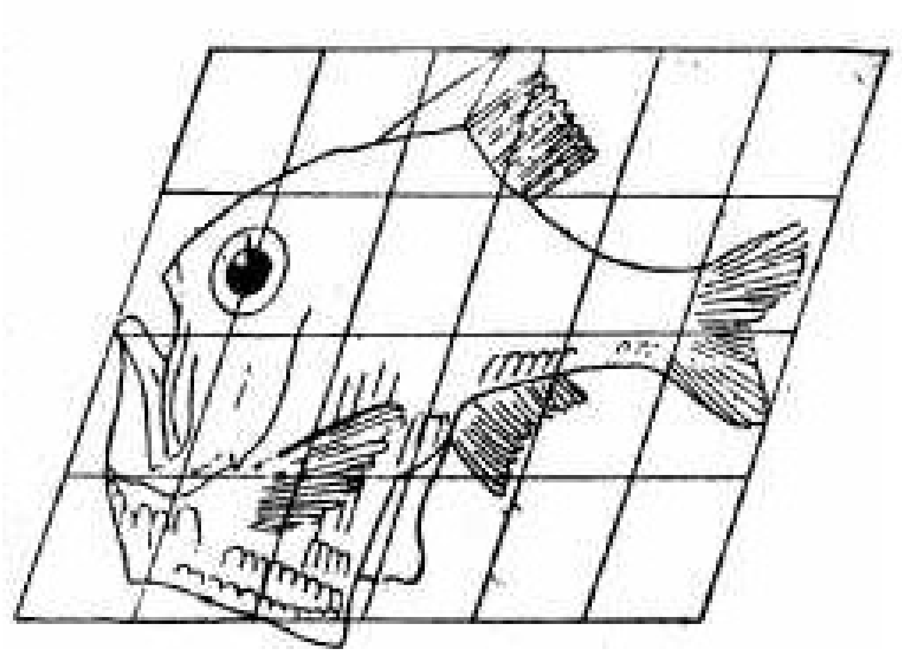


MATH 223 Linear Algebra

Wucheng Zhang

http://blogs.ubc.ca/wucheng/

[1]

[1] The cover image is from D'Arcy Thompson

This work is licensed under a Creative Commons "Attribution-NonCommercial-ShareAlike 4.0 International" license.

Contents

Sets and Maps

A. Set

1. **Definition**
   (Cantor) A *set* is a collection into a whole of definition distinct objects of our intuition or our thought, which called elements of the set.
   Note. According to the definition, if we know the elements, we know the set.

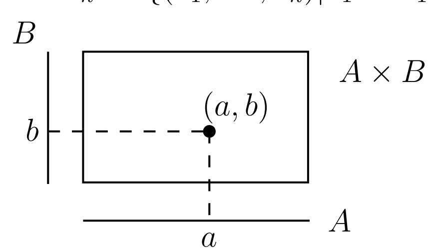
2. **Examples**

$\mathbb{N}$	natural numbers	$\mathbb{N}_0$	non-negative integers
$\mathbb{Z}$	integers	$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers	$\emptyset$	empty set

3. **Notation**
   (a) $x \in M$: x is an element of M.
   (b) $y \notin M$: y is not an element of M.
   (c) $\{\,\}$: used to specify an set by listing its elements.
   (d) $\{x|\text{the properties of } x\}$ or $\{x \in X|\text{the properties of } x\}$

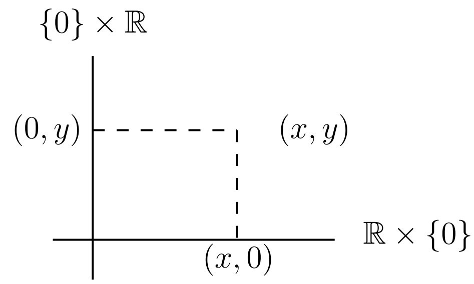
4. **Subset** A is subset of B, written $A \subset B$, if every element of A is an element of B, i.e. $x \in A \implies x \in B$.
   e.g. $\emptyset \subset \mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$

5. **Operations**
   (a) $A \cap B$: $x \in A \cap B \iff x \in A$ and $x \in B$
   (b) $A \cup B$: $x \in A \cup B \iff x \in A$ or $x \in B$
   (c) $A \setminus B$: $A \setminus B = \{x \in A|x \notin B\}$
   (d) **Cartesian Product**
      i. **Pair and Tuple:** An ordered *pair* (a, b) is a pair of objects in which the order is significant ($(a,b) \neq (b,a)$). A pair could also be treated a 2-tuple. A tuple is a finite ordered list (sequence) of elements. An n-tuple is a sequence (or ordered list) of n elements, where n is a non-negative integer.
      ii. **Definition:** $A \times B := \{(a,b)|a \in A$ **and** $b \in B\}$
      Note. In general: $A_1 \times \cdots \times A_n := \{(a_1,\ldots,a_n)|a_1 \in A_1,\ldots,a_n \in A_n\}$



iii. **Examples**

- $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$



- $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R}}_{n}$

B. Map

1. **Definition**
Let X, Y to be sets. A map f from X to Y is a rule which to each $x \in X$ assigns precisely one element of f.

2. **Notation** We denote the map as

$$\begin{aligned} f: \quad & X \longrightarrow Y \\ & x \mapsto f(x) \end{aligned}$$

where f is the name of the map, X is the *domain*, Y is the *codomain* (*target*) and $x \mapsto f(x)$ is the rule.

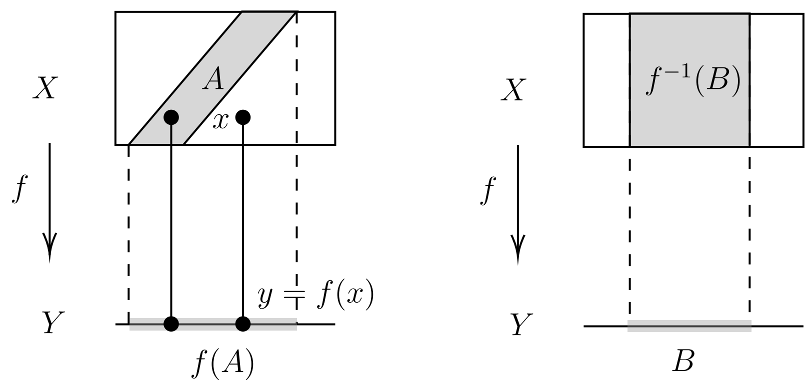
3. **Examples**

- $\begin{aligned} f: \quad & \mathbb{Z} \longrightarrow \mathbb{N}_0 \\ & n \mapsto n^2 \end{aligned}$
- $\begin{aligned} t: \quad & \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \\ & (a, b) \mapsto a + b \end{aligned}$
- **Identity:** $\begin{aligned} id_M: \quad & M \longrightarrow M \\ & x \mapsto x \end{aligned}$
- **First Projection:** $\begin{aligned} \pi_1: \quad & A \times B \longrightarrow A \\ & (a, b) \mapsto a \end{aligned}$
- **Constant Map:** $\begin{aligned} f: \quad & X \longrightarrow Y \\ & x \mapsto y_0 \end{aligned}$
- **Dirichlet Function:** $\begin{aligned} f: \quad & \mathbb{R} \longrightarrow \mathbb{R} \\ & x \mapsto \begin{cases} 1 & x \in \mathbb{R} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} \end{aligned}$

4. **Non-examples**

- $f: \begin{array}{l} \mathbb{Z} \longrightarrow \mathbb{N} \\ n \mapsto n^2 \end{array}$ is not well defined at 0.
- $t: \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ x \mapsto \log(x) \end{array}$ is not well defined for $x \leq 0$.
- $g: \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ x \mapsto y \text{ such that } y^2 = x \end{array}$ is not well defined since y is not unique.

5. **Image and Preimage**
   Let $f : X \longrightarrow Y$ be a map, $A \subset X$, $B \subset Y$ be subsets. We say $f(A) := \{f(x) | x \in A\} = y \in Y | \exists x \in A : f(x) = y$ is the *image* of A and $f^{-1}(B) := \{x | f(x) \in B\} = \{x \in X | f(x) \in B\}$ is the *preimage* of B.
   Note. We have not assigned any meaning to the symbol f^{-1} and it may not be an inverse. For example, let's check out π_1.



   It can happen that $x \notin A$ but $f(x) \in f(A)$.

6. **Injective, Surjective and Bijective**
   (a) $f : X \longrightarrow Y$ is *injective* (one-to-one) if $\forall x, x' \in X$. $f(x) = f(x')$ implies $x = x'$.
   (b) $f : X \longrightarrow Y$ is *surjective* (one-to-one) if $\forall y \in Y$, $f(x) = y$. (Equally, $f(X) = Y$.)
   (c) $f : X \longrightarrow Y$ is *bijective* if it is both injective and surjective.
   Remark 1: If $f : X \longrightarrow Y$ is bijective, there is a well-defined map

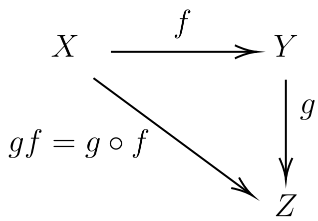
$$f^{-1} = g : \begin{array}{l} Y \longrightarrow X \\ y \mapsto x \text{ such that there is a unique } x \in X \text{ such that } f(x) = y \end{array}$$

   The rule applies to all $y \in Y$ because f is surjective and the rule is unambiguous because f is injective. This map is the inverse of f.
   Remark 2: If f is bijective and $B \subset Y$ then $f^{-1}(B) = f^{-1}(B)$ where the left side is the preimage of B under f and the left side is image of B under f^{-1}.

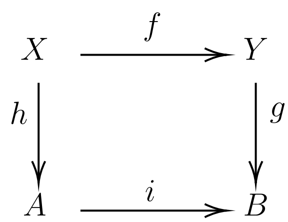
7. **Map Composition**
   Let $X \xrightarrow{f} Y \xrightarrow{g} Z$. Then we define the composition

$$g \circ f \quad \begin{array}{l} X \longrightarrow Z \\ x \mapsto g(f(x)) \end{array}$$

In this case we say the diagram commutes.



8. **Diagram Commutation**
   A diagram of sets and maps is commutative if any of two sets in the diagram, all composition of maps from one to the other equal.
   For example,



   this diagram commutes if and only if $gf = ih$.

9. **Propositions**
   (a) Let $f : X \longrightarrow Y$ be a map, A, B be subsets, then $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.
   (b) $f : X \longrightarrow Y$ be a map, A, B be subsets, then $f(A \cup B) \subset f(A) \cup f(B)$.
   (c) Let $f : X \to Y$ be a map, $A, A' \subset X$ subsets of X, and $B, B' \subset Y$ subsets of Y. then $f(A \cup A') = f(A) \cup f(A')$, and $f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B')$.

