y}_2$ can be anything. As A is invertible, then $A\vec{x} = \vec{0}$ only has one trivial solution $\vec{x} = \vec{0}$. Also, Q is invertible, hence $\vec{y}$ has only one trivial solution $\vec{0}$, i.e., $\vec{y}_2 = \vec{0}$. Hence we have $n - r = 0$ due to the size of $\vec{y}_2$ being 0. Hence the rank of A is n.

If $\text{rank}(A) = n$, then the canonical form of A is

$$\begin{bmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} = \begin{bmatrix} I_{n\times n} & 0_{n\times(n-n)} \\ 0_{(m-n)\times n} & 0_{(m-n)\times(n-n)} \end{bmatrix} = \begin{bmatrix} I_{n\times n} \\ 0_{(m-n)\times n} \end{bmatrix} = \begin{bmatrix} I_n \\ 0 \end{bmatrix}$$

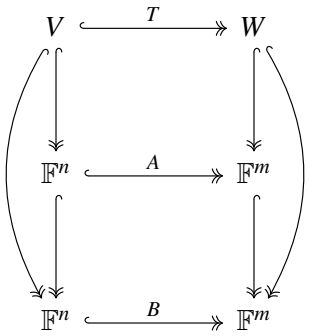
If the canonical form of A is $\begin{bmatrix} I_n \\ 0 \end{bmatrix}$, then we consider $PAQ^{-1} = C$. Also, $A = P^{-1}CQ$. We construct a candidate for left inverse $D = [I_n\ 0]$. Then we have $DC = [I_n\ 0]\begin{bmatrix} I_n \\ 0 \end{bmatrix} = I_n$. Then the left inverse of A is $L = QDP^{-1}$. Then we check, $LA = QDP^{-1}A = QDP^{-1}PCQ^{-1} = I_n$. Hence, A indeed has a left inverse. ■

Proposition 2.12 Let A be an $m \times n$ matrix over a field $\mathbb{F}$. Then the following statements are equivalent:

$$A \text{ has a right inverse} \iff A \text{ is surjective} \iff \text{rank}(A) = m \iff \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_m & 0 \end{bmatrix}$$

Proposition 2.13 For every $\vec{b}$, $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\vec{x} = \vec{b}$ has a unique solution.

Linear Algebra is the study of linear map between two finite dimensional vector spaces.



where $B = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, $\dim V = n$ and $\dim W = m$. The bended arrows denote the trivialisation of the vector spaces and the bottom arrow denotes the canonical matrix representation of the linear map T.

3. Linear Spaces

"Babies have to survive, so they have the strong desire to learn stuffs. You think you are not good at math because you don't have the strong desire to learn math."

Guowu Meng

3.1 Linear Subspaces, Kernels and Images

Here, we discuss linear spaces with more in depth terms.

Definition 3.1 — Linear Subspaces. Let W be a linear space over $\mathbb{F}$ and V is a subset of W, denoted as $V \subset W$. V is a *linear subspace* of W if V, with $+$ and $\cdot$ inherited from those of W, is a linear space.

Proposition 3.1 Let $V \subset W$. V is a subspace of W if and only if V is not empty and V is closed under $+$ and $\cdot$.

Proof. If V is a subspace of W, then V is non-empty as a linear space must contain a zero vector by definition, as V is also a linear space. Also, the other two are due to the axioms of linear space.

If V is not empty and closed under $+$ and $\cdot$, we just have to check the each axiom. ■

Definition 3.2 — Kernels. Let $f : V \to W$ be a linear map. The *kernel* of f, denoted as $\ker(f)$, is defined as

$$\ker(f) \stackrel{\text{def}}{=\!=} f^{-1}(0_W) = \{v \in V \mid f(v) = 0_W\}$$

▪ **Example 3.1** Let $f : V \to W$ be a linear map. $\ker(f)$ is a subspace of domain of f, i.e., V.

First, we have $0_V \in \ker(f)$, as $f(0_V) = 0_W$, so $\ker(f)$ is not empty.

