1 Logics and Proofs

1.1 Logical Notations and Their properties

1.1.1 Notations

- p, q : propositions (sentences)
- $-p$: negation of p (It is not the case of p.)
- $p \wedge q$: p and q
- $p \vee q$: p or q
- $p \Rightarrow q \equiv -p \vee q$: If p, then q.
  - p only if q.
  - p implies q.
  - p is sufficient for q.
  - q is necessary for p.
- $p \Leftrightarrow q \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$: p if and only if q. (p iff q)
  - To show $p \Leftrightarrow q$, we have to show both $p \Rightarrow q$ and $q \Rightarrow p$.

1.1.2 Truth Table

- Negation:

p	$-p$
T	F
F	T

- Conjunction:

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	T	F

Remark 1 *$p \wedge -p$ is always false.*

- Disjunction:

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	T	F

- Implication:

p	q	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	T	T

1.1.3 Properties

- $-(-p) = p$
- $-(p \wedge q) = -p \vee -q$
- $-(p \vee q) = -p \wedge -q$
- $-(p \Rightarrow q) = p \wedge -q$
- $p \Rightarrow q = -q \Rightarrow -p$ (contrapositive of $p \Rightarrow q$)

1.1.4 Quantifiers

- $\forall$: universal quantifier ("for every" quantifier)
- $\exists$: existential quantifier ("there exists" quantifier)
- $\exists! x$: There is exactly one x such that ....

Negation of a proposition with quantifiers

- $-(\forall x, p(x)) = \exists x, p(x)$
- $-(\exists x, p(x)) = \forall x, -p(x)$
- $-(\forall x, \exists y, p(x, y)) = \exists x, \forall y, -p(x, y)$

Note: $(\forall x, \exists y, p(x, y))$ and $(\exists y, \forall x, p(x, y))$ are different.

1.2 Proofs

1.2.1 Direct Proof

To show: $p \Rightarrow q$

Proof.

$$p \Rightarrow p_1 \Rightarrow p_2 \Rightarrow ... \Rightarrow q$$

Example 1 *If m is and even integer, for any integer p, mp is an even integer.*

1.2.2 Indirect Proof (Proof by Contradiction)

To show: $p \Rightarrow q$

Proof.

Suppose $-(p \Rightarrow q)$. (i.e., $(p \wedge -q)$ is true.)

$(p \wedge -q) \Rightarrow ... \Rightarrow -p$.

That is, $(p \wedge -p)$ is true. Contradiction

Example 2 *Let a be an integer. Show the following statement: a^2 is odd$\Rightarrow$ a is odd.*

1.2.3 Mathematical Induction

Consider a sequence of statements indexed by the natural numbers: $p_1, p_2, ..., p_n, ...$.

To show: p_n is true for all $n = 1, 2, ...$

Proof.

(1) p_1 is true.

(2) If p_k is true, then p_{k+1} is true.

Example 3 *Show* $1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$.

2 Sets, Functions, and Numbers

2.1 Sets

2.1.1 Basic Notation

- A : set
    - $a \in A$: a is an element of A.
    - $a \notin A$: a is not an element of A.
- Examples:

$$\begin{array}{rcl} \mathbb{N} & = & \{1, 2, 3, ...\} \\ \mathbb{Z} & = \{ & ..., -2, -1, 0, 1, 2, ...\} \\ \mathbb{Q} & = & \{x = \frac{p}{q} : p, q \in \mathbb{Z} \text{ and } q \neq 0\} \\ \mathbb{R} & : & \text{the set of all real numbers} \\ \mathbb{R}_+ & = & \{x \in \mathbb{R} : x \geq 0\} \\ \mathbb{R}_{++} & = & \{x \in \mathbb{R} : x > 0\} \\ \phi & : & \text{empty set (the set having no elements)} \end{array}$$

- $A \subset B \underset{def}{\Leftrightarrow} \forall a \in A, a \in B$ (Every element of A is also an element of B.)
- $A = B \underset{def}{\Leftrightarrow} A \subset B$ and $B \subset A$
- $A \subsetneq B \underset{def}{\Leftrightarrow} A \subset B$ and $A \neq B$.

2.1.2 Binary Operation of Sets

- Union:

$$A \cup B \equiv \{x : x \in A \text{ or } x \in B\}$$

- Intersection:

$$A \cap B \equiv \{x : x \in A \text{ and } x \in B\}$$

Definition 1 *If $A \cap B = \phi$, A and B are disjoint.*

- Difference:

$$A \setminus B \equiv \{x : x \in A \text{ and } x \notin B\}$$

    - Complement:

$$\begin{array}{rcl} X & : & \text{universal set} \\ A^c & \equiv & X \setminus A \end{array}$$

Properties

1. $$A \cap A = A, A \cup A = A, A \cup \phi = A, A \cap \phi = \phi.$$

2. Commutative laws:
$$\begin{aligned} A \cap B &= B \cap A; \\ A \cup B &= B \cup A. \end{aligned}$$

3. Associative laws:
$$\begin{aligned} (A \cap B) \cap C &= A \cap (B \cap C); \\ (A \cup B) \cup C &= A \cup (B \cup C). \end{aligned}$$

4. Distributive laws:
$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C); \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned}$$

5. $$(A \cap B) \cap (A \setminus B) = \phi.$$

6. DeMorgan's laws:
$$\begin{aligned} A \setminus (B \cup C) &= (A \setminus B) \cap (A \setminus C); \\ A \setminus (B \cap C) &= (A \setminus B) \cup (A \setminus C). \end{aligned}$$

Note: If $A = X$ (universal set), then DeMorgan's laws imply
$$\begin{aligned} (B \cup C)^c &= B^c \cap C^c; \\ (B \cap C)^c &= B^c \cup C^c. \end{aligned}$$

Notations

- For a sequence of sets, $A_1, A_2, ..., A_n, ...,$
$$\begin{aligned} \bigcup_{i=1}^{n} A_i &\equiv \{x : x \in A_i \text{ for some } i = 1, 2, ..., n\}; \\ \bigcap_{i=1}^{n} A_i &\equiv \{x : x \in A_i \text{ for all } i = 1, 2, ..., n\}; \\ \bigcup_{i=1}^{\infty} A_i &\equiv \{x : x \in A_i \text{ for some } i \in \mathbb{N}\}; \\ \bigcap_{i=1}^{\infty} A_i &\equiv \{x : x \in A_i \text{ for all } i \in \mathbb{N}\}. \end{aligned}$$