Vector Space

A. Field

1. **Definition**

   A field is a triple $(\mathbb{F}, +, \cdot)$ where

$$+ : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F} \quad \text{addition}$$
$$\cdot : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F} \quad \text{multiplication}$$

   if the following axioms are satisfied:

   - $\forall \alpha, \beta, \gamma \in \mathbb{F}, (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
   - $\forall \alpha, \beta \in \mathbb{F}, \alpha + \beta = \beta + \alpha$
   - $\exists 0 \in \mathbb{F}, \forall \alpha \in \mathbb{F}, \alpha + 0 = \alpha$
   - $\forall \alpha \in \mathbb{F}, \exists - \alpha \in \mathbb{F}, \alpha + (-\alpha) = 0$
   - $\forall \alpha, \beta, \gamma, (\alpha\beta)\gamma = \alpha(\beta\gamma)$
   - $\forall \alpha, \beta \in \mathbb{F}, \alpha\beta = \beta\alpha$
   - $\exists 1 \in \mathbb{F}, \forall \alpha \in \mathbb{F}, \alpha \cdot 1 = \alpha$
   - $\forall \alpha \in \mathbb{F}, \exists - \alpha \in \mathbb{F}, \alpha \cdot \alpha^{-1} = 1$
   - $\forall \alpha, \beta, \gamma \in \mathbb{F}, (\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$

   Note: The property can be explained by an abelian addition group (closure, associativity, identity elementary, inverse elementary, commutativity) and an abelian multiplication group and a distributive law.

   e.g.

   − $(\mathbb{R}, +, \cdot)$
   − $(\mathbb{Q}, +, \cdot)$
   − $(\mathbb{C}, +, \cdot)$
   − $(\mathbb{Z}, +, \cdot)$ is not a field, lack of multiplication inverses

2. **Facts:**

   (a) 0 is unique.
   Proof: Assume $0, 0' \in \mathbb{F}$ satisfy the axiom 3, then $0 = 0 + 0' = 0' + 0 = 0'$.

   (b) $-\alpha$ is unique.
   Proof: Assume $\gamma, \delta \in \mathbb{F}$ satisfy $\alpha+\gamma = 0$ and $\alpha+\delta = 0$, then $\gamma = \gamma+0 = \gamma+(\alpha+\delta) = (\gamma + \alpha) + \delta = (\alpha + \gamma) + \delta = 0 + \delta = \delta + 0 = \delta$.

   (c) 1 is unique.
   Proof: Assume $1, 1' \in \mathbb{F}$ satisfy the axiom 7, then $1 = 1 \cdot 1' = 1' \cdot 1 = 1'$.

   (d) α^{-1} is unique.
   Proof: Assume $\gamma, \delta \in \mathbb{F}$ satisfy $\alpha \cdot \gamma = 1$ and $\alpha \cdot \delta = 1$, then $\gamma = \gamma \cdot 1 = \gamma \cdot (\alpha \cdot \delta) = (\gamma \cdot \alpha) \cdot \delta = (\alpha \cdot \gamma) \cdot \delta = 1 \cdot \delta = \delta \cdot 1 = \delta$.

   (e) $\forall \lambda \in \mathbb{F}, 0\lambda = 0$.
   Proof: Let $\lambda \in \mathbb{F}$, we have $0 \cdot \lambda = (0+0) \cdot \lambda = 0 \cdot \lambda + 0 \cdot \lambda$, then $0 = 0 \cdot \lambda + (-(0 \cdot \lambda)) = 0 \cdot \lambda + 0 \cdot \lambda + (-(0 \cdot \lambda)) = 0 \cdot \lambda + (0 \cdot \lambda + (-(0 \cdot \lambda))) = 0 \cdot \lambda + 0 = 0 \cdot \lambda$

(f) $\forall \lambda \in \mathbb{F}$, $(-1)\lambda = -\lambda$.
(g) $(-1)(-1) = 1$.
(h) $\forall \lambda, \mu \in \mathbb{F}, (\lambda\mu)^{-1} = \lambda^{-1}\mu^{-1}$.
(i) $\forall \lambda, \mu \in \mathbb{F}, \lambda\mu = 0 \iff \lambda = 0$ or $\mu = 0$.

3. **Important case: complex number**
   (a) Definition: the field $(\mathbb{C}, +, \cdot)$ is defined as:
   - a set of $\mathbb{C} = \mathbb{R}^2$
   - addition(addition in $\mathbb{R}^2$)

   $$\begin{aligned} + :\ & \mathbb{C} \times \mathbb{C} \to \mathbb{C} \\ & (a, b) + (c, d) \mapsto (a + c, b + d) \end{aligned}$$

   - multiplication

   $$\begin{aligned} \cdot :\ & \mathbb{C} \times \mathbb{C} \to \mathbb{C} \\ & (a, b) \cdot (c, d) \mapsto (ac - bd, ad + bc) \end{aligned}$$

   [rationale: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$]
   (b) Axioms checking:
   (1) - (4) sames as $(\mathbb{R}^2, +)$
   (5) - (9) checked
   (c) Abbreviation $(1, 0) = 1$, $(0, 1) = i$

4. **Characteristic**
   (a) Definition: $\mathbb{F}$ is a field. $\forall n \in \mathbb{N}$, if $n \cdot 1 = \underbrace{1 + \cdots + 1}_{n} \neq 0$, then $\mathbb{F}$ has characteristic 0; else the smallest prime number $p \in \mathbb{N}$ such that $p \cdot 1 = \underbrace{1 + \cdots + 1}_{p} = 0$ is characteristic of $\mathbb{F}$.
   Notes:
   - 1 can be an element in $\mathbb{F}$ or $\mathbb{N}$. $1_{\mathbb{N}} \cdot 1_{\mathbb{F}} = 1_{\mathbb{F}}$
   - $-n\lambda = -(n\lambda) = -(\underbrace{\lambda + \cdots + \lambda}_{n})$

     $n\lambda = \underbrace{\lambda + \cdots + \lambda}_{n}$

   (b) Remark: if $p = \text{Char}\mathbb{F} > 0$, then p is prime.
   Proof: Assume p is not prime. (1) $p = 1$ is not possible because $p \cdot 1 = 1 \neq 0$. (2) Then let $p = p_1 p_2$ where $p_1, p_2 > 1$, then $p \cdot 1 = 0$ which is $p \cdot 1 = (p_1 p_2 \cdot 1) = (p_1 \cdot 1)(p_2 \cdot 1) = 0$ implies $p_1 \cdot 1 = 0$ or $p_2 \cdot 1 = 0$. Then p is not minimal with property $p \cdot 1 = 0$ which gives a contradiction. So p is prime.
   (c) Example:
   - $\text{Char}\mathbb{Q} = 0$
   - $\text{Char}\mathbb{R} = 0$
   - $\text{Char}\mathbb{C} = 0$
   - $\text{Char}\mathbb{F}_p = p$

B. Vector Space

1. **Definition** A triple $(V, +, \dot{)}$ where V is a set and $+ : V \times V \longrightarrow V$ and $\cdot : \mathbb{F} \times V \longrightarrow V$ which is $(\lambda, x) \mapsto \lambda x$ are maps is called vector space if
   - $\forall x, y, z \in V, (x + y) + z = x + (y + z)$
   - $\forall x, y \in V, x + y = y + x$.
   - $\exists 0 \in V$ such that $x + 0 = x$ for $\forall x$.
   - $\forall x \in V, \exists \tilde{x}$ such that $x + = 0$. (Notation: $\tilde{x} = -x$ and $x + (-y) = x - y$)
   - $\forall \lambda, \mu \in \mathbb{F}, x \in V, \lambda(\mu x) = (\lambda\mu)x$.
   - $\forall x \in V, 1x = x$.
   - $\forall \lambda \in \mathbb{F}, x, y \in V, \lambda(x + y) = \lambda x + \lambda y$.
   - $\forall \lambda, \mu \in \mathbb{F}, x \in V, (\lambda + \mu)x = \lambda x + \mu x$

   Note: (1) If $\mathbb{F} = \mathbb{R}$, we call it real vector space or vector space over $\mathbb{R}$. (2) If $\mathbb{F} = \mathbb{C}$, we call it complex vector space or vector space over $\mathbb{C}$.

2. **Remarks**
   (a) 0 is unique.
   (b) $\forall x \in V$, the $\tilde{x}$ is unique.
   (c) $\forall x \in V, 0x = 0$.
   (d) $\forall x \in V, (-1)x = -x$.
   (e) $-0 = 0$
   (f) $\lambda 0 = 0$
   (g) $\lambda x = 0 \Longleftrightarrow \lambda = 0$ or $x = 0$

3. **Examples**
   (a) $\mathbb{R}^n$
   (b) $M = \{f | f : [0, 1] \longrightarrow \mathbb{R}\}$

C. Subspace

1. **Definition**
   Let V be a vector space over $\mathbb{F}$, $U \subset V$. Then U is called a (vector) subspace of V, if:
   - $U \neq \emptyset$.
   - $\forall x, y \in U, x + y \in U$.
   - $\lambda \in \mathbb{F}, x \in U, \lambda x \in U$
2. **Remarks**
   Assume $U \subset V$ is a subspace, then:
   - $0 \in V$ is contained in U.
   - $\forall x \in U, -x \in U$.

   Proof: $\forall x \in U$, $0 = 0 \cdot x \in U$ and $-x = (-1)x \in U$.
   Corollary If U is a subspace of V, then U together with the addition and scalar multiplication inherited from $(V, +, \cdot)$ is itself a vector over $\mathbb{F}$.

3. **Facts**

(a) If U_1, U_2 are vector subspaces of V, then $U_1 \bigcap U_2$ is also a vector subspace of V.

(b) Let V be vector space over $\mathbb{F}$ and U_1, U_2 be subspace of V. Then that if $U_1 \bigcap U_2 = V$, then $U_1 = V$ or $U_2 = V$ or both.

Dimensions

A. Linear Combination

1. **Definition**
   Let $v_1, \ldots, v_r \in V$. The set

$$L(v_1, \ldots, v_r) := \{\lambda_1 v_1 + \cdots + \lambda_r v_r | v_i \in \mathbb{F}\} \subset V$$

   of all *linear combinations* of $v_1, \ldots, v_r$ is called the *linear null* of the r-tuple $(v_1, \ldots, v_r)$ of vectors. For the "0-tuple" consisting of no vectors and denoted by $\emptyset$, we write $L(\emptyset) = \{0\}$. $(\sum_{r=1} \lambda_r v_r = 0)$

2. **Remarks**
   (a) $L(v_1, \ldots v_r) \subset V$ is a subspace.
   (b) Let V be an $\mathbb{F}$-vector space.Let $v_1, \ldots, v_n \in V$. Then $L(v_1, \ldots, v_n)$ is the smallest subspace of V containing $v_1, \ldots, v_n$. More precisely, $U \subset V$ is a subspace and $v_1, \ldots, v_n \in U$, then $L(v_1, \cdots, v_n) \subset U$.