Then we consider $\alpha_1, \alpha_2 \in \mathbb{F}$ and $v_1, v_2 \in \ker(f)$, we have

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2) = \alpha_1(0_W) + \alpha_2(0_W) = 0_W$$

The first equality due to the linearity of f and the second is due to $v_i \in \ker(f)$. ▪

Definition 3.3 — Images. Let $f : V \to W$ be a linear map. The *image* of f, denoted by $\mathrm{Im}(f)$, is defined as

$$\mathrm{Im}(f) \stackrel{\text{def}}{=\!=} \{f(v) \mid v \in V\} \subset W$$

■ **Example 3.2** Let $f : V \to W$ be a linear map. $\mathrm{Im}(f)$ is a subspace of codomain of f, i.e., W.
First, we have $f(0_V) = 0_W \in \mathrm{Im}(f)$, so $\mathrm{Im}(f)$ is not empty.
Then we consider $\alpha_1, \alpha_2 \in \mathbb{F}$ and $f(v_1), f(v_2) \in \mathrm{Im}(f)$. We have

$$\alpha_1 f(v_1) + \alpha_2 f(v_2) = f(\alpha_1 v_1 + \alpha_2 v_2) \in \mathrm{Im}(f)$$

The equality is due to the linearity of f. ■

■ **Example 3.3** Let W be a linear space over a field $\mathbb{F}$ and $\{V_\alpha\}_{\alpha \in I}$ be the family of subspaces of W indexed by the element in the index set I. Then $\bigcap_{\alpha \in I} V_\alpha$ is also a subspace of W.
First, we have $0_W \in V_\alpha$ for all $\alpha \in I$, so $0_W \in \bigcap_{\alpha \in I} V_\alpha$. Thus, $\bigcap_{\alpha \in I} V_\alpha$ is not empty.
Then we consider $\alpha_1, \alpha_2 \in \mathbb{F}$ and $v_1, v_2 \in \bigcap_{\alpha \in I} V_\alpha$. We have $v_1, v_2 \in V_\alpha$ for all $\alpha \in I$. Thus, $\alpha_1 v_1 + \alpha_2 v_2 \in V_\alpha$ for all $\alpha \in I$. This shows that $\alpha_1 v_1 + \alpha_2 v_2 \in \bigcap_{\alpha \in I} V_\alpha$. ■

Then we consider the duality of the intersection and union of subspaces. Whether the union of two subspaces is still a subspace? Unfortunately, the answer is no in general case. However, we have the following proposition.

Proposition 3.2 Let W be a linear space over a field $\mathbb{F}$ and consider the family of subspaces $\{V_\alpha\}_{\alpha \in I}$. Then $\overline{\bigcup_{\alpha \in I} V_\alpha}$ is a subspace of W where $\overline{\bigcup_{\alpha \in I} V_\alpha}$ is the completion of $\bigcup_{\alpha \in I} V_\alpha$ under linear combinations. We call $\overline{\bigcup_{\alpha \in I} V_\alpha}$ the *sum* of the subspaces $\{V_\alpha\}_{\alpha \in I}$, denoted by $\sum_{\alpha \in I} V_\alpha$.

3.2 Linear Span and Linear Independence

Definition 3.4 — Linear Span. Let V be a linear space over a field $\mathbb{F}$ and $S \subset V$. The *linear span* of S, denoted by $\text{span}_{\mathbb{F}}(S)$ or simply $\text{span}(S)$ or $\overline{S}$ or $\langle S \rangle$, is defined as the completion of S inside V under linear combinations.

Corollary 3.1 The linear span of S can also be defined as the intersection of all subspaces of V containing S, which is the smallest linear subspace of V containing S. It can be written as:

$$\text{span}(S) = \bigcap_{\alpha \in I} V_\alpha \subset V \quad \text{where } I = \{V_\alpha \subset V \mid V_\alpha \text{ is a subspace of } V \text{ and } S \subset V_\alpha\}$$

R Note that I is not empty as $V \in I$. Thus, $\text{span}(S)$ is well-defined. V is the largest subspace of itself and $\{0_V\}$ is the smallest subspace of V.