2.2 Functions

Definition 2 *f is a functions from a set A to a set B, if f assigns a unique point $f(x) \in B$ for each $x \in A$. $(f : A \to B)$*

- Given a function $f : A \to B$,

Definition 3

$$\begin{aligned} D(f) &\equiv A \text{ is the domain of } f\text{; and} \\ R(f) &\equiv \{y \in B : \exists x \in A, y = f(x)\} \text{ is the range of } f. \end{aligned}$$

Definition 4 *Given a set $C \subset A$, the image of C under f is defined by*

$$f(C) \equiv \{y \in B : \exists x \in A, y = f(x)\}.$$

Definition 5 *Given a set $E \subset B$, the inverse image of E under f is defined by*

$$f^{-1}(E) \equiv \{x \in A : \exists y \in E, y = f(x)\}.$$

Definition 6 *A function f is said to be onto if $\forall y \in B, \exists a \in A, f(x) = y$.*

Definition 7 *A function f is said to be one-to-one if $\forall x, x' \in A, x \neq x' \Rightarrow f(x) \neq f(x')$.*

Definition 8 *Given $f : X \to Y$ and $g : Y \to Z$, $g \circ f : X \to Z$ is defined by*

$$g \circ f(x) = g(f(x)).$$

2.3 Real Numbers

- $\mathbb{R}$: the set of all real numbers
  - $\mathbb{R}$ includes both rationals and irrationals.
- Given $z \in \mathbb{R}$,

$$|z| = \begin{cases} z \text{ if } z \geq 0 \\ -z \text{ if } z < 0 \end{cases}.$$

- Properties: Given $z \in \mathbb{R}$,

$$\begin{aligned} \text{(i) } |z| &> 0 \text{ if } z \neq 0; \text{ and} \\ |z| &= 0 \text{ if } z = 0. \\ \text{(ii) } |az| &= |a|\,|z|\,, \forall a \in \mathbb{R}. \\ \text{(iii) } |z+w| &\leq |z| + |w| \end{aligned}$$

- Euclidian distance between x and y in $\mathbb{R}$: $|x - y|$

- Properties: Given $x, y \in \mathbb{R}$,

$$\begin{aligned} \text{(i) } |x-y| &\geq 0; \text{ and} \\ |x-y| &= 0 \Leftrightarrow x = y. \\ \text{(ii) } |x-y| &= |y-x|. \\ \text{(iii) } |x-y| &\leq |x-z| + |z-y|, \forall z \in \mathbb{R}. \end{aligned}$$

2.4 n-Dimensional Euclidian Space

- Cartesian product of A and B:

$$\begin{aligned} A \times B &\equiv \{(a,b) : a \in A, b \in B\}; \\ \prod_{i=1}^{n} A_i &= \{(a_1, a_2, ..., a_n) : a_i \in A_i, i = 1, 2, ..., n\}. \end{aligned}$$

- $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n} = \{x = (x_1, x_2, ..., x_n) : x_i \in \mathbb{R}, i = 1, 2, ..., n\}$

- Notations: For $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$,

$$\begin{aligned} x + y &= (x_1 + y_1, x_2 + y_2, ..., x_n + y_n); \\ \alpha x &= (\alpha x_1, \alpha x_2, ..., \alpha x_n); \\ x &= y \text{ if } x_i = y_i \text{ for all } i = 1, 2, ..., n; \\ x &\geq y \text{ if } x_i \geq y_i \text{ for all } i = 1, 2, ..., n; \\ x &> y \text{ if } x_i \geq y_i \text{ for all } i = 1, 2, ..., n, \text{ and } x \neq y; \\ x &\gg y \text{ if } x_i > y_i \text{ for all } i = 1, 2, ..., n; \\ \mathbb{R}^n_+ &\equiv \{x \in \mathbb{R}^n : x \geq 0\}; \\ \mathbb{R}^n_{++} &\equiv \{x \in \mathbb{R}^n : x \gg 0\}. \end{aligned}$$

2.4.1 Inner Product, Norm, and Metric in $\mathbb{R}^n$

For $x, y \in \mathbb{R}^n$,

- Euclidian inner product:

$$x \cdot y \equiv \sum_{i=1}^{n} x_i y_i.$$

- Euclidian norm of x:

$$\|x\| \equiv |x \cdot x|^{1/2} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

Theorem 1 (Cauchy-Schwartz Inequality) *For any* $x, y \in \mathbb{R}^n$,

$$|x \cdot y| \leq \|x\| \, \|y\| .$$

Theorem 2 *For* $x, y \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$,

$$\begin{array}{rcl} \textit{(i)}\ \|x\| & \geq & 0;\ \textit{and} \\ \|x\| & = & 0 \Leftrightarrow x = 0. \\ \textit{(ii)}\ \|\alpha x\| & = & |\alpha| \, \|x\| . \\ \textit{(iii)}\ \|x - y\| & \leq & \|x\| + \|y\| \end{array}$$

- Euclidian Distance in $\mathbb{R}^n$:

$$d(x, y) \equiv \|x - y\| .$$

Theorem 3 *For* $x, y \in \mathbb{R}^n$,

$$\begin{array}{rcl} \textit{(i)}\ d(x,y) & \geq & 0;\ \textit{and} \\ d(x,y) & = & 0 \Leftrightarrow x = y. \\ \textit{(ii)}\ d(x,y) & = & d(y,x). \\ \textit{(iii)}\ d(x,y) & \leq & d(x,z) + d(z,y). \end{array}$$

3 Sequences

3.1 Sequences and Limits

Definition 9 *A sequence in* $\mathbb{R}^n$ *is a function* $f : \mathbb{N} \to \mathbb{R}^n$*.*

Definition 10 *A sequence* $\{x_k\}$*converges to* x ($x_k \to x$) *if* $d(x_k, x) \to 0$ *as* $k \to \infty$ (*i.e.,* $\forall \varepsilon > 0, \exists N(\varepsilon) \in \mathbb{N} : \forall k \geq N(\varepsilon), d(x_k, x) < \varepsilon$).

Theorem 4 *A sequence have at most one limit.*

Definition 11 *A sequence* $\{x_k\}$ *is bounded if* $\exists M \in \mathbb{R}_+ : \|x_k\| \leq M \quad \forall k = 1, 2, ...$

Theorem 5 *Every convergent sequence in* $\mathbb{R}^n$ *is bounded.*

- **Note:** There is a bounded sequence which does not converge.
  - Example: $x_k = (-1)^k \ k = 1, 2, ...$

Theorem 6 $x_k \to x$ *in* $\mathbb{R}^n \Leftrightarrow x_k^i \to x^i$ *for all* $i = 1, 2, ..., n$*.*

Theorem 7 *Let* $x_k \to x$*. Then,*

$$a \leq x_k \leq b \text{ for all } k = 1, 2, ... \Rightarrow a \leq x \leq b.$$

- **Note:** For a convergent sequence $x_k \to x$, it is possible that $a < x_k$ for all $k = 1, 2, ...$, but $a = x$.