B. Linear independence and dependence

1. **Definition**
   Let V be an $\mathbb{F}$-vector space, $v_1, \ldots, v_r \in V$ is linear independent if and only if $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0$ implies $(\alpha_1, \ldots, \alpha_r) = (0, \ldots, 0)$; otherwise $(v_1, \ldots, v_r)$ is dependent.

2. **Remarks**
   (a) $(v_1, \ldots, v_r)$ is linearly independent if and only if none of the vector v_i is a linear combination of the others. Alternatively, $(v_1, \ldots, v_r)$ is linearly dependent if and only if $\exists i \in \{1, \ldots, r\}$ such that $v_i \in L(v_1, \ldots, \hat{v}_i, \ldots, v_r)$
   (b) If 0 is among the v_i or if there is a repeated vector among the v_i then $(v_1, \ldots, v_r)$ is linearly dependent.

C. Basis

1. **Definition**
   $(v_1, \cdots, v_r)$ is a *basis* of V if:
   - $(v_1, \ldots, v_r)$ is linear dependent
   - $V = L(v_1, \ldots, v_r)$

   Note: *Canonical basis* $(e_1, \ldots, e_n)$ of $\mathbb{F}^n$ where $e_1 = (1, 0, \ldots, 0)$, $e_1 = (0, 1, \ldots, 0)$,..., $e_1 = (0, 0, \ldots, 1)$.

2. **Remark**
If $(v_1, \ldots, v_n)$ is a basis of V, then $\forall v \in V$, there exists exactly one $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ with $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

Corollary Hence

$$\mathbb{F}^r \to V$$
$$(\alpha_1, \ldots, \alpha_n) \mapsto \sum_{i=1}^{r} \alpha_i v_i$$

is bijective. In another view, $(v_1, \ldots, v_r)$ as basis determines V.

3. **Basis Extension Theorem**
Suppose V is a vector space over $\mathbb{F}$. Assume $(v_1, \ldots, v_r)$ is a linearly independent family of vectors. And that $(v_1, \ldots, v_r, w_1, \ldots, w_s)$. Then by suitably choosing vectors from $(w_1, \ldots, w_s)$, one can extend $(v_1, \cdots, v_r)$ to a basis of V.

Corollary

- $(r = 0)$ Given $V = L(w_1, \ldots, w_s)$ then suitably choosing elements from $w_1, \cdots, w_s$ to form a basis of V since $L(\emptyset) = \{0\}$ counts as linearly independent.
- Every finite dimensional vector space has a basis.

4. **Exchange Lemma**
V is a vector space over $\mathbb{F}$. If $(v_1, \ldots, v_r)$ is linear independent and $(w_1, \ldots, w_s)$ spans V. Then $\forall k \in \{1, \ldots, r\}$, $\exists l \in \{1, \ldots, s\}$ such that $(v_1, \ldots, \hat{v_k}, \ldots, v_k, w_l)$ is independent.
Alternative Version: If $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$ are basis of V, then for each v_i there exists some w_j so that on replacing v_i by w_j in $(v_1, \ldots, v_n)$ we can still have a basis.

Corollary

- If $(v_1, \ldots, v_r)$ is linear independent and $(w_1, \ldots, w_s)$ spans V, then $r \leq s$
- If $(v_1, \ldots, v_r)$ and $(w_1, \ldots, w_s)$ are basis of V, then $r = s$.

D. Dimension

1. **Definition**
If the the vector space V over $\mathbb{F}$ has a basis $(v_1, \ldots, v_n)$, then n is called the *dimension* of V, which is $\dim V := \text{Length}(v_1, \ldots, v_n)$.
Note: $\dim V$ is the maximum number of linearly independent vectors in V and the minimum number of spanning vectors for V.
Note: $\dim \mathbb{F}^n = n$

2. **Finite Dimension And Infinite Dimension**
   (a) The $\mathbb{F}$ vector space is called finite-dimensional if there exists a family $v_1, \ldots, v_n \in V$, $n \in \mathbb{N}_0$ such that $V = L(v_1, \ldots, v_n)$.

(b) If V possesses no basis $(v_1, \ldots, v_n)$ for $0 \leq n < \infty$, then V is called an finite-dimensional vector space and we write $\dim V = \infty$

3. **Remarks**
   (a) Let $v_1, \ldots, v_r$ be vectors in V and $r > \dim V$. Then $(v_1, \ldots, v_r)$ is linear dependent.
   (b) If V is finite-dimensional and $U \subset V$ is a subspace, then U is also finite-dimensional.
   (c) If U is a subspace of finite-dimensional vector space V, then $\dim U < \dim V$ is equivalent to $U \neq V$.

   Note: By remark (b) and (c), if $U \subset V$, then a basis $(v_1, \ldots, v_r)$ of U can always be extended to a basis of V: just applying the basis extension theorem to $(v_1, \ldots, v_r, w_1, \ldots, w_n)$ where $(w_1, \ldots, w_n)$ is a basis of V. Here if $U \subsetneq V$, the basis $(v_1, \ldots, v_r)$ is genuinely lengthened.

4. **Dimensional formula of subspaces**
   Let U_1 and U_2 be finite-dimensional subspaces of V, then $\dim(U_1 \bigcap U_2) + \dim(U_1 + U_2) = \dim U_1 + \dim U_2$.

E. Vector Space operations

1. $U_1 \bigcap U_2 := \{u | u \in U_1 \text{ and } u \in U_2\}$.
2. $U_1 \bigcup U_2 := \{u | u \in U_1 \text{ or } u \in U_2\}$.
3. $U_1 + U_2 := \{x + y | x \in U_1, \, y \in U_2\} \in V$.
4. $U_1 \oplus U_2 = V$
   $\Longleftrightarrow U_1 + U_2 = V$ and $U_1 \bigcap U_2 = \{0\}$ - complementary subspaces
   $\Longleftrightarrow \dim(U_1 + U_2) = \dim U_1 + \dim U_2$
   $\Longleftrightarrow \forall v \in V$ can be written uniquely as $v = u_1 + u_2$ where $u_1 \in U_1$, $u_2 \in U_2$

Linear Maps

A. Definitions

1. **Linear Map**
   Let V and W be vector spaces over $\mathbb{F}$. A map $f : V \longrightarrow W$ is called *linear* or *homomorphism* if for all $x, y \in V$, $\lambda \in \mathbb{F}$, we have:
   - $f(x + y) = f(x) + f(y)$
   - $f(\lambda x) = \lambda f(x)$

2. **Kernel** $f : V \longrightarrow W$
   $\text{Ker} f := \{v \in V | f(v) = 0\}$

3. **Image** $f : V \longrightarrow W$
   $\text{Im} f := f(V)$

4. **morphisms** $f : V \longrightarrow W$
   (a) monomorphism: injective.
   (b) epimorphism: surjective.
   (c) isomorphism: bijective.
   (d) endomorphism: $V = W$.
   (e) automorphism: bijective and $V = W$.

B. Facts And Remarks

1. **Linear Map**
   (a) Id_V is a linear map.
   (b) If $V \xrightarrow{f} W \xrightarrow{g} Y$ are linear maps, then $gf : V \longrightarrow Y$ is also a linear map.
   (c) If $f : V \longrightarrow W$ and $g : V \longrightarrow W$ are linear map and $\lambda \in \mathbb{F}$, then λf and $f + g$ are linear maps.
   Corollary $\text{Hom}(V, W)$ is vector space over $\mathbb{F}$.
   (d) If f is a linear map, then $f(0) = 0$.

2. **Linear Map Related to Kernel**
   $f : V \longrightarrow W$ is linear then f is injective if and only if $\text{Ker}(f) := \{0\}$.

3. **Linear Map Related to Isomorphism**
   If $f : V \longrightarrow W$ is an isomorphism, then $f^{-1} : W \longrightarrow V$ is also an isomorphism.
   Corollary The importance of isomorphism: $\varphi : V \longrightarrow W$

- Isomorphism φ applies some structure (i.e. linear properties) of V to W. Here linear property means the formulated interns of vector space's set, addition and scalar multiplication. For example, $\dim(U_1 \bigcap U_2) = \dim(\varphi(U_1) \bigcap \varphi(U_2))$. However, non-linear property means the property is not linear. For example, $x \in V = \mathbb{R}^2$ which is a pair, but $\varphi(U)$ need not to be a circle if $U \subset \mathbb{R}^2$ is a circle.
- Isomorphism can relate linear map with one another. In the diagram below,

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle\varphi_1}\downarrow & & \downarrow{\scriptstyle\varphi_2} \\ V' & \xrightarrow{f'} & W' \end{array}$$

  φ_1 and φ_2 are isomorphism, then a complicated f and be simplified to $f' = \varphi_2^{-1} f \varphi_1$.

4. **Linear Map Related to Basis and Dimension**

   (a) (Universal Mapping Properties) V, W are vector spaces over $\mathbb{F}$ and $(v_1, \ldots, v_n)$ be a basis of V. $\forall (w_1, \ldots, w_n) \in W$, there exists a unique linear map $f : V \longrightarrow W$ such that $f(v_i) = w_i$ where $i = 1, \ldots, n$.

   (b) Let V and W be vector spaces over $\mathbb{F}$ and let $(v_1, \ldots, v_n)$ be a basis of V. A linear map $f : V \longrightarrow W$ is an isomorphism if and only if $(f(v_1), \ldots, f(v_n))$ is a basis of W.

   (c) Any two n-dimensional vector spaces over $\mathbb{F}$ are isomorphic (There exists a linear map f such that $f : V \cong W$).

   (d) (Dimension Formula for Linear Maps) V is a finite-dimensional vector space and $f : V \longrightarrow W$ is a linear map. Then

$$\dim \mathrm{Ker} f + \dim \mathrm{Im} f = \dim V$$

   (e) A linear map between two spaces of the same dimension is surjective if and only if it is injective.

   (f) Sylvester Inequality: $\mathrm{rk}A + \mathrm{rk}B - n \leq \mathrm{rk}AB \leq \min\{\mathrm{rk}A, \mathrm{rk}B\}$.

Matrix

A. Definitions

1. **Matrix**
   An $m \times n$ matrix $(m, n \in \mathbb{N})$ with $m \times n$ entries in the field $\mathbb{F}$ is a map

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

$$(i, j) \longmapsto a_{ij}$$

   where a_{ij} is the (i, j) entry of A where i is the row index and j is the column index. We think A as an array

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix}$$

   We always think of elements of $\mathbb{F}^m$ as columns $\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{F}^m$.

   Then we can write $A = (A_1, \ldots, A_n)$ where $A_j \in \mathbb{F}^m$ is the j^{th} column. Here $(A_j)_i = a_{ij}$ - the i^{th} entry of the j^{th} column. And also we can write $A = (a_{ij})$.
   Note: $M(m \times n, \mathbb{F})$, the set of all $m \times n$ matrices with entries in $\mathbb{F}$, is an $\mathbb{F}$-vector space where $+$ and $\cdot$ are entry wise. And $\dim M(m \times n, \mathbb{F}) = mn$.