Proposition 3.3 Let W be a linear space over a field $\mathbb{F}$ and $S \subset W$. Then

$$\text{span}(S) = \left\{ \sum_{i=1}^{n} \alpha_i s_i \mid n \in \mathbb{N}, \alpha_i \in \mathbb{F}, s_i \in S \right\}$$

Note that the summation is a finite summation.

Definition 3.5 — Linear Independences. Let W be a linear space over a field $\mathbb{F}$ and $V_1, \cdots, V_k$ be subspaces of W. The subspaces $V_1, \cdots, V_k$ are said to be *linearly independent* if $V_i \neq \{0_W\}$ for all i and there is one and only one way to split $0_W \in W$ as a sum of vectors from each V_i, i.e., if $v_i \in V_i$ for all i and $\sum_{i=1}^{k} v_i = 0_W$, then $v_i = 0_W$ for all i.

Vectors $v_1, v_2, \cdots, v_k \in W$ are said to be independent if the subspaces $\text{span}(v_1)$, $\text{span}(v_2)$, $\cdots$, $\text{span}(v_k)$ are linearly independent.

Proposition 3.4 $v_1, v_2, \cdots, v_k \in W$ are linearly independent if and only if there is one and only one way to write $0_W \in W$ as the combination of $v_1, \cdots, v_k$ with coefficients in $\mathbb{F}$, i.e., the equation

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0_W$$

has only the trivial solution, i.e., $\alpha_i = 0$ for all i.

3.3 Linearly Independent Sets and Spanning Sets

If we consider a set, what does it mean by being linearly independent? Is there any properties for spanning if the set spans the whole codomain?

Definition 3.6 — Linearly Independent Sets. Let V be a linear space over a field $\mathbb{F}$. A subset $S \subseteq V$ is said to be a *linearly independent set* of vectors in V if no elements in S can be expressed as a linear combination of the finitely many other elements in S.

Definition 3.7 — Spanning Sets. Let V be a linear space over a field $\mathbb{F}$. A subset $S \subseteq V$ is said to be a *spanning set* of V if $\text{span}(S) = V$.

■ **Example 3.4** Let $V = \mathbb{F}^3$ and consider the three vectors $\vec{e}_1$, $\vec{e}_2$ and $\vec{e}_3$.

Then the set $S = \{\vec{e}_1, \vec{e}_2, \vec{e}_1 + \vec{e}_2\}$ is not a spanning set of V as $\text{span}(S) = \text{span}\{\vec{e}_1, \vec{e}_2\} \neq V$. If we consider the $\text{span}\{\vec{e}_1, \vec{e}_2\} = W$, then $\{\vec{e}_1, \vec{e}_2\}$ is a minimal spanning set of W.

The set $S = \{\vec{e}_1, \vec{e}_1 + \vec{e}_2, \vec{e}_1 + \vec{e}_2 + \vec{e}_3\}$ is a spanning set of V. ■

R If we consider the matrix of $\{\vec{e}_1, \vec{e}_2, \vec{e}_1 + \vec{e}_2\}$ with respect to the standard basis of $\mathbb{F}^3$, we have:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Then we have $\text{rank}(A) = 2 < 3$. Thus, the set is not a spanning set of $\mathbb{F}^3$.

■ **Example 3.5** Consider the subset $S = \{1, t, t^2, \cdots\} \subset \mathbb{F}[[t]]$. Then $\text{span}(S) = \mathbb{F}[t]$ which is a proper subspace of $\mathbb{F}[[t]]$. As the linear combination of finitely many elements in S is a polynomial, but an element in $\mathbb{F}[[t]]$ can be a power series. ■

Definition 3.8 — Minimal Spanning Sets. Let V be a linear space over a field $\mathbb{F}$. A spanning set $S \subseteq V$ is said to be a *minimal spanning set* of V if no proper subset of S is a spanning set of V, i.e., $S' \subset S \implies \text{span}(S') \subset \text{span}(S) = V$ where $\text{span}(S') \neq V$.

The following is also the equivalence definition of linearly independent sets, spanning sets and minimal spanning sets.