3.2 Subsequences and limit points

Definition 12 *Given a sequence* $\{x_k\}$*, consider a strictly increasing function* $K : \mathbb{N} \to \mathbb{N}$*. Then, the sequence* $\{x_{K(m)}\}_{m=1}^{\infty}$ *is called a subsequence of* $\{x_k\}$*.*

Definition 13 *If a subsequence of* $\{x_k\}$ *converges, the limit of the subsequence is called a limit point of* $\{x_k\}$*.*

- **Note:** A sequence may have multiple limit points.

Theorem 8 $\{x_k\}$ *converges to* x $\Leftrightarrow$*Every subsequence of* $\{x_k\}$ *converges to* x*.*

3.3 Cauchy Sequences and Completeness

Definition 14 *A sequence* $\{x_k\}$ *satisfies the Cauchy criterion if* $\forall \varepsilon > 0, \exists N(\varepsilon) \in \mathbb{N} : \forall m, l \geq K(\varepsilon), d(x_m, x_l) < \varepsilon$.

Definition 15 *A sequence that satisfies the Cauchy criterion is called a Cauchy sequence.*

Definition 16 *Two Cauchy sequences* $\{x_k\}$ *and* $\{y_k\}$ *are equivalent if* $\forall \varepsilon > 0, \exists N(\varepsilon) \in \mathbb{N} : \forall k \geq N(\varepsilon), d(x_k, y_k) < \varepsilon$.

Theorem 9 *Suppose* $\{x_k\}$ *and* $\{y_k\}$ *are equivalent Cauchy sequences. Then,*

$$x_k \to x \Rightarrow y_k \to y.$$

- **Note:** The real number system $\mathbb{R}$ is the set of all "limits" of Cauchy sequences of rationals. (Each element in $\mathbb{R}$ is the limit of some Cauchy sequence of rationals.)

Theorem 10 $\{x_k\}$ *in* $\mathbb{R}^n$ *is a Cauchy sequence* $\Leftrightarrow$ $\{x_k^i\}$ *in* $\mathbb{R}$ *is a Cauchy sequence for all* $i = 1, 2, ..., n$.

Theorem 11 $\{x_k\}$ *in* $\mathbb{R}^n$ *is a Cauchy sequence* $\Leftrightarrow$ $\{x_k\}$ *is a convergent sequence in* $\mathbb{R}^n$.

Corollary 1 $\{x_k\}$ *in* $\mathbb{R}^n$ *is a Cauchy sequence*$\Rightarrow$ $\{x_k\}$ *is a bounded sequence.*

Definition 17 S *is a complete space if every Cauchy sequence in* S *converges to an element in* S.

Corollary 2 $\mathbb{R}^n$ *is complete.*

3.4 Suprema, Infima, Maxima, and Minima

Let A be a nonempty set in $\mathbb{R}$.

- The set of upper bound of A :

$$U(A) \equiv \{u \in \mathbb{R} : u \geq a \quad \forall a \in A\}$$

- The set of lower bound of A :

$$L(A) \equiv \{l \in \mathbb{R} : l \leq a \quad \forall a \in A\}$$

- A is called bounded above if $U(A) \neq \phi$.
- A is called bounded below if $L(A) \neq \phi$.
- A is called bounded if A is both bounded below and above.
- The least upper bound of A (supremum of A):

$$\sup A \equiv \{a^* \in U(A) : a^* \leq u \quad \forall u \in U(A)\}.$$

- The greatest lower bound of A (infimum of A):

$$\inf A \equiv \{\widehat{a} \in L(A) : \widehat{a} \geq l \quad \forall l \in L(A)\}.$$

- Maximum of A :

$$\max A \equiv \{z \in A : z \geq a \quad \forall a \in A\}.$$

- Minimum of A :

$$\min A \equiv \{w \in A : w \leq a \quad \forall a \in A\}.$$

Theorem 12 *Suppose* $\sup A$ *is finite. Then,*

$$\forall \varepsilon > 0, \exists a \in A : a > \sup A - \varepsilon.$$

Theorem 13 (Least Upper Bound Property)

$$\begin{aligned} U(A) &\neq \phi \Rightarrow \exists a^* \in U(A) : a^* \leq u \quad \forall u \in U(A); \\ L(A) &\neq \phi \Rightarrow \exists \widehat{a} \in L(A) : \widehat{a} \leq l \quad \forall l \in L(A). \end{aligned}$$

Theorem 14 (Archimedian Property)

$$x, y \in \mathbb{R}, x > 0 \Rightarrow \exists n \in \mathbb{N} : nx > y.$$

- **Note:** Taking $y = 1$, this theorem implies the following: For any (small) x, we can find a (large) natural number n such that $x > \frac{1}{n}$.

Theorem 15 (Denseness of Rationals) *Any real number can be approximated by a rational number:*

$$\forall x \in \mathbb{R}, \forall x \in \mathbb{N}, \exists y \in \mathbb{Q} : |x - y| < \frac{1}{n}.$$

3.5 Monotone Sequences in $\mathbb{R}$

Definition 18 *$\{x_k\}$ in $\mathbb{R}$ is a monotone increasing sequence if $x_{k+1} \geq x_k$ for all $k \in \mathbb{N}$.*

Definition 19 *$\{x_k\}$ in $\mathbb{R}$ is a monotone decreasing sequence if $x_{k+1} \leq x_k$ for all $k \in \mathbb{N}$.*

Definition 20 *$\{x_k\}$ in $\mathbb{R}$ diverges to ∞ if $\forall p \in \mathbb{N}, \exists K(p) \in \mathbb{N} : \forall k \geq K(p), x_k \geq p$.*

Definition 21 *$\{x_k\}$ in $\mathbb{R}$ diverges to -∞ if $\forall p \in \mathbb{N}, \exists K(p) \in \mathbb{N} : \forall k \geq K(p), x_k \leq -p$.*

- **Note:**

$$\{x_k\} \text{ diverges } \overset{\Rightarrow}{\not\Leftarrow} \{x_k\} \text{ is unbounded}$$

Theorem 16 *Let $\{x_k\}$ be a monotone increasing (decreasing) sequence in $\mathbb{R}$. Then, (i) if $\{x_k\}$ is unbounded, it diverges to ∞ $(-\infty)$; and (ii) if $\{x_k\}$ is bounded, it converges.*

3.6 Lim Sup and Lim Inf

- $\pm\infty$ is allowed as "limit points" of a sequence.