2. **Matrix-Vector Product**
   (a) **Special case**
   Let $f : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ be a linear. f is completely determined by $f(e_1), \ldots, f(e_n) \in \mathbb{F}^n$. $[f] = (f(e_1), \ldots, f(e_n)) \in M(m \times n, \mathbb{F})$ is a matrix of f. Let $A = [f]$ and $A_j = f(e_j)$, then we would have an isomorphism by universal mapping properties.

$$\text{Hom}(\mathbb{F}^n, \mathbb{F}^m) \longrightarrow M(m \times n, \mathbb{F})$$

$$f \mapsto [f]$$

$$L_A \leftarrow A$$

   Then L_A is the unique linear map such that $L_A(e_j) = A_j$.
   We can rewrite this into a **theorem**: Let $A \in M(m \times n, \mathbb{F})$. Then the map

$$\mathbb{F}^n \longrightarrow \mathbb{F}^m$$

$$x \longmapsto Ax$$

   is linear. And conversely, if $\mathbb{F}^n \longrightarrow \mathbb{F}^m$ is a linear map, there exists a unique matrix $A \in M(m \times n, \mathbb{F})$ with $f(x) = Ax$ for $\forall x \in \mathbb{F}^n$.

Then we can give the *definition* of the matrix-vector product. Let $A \in M(m \times n, \mathbb{F}), x \in \mathbb{F}^n$, we define the matrix-vector product by $Ax := L_A(x) \in \mathbb{F}^m$.

$$M(m \times n, \mathbb{F}) \times \mathbb{F}^n \longrightarrow \mathbb{F}^m$$

$$(A, x) \longmapsto Ax$$

Let $A = (A_1, \ldots, A_n)$. In terms of column view, it is the linear combination of the columns of A with coefficient given by x:

$$Ax = L_A(x) = L_A(\sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j L_A(e_j) = \sum_{j=1}^{n} x_j A_j$$

In terms of entry view, it would be

$$(Ax)_i = (\sum_{j=1}^{n} x_j A_j)_i = \sum_{j=1}^{n} x_j (A_j)_i = \sum_{j=1}^{n} x_j a_{ij} = \sum_{j=1}^{n} a_{ij} x_j$$

(b) **General Case**
Consider the following diagram:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \Phi_{\mathcal{B}} \uparrow & & \Phi_{\mathcal{C}} \uparrow \\ \mathbb{F}^n & \xrightarrow{[f]_{\mathcal{C}}^{\mathcal{B}}} & \mathbb{F}^m \end{array}$$

where $f : V \longrightarrow W$ is a linear map, $\mathcal{B} = (v_1, \ldots, v_n)$ is the basis of V, $\mathcal{C} = (w_1, \ldots, w_m)$ is the basis of W and Φ is the *canonical basis isomorphism*:

$$K^n \xrightarrow{\cong} V$$

$$(\lambda_1, \ldots, \lambda_n) \mapsto \lambda_1 v_1 + \cdots + \lambda_n v_n$$

respect to the the basis $(v_1, \ldots, v_n)$. Then we have a unique $m \times n$-matrix making diagram commute

$$[f]_{\mathcal{C}}^{\mathcal{B}} = \Phi_{\mathcal{C}} f \Phi_{\mathcal{B}}^{-1}$$

And in column view, the j^{th} column of $[f]_{\mathcal{C}}^{\mathcal{B}}$ is the coordinate vector of $f(v_j)$ with respect to basis $\mathcal{C}$.

$$[f]_{\mathcal{C}}^{\mathcal{B}}(e_j) = [f(v_j)]_{\mathcal{C}}$$

(c) **Change of basis**
Consider a endomorphism $f : V \to V$ with dimension, with canonical basis, then we want to transform it into a basis $\mathcal{B}$. Consider the diagram

$$\begin{array}{ccc} \mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^n \\ \Phi_{\mathcal{B}} \uparrow & & \Phi_{\mathcal{B}} \uparrow \\ \mathbb{F}^n & \xrightarrow{[A]_{\mathcal{B}}^{\mathcal{B}}} & \mathbb{F}^n \end{array}$$

then we have

$$[A]_{\mathcal{B}}^{\mathcal{B}} = \Phi_{\mathcal{B}} A \Phi_{\mathcal{B}}^{-1}$$

B. Matrix Multiplication

1. **Definition**

   Let $A \in M(m \times n, \mathbb{F})$ and $B \in M(n \times p, \mathbb{F})$, then $AB \in M(m \times p, \mathbb{F})$ is defined to the unique matrix such that

$$\begin{array}{ccc} \mathbb{F}^p & \xrightarrow{B} & \mathbb{F}^n \\ & \searrow^{AB} & \downarrow A \\ & & \mathbb{F}^m \end{array}$$

   commutes. This also means $L_{AB} = L_A \circ L_B$ and $(AB)x = A(Bx)$. (Also it is $[fg]^{\mathcal{A}}_{\mathcal{C}} = [f]^{\mathcal{B}}_{\mathcal{C}}\,[g]^{\mathcal{A}}_{\mathcal{B}}$.) Therefore AB is only defined if the number of rows of B equal to the number of the columns of A. Then we can find the k^{th} column of AB is

$$(AB)_k = AB(e_k) = A(B_k) = \sum_{j=1}^{n} (B_k)_j A_j = \sum_{j=1}^{n} b_{jk} A_j$$

   and the l^{th} entry of the j^{th} column is

$$(AB)_{lk} = (\sum_{j=1}^{n} b_{jk} A_j)_l = \sum_{j=1}^{n} b_{jk} (A_j)_l = \sum_{j=1}^{n} b_{jk} a_{lj}$$

   Note: Another way to find basis transformation, $[A]^{\mathcal{S}}_{\mathcal{S}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{S}}\,[A]^{\mathcal{B}}_{\mathcal{B}}\,[\mathrm{id}]^{\mathcal{S}}_{\mathcal{B}}$

2. **Properties**

   (a) (Non-commutative) $AB \neq BA$.

   (b) (Associative) $A(BC) = (AB)C$.

   (c) (Distributive) $A(B+C) = AB + AC$, $(A+B)C = AC + BC$.

C. Rank

1. **Definition**

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \Phi_{\mathcal{B}} \uparrow & & \Phi_{\mathcal{C}} \uparrow \\ \mathbb{F}^n & \xrightarrow{A=[f]^{\mathcal{B}}_{\mathcal{C}}} & \mathbb{F}^m \end{array}$$

   (a) For linear map f, the *rank* of f is $\mathrm{rk} f = \dim \mathrm{Im} f$.

   (b) For the corresponding matrix A, the *rank* of A is $\mathrm{rk} A = \dim \mathrm{Im} A$.
   The *column rank* of A, col $\mathrm{rk}A$, is the maximum number of linearly independent columns of A and the *row rank* of A, row $\mathrm{rk}A$, is the maximum number of linearly independent rows of A.

   Note: By using the dimension formula of linear map, we can get $\dim \mathrm{Ker} f + \mathrm{rk} f = n$ or $\dim \mathrm{Ker} A + \mathrm{rk} A = n$.

2. **Proposition**

   row $\mathrm{rk}A$ = col $\mathrm{rk}A$ = $\mathrm{rk}A$

D. Row Echelon Form

1. **Definition**
   A matrix $A \in M(m \times n, \mathbb{F})$ is *row echelon form* (REF) if:
   - Every zero row is below every non-zero row.
   - The first non-zero entry of every non-zero row is to the right of the first non-zero entry of every row above.

   Note: The leading entry of the non-zero row is called *pivot*. The column with a pivot is called *pivot column* and the row with a pivot is called *pivot row*.

2. **Proposition**
   Let A be in REF then:
   - the rows of A are linearly independent,
   - the pivot columns of A are linearly independent.

   Corollary If A is in REF, then:
   - The pivot rows of A form a basis of row space of A.
   - The pivot columns of A are a basis of column space of A.

E. Elementary Row Operation

1. **Definition And the Corresponding Matrices**
   (a) **R1** Swap row i and row j of the matrix $A \in M(m \times n, \mathbb{F})$. And the corresponding matrix $T_{ij} \in M(m \times m, \mathbb{F})$ where $T_{ij}A$ is the required matrix would be

$$T_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

   (b) **R2** Multiply row i of the matrix $A \in M(m \times n, \mathbb{F})$ by a non zero scalar λ. And the corresponding matrix $D_i(\lambda) \in M(m \times m, \mathbb{F})$ where $D_i(\lambda)A$ is the required matrix would be

$$D_i(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

(c) **R3** Replace row A^i f the matrix $A \in M(m \times n, \mathbb{F})$ by $A^i + \lambda A^j$ for some $i \neq j$ and $\lambda \in \mathbb{F}$. And the corresponding matrix $L_{ij}(\lambda) \in M(m \times m, \mathbb{F})$ where $L_{ij}(\lambda)A$ is the required matrix would be

$$L_{ij}(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \lambda & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Note: $T_{ij}, D_i(\lambda), L_{ij}(\lambda)$ are all invertible.

2. **Proposition**
Every matrix can be converted to a matrix in REF by applying a sequence of elementary row operations.

3. **Row equivalent**
   (a) **Definition**
   A, A' are row equivalent if they can be obtained from each other by a sequences of elementary row operations.
   (b) **Remarks**
   If A, A' are row equivalent:
   - Row space of A and row space of A' are the same.
   - $\text{rk}A = \text{rk}A'$.
   - Null space of A and null space of A' are the same.
   - If A' is in REF, then the columns of A corresponding to the pivot column of A' forms a basis of column space of A. (Since the pivot columns of A' form a max independent subfamily of columns of A'.)

   Corollary (Algorithm) Put matrix A into REF, we can:
   - find a basis of row space of A by taking the pivot rows,
   - find $\text{rk}A$,
   - find a basis for the column space of A.

F. Inverse Matrix

1. **Definition**
The inverse matrix A^{-1} of A is the matrix of inverse map associated to A.
Note: Obviously the map associated to A should be isomorphism.

2. **Remarks**

(a) A matrix is invertible only if it is a square matrix.

(b) A matrix A is invertible if and only if the REF of A has a pivot in every row and every column.

(c) If A is invertible, $(A^{-1})^{-1} = A$.

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

(e) $AB = I_n \iff BA = I_n \iff B = A^{-1}$

3. **Algorithm**

If A is invertible, by applying row elementary operations, we have $E_k \cdots E_1 A = I_n$. Then we have $A^{-1} = E_k \cdots E_1 I_n$ since $E_k \cdots E_1 A A^{-1} = I_n A^{-1}$.

Determinant

A. Definition

1. **Theorem**
   There exits a unique map
   $$\det : M(n \times n, \mathbb{F}) \longrightarrow \mathbb{F}$$
   with the following properties:
   - det is linear in each row,
   - If the (row) rank is smaller than n, then $\det A = 0$.
   - $\det I_n = 1$, where I_n is an $n \times n$ identity matrix.