Given a linear space V over a field $\mathbb{F}$. We define the order set $S := \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\} \subseteq V$. The order set S forms a linear map $\phi_S : \mathbb{F}^n \to V$ defined by:

$$\phi_S(\vec{x}) = \phi_S\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \sum_{i=1}^{n} x_i\vec{v}_i$$

Proposition 3.5 The order set $S := \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\} \subseteq V$ is said to be *linearly independent* if and only if the linear map $\phi_S : \mathbb{F}^n \to V$ defined above is injective.

Proposition 3.6 The order set $S := \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\} \subseteq V$ is said to be a *spanning set* of V if and only if the linear map $\phi_S : \mathbb{F}^n \to V$ defined above is surjective.

Proposition 3.7 The order set $S := \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\} \subseteq V$ is said to be a *minimal spanning set* of V if and only if the linear map $\phi_S : \mathbb{F}^n \to V$ defined above is bijective.

R A order minimal spanning set is regarded as *basis*.

■ **Example 3.6** Let X be a set, $\mathbb{F}[[X]]$ be the set of all functions $f : X \to \mathbb{F}$ and $\mathbb{F}[X]$ be the set of all finite support functions $f : X \to \mathbb{F}$. For each $x \in X$, we define the delta function $\delta_x : X \to \mathbb{F}$ at

point x by

$$\delta_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{if } y \neq x \end{cases}$$

Clearly, δ_x has finite support, thus $\delta_x \in \mathbb{F}[X]$.

Then we have a set $\delta_X = \{\delta_x \mid x \in X\} \subset \mathbb{F}[X]$. We have $\text{span}(\delta_X) = \mathbb{F}[X]$ as any finite support function $f : X \to \mathbb{F}$ can be written as a linear combination of finitely many delta functions. Thus, δ_X is a spanning set of $\mathbb{F}[X]$.

However, δ_X is a linearly independent set. Assume that there exists a finite linear combination of other delta functions such that $\delta_x = \sum \alpha_y \delta_y$. Then we have $\delta_x(x) = 1 = \sum \alpha_y \delta_y(x) = 0$. This is a contradiction. Thus, δ_X is a linear independent set. ∎

3.4 Group Actions

Next, we discuss quotient space. However, before introducing quotient space, we have to understand what group actions are.

Definition 3.9 — Group Actions. Let G be a group and X be a set. A *left group action* of G on X is a map $\cdot : G \times X \to X$, $(g,x) \mapsto g \cdot x$, such that for all $g_1, g_2 \in G$ and $x \in X$, the following properties hold:

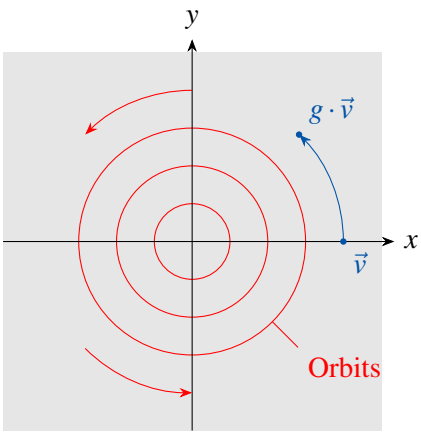
1. Compatibility: $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$.
2. Identity: $e \cdot x = x$ where e is the identity element of G.

Same for the right group action of G on X, just think it dually.

Consider rotation on a plane. It is a group action of the group $SO(2)$ on the set $\mathbb{R}^2$.

$$g = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Then we have the following group action:



Definition 3.10 — Orbits. Let G be a group acting on a set X. The *orbit* of the action through a point $x \in X$, denoted as $G \cdot x$, is defined as the set of points in X that can be reached from x by the action of elements of G, i.e.,

$$G \cdot x = \{g \cdot x \mid g \in G\}$$

There is only two situation for the orbits, either the origin or a circle.

In the following section, we may regard the orbits $G \cdot x$ as a *coset*.

Definition 3.11 — Partition. A *partition* of a set X is a collection of non-empty, disjoint subsets of X whose union is X. The partition of the set X is the same as an equivalence relation on X.

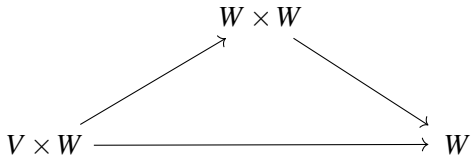
Orbits give a partition of the set X, i.e., X can be expressed as the disjoint union of its orbits. The orbits of the action are the equivalence classes of the equivalence relation.