Definition 22 *Given a sequence $\{x_k\}$,*

$$\begin{aligned} \limsup_{k\to\infty} x_k &\equiv \lim_{k\to\infty} a_k \\ \textit{where } a_k &= \sup\{x_k, x_{k+1}, ...\} \\ &= \sup_{m\geq k} x_m, \end{aligned}$$

and

$$\begin{aligned} \liminf_{k\to\infty} x_k &= \lim_{k\to\infty} b_k \\ \textit{where } b_k &= \inf\{x_k, x_{k+1}, ...\} \\ &= \inf_{m\geq k} x_m. \end{aligned}$$

Theorem 17 *Let $\{x_k\}$ be a sequence in $\mathbb{R}$. Then, there exist subsequences $\{x_{K(m)}\}$ and $\{x_{K'(m)}\}$ of $\{x_k\}$ such that*

$$\begin{aligned} x_{K(m)} &\rightarrow \limsup_{k\to\infty} x_k \quad \text{as } m \to \infty; \text{ and} \\ x_{K'(m)} &\rightarrow \liminf_{k\infty} x_k \quad \text{as } m \to \infty. \end{aligned}$$

Theorem 18 *Let A be the set of all limit points of $\{x_k\}$ (including $\pm\infty$ if $\{x_k\}$ has divergent subsequences). Then,*

$$\begin{aligned} \limsup_k x_k &= \sup A; \text{ and} \\ \liminf_k &= \inf A. \end{aligned}$$

Theorem 19

$$x_k \to x \Leftrightarrow \limsup_{k\to\infty} x_k = \liminf_{k\to\infty} x_k = x.$$

3.7 Series

- Given $\{x_k\}$ in $\mathbb{R}$, we can define another sequence by

$$S_n = \sum_{k=1}^{n} x_k.$$

- If S_n converges to S, we write

$$S = \sum_{k=1}^{\infty} x_k = \lim_{n\to\infty} \sum_{k=1}^{n} x_k.$$

- If $x_k > 0$ for all k, then $\{S_n\}$ is monotone increasing.

4 Basic Topology

4.1 Open Sets and Closed Sets

4.1.1 Definitions

- Consider $\mathbb{R}^n$ with the Euclidian distance $d(x,y) = \|x-y\|$.

Definition 23 (Open Balls) *Let $x \in \mathbb{R}^n$. The open ball $B(x,r)$ with center x and radius $r > 0$ is defined by*

$$B(x,r) = \{y \in \mathbb{R}^n : d(x,y) < r\}.$$

Definition 24 (Open Sets) *A set $S \in \mathbb{R}^n$ is called an open set if*

$$\forall x \in S, \exists r > 0 : B(x,r) \subset S.$$

Definition 25 (Closed Sets) *A set $S \in \mathbb{R}^n$ is called an closed set if $S^c = \mathbb{R}^n \setminus S$ is open.*

Theorem 20 *Every open ball is open.*

Theorem 21 *$S \subset \mathbb{R}^n$ is closed $\Leftrightarrow$ for all convergent sequences $\{x_k\}$ in $S, x_k \to x \in S$. (The limit lies in S.)*

- **Note:** There is a set which is neither open nor closed.
  - Example: $[0,1)$ in $\mathbb{R}$ is neither open nor closed.

4.1.2 Properties

Let A be a set and $(G_\alpha)_{\alpha \in A}$ be a collection of sets in $\mathbb{R}^n$.

- If G_α is open for all $\alpha \in A$, then $\bigcup_{\alpha \in A} G_\alpha$ is also open.
- If G_α is closed for all $\alpha \in A$, then $\bigcap_{\alpha \in A} G_\alpha$ is also closed.

Now suppose A is a finite set: $A = \{1, 2, 3, ..., n\}$.

- If G_i is open for all $i = 1, 2, ..., n$, then $\bigcap_{i=1}^{n} G_i$ is also open.
- If G_i is closed for all $i = 1, 2, ..., n$, then $\bigcup_{i=1}^{n} G_i$ is also closed.

4.1.3 Closure of Sets

Definition 26 *Given a set $A \subset \mathbb{R}^n$. The closure of A (write $\overline{A}$), is the intersection of all closed sets containing A :*

$$\begin{aligned} \overline{A} &\equiv \bigcap_{F \in \mathcal{F}} F, \\ \textit{where } \mathcal{F} &= \{F \subset \mathbb{R}^n : A \subset F \textit{ and } F \textit{ is closed}\}. \end{aligned}$$

Theorem 22 *For any set $A \subset \mathbb{R}^n$, $\overline{A}$ is closed. (The closure of a set is closed.)*

- The closure of A is the smallest closed set containing A.

Theorem 23 *Let $A \subset \mathbb{R}^n$. Then,*

$$A = \overline{A} \Leftrightarrow A \textit{ is closed.}$$

Theorem 24 *Let A be a nonempty subset of $\mathbb{R}$. If A is bounded above, then* $\sup A \in \overline{A}$.

Theorem 25 $\overline{\mathbb{Q}} = \mathbb{R}$

4.2 Compact Sets

Definition 27 *A set K is compact if every sequence $\{x_k\}$ in K has a convergent subsequence$\{x_{k(m)}\}$ such that $x_{k(m)} \to x \in K$.*

Example 4 *Any finite set in $\mathbb{R}^n$ is compact.*

Example 5 $K = \{\frac{1}{n} : n \in \mathbb{N}\}$ *is not compact.*

Theorem 26 *Every compact set in $\mathbb{R}^n$ is closed.*

Theorem 27 *Every closed subset of a compact set is compact.*

Theorem 28 *If $F \subset \mathbb{R}^n$ is closed and $K \subset \mathbb{R}^n$ is compact, then $F \cap K$ is compact.*

Theorem 29 *Let S_1 and S_2 be both compact in $\mathbb{R}^n$. Then,*

$$S_1 + S_2 \equiv \{x \in \mathbb{R}^n : x = x_1 + x_2, x_1 \in S_1, x_2 \in S_2\}$$

is also compact.