   Note:
   - For the first property, it means, for example, in the row k with the following fix rows $A^1, \ldots, A^{k-1}, A^{k+1}, \ldots, A^n$, the function $\mathbb{F}^n \longrightarrow \mathbb{F}$, given by
     $$x = (x_1, \ldots, x_n) \mapsto \det \begin{pmatrix} A^1 \\ \vdots \\ A^{k-1} \\ x \\ A^{k+1} \\ \vdots \\ A^n \end{pmatrix}$$
     is linear.
   - For the second property, $\mathrm{rk} A < n \Longleftrightarrow \det A = 0$
2. **Proof of the Theorem**
   Lemma:
   - Exchange Two rows: $\det A' = -\det A$
   - Scale one row by λ: $\det A' = \lambda \det A$
   - Add one row by another times λ: $\det A' = \det A$
3. **Definition**
   The map $\det : M(m \times n, \mathbb{F}) \longrightarrow \mathbb{F}$ is called the *determinant* and the number $\det A \in \mathbb{F}$ is called the determinant of A.

B. Determinantal Formula for the Inverse Matrix

1. **Theorem**
   If $A \in M(n \times n, \mathbb{F})$ is invertible, then
   $$A^{-1} = \frac{1}{\det A} \tilde{A}$$
   where $\tilde{A} \in M(n \times n, \mathbb{F})$ is the adjugate matrix of A defined by $\tilde{a}_{ij} := (-1)^{i+j} \det A_{ji}$.

2. **Application on 2×2 Matrix**

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

C. Determinant of the Transposed Matrix

1. **Definition**
   Let $A = (a_{ij}) \in M(n \times n, \mathbb{F})$, then the *transpose matrix* of A is

$$A^t = (a^t_{ji}) \in M(n \times n, \mathbb{F})$$

2. **Remark**
   Let $A \in M(m \times n, \mathbb{F})$ and $B \in M(n \times p, \mathbb{F})$,

$$(AB)^t = B^t A^t$$

3. **Theorem**
   For any square matrix A, $\det A = \det A^t$.

   Corollary Row (Laplace) Expansion

D. Determinant of the Matrix Product

1. **Theorem** Let $A, B \in M(n \times n, \mathbb{F})$, we have

$$\det AB = \det A \cdot \det B$$

2. **Corollary**
   (a) If A is invertible, then $\det A^{-1} = \dfrac{1}{\det A}$.
   (b) We can define $\det F$ for any endomorphism

$$F : V \longrightarrow V$$

   where V is a finite dimensional vector space over $\mathbb{F}$, with the following properties:
   - $\det F \neq 0 \Longleftrightarrow f$ is an isomorphism.
   - $\det gf = \det f \det g$.
   - $\det Id_V = 1$.

System of Linear Equations

A. Basic Definitions

1. **Systems of Linear Equations**
   We write the *system of linear equations* with all coefficients $a_{ij}, b_k \in \mathbb{F}$,

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

   as $Ax = b$, where $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ as *unknowns* and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

2. **Homogeneous and Inhomogeneous**
   If $b = 0$, then this linear equation system is *homogeneous*.
   If $b \neq 0$, then this linear equation system is *inhomogeneous*. ($Ax = 0$ is the associated equation system.)

3. **Solution set**
   The *solution set* of the system of equations associated to (A, b) is defined to be $\mathrm{Sol}(A, b) := \{x \in \mathbb{F}^n | Ax = b\} = A^{-1}\{b\}$.
   Note: If $\mathrm{Sol}(A, b) \neq \emptyset$, the system of equations is solvable.

B. Algorithm

1. **Reduced Row Echelon Form** (RREF)
   A *reduced row echelon form* is a row echelon form such that:
   - Every pivot is 1.
   - Every pivot is the only non-zero entry in its column.

2. **Theorem**
   Every matrix is row equivalent to a *unique* matrix in RREF.

3. **Algorithm**
   (a) Transform augmented matrix $(A|b)$ into RREF $(\tilde{A}|\tilde{b})$.
   (b) We pick non-pivot column in the RREF $(\tilde{A}|\tilde{b})$ as free variables.
   (c) Write out the vector form of the solution:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + x_{j_1} \begin{pmatrix} d_{11} \\ \vdots \\ d_{1n} \end{pmatrix} + \cdots + x_{j_k} \begin{pmatrix} d_{k1} \\ \vdots \\ d_{kn} \end{pmatrix}$$

C. The Criterion for Solvablity and The Structure of the Solution Set

1. **Homogeneous System**
   We have $Ax = 0$, then $\text{Sol}(A, 0) = A^{-1}\{0\} = \text{Ker}A$. To find $\text{Ker}A$, is to find the basis of $\text{Ker}A$. We can first consider the dimension of $\text{Ker}A$. $\dim \text{Ker}A = \dim \text{Ker}\tilde{A}_{RREF} = n - \text{rk}\tilde{A} = n - \#$ of pivots $= \#$ of free variables. We can see that the i^{th} basis vector of $\text{Ker}A$ is obtained by setting the i^{th} free variable to 1 and others to 0.

   In other word, the general solution is the sum of the free variables multiplying the parameter vector:
   $$\vec{x} = x_{j_1}\vec{v_1} + \cdots + x_{j_k}\vec{v_k}$$
   where $j_1, \ldots, j_k$ are the non-pivot column and $\vec{v_1}, \ldots, \vec{v_k}$ are the basis vectors of $\text{Ker}A$.

2. **Inhomogeneous System**
   We have $Ax = b$ where $b \neq 0$. Then we transform augmented matrix $(A|b)$ into RREF $(\tilde{A}|\tilde{b})$. Since $Ax = b$ is equivalent to $EAx = Eb$ which is $\tilde{A}x = \tilde{b}$, where E is elementary row operation, $\text{Sol}(A, b) = \text{Sol}(\tilde{A}, \tilde{b})$.

   If the system $Ax = b$ is consistent, which means there is no pivot in augmented column of $(\tilde{A}|\tilde{b})$, then $\text{Sol}(A, b) \neq \emptyset$. If the system $Ax = b$ is inconsistent, which means there is pivot in augmented column of $(\tilde{A}|\tilde{b})$, then $\text{rk}(A|b) > \text{rk}A$ and $\text{Sol}(A, b) = \emptyset$.

3. **Criterion**
   (a) $Ax = b$ is *solvable* if and only if $\text{rk}A = \text{rk}(A|b)$.
   (b) $Ax = b$ is *uniquely solvable* if and only if $\text{Ker}A = 0$, which is $\text{rk}A = n$.
   (c) If A is a square matrix, $Ax = b$ is *uniquely solvable* if and only if $\det A \neq 0$.

4. **Structure of the Solution**
   If x_0 is a solution of $Ax = b$, that is $Ax_0 = b$, then $\text{Sol}(A, b) = x_0 + \text{Ker}A := \{x + x_0 | x \in \text{Ker}A\}$.

Euclidean Vector Space

A. Definition

1. **Inner Product**
   Let V be a real vector space. An inner product on V is
   $$\begin{aligned} \langle , \rangle : \quad & V \times V \to \mathbb{R} \\ & (x, y) \mapsto \langle x, y \rangle \end{aligned}$$
   with the following properties:
   - bilinear, i.e. if $x \in V$ then
     $$\begin{aligned} \langle x, \cdot \rangle : \quad & V \to \mathbb{R} \\ & v \mapsto \langle x, v \rangle \end{aligned}$$
     and
     $$\begin{aligned} \langle \cdot, y \rangle : \quad & V \to \mathbb{R} \\ & v \mapsto \langle v, y \rangle \end{aligned}$$
     are linear.
   - symmetric, i.e. $\forall x, y \in V, \langle x, y \rangle = \langle y, x \rangle$.
   - positive definite, i.e. $\forall x \neq 0$, $\langle x, x \rangle > 0$.

2. **Examples**
   (a) Let $A \in M(n \times n, \mathbb{R})$ and $V = \mathbb{R}^n$, $\langle x, y \rangle_A = x^t A y$ where A is symmetric and positive definite.
   (b) (Standard Inner Product) Let $V = \mathbb{R}^n$, $\langle x, y \rangle = x^t y$ $(= \langle x, y \rangle_{I_n})$.
   (c) Let $V = \text{cont Map}([0,1], \mathbb{R})$ which a real vector space, $\langle f, g \rangle = \displaystyle\int_0^1 f(t)g(t)\, dt$.

3. **Euclidean Vector Space**
   A *euclidean vector space* is a pair $(V, \langle , \rangle)$, where V is a real vector space and $\langle , \rangle$ is an inner product on V.

B. Norm And Angle

1. **Norm**
   Let $(V, \langle , \rangle)$ be a Euclidean vector space and $x \in V$, then $||x|| = \sqrt{\langle x, x \rangle}$ is the *norm* (length) of x.

2. **Theorem**
   (Cauchy-Schwartz Inequality) Let $(V, \langle , \rangle)$ be a Euclidean vector space, then $\forall x, y \in V$, $|\langle x, y \rangle| \leq ||x|| \cdot ||y||$.

3. **Remarks**
   Let $(V, \langle , \rangle)$ be a Euclidean vector space, then:

(a) $\forall x \in V, ||x|| \geq 0$.
(b) $\forall x \in V, ||x|| = 0 \Longleftrightarrow x = 0$.
(c) $\forall x \in V, \lambda \in \mathbb{R}, ||\lambda x|| = |\lambda| \cdot ||x||$.
(d) $\forall x, y \in V$, $||x + y|| \leq ||x|| + ||y||$.
(e) $||a + b||^2 - ||a - b||^2 = 4\langle a, b\rangle$. Then we have $\langle a, b\rangle = 0 \Longleftrightarrow ||a + b|| = ||a - b|| \Longleftrightarrow \overrightarrow{Oa} \perp \overrightarrow{Ob} \Longleftrightarrow$ The parallelogram is a rectangle.

4. **Angle**
Let $(V, \langle,\rangle)$ be a Euclidean vector space, $x, y \in V$, $x, y \neq 0$, then there exits a unique $\alpha \in [0, \pi]$ such that

$$\cos\alpha = \frac{\langle x, y\rangle}{||x||\,||y||}$$

because by Cauchy-Schwartz inequality $-1 \leq \dfrac{\langle x, y\rangle}{||x||\,||y||} \leq 1$ and $\cos : [0, \pi] \to [-1, 1]$ is bijective. The the α is called the angle between x and y.

Note: With the definition of length and angle, we can study Euclidean geometry on vector space.

C. Orthogonal Vector

1. **Orthogonality**
Two elements v, w of a Euclidean vector space are said to be *orthogonal* or *perpendicular* to each other (written $v \perp w$), if $\langle v, w\rangle = 0$.