Let $f : X \to Y$ be a map between two sets X and Y. Then f defines a partition of X by the equivalence relation. The equivalence classes are the preimages of points in Y, i.e., $f^{-1}(y)$ for each $y \in Y$.

3.5 Quotient Spaces

Let V be a subspace of a linear space W over a field $\mathbb{F}$. We know $(V,+)$ is an abelian group. Then we have the group action of V on W defined by: $(v,w) \mapsto v \cdot w$ for all $v \in V, w \in W$. $v \cdot w$ is defined as $v+w$ where $+$ is the addition operation in W. We know that $(v_1+v_2)+w = v_1+(v_2+w)$ and $0_V + w = w$ for all $v_1, v_2 \in V$ and $w \in W$. Thus, it is a group action.

The following commutative diagram illustrates the group action, where the associative and identity properties are inherited from the addition operation in W, i.e., we need not prove the group action as above.



This group action defines the following equivalence relation on W, where V is the acting group:

$$\begin{aligned} w_1 \sim w_2 &\implies \exists v \in V \text{ such that } w_2 = v + w_1 \\ &\iff w_2 - w_1 \in V \end{aligned}$$

Definition 3.12 — Quotient Spaces. Let W be a linear space over a field $\mathbb{F}$ and V be a subspace of W. The *quotient space* of W by V, denoted by W/V, is defined as the set of orbits of the group action of V on W, or the set of V-equivalence classes in W with the equivalence relation defined above, i.e.,

$$W/V = \{V \cdot w \mid w \in W\} = \{w+V \mid w \in W\}$$

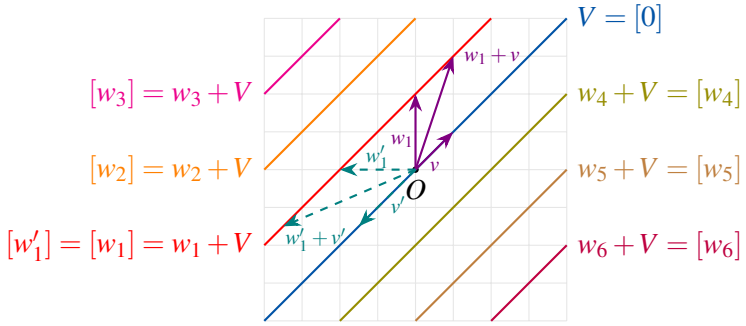
where $V \cdot w = w+V = \{w+v \mid v \in V\}$ is called the *coset* of V in W containing w.

Definition 3.13 — Quotient Map. The natural surjective map $\pi : W \to W/V$ defined by $\pi(w) = w+V$ for all $w \in W$ is called the *quotient map* or *projection map*. Note that $w+V$ can be written as $\overline{w}$ or $[w]$.

In general, if a group G acts on a set X, then the quotient set X/G is defined as the set of orbits of the action, i.e.,

$$X/G = \{G \cdot x \mid x \in X\}$$

Similarly, there is a natural surjective map $\pi : X \to G$ defined by $\pi(x) = G \cdot x$ for all $x \in X$. The following is a graphical illustration of the quotient space.



We can see that each line parallel to V represents a coset of V in W. The quotient space W/V is the set of all such lines. We may consider each line as an orbit of the group action of V on W. Note that there is not only one unique way to represent the coset $w+V$. Just like the illustration above, w_1 and w_1' are two different representatives of the same coset $w_1+V=w_1'+V$. Note that their difference is an element in V, i.e., $w_1-w_1'\in V$.

Note that we now do not know whether W/V is a linear space or not. We will show that it is indeed a linear space by using the following proposition.

Proposition 3.8 There is a unique linear structure on W/V such that the quotient map $\pi: W\to W/V$ is a linear map.