Theorem 30 *Let $K \subset \mathbb{R}^n$. Then,*

$$K \textit{ is compact} \Leftrightarrow K \textit{ is closed and bounded.}$$

4.3 Connected Sets

Definition 28 *Let A and B be subsets of $\mathbb{R}^n$. A and B are said to be separated if*

$$A \cap \overline{B} = \phi \text{ and } \overline{B} \cap B = \phi.$$

- $A = (0,1)$ and $B = [1,2)$ are not separated.
- $A = (0,1)$ and $B = (1,2)$ are separated.

Definition 29 *$E \subset \mathbb{R}^n$ is said to be connected if E is not a union of two nonempty separated sets.*

- Any interval in $\mathbb{R}$ is connected.

Theorem 31 *$E \subset \mathbb{R}$ is connected* $\Leftrightarrow [x \in E, y \in E, x < z < y \Rightarrow z \in E]$.

Theorem 32 *$\mathbb{R}^n$ is connected.*

4.4 Convex Sets

$x_1, x_2, ..., x_k \in \mathbb{R}^n$

- $z = \sum_{i=1}^{k} \lambda_i x_i$ with $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0$ for all $i = 1, ..., k$, is called a convex combination of $(x_1, ..., x_k)$.

Definition 30 *$S \subset \mathbb{R}^n$ is convex if the convex combination of any two points in S is also lies in S :*

$$\forall x, y \in S, \forall \alpha \in [0,1], \alpha x + (1-\alpha) y \in S.$$

Theorem 33 *$\{S_\alpha\}_{\alpha \in A}$ is a collection of convex sets in $\mathbb{R}^n \Rightarrow \cap_{\alpha \in A} S_\alpha$ is also convex.*

Theorem 34 *If S_1 and S_2 are convex, then $S_1 + S_2$ is also convex.*

5 Continuity

5.1 Continuous Function

$f : S \to T, \quad S \subset \mathbb{R}^n, T \subset \mathbb{R}^l$

Definition 31 *A function* f *is continuous at* x *if* $\forall \varepsilon > 0, \exists \delta > 0 : \forall y \in S$ *with* $d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$.

Definition 32 f *is continuous on* S *if* f *is continuous at all* x *in* S.

Theorem 35 f *is continuous at* $x \Leftrightarrow [x_n \to x \Rightarrow \lim_{x_n \to x} f(x_n) = f(x)]$.

- Elementary functions: The following functions are all continuous.

$$\begin{aligned}
f(x) &= \alpha x, \quad \alpha \in \mathbb{R} \\
f(x) &= x^\alpha, \quad \alpha \in \mathbb{R} \\
f(x) &= \alpha^x, \quad \alpha \in \mathbb{R} \\
f(x) &= \log_\alpha x, \quad \alpha \in \mathbb{R} \\
f(x) &= \sin x, \cos x, \tan x
\end{aligned}$$

Theorem 36 *Given two functions* $f : X \to Y$ *and* $g : Y \to Z$. *If* f *is continuous at* $x \in X$ *and* g *is continuous at* $f(x) \in Y$, *then* $g \circ f$ *is continuous at* x.

Theorem 37 *Given two functions* $f : X \to Y$ *and* $g : X \to Y$. *If both functions are continuous at* $x \in X$, *then* $f + g$ *and* $f \cdot g$ *are both continuous at* x.

- Example: Continuous function from $\mathbb{R}^n$ to $\mathbb{R}^l$.
  - $f(x, y) = (2x + y, x + 3y)$ is continuous at $(a, b) \in \mathbb{R}^2$ for any a and b.

Theorem 38 *Let* $f_1, f_2, ..., f_l$ *be functions from* $\mathbb{R}^n \to \mathbb{R}$ *and let* f *be the function from* $\mathbb{R}^n$ *to* $\mathbb{R}^l$ *defined by*

$$f(x) = (f_1(x), f_2(x), ..., f_n(x)).$$

Then, f *is continuous at* $x \Leftrightarrow$ *each* f_i *is continuous at* x.

Theorem 39 *A function f is continuous at $x \Leftrightarrow$ for any open set $V \subset \mathbb{R}^l$ with $f(x) \in V$, there exists an open set $U \subset \mathbb{R}^n$ such that $x \in U$ and $f(z) \in V$ for all $z \in U \cap S$.*

Definition 33 *Suppose $X \subset \mathbb{R}^n$. A subset E of X is open relative to X if $E = X \cap G$ for some open subset G in $\mathbb{R}^n$.*

Corollary 3 *Let f be a function from $S \subset \mathbb{R}^n$ to $\mathbb{R}^l$. Then, f is continuous on $S \Leftrightarrow$ for any open set $V \subset \mathbb{R}^l$, $f^{-1}(V) = \{x \in S : f(x) \in V\}$ is open relative to S.*

Corollary 4 *Let f be a function from $S \subset \mathbb{R}^n$ to $\mathbb{R}^l$. Then, f is continuous on $S \Leftrightarrow$ for any closed set $C \subset \mathbb{R}^l$, $f^{-1}(C) = \{x \in S : f(x) \in C\}$ is closed relative to S.*

- Notation:

$$\begin{aligned} \lim_{y \to x} f(y) &= q \underset{def}{\Leftrightarrow} \\ \forall \varepsilon &> 0, \exists \delta > 0 \quad s.t. \quad d(x,y) < \delta \Rightarrow d(q, f(y)) < \varepsilon \end{aligned}$$

5.2 Continuity and Compactness

$f : \mathbb{R}^n \to \mathbb{R}^l$

Theorem 40 *Suppose f is continuous. Then, $X \subset \mathbb{R}^n$ is compact $\Rightarrow f(X)$ is compact.*

Theorem 41 (Weierstrass Theorem) *$f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$. If f is continuous on S, and if S is compact, then there exist points $p, q \in S$ such that*

$$\begin{aligned} f(p) &= \sup_{x \in S} f(x) = \max_{x \in S} f(x) \\ f(q) &= \inf_{x \in S} f(x) = \min_{x \in S} f(x) \end{aligned}$$

5.3 Continuity and Connectedness

$f : \mathbb{R}^n \to \mathbb{R}^l$

Theorem 42 *Suppose f is continuous. $X \subset \mathbb{R}^n$ is connected $\Rightarrow f(X)$ is connected.*

Theorem 43 (Intermediate Value Theorem) $f : [a, b] \to \mathbb{R}$. *If f is continuous on $[a, b]$ and if there exists c with $f(a) < c < f(b)$, then there exists a point $x \in [a, b]$ such that $f(x) = c$.*