2. **Orthogonal Complement**
Let $(V, \langle,\rangle)$ be a Euclidean vector space, $M \subset V$, then the orthogonal complement of M is $M^\perp := \{v \in V | \forall x \in M, \langle x, v\rangle\} = \bigcap_{x \in M} \mathrm{Ker}\langle x, \cdot\rangle$.

3. **Remarks**
(a) $M^\perp$ is a subspace of V.
(b) If $M = L(v_1, \ldots, v_n)$, then $M^\perp = \{v_1, \ldots, v_r\}^\perp$

4. **orthonormal System**
$(v_1, \ldots, v_n)$ is an *orthonormal system* (family) if:
- $\forall i = 1, \ldots, r$, $||v_i|| = 1$.
- $\forall 1 \leq i < j \leq r$, $v_i \perp v_j$. (Alternatively, $\forall i, j \in \{1, \ldots, r\}$, $\langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$.)

5. **Lemmas**
(a) Every orthonormal system is linear independent.

(b) Let $(v_1, \ldots, v_n)$ be an orthonormal basis for V. Then $\forall x \in V$,

$$x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$$

D. Orthonormalization

1. **Lemma**
(QR-Factorization) If $v_1, \ldots, v_r \in V$ is an orthonormal basis and $U = L(v_1, \ldots, v_r)$. Then $V = U \oplus U^{\perp}$, i.e. $\forall v \in V$, $\exists! u \in U, w \in U^{\perp}$ such that $v = u + w$. We denote $w = \mathrm{Proj}_U(v)$, the projection of v onto u. Moreover $u = \sum_{i=1}^{r} \langle v, v_i \rangle v_i$.

2. **Gram-Schmidt Orthonormalization Process**
If $(v_1, \ldots, v_n)$ is linearly independent, $U_k = L(v_1, \ldots, v_k)$ where $\dim U_k = k \in \{1, \ldots, n\}$ and $\{0\} = U_0 \subset U_1 \subset \cdots \subset U_n \subset V$. Then there exits a unique an orthonormal system $(\tilde{v}_1, \ldots, \tilde{v}_n)$ such that:
   - $L(\tilde{v}_1, \ldots, \tilde{v}_k) = U_k$
   - $\langle \tilde{v}_k, v_k \rangle > 0$

3. **Formula of Gram-Schmidt Orthonormalization**
If $(v_1, \ldots, v_n)$ is the basis of V, we have

$$\tilde{v}_1 = \frac{v_1}{||v_1||}$$

and by the recursion formula

$$\tilde{v}_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, \tilde{v}_i \rangle \tilde{v}_i}{\left|\left| v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, \tilde{v}_i \rangle \tilde{v}_i \right|\right|}$$

we can find the orthonormal basis $(\tilde{v}_1, \ldots, \tilde{v}_n)$ of V.

4. **Corollary**
   (a) Let $(V, \langle , \rangle)$ be a Euclidean vector space and U is a finite dimensional subspace of V. There exists a unique linear map $\mathrm{Proj}_U : V \to U$ with $\mathrm{Proj}_U | U = Id_U$ and $\mathrm{Ker}(\mathrm{Proj}_U) = U^{\perp}$. This map is called the orthogonal projection onto U.
   (b) If U is a subspace of a finite-dimensional Euclidean vector space and $U^{\perp} = 0$, then $U = V$.

E. Orthogonal Maps

1. **Definition**
Let V, V' be Euclidean vector spaces. A linear map $\varphi : V \to V'$ is *orthogonal* or *isometric* if $\forall x, y \in V, \langle x, y \rangle = \langle \varphi(x), \varphi(y) \rangle$.

2. **Remarks**

(a) Every orthogonal linear maps is injective.

(b) If $\varphi : V \to V$ is an orthogonal endomorphism and $\dim V < \infty$, then φ is an isomorphism and φ^{-1} is also orthogonal.

(c) Let V, V' to be Euclidean vector spaces and $(v_1, \ldots, v_n)$ be an orthonormal basis of V. Then a linear map: $V \to V'$ is orthogonal if and only if $(f(v_1), \ldots, f(v_n))$ is an orthonormal system in V'.

3. **Orthogonal Matrix**

Let $O(V)$ be the set of orthogonal isomorphisms on Euclidean vector space V. If $V = \mathbb{R}^n$ with the standard inner product, we use $O(n)$ to represent $O(V)$. Usually we consider them to be matrices, $O(n) \subset M(n \times n, \mathbb{R})$, called *orthogonal matrices.*

4. **Remarks**

$A \in O(n)$
$\iff Ae_1, \ldots, Ae_n$ is an orthonormal basis of $\mathbb{R}^n$
$\iff$ the columns of Aform and orthonormal basis of$\mathbb{R}^n$
$\iff A^tA = I_n$
$\iff A$ is invertible and $A^{-1} = A^t$
$\iff AA^t = I_n$
$\iff$ the rows of A form an orthonormal basis of $\mathbb{R}^n$

Note: If $A \in O(n)$, by $\det A \det A^t = \det I_n$, we have $\det A = \pm 1$. If $\det A = 1$, A is called special orthogonal matrix. We denote the set of them as $SO(n)$.

F. Group

1. **Group**

A *group* is a pair $(G, \cdot)$ where G is a set and $\cdot$ is an operation

$$\begin{aligned} \cdot : \quad & G \times G \to G \\ & (g, h) \mapsto g \cdot h = gh \end{aligned}$$

subject to:

- (Associativity) $\forall g, h, k \in G, (gh)k = g(hk)$.
- (Neutral Element) $\forall e \in G$ such that $\forall g \in G, ge = eg = g$.
- (Inverse Element) $\forall g \in G$, $\exists h \in G$ such that $gh = hg = e$. (We denote h as g^{-1}).

2. **Remarks**

(a) The neutral/identity element e is unique.

(b) The inverse element g^{-1} of g is unique.

3. **Abelian Group**

If a group $(G, \cdot)$ has the additional property of being "commutative," i.e. $\forall g, h \in G, gh = hg$, then $(G, \cdot)$ is called an abelian group.

4. **Subgroup**
   Let $(G, \cdot)$ be a group and $H \subset G$. H is a *subgroup* if:
   - $\forall g, h \in H, gh \in H$. ($H$ is closed under the group operation in G).
   - $e \in H$.
   - $\forall g \in H, g^{-1} \in H$. ($H$ is closed under inverse.)

Eigenvalues and Dynamic Systems

A. Basic Definitions

1. **Eigenvalues and Eigenvectors**
   Let V be a vector space over $\mathbb{F}$ and $f : V \to V$ be an endomorphism. If there exits a $\lambda \in \mathbb{F}$ and a vector $v \in V$ and $v \neq 0$ such that $f(v) = \lambda v$, then we call the coefficient λ *eigenvalues* and the vector v *eigenvector*.

2. **Eigenspaces and Geometric Multiplicity**
   Assume $\lambda \in \mathbb{F}$ and $v \in V, v \neq 0$ such that $f(v) = \lambda v$. This is equivalent to $(f - \lambda \mathrm{Id})v = 0$, which means $v \in \mathrm{Ker}(f - \lambda \mathrm{Id})$. By definition, it is easy to prove $\mathrm{Ker}(f - \lambda \mathrm{Id})$ is a subspace of V. Therefore, if λ is an eigenvalue of f, the subspace

   $$E_\lambda := \mathrm{Ker}(f - \lambda \mathrm{Id})$$

   is called the *eigenspace* of f for the eigenvalue λ, and its dimension $\dim \mathrm{Ker}(f - \lambda \mathrm{Id})$ is called *geometric multiplicity* of the eigenvalue.

B. Characteristic Polynomial

1. **Definition**
   By the definition of eigenvector, we hope $\mathrm{Ker}(f - \lambda \mathrm{Id}) \neq \{0\}$. This is equivalent to $\det(f - \lambda \mathrm{Id}) = 0$. Consider the commutative diagram below,

   $$\begin{array}{ccc} V & \xrightarrow{\ f\ } & V \\ {\scriptstyle \Phi_{\mathcal{B}}}\uparrow & & {\scriptstyle \Phi_{\mathcal{B}}}\uparrow \\ \mathbb{F}^n & \xrightarrow{A=[f]_{\mathcal{B}}^{\mathcal{B}}} & \mathbb{F}^n \end{array}$$

   we would have $\det(f - \lambda \mathrm{Id}) = \det(A - \lambda I_n)$. And we have

   $$\begin{aligned} \det(A - \lambda I_n) &= \det(A^1 - \lambda I_n^1, \dots, A^n - \lambda I_n^n) \\ &= \sum_{k=1}^{n} a_k (-1)^k \lambda^k \end{aligned}$$

   where $a_k = \sum_{i_1 < \cdots < i_k} \det(A^1, \dots, I_n^{i_1}, \dots, I_n^{i_k}, A^n)$. Then here comes the definition. If $f : V \to V$ is an endomorphism of an n-dimensional vector space over $\mathbb{F}$, then there exits coefficients $a_0, \dots, a_{n-1} \in \mathbb{F}$ with

   $$\det(f - \lambda \mathrm{Id}) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1 \lambda + a_0 =: P_f(\lambda)$$

   for all $\lambda \in \mathbb{F}$. (And actually $a_0 = (-1)^n \det A$.) P_f is called the *characteristic polynomial* of f. Since we want $P_f(\lambda) = \det(f - \lambda \mathrm{Id}) = 0$, the eigenvalues are the zeros of the characteristic polynomial.

Note: λ being an eigenvalue of f is independent for the choice of basis, so the roots of the characteristic polynomial are independent of the choice of basis.

2. **Polynomial And Zeros**

(a) **Definition**

Let λ be a formal symbol. An expression such that as

$$P(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$$

where $a_0, \ldots, a_n \in \mathbb{F}$ is called a *polynomial* with coefficients in $\mathbb{F}$. And if $a_n \neq 0$, the *degree* of P is n.

Let's call the set of these polynomials $P_{n,\mathbb{F}}$. Since $a_n\lambda^n + \cdots + a_1\lambda + a_0 = (a_n, \ldots, a_0)$, we can say $P_{n,\mathbb{F}} = \mathbb{F}^{n+1}$. Therefore, $P_{n,\mathbb{F}}$ is a vector space over $\mathbb{F}$ with $\dim P_{n,\mathbb{F}} = n+1$.