Proof. Assume that such a linear structure exists. Then for all $w_1,w_2\in W$ and $\alpha_1,\alpha_2\in\mathbb{F}$, we have

$$\pi(\alpha_1 w_1+\alpha_2 w_2)=\overline{\alpha_1 w_1+\alpha_2 w_2}=\alpha_1\overline{w_1}+\alpha_2\overline{w_2}=\alpha_1\pi(w_1)+\alpha_2\pi(w_2)$$

This suggests that $\alpha_1\overline{w_1}+\alpha_2\overline{w_2}$ should be defined as $\overline{\alpha_1 w_1+\alpha_2 w_2}$ if π is linear. As there is only one formula, this proves the uniqueness of the linear structure on W/V.

Then we consider whether the linear combination on W/V is well-defined. Assume that $\overline{w_1}=\overline{w_1'}$ and $\overline{w_2}=\overline{w_2'}$, i.e., $w_1-w_1'\in V$ and $w_2-w_2'\in V$. Then we have

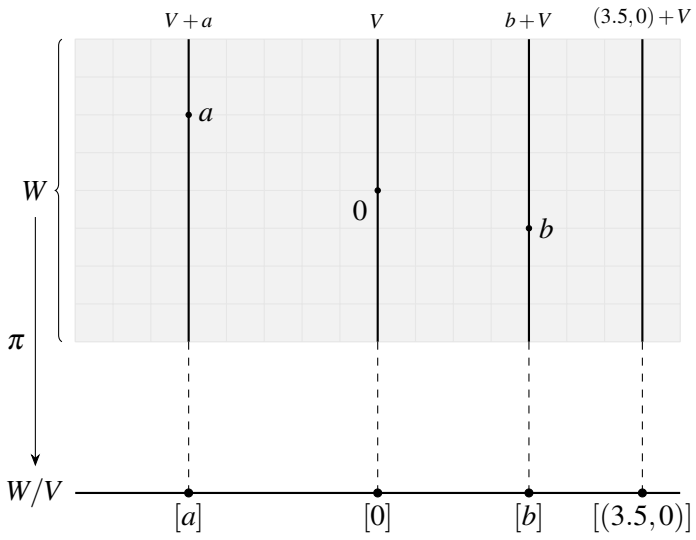
$$(\alpha_1 w_1+\alpha_2 w_2)-(\alpha_1 w_1'+\alpha_2 w_2')=\alpha_1(w_1-w_1')+\alpha_2(w_2-w_2')\in V$$

which means $\overline{\alpha_1 w_1+\alpha_2 w_2}=\overline{\alpha_1 w_1'+\alpha_2 w_2'}$. This means that the linear combination is independent of the choice of representatives. Thus, the linear combination is well-defined. ■

In normal procedure, we first define the operations and then check whether the set is closed under the operations and zero exists. Then we check whether the map preserves the structure and show the uniqueness of the structure. However, in this case, we first assume that such a structure exists and then derive the operations from the assumption. Then we check whether the operations are well-defined.

In the first part, we show that there is only one possible way to define the operations if the quotient map is linear. Also, during the definition, it ensures the preservation of the linear structure. In the second part, we show that the operations on the set W/V are well-defined.

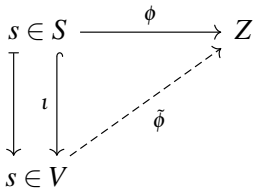
If we consider the graphical representation of the quotient space W/V and the quotient map π, we may use the following diagram:



3.6 Universal Properties

Proposition 3.9 Let V be a linear space over a field $\mathbb{F}$ and S be a minimal spanning set of V. Then for any set map $\phi : S \to Z$, where Z is any linear space over $\mathbb{F}$, there is a unique linear map $\tilde{\phi} : V \to Z$ such that $\tilde{\phi}|_S = \phi$.

In other words, the following diagram commutes:



Proof. Assume the existence of such a linear map $\tilde{\phi}$. Then for all $s \in S$, we have $\tilde{\phi} \circ \iota(s) = \tilde{\phi}(s) = \phi(s)$.