- **Note:** The intermediate value theorem implies the following statement (Fixed Point Theorem):

  For a continuous function f from $[a, b]$ to $\mathbb{R}$, define $g(x) = f(x) - x$. Then,

$$\begin{aligned} \exists x_1, x_2 &\in [a, b] : x_1 > x_2, g(x_1) > 0 \text{ and } g(x_2) < 0 \\ &\Rightarrow \exists \bar{x} \in (x_1, x_2) : g(\bar{x}) = 0, \text{ (i.e., } f(\bar{x}) = \bar{x}). \end{aligned}$$

6 Differentiation

6.1 Differentiation of a Function from $\mathbb{R}$ to $\mathbb{R}$

$f : (a, b) \to \mathbb{R}$

Definition 34 *A function* f *is differentiable at* $x \in (a, , b)$ *if* $\lim_{t\to x} \frac{f(t)-f(x)}{t-x}$ *exists.*

Definition 35 *If* f *is differentiable at* x*, we call* $\lim_{t\to x} \frac{f(t)-f(x)}{t-x}$ *the derivative of* f *at* x *and write* $f'(x)$*.*

Theorem 44 *If* f *is differentiable at* $x \in (a, b)$*, then* f *is continuous at* x*.*

Theorem 45 *Let* $f : (a, b) \to \mathbb{R}$ *and* $g : (a, b) \to \mathbb{R}$ *be both differentiable at* $x \in (a, b)$*. Then,* $f + g$*,* $f \cdot g$*, and* $\frac{f}{g}$ *are all differentiable at* x*.*

Theorem 46 *Let* $f : (a, b) \to \mathbb{R}$ *be continuous in* (a, b) *and differentiable at* $x \in (a, b)$*. Let* $g : f((a, b)) \to \mathbb{R}$ *be differentiable at* $f(x)$*. Then, the function* $h(x) \equiv g(f(x))$ *is differentiable at* x *and the derivative of* h *at* x *is given by*

$$h'(x) = g'(f(x))f'(x).$$

- **Note:** If a function $f : (a, b) \to \mathbb{R}$ is differentiable in (a, b), then f' is also a function from (a, b) to $\mathbb{R}$. If f' is continuous on (a, b), we call f is continuously differentiable (or f is C^1).

6.2 Mean Value Theorems

$f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$.

Definition 36 $x \in S$ *is a local maximum (minimum) of* f *if* $\exists\delta > 0$ *such that* $\forall y \in S$ *with* $d(x, y) < \delta$*,* $f(y) \leq (\geq) f(x)$*.*

$f : [a, b] \to \mathbb{R}$

Theorem 47 *If* f *has a local maximum at* $x \in [a, b]$ *and if* $f'(x)$ *exists, then* $f'(x) = 0$*.*

Theorem 48 (Rolle's Theorem) *Let f be continuous on$[a,b]$ and differentiable on (a,b). If $f(a)=f(b)=0$, then $\exists x\in(a,b)$ such that $f'(x)=0$.*

Theorem 49 (The Mean Value Theorem) *Let f be continuous on $[a,b]$ and differentiable on (a,b). Then, $\exists x\in(a,b)$ such that*

$$f(b)-f(a)=f'(x)(b-a).$$

- **Note:**

$$\begin{array}{rcl} f(b)-f(a) & = & f'(x)(b-a) \\ & \Leftrightarrow & \\ \dfrac{f(b)-f(a)}{b-a} & = & f'(x) \end{array}$$

- **Note:**

If $f(a)=f(b)=0$, the Mean Value Theorem implies

$$f'(x)=0.$$

Theorem 50 (Cauchy Mean Value Theorem) *Let $f:[a,b]\to\mathbb{R}$ and $g:[a,b]\to\mathbb{R}$ be both continuous on $[a,b]$ and differentiable on (a,b). Then, $\exists x\in(a,b)$ such that*

$$f'(x)[g(b)-g(a)]=g'(x)[f(b)-f(a)].$$

- **Note:**

If $g(x)=x$, then Cauchy Mean Value Theorem implies

$$f'(x)(b-a)=(f(b)-f(a)).$$

Corollary 5 *Let $f:(a,b)\to\mathbb{R}$ be differentiable in (a,b). Then,*

$$\begin{array}{rcl} \textit{(i) } f'(x) & \geq & 0, \forall x\in(a,b)\Rightarrow f \textit{ is monotonically increasing in } (a,b); \\ \textit{(ii) } f'(x) & = & 0, \forall x\in(a,b)\Rightarrow f \textit{ is constant in } (a,b); \\ \textit{(iii) } f'(x) & \leq & 0, \forall x\in(a,b)\Rightarrow f \textit{ is monotonically decreasing in } (a,b). \end{array}$$

6.3 L'Hospital's Rule

Theorem 51 *Let $f:(a,b)\to\mathbb{R}$ be differentiable in (a,b) and $g:(a,b)\to\mathbb{R}$ be differentiable in (a,b) with $g'(x)\neq 0$ for all $x\in(a,b)$. (Note: $-\infty\leq a<b\leq\infty$.) Suppose $\frac{f'(x)}{g'(x)}\to A$ as $x\to a$. If $f(x)\to 0$ and $g(x)\to 0$ as $x\to a$, or if $g(x)\to+\infty(-\infty)$ as $x\to a$, then*

$$\frac{f(x)}{g(x)}\to A \text{ as } x\to a.$$

6.4 Higher Order Derivatives and Taylor's Theorem

Definition 37 *Let $f : (a,b) \to \mathbb{R}$ be differentiable in (a,b). If $f' : (a,b) \to \mathbb{R}$ is differentiable at $x \in (a,b)$, we denote the derivative of f' at x by $f''(x)$ and call it the second order derivative of f at x. Similarly, we can define the nth order derivative of f at x and denote it by $f^{(n)}(x)$.*

Theorem 52 (Taylor's Theorem) *Let $n \in \mathbb{N}$ and let $f : (a,b) \to \mathbb{R}$ be a function such that $f^{(n-1)}$ is continuous on (a,b) and $f^{(n)}(t)$ exists for every $t \in (a,b)$. Then, for any $\alpha, \beta \in (a,b)$, there exists $\gamma \in (a,b)$ such that*

$$f(\beta) = f(\alpha)+\frac{f'(\alpha)}{1!}(\beta-\alpha)+\frac{f''(\alpha)}{2!}(\beta-\alpha)^2+\cdots+\frac{f^{(n-1)}(\alpha)}{(n-1)!}(\beta-\alpha)^{n-1}+\frac{f^{(n)}(\gamma)}{n!}(\beta-\alpha)^n.$$

6.5 Differentiation of Functions from $\mathbb{R}^n$ to $\mathbb{R}^l$

$f : S \to \mathbb{R}^l$, $S \subset \mathbb{R}^n$.