We can treat it as a map $P : \mathbb{F} \to \mathbb{F}$ determined by the coefficients. Let $\text{Map}(\mathbb{F}, \mathbb{F}) = \mathbb{F}^{\#\mathbb{F}}$ be the vector space over $\mathbb{F}$ of all the maps $\mathbb{F} \to \mathbb{F}$ where $\#\mathbb{F}$ is the number of elements in $\mathbb{F}$ and $\dim \text{Map}(\mathbb{F}, \mathbb{F}) = \#\mathbb{F}$. Then we can find a map $\mathscr{P}$ such that

$$\mathscr{P} : P_{n,\mathbb{F}} \longrightarrow \text{Map}(\mathbb{F}, \mathbb{F})$$

$$\sum_{i=0}^{n} a_i t^i \mapsto \begin{cases} \mathbb{F} \to \mathbb{F} \\ \lambda \mapsto \sum_{i=0}^{n} a_i t^i \end{cases}$$

If $\#\mathbb{F} < n+1$, $\mathscr{P}$ cannot be injective.
If $\#\mathbb{F} \geq n+1$, $\mathscr{P}$ is injective, we can identify $P_{n,\mathbb{F}}$ with its image in $\text{Map}(\mathbb{F}, \mathbb{F})$. Moreover, if $\mathbb{F}$ is infinite such as $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, then $\mathscr{P}$ is injective for any n. From now on, we assume $\mathbb{F}$ is infinite.

Corollary A polynomial of degree n can have no more than n zeros.

(b) **Fundamental Theorem of Algebra** Each complex polynomial of degree $n \geq 1$, that is, each map $P : \mathbb{C} \to \mathbb{C}$ of the form

$$P(z) = c_n z^n + \cdots + c_1 z + c_0$$

with $n \geq 1$, $c_0, \ldots, c_n \in \mathbb{C}$ and $c_n \neq 0$ has least one zero.

(c) **Lemma** If $P(\lambda)$ is a polynomial of degree n, $n \geq 1$ and $\lambda_0 \in \mathbb{F}$ is one of the zeros, then

$$P(\lambda) = (\lambda - \lambda_0)Q(\lambda)$$

for some well-determined polynomial Q of degree of $n-1$.

Corollary Each complex polynomial P splits into linear factors, that is, if $P(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$ with $a_n \neq 0$ and $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ are the pairwise distinct zeros of P, we have

$$P(\lambda) = c_n \prod_{i=1}^{r} (\lambda - \lambda_i)^{m_i}$$

with well-determined exponents $m_i \geq 1$, called the multiplicities of the zeros. In particular, if P is a characteristic polynomial of $f : V \to V$, the m_i is called the *algebraic multiplicities* of the eigenvalues λ_i.

C. Diagonalization

1. **Definition**
Let V be a finite dimensional vector space over $\mathbb{F}$ and $f : V \to V$ is an endomorphism. Suppose $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of V consisting eigenvector for f, corresponding to the eigenvalue $\lambda_1, \ldots, \lambda_n$ such that $f(v_i) = \lambda_i v_i$ and $v_i \neq 0$. Consider the diagram below,

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \Phi_{\mathcal{B}}\uparrow & & \Phi_{\mathcal{B}}\uparrow \\ \mathbb{F}^n & \xrightarrow{A=[f]_{\mathcal{B}}^{\mathcal{B}}} & \mathbb{F}^n \end{array}$$

we would have $A(e_i) = \Phi^{-1} f \Phi(e_i) = \lambda_i e_i$. The $A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ is diagonal. Conversely, if $f : V \to V$ and there exits a basis $\mathcal{B}$ such that $A = [f]_{\mathcal{B}}$ is diagonal. Then $A(e_i) = \lambda_i e_i$ where $A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Then $\Phi^{-1} f \Phi(e_i) = \lambda_i e_i$, $f(\Phi(e_i)) = \lambda_i \Phi(e_i)$. So $\mathcal{B} = (\Phi(e_1), \ldots, \Phi(e_n))$ is a basis consisting eigenvalues of f. Therefore, it is reasonable to give the definition of diagonalizability.

Definition: $f : V \to V$ where V is a finite dimensional vector space over $\mathbb{F}$ is *diagonalizable* if there exists a basis of V consisting eigenvalues of f.

2. **Remark**
$A \in M(n \times n, \mathbb{F})$ is diagonal is equivalent to that there exists a matrix $P \in M(n \times n, \mathbb{F})$ such that $D = P^{-1}AP$ is diagonal. Then $A = PDP^{-1}$ where P is a change basis matrix whose columns are eigenvectors of A.

3. **Lemma**
Suppose $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of f and $E_1, \ldots, E_k$ are the corresponding eigenspaces with bases $\mathcal{B}_1, \ldots, \mathcal{B}_k$. Then $(\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is linear independent.

Corollary The sum of the geometric multiplicities of the eigenvalues of f is at most $\dim V$. And the equality meets if and only if f is diagonalizable.

Corollary (Criteria 1) To determine if f is diagonalizable:
- find all basis of E_λ of eigenvalues,
- for each basis of E_λ of eigenvalues, put them together to see if there are $\dim V$ many eigenvectors.

4. **Fact**
   Let λ to be an eigenvalues of $f: V \to V$, then geometric multiplicity of λ is smaller or equal to the algebraic multiplicity.

   Corollary (Criteria 2) Diagonalization can go wrong in two ways:
   - Characteristic polynomial may not have enough roots.
   - Any of the eigenvalues could be deficient. (An eigenvalue is *deficient* if geometric multiplicity is smaller to the algebraic multiplicity.)

   Note: If we consider $\mathbb{F} = \mathbb{C}$, then the two way of diagonalization will be:
   - impossible,
   - if the $\mathbb{F} = \mathbb{R}$ and λ is deficient, it will help to pass to $\mathbb{C}$ because $\dim(f - \lambda \mathrm{Id})$ is determined by number of pivots and if a matrix in $M(n \times n, \mathbb{R})$ is REF, it is still in REF if it is considered as a matrix in $M(n \times n, \mathbb{C})$.

D. Discrete Dynamic System

1. **General Solution**
   We have a dynamic system involving two variables

$$x_{n+1} = a_{11}x_n + a_{12}y_n$$

$$y_{n+1} = a_{21}x_n + a_{22}y_n$$

   Let $r_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ to be the state vector and $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Then the system can be abbreviated to $r_{n+1} = Ar_n$. To find $r_n = A^n r_0$ where $r_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, we need to know A^n. Assume A is diagonalizable with eigenvectors (v_1, v_2) corresponding to the eigenvalues λ_1, λ_2. An easy way is to diagonalize $A = PDP^{-1}$ where $D = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$ is diagonal, then $A^n = PD^nP^{-1}$ where $D^n = \begin{pmatrix} \lambda_1^n & \\ & \lambda_2^n \end{pmatrix}$. Another way is to assume $r_0 = c_1v_1 + c_2v_2$ where c_1, c_2 are undetermined (If r_0 is known, we can solve c_1, c_2). Then $x^n = A^n x_0 = c_1A^nv_1 + c_2A^nv_2 = c_1\lambda_1^n v_1 + c_2\lambda_2^n v_2$ would be the general solution of the dynamic system.

2. **Equilibrium State**
   In general, if A diagonalizable, λ_1 is the eigenvalue such that $\lambda_1 > |\lambda_i|$ to all eigenvalues, then the limiting growing rate $\lim_{n\to\infty} \frac{y_{n+1}}{y_n} = \lambda_1$ and $\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = \lambda_1$. And λ_1 determine the largest tern behaviour of the state vector. Also $\lim_{n\to\infty} \frac{x_n}{y_n} = 1$.

   If $1 > \lambda_1 > |\lambda_i|$, the state would converge to origin.

The Principle Axes Transformation

A. Self-adjoint Operator

1. **Definition**
   Let $(V, \langle , \rangle)$ be a Euclidean vector space. An operator or endomorphism $f : V \to V$ is said to be *self-adjoint* if $\langle f(v), w\rangle = \langle v, f(w)\rangle$ for all $v, w \in V$.

2. **Facts**
   (a) Any two eigenvectors v and w of a self-adjoint operator f corresponding to distinct eigenvalues $\lambda \neq \mu$ are orthogonal to each other, since $(f(v), w) = (v, f(w))$ implies $\langle \lambda v, w\rangle = \langle v, \mu w\rangle$, thus $(\lambda - \mu)(v, w) = 0$.
   (b) If v is an eigenvector of the self-adjoint operator $f : V \to V$, then the subspace $v^\perp := \{w \in V | w \perp v\}$ is invariant under f, that is, $f(v^\perp) \subset v^\perp$, since $\langle f(w), v\rangle = \langle w, f(v)\rangle = \langle w, \lambda v\rangle = 0$.

B. Symmetric Matrix

1. **Definition**
   A matrix A is *symmetric* if $A^t = A$.

2. **Remark**
   If $(V, \langle , \rangle)$ is a Euclidean vector space and $(v_1, \ldots, v_n)$ is an orthonormal basis of V, the matrix A of an endomorphism is given by $a_{ij} = \langle v_i, f(v_j)\rangle$.

   Corollary (Spectral Theorem) If $(v_1, \ldots v_n)$ is an orthonormal basis for the Euclidean vector space V, an operator $f : V \to V$ is self-adjoint if and only if its matrix A with respect to $(v_1, \ldots, v_n)$ is symmetric.

C. The Principle Axes Transformation for Self-adjoint Endomorphism

1. **Theorem** If $(V, \langle , \rangle)$ is a finite-dimensional Euclidean vector space and $f : V \to V$ is a self-adjoint endomorphism, there exists an orthonormal basis of eigenvectors of f.

   Lemma Each self-adjoint endomorphism of an n-dimensional Euclidean vector space V with $n > 0$ has an eigenvector.

2. **Corollary**
   (a) (For self-adjoint operator) Given a self-adjoint endomorphism $f : V \to V$ of an n-dimensional Euclidean vector space, it is always possible to find an orthogonal transformation

$$P : \mathbb{R}^n \xrightarrow{\cong} V$$

("principal axes transformation"), which reduces f to a diagonal matrix $D := P^{-1}DP$ the form

$$D = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & \\ & & & & & \ddots & \\ & & & & & & \lambda_r \end{pmatrix}$$

Here $\lambda_1, ..., \lambda_r$ are the distinct eigenvalues of f, the number of each appearing on the diagonal being equal to the geometric multiplicity.

(b) (For Symmetric real matrices) If A is a symmetric real n x n matrix, there is an orthogonal transformation $P \in O(n)$, such that $D := P^{-1}AP$ is a diagonal matrix with the eigenvalues of A as diagonal entries, each appearing with its geometric multiplicity. The way to find P is normalize every basis of eigenspace E_{λ_i} and put all the basis together.

(c) (Spectral Decomposition) If $f : V \to V$ is a self-adjoint endomorphism of a finite-dimensional Euclidean vector space, $\lambda_1, ..., \lambda_r$ its distinct eigenvalues, and $P_k : V \to V$ the orthogonal projection onto the eigenspace E_{λ_k}, then

$$f = \sum_{k=1}^{r} \lambda_k P_k$$

Quadratic Form

A. Definition

1. **Quadratic Form**
   Let $\mathbb{F}$ be a field, $\text{char}(\mathbb{F}) \neq 2$. Let V be an $\mathbb{F}$-vector space. $\beta : V \times V \to \mathbb{F}$ is a symmetric bilinear form. Then

$$\begin{aligned} q: \quad & V \to \mathbb{C} \\ & x \mapsto \beta(x, x) \end{aligned}$$

   is the *associated quadratic form.* We can recover β from q,

$$\beta(x, y) = \frac{1}{2}\left(q(x+y) - q(x) - q(y)\right)$$

   Then β is also the the bilinear form associated to q.