Since S is a minimal spanning set of V, for any $v \in V$, we have a unique way to write v as a linear combination of finitely many elements in S, i.e., $v = \sum_{i=1}^{n} \alpha_i s_i$ where $\alpha_i \in \mathbb{F}$ and $s_i \in S$ are distinct. Then we have

$$\tilde{\phi}(v) = \tilde{\phi}\left(\sum_{i=1}^{n} \alpha_i s_i\right) = \sum_{i=1}^{n} \alpha_i \tilde{\phi}(s_i) = \sum_{i=1}^{n} \alpha_i \phi(s_i) = \phi\left(\sum_{i=1}^{n} \alpha_i s_i\right) = \phi(v)$$

This shows that $\tilde{\phi}$ agrees with ϕ on all of V, and thus $\tilde{\phi}$ is uniquely determined by ϕ. ■

Note that we first define the map on the spanning set and then extend it to the whole space. The uniqueness is due to the fact that there is only one way to write each element in V as a linear combination of elements in S and the existence is due to the fact that we can always define the map on V by using the linear combination.

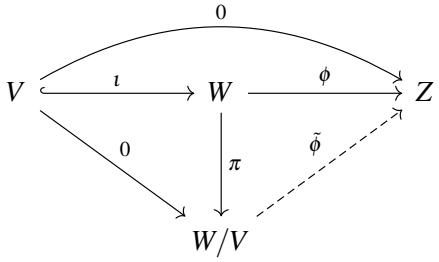
This proposition shows thats a linear space with a minimal spanning set has the following universal property: any set map from the minimal spanning set to another linear space can be uniquely extended to a linear map from the whole space to that linear space.

$$\phi \longmapsto \tilde{\phi}$$

$$\mathrm{Map}(S, Z) \cong \mathrm{Hom}(V, Z)$$

$$\tilde{\phi} \circ \iota \longmapsto \tilde{\phi}$$

Proposition 3.10 Let W be a linear space over a field $\mathbb{F}$ and V be a subspace of W. Then we have the following commutative diagram:



where Z is any linear space over $\mathbb{F}$ and $\phi : W \to Z$ is any linear map such that $\phi(v) = 0_Z$ for all $v \in V$. Then there is a unique linear map $\tilde{\phi} : W/V \to Z$ such that $\tilde{\phi} \circ \pi = \phi$.

Proof. Assume the existence of such a linear map $\tilde{\phi}$. Then for all $w \in W$, we have $\tilde{\phi}(\overline{w}) = \phi(w)$. However, this may not be well-defined. Then, we check whether it is well-defined. Assume that $\overline{w} = \overline{w'}$, then we have $\tilde{\phi}(\overline{w'}) = \phi(w')$. Note that $w - w' \in V$. Thus, we have $\phi(w' - w) = 0_Z$. This means that $\phi(w') - \phi(w) = 0_Z$, i.e., $\phi(w') = \phi(w)$. This shows that $\tilde{\phi}(\overline{w'}) = \tilde{\phi}(\overline{w})$. Thus, $\tilde{\phi}$ is well-defined.

Then we consider the linearity of $\tilde{\phi}$. For all $\overline{w_1}, \overline{w_2} \in W/V$ and $\alpha_1, \alpha_2 \in \mathbb{F}$, we have

$$\begin{aligned}
\tilde{\phi}(\alpha_1\overline{w_1} + \alpha_2\overline{w_2}) &= \tilde{\phi}(\overline{\alpha_1 w_1 + \alpha_2 w_2}) \\
&= \phi(\alpha_1 w_1 + \alpha_2 w_2) \\
&= \alpha_1\phi(w_1) + \alpha_2\phi(w_2) \\
&= \alpha_1\tilde{\phi}(\overline{w_1}) + \alpha_2\tilde{\phi}(\overline{w_2})
\end{aligned}$$

This shows that $\tilde{\phi}$ is linear. ■

R Note that $[0] = V$. If $v \in V$, then $[v] = v + V = \{v + v' \mid v' \in V\} = \{v'' \mid v'' \in V\} = V = [0]$. Thus, $\pi(v) = [v] = [0]$ for all $v \in V$. So the map from $V \to W/V$ is the zero map. Thus, the triangle commutes. Also, the map from v to Z is defined as the zero map, making the construction of $\tilde{\phi}$ is possible, as the key step is that $\phi(w' - w) = 0_Z$ for all $w' - w \in V$.