Definition 38 *f is differentiable at $x \in S$ if there exists an $l \times n$ matrix A such that*

$$\lim_{y \to x} \frac{\|f(y) - f(x) - A(y-x)\|}{\|y-x\|} = 0.$$

The matrix A is called the derivative of f at x and denoted $Df(x)$.

Theorem 53 *Let $f : \mathbb{R}^n \to \mathbb{R}^l$ be differentiable at $x \in \mathbb{R}^n$ and $g : \mathbb{R}^l \to \mathbb{R}^m$ be differentiable at $f(x) \in \mathbb{R}^m$. Then, $g \circ f$ is differentiable at x and*

$$D(g \circ f)(x) = Dg(f(x))Df(x).$$

Theorem 54 *Let $f = (f_1, ..., f_l)$ and $f_i : \mathbb{R}^n \to \mathbb{R}$. f is differentiable at $x \Leftrightarrow f_i$ is differentiable for all $i = 1, 2, ..., l$.*

6.6 Partial Derivatives and Differentiability

6.6.1 Partial Derivatives

$f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$.

- Define $e_j \in \mathbb{R}^n$:

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow j\text{th component} \\ \\ \\ \\ \end{matrix}$$

Definition 39 *The j-th partial derivative of f exists at x if*

$$\lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t}$$

exists and denote

$$\frac{\partial f}{\partial x_j}(x) = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t}.$$

Theorem 55 *f is differentiable at $x \Rightarrow$ all partials $\frac{\partial f}{\partial x_j(x)}$ exist at x and*

$$Df(x) = \left(\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_n}(x) \right).$$

- **Note:** f is differentiable $\underset{\not\Leftarrow}{\Rightarrow}$ f has all partials.

Theorem 56 *If all partials of f exists at x, and f is continuous at x, then $Df(x)$ exists and*

$$Df(x) = \left(\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_n}(x) \right).$$

6.6.2 Directional Derivatives

$f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$.

Definition 40 *The directional derivative of f at x in the direction h (with $\|h\| = 1$) is defined as*

$$Df(x; h) = \lim_{t \to +0} \frac{f(x + th) - f(x)}{t}.$$

- **Note:** When the condition $t \to +0$ is replaced with $t \to 0$, we obtain "two-sided directional derivatives", and the two-sided directional derivative of f in the direction $h = e_j$ is equal to the jth partial derivative of f.

6.7 Higher Order Derivatives of Functions from $\mathbb{R}^n$ to $\mathbb{R}$

- $f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$ is open
- If f is differentiable in S, then the derivative

  $$Df = \left(\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}\right)$$

  is a function from S to $\mathbb{R}^n$.
- If $\frac{\partial f}{\partial x_i}$ is differentiable, we can denote the partial of $\frac{\partial f}{\partial x_i}$ in the direction e_j at x by $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$.
- If Df is differentiable, the second-order derivative of f at x is

  $$D^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix} : \text{Hessian matrix of } f \text{ at } x.$$
- When $\frac{\partial^2 f}{\partial x_i \partial x_j} : S \to \mathbb{R}$ is continuous for each i and j, we say that f is twice continuously differentiable or f is C^2.

Theorem 57 *f is $C^2 \Rightarrow D^2 f$ is symmetric, i.e.,*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x), \text{ for all } i, j \text{ and for all } x \in S.$$

Theorem 58 (Taylor's Theorem in $\mathbb{R}^n$)

$$\begin{aligned} \textit{(i) } f \textit{ is } C^1 &\Rightarrow \\ f(y) &= f(x) + Df(x) \cdot (y - x) + R_1(x, y), \\ \textit{where } \lim_{y \to x} \frac{R_1(x, y)}{\|y - x\|} &= 0. \end{aligned}$$

$$\begin{aligned} \textit{(ii) } f \textit{ is } C^2 &\Rightarrow \\ f(y) &= f(x) + Df(x) \cdot (y - x) + \frac{1}{2}(y - x)' \cdot D^2 f(x) \cdot (y - x) + R_2(x, y), \\ \textit{where } \lim_{y \to x} \frac{R_2(x, y)}{\|y - x\|} &= 0. \end{aligned}$$

6.8 Homogeneous Functions and the Euler Theorem

6.8.1 Homogeneous Functions

$f : \mathbb{R}^n \to \mathbb{R}$

Definition 41 *A function f is homogeneous of degree $r \in \mathbb{R}_+$ if for any $t > 0$ we have $f(tx_1, tx_2, ..., tx_n) = t^r f(x_1, ..., x_n)$.*

- $f(x_1, x_2) = \frac{x_1}{x_2}$ is homogeneous of degree 0.
- $f(x_1, x_2) = x_1 x_2$ is homogeneous of degree 2.

Theorem 59 *If $f(x_1, ..., x_n)$ is homogeneous of degree r and differentiable in $\mathbb{R}^n$, then for any $i = 1, 2, ..., n$, the i-th partial derivative*

$$\frac{\partial f}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$$

is homogeneous of degree $r - 1$.

Theorem 60 *f is homogeneous of any degree and*

$$f(x_1, ..., x_n) = f(x'_1, ..., x'_n).$$

Then,

$$f(tx_1, ..., tx_n) = f(tx'_1, ..., tx'_n).$$

Definition 42 *A level set of a function is the set $\{x \in \mathbb{R}^n : f(x) = k\}$ for some $k \in \mathbb{R}$.*

Theorem 61 *Let f be homogeneous of degree r. Then, the slope of the level sets of f is unchanged along any ray through the origin.*

6.8.2 Homothetic Functions

Definition 43 *Let f be homogeneous of degree r and h be a increasing function from $\mathbb{R}$ to $\mathbb{R}$. Then, $h \circ f$ is called a homothetic function.*

- $h(y) = \log y$ and $f(x_1, x_2) = x_1 x_2$. Then, $h \circ f(x_1, x_2) = \log(x_1 x_2)$ is homothetic but not homogeneous.