2. **Example**
   (a) Let $\text{Sym}(n \times n, \mathbb{F}) = \{A \in M(n \times n, \mathbb{F}) | A = A^t\}$. Then define quadratic form q_A associated to $\beta_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ where $\beta_A(x, y) = x^t Ay$. Since

$$\text{Sym}(n \times n, \mathbb{F}) \to \{\text{quadratic forms } \mathbb{F}^n \to \mathbb{F}^n\}$$

$$A \mapsto q_A$$

   is bijective, given the quadratic form q associated to $\beta : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ define the matrix A such that $a_{ij} = \beta(e_i, e_j)$.

   (b) Hessian form $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(P)\right)$.

   (c) Conic sections: if $q : \mathbb{R}^n \to \mathbb{R}$ is a quadratic form, then $q(x) = 1$ is the associated conic section.

3. **The Matrix of the Quadratic Form**
   Let $q : V \to \mathbb{F}$ be a quadratic form corresponding to symmetric bilinear form β. $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of V. Then *matrix of the quadratic form* $[q]_{\mathcal{B}} \in \text{Sym}(n \times n, \mathbb{F})$ is defined by $([q]_\beta)_{ij} = \beta(v_i, v_j)$. Then, assuming $v = \sum_{i=1}^n x_i v_i$ and $w = \sum_{j=1}^n y_j v_j$. Then we have

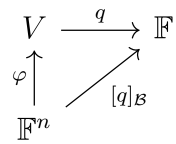
$$\beta(v, w) = \beta(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j) = \sum_{1 \leq i,j \leq n} x_i y_j \beta(v_i, v_j) = x^t [q]_{\mathcal{B}} y = [v]_{\mathcal{B}}^t [q]_{\mathcal{B}} [w]_{\mathcal{B}}$$

   Therefore, the map

$$\{\text{quadratic forms } q : V \to \mathbb{F}^n\} \longrightarrow \text{Sym}(n \times n, \mathbb{F})$$

$$q \mapsto [q]_{\mathcal{B}}$$

is bijective. Put it into a diagram we can see



4. **Definition** Let $q : V \to \mathbb{F}$ to be a quadratic form, β to be the corresponding symmetric bilinear form. The kernel of q is $\text{Ker}(q) := \{x \in V | \beta(\cdot, x) : V \to \mathbb{F}\text{is the zero map}\}$. The *rank* of q is $\dim V - \dim \text{Ker}q$. We call q is *non-degenerate* if $\text{Ker}q = \{0\}$.

5. **Theorem** Let $\dim V = n < \infty$, q is a quadratic form associated to the symmetric bilinear form β. Then there exists a basis $\mathcal{B}$ of V such that $[q]_{\mathcal{B}}$ is diagonal.

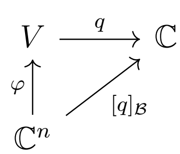
6. **Proposition**
   $\text{rk}(q)$ is the max dimension of a subspace $U \subset V$ such that $q|U$ is non-degenerate.

B. Quadratic Forms Over $\mathbb{F}$

1. **Quadratic Form Over $\mathbb{C}$**
   Let $\mathbb{F} = \mathbb{C}$, $\dim V = n < \infty$ and $q : V \to \mathbb{C}$ is a quadratic form with $\text{rk}q = k$. Since $\mathbb{F} = \mathbb{C}$, $\forall x \in \mathbb{C}$, $\exists y$ such that $y^2 = x$ call any of these $\sqrt{x}$. Arrange an "orthonormal basis" such that $q(v_1), \ldots, q(v_k) \neq 0$ and $q(v_{k+1}), \ldots, q(v_n) = 0$. Then replace v_i by $\dfrac{1}{\sqrt{q(v_i)}}$ for $i \leq l$ so that $q(\dfrac{1}{\sqrt{q(v_i)}}v_i) = \beta(\dfrac{1}{\sqrt{q(v_i)}}v_i, \dfrac{1}{\sqrt{q(v_i)}}v_i) = \dfrac{1}{q(v_i)}q(v_i) = 1$. Then we have $[q]_{\mathcal{B}} =$

$$\left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0 \end{array}\right)$$

   Consider it in a diagram, we have



   where

$$[q]_{\mathcal{B}} : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1^2 + \cdots + x_k^2$$

2. **Quadratic Form Over $\mathbb{R}$**

(a) **Sylvester Inertia Theorem** Let $\mathbb{F} = \mathbb{R}$, $\dim V = n < \infty$ and $q : V \to \mathbb{R}$ is a quadratic form. There exists a basis $\mathcal{B}$ such that $[q]_{\mathcal{B}} =$

$$\left(\begin{array}{c|c|c} I_r & 0 & 0 \\ \hline 0 & -I_s & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

where $r, s \geq 0$ and $r + s \leq n$ (It is obvious that $[q]_{\mathcal{B}}$ is unique). We call $\mathcal{B}$ *Sylvester basis*, $r + s$ is the *rank* of q and $r - s$ is the *signature* of q. Consider it in a diagram, we have

$$\begin{array}{ccc} V & \xrightarrow{q} & \mathbb{R} \\ {\scriptstyle \varphi}\uparrow & \nearrow {\scriptstyle [q]_{\mathcal{B}}} & \\ \mathbb{R}^n & & \end{array}$$

where

$$[q]_{\mathcal{B}} : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2$$

(b) **Method of finding Sylvester Basis**

- Follow the proof of the existence of a Sylvester basis. The inductive proof converts into a recursive algorithm. Start by finding a vector v_1, such that $q(v_1) \neq 0$. Then find v_2 in $v_1^{\perp}$, such that $q(v_2) \neq 0$. Then find v_3 in $v1, v2^{\perp}$, such that $q(v_3) \neq 0$, etc.
- Find an orthonormal basis consisting of eigenvectors of A. Then normalize each eigenvector by rescaling it by the inverse of the square root of the absolute value of the corresponding eigenvalue.
- Follow the algorithm explained in the box on page 183 of the textbook. This means to do elementary symmetric operations on A, until A is in Sylvester form, and do the same column operations on the identity matrix to get the matrix P, such that $P^t A P = S$, which is the matrix which contains the Sylvester basis as columns. An elementary symmetric operation is simply an elementary row operation followed by the same column operation (or the other way around, it gives the same answer).

Note: Sylvester basis is not unique.

(c) **Conic Section** Conic section is $q(x) = 1$ where $q : V \to \mathbb{R}$ is a quadratic form. By Sylvester Inertia Theorem, we could have $\lambda_1^2 x_1'^2 + \cdots + \lambda_n x_n'^2 = 1$. Then $L(x_1')$ are coordinate axes: principle axes $q(x) = 1$. If $\lambda_i > 0$, then the intersection of $q(x) = 1$ with $L(x_i')$ would be $\lambda_i^2 x_i'^2 = 1$ which gives $x_i = \pm\dfrac{1}{\sqrt{\lambda_i}}$ and $\dfrac{1}{\sqrt{\lambda_i}}$ would be the i^{th} semi-axis. If $\lambda_i < 0$ there would be no intersection with i^{th} principle axis.

Classification of Matrices

A. Class And Classification

1. **Equivalence**
   Let M be a set. By an equivalence relation on M ($x \sim y$, $x, y \in M$) one understands a relation which satisfies the following axioms: reflexivity, symmetry and transitivity.

2. **Equivalence Class**
   Let $\sim$ be a equivalence relation on M. For $x \in M$ the subset $[x] := \{y|x \sim y\} \subset M$ is called the equivalence class of x with respect to $\sim$.

3. **Canonical Projection**
   Let $M/\sim := \{[x]|x \in M\}$ to be set of equivalence classes (the *quotient* of M by $\sim$). Then the map $\pi : M \to M/\sim$ mapping x to $[x]$ is called the *canonical projection* of the equivalence relation $\sim$.

4. **Classification** Classification up to $\sim$ is to find the canonical projection and the corresponding $M/\sim$. There are two ways to do:
   (a) Classification by means of characteristic data: find a well-known set D with a surjective map $c : M \to D$ with the property $x \sim y \Longleftrightarrow c(x) = c(y)$. Then the map $M/\sim \to D$ is well-defined and bijective. In this case, we call $c(x)$ is a *characterizing datum* for x respect to $\sim$.
   (b) Classification by representatives: find an easily-understood subset $M_0 \subset M$ such that $\pi|M_0 : M_0 \to M/\sim$ is bijective. Then for each $x \in M$ there exists a unique representative $x_0 \in M_0$ with $x \sim x_0$. So M_0 contains exactly one sample from each equivalence class.

5. **Matrix equivalence**
   Two $m \times n$ matrices $A, B \in M(m \times n, \mathbb{F})$ are said to be *equivalent*, written $A \sim B$, if there exists invertible matrices $P \in M(m \times n, \mathbb{F})$ and $Q \in M(m \times m)$ with $B = Q^{-1}AP$.

$$\begin{array}{ccc} \mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^m \\ {\scriptstyle P}\uparrow & & {\scriptstyle Q}\uparrow \\ \mathbb{F}^n & \xrightarrow{B} & \mathbb{F}^m \end{array}$$

6. **Similarity**
   Two $n \times n$ matrices A, B are said to be *similar* if there exists an invertible $n \times n$ matrix P such that $B = P^{-1}AP$.

$$\begin{array}{ccc} \mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^n \\ {\scriptstyle P}\uparrow & & {\scriptstyle P}\uparrow \\ \mathbb{F}^n & \xrightarrow{B} & \mathbb{F}^n \end{array}$$

B. Classification by Four Fundamental Theorem in Linear Algebra

1. **Rank Theorem** General matrices over $\mathbb{F}$, $B = P^{-1}AQ$
   (a) A is equivalent to $B \iff \mathrm{rk}A = \mathrm{rk}B$.
   (b) Normal form is

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

   where r is the rank of the matrix.

2. **Sylvester Inertia Theorem** Symmetric real matrix, $B = P^tAP$
   (a) A is symmetric equivalent to $B \iff \mathrm{rk}A = \mathrm{rk}B$ and $\mathrm{sgn}A = \mathrm{sgn}B$
   (b) The normal form is

$$\left(\begin{array}{c|c|c} I_r & 0 & 0 \\ \hline 0 & -I_s & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

3. **Principal Axes Theorem** Symmetric real matrix, $B = P^{-1}AP = P^tAP, P \in O(n)$
   (a) A is orthonormal similar to $B \iff$ eigenvalues are equal.
   (b) The normal form is

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

4. **Jordan Normal Form** Complex square matrix
   (a) A is similar to $B \iff A, B$ have the same Jordan normal form