Theorem 62 *The family of level sets of $h \circ f$ coincides with the family of level sets of f.*

Theorem 63 *For any homothetic function, the slope of the level sets is unchanged along rays through the origin.*

6.8.3 Euler's Formula

Theorem 64 (Euler's Formula) *$f(x_1, ..., x_n)$ is homogeneous of degree r and differentiable. Then, at any $\bar{x} = (\bar{x}_1, ..., \bar{x}_n) \in \mathbb{R}^n$,*

$$\sum_{i=1}^{n} \frac{\partial f\left(\bar{x}_1, ..., \bar{x}_n\right)}{\partial x_i} \bar{x}_i = r f(\bar{x}_1, ..., \bar{x}_n).$$

7 Integral

7.1 Definitions and Existence

Definition 44 *Given $[a, b]$. A partition P of $[a, b]$ is a finite set of points $\{x_0, x_1, ..., x_n\}$, where $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$.*

- $\Delta x_i = x_i - x_{i-1}$
- $f : [a, b] \to \mathbb{R}$, bounded function

Definition 45 (Rieman Integral) *Given a partition P. Let*

$$\begin{aligned}
M_i &= \sup\{f(x) : x_{i-1} \leq x \leq x_i\}; \\
m_i &= \inf\{f(x) : x_{i-1} \leq x \leq x_i\}; \\
U(P, f) &= \sum_{i=1}^{n} M_i \Delta x_i; \\
L(P, f) &= \sum_{i=1}^{n} m_i \Delta x_i; \\
\overline{\int_a^b} f dx &= \inf_P \{U(P, f) : P \textit{ is a partition of } [a, b]\}\textit{: upper Rieman integral;} \\
\underline{\int_a^b} f dx &= \sup_P L(P, f) : P \textit{ is a partition of } [a, b]\}\textit{: lower Rieman integral.}
\end{aligned}$$

If $\overline{\int}_a^b f dx = \underline{\int}_a^b f dx$, we say that f is Rieman integrable on $[a, b]$, and write $f \in \mathcal{R}$. Denote

$$\int_a^b f dx \equiv \overline{\int_a^b} f dx \left(= \underline{\int_a^b} f dx \right).$$

- **Note:** f is bounded $\Rightarrow c \leq f(x) \leq d$ for any $x \in [a, b]$. Thus, for any partition
$$c(b - a) \leq L(P, f) \leq U(P, f) \leq d(b - a).$$

Definition 46 (Rieman-Stieltjes Integral) *Let α be a monotonically increasing function on $[a, b]$. Given a partition P,*

$$\Delta \alpha_i = \alpha(x_i) - \alpha(x_{i-1}) > 0.$$

For any bounded function f, define

$$\begin{aligned} U(P,f,\alpha) &= \sum_{i=1}^{n} M_i \Delta\alpha_i; \\ L(P,f,\alpha) &= \sum_{i=1}^{n} m_i \Delta\alpha_i; \\ \overline{\int_a^b} f d\alpha &= \inf_P U(P,f,\alpha); \\ \underline{\int_a^b} f d\alpha &= \sup_P L(P,f,\alpha). \end{aligned}$$

If $\overline{\int_a^b} f d\alpha = \underline{\int_a^b} f d\alpha$, we say that f is integrable with respect to α and write $f \in \mathcal{R}(\alpha)$ and denote

$$\int_a^b f d\alpha \equiv \overline{\int_a^b} f d\alpha.$$

- Example of $\alpha(x)$: Distribution function

$$\alpha(x) = F(x) = \Pr[z \le x].$$

Definition 47 $f : S \to \mathbb{R}^l$, $S \subset \mathbb{R}^n$. *f is uniformly continuous on S if* $\forall \varepsilon > 0, \exists \delta > 0$ *such that* $d(x,y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$.

- $f(x) = x^2$ is not uniformly continuous on $\mathbb{R}$.

Theorem 65 *If f is continuous on a compact subset X of $\mathbb{R}^n$, then f is uniformly continuous on X.*

Lemma 66 *$f \in \mathcal{R}(\alpha)$ on $[a,b]$ if and only if for any $\varepsilon > 0$ there exists a partition P such that*

$$U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon.$$

Theorem 67 *If f is continuous on $[a,b]$, then $f \in \mathcal{R}(\alpha)$ on $[a,b]$.*

Theorem 68 *If f is monotonic on $[a,b]$ and α is continuous on $[a,b]$, then $f \in \mathcal{R}(\alpha)$ on $[a,b]$.*

Theorem 69 *If f is bounded on $[a,b]$, f has only finitely many points of discontinuity on $[a,b]$, and α is continuous at every point at which f is discontinuous, then $f \in \mathcal{R}(\alpha)$ on $[a,b]$.*

- **Note:** The following function is not (Riemann) integrable:

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{otherwise} \end{cases}.$$

7.2 Properties of Integral

Theorem 70 *$f_1, f_2 \in \mathcal{R}(\alpha)$ on $[a,b] \Rightarrow f_1 + f_2 \in \mathcal{R}(\alpha)$ on $[a,b]$ and*

$$\int_a^b (f_1 + f_2) d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha.$$

Theorem 71 *$f \in \mathcal{R}(\alpha)$ on $[a,b]$ and $c \in \mathbb{R} \Rightarrow cf \in \mathcal{R}(\alpha)$ on $[a,b]$ and*

$$\int cf d\alpha = c \int f d\alpha.$$

Theorem 72 *$f_1(x) \leq f_2(x)$ for all $x \in [a,b] \Rightarrow \int_a^b f_1 d\alpha \leq \int_a^b f_2 d\alpha$.*

Theorem 73 *$f \in \mathcal{R}(\alpha)$ on $[a,b]$, $a < c < b \Rightarrow \int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha$*

Theorem 74 *$f \in \mathcal{R}(\alpha_1)$ and $f \in \mathcal{R}(\alpha_2)$ on $[a,b] \Rightarrow f \in \mathcal{R}(\alpha_1 + \alpha_2)$ on $[a,b]$ and*

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2.$$

Theorem 75 *$f \in \mathcal{R}(\alpha)$ on $[a,b]$ and $c \in \mathbb{R} \Rightarrow f \in \mathcal{R}(c\alpha)$ on $[a,b]$ and*

$$\int_a^b f d(c\alpha) = c \int_a^b f d\alpha.$$

Theorem 76 *Assume $\alpha' \in \mathcal{R}$ on $[a,b]$. Then, $f \in \mathcal{R}(\alpha)$ if and only if $f\alpha' \in \mathcal{R}$. In that case*

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x) dx.$$

Theorem 77 (Change of Variable) *Suppose φ is a strictly increasing continuous function that maps an interval $[A,B]$ onto $[a,b]$. Suppose $f \in \mathcal{R}(\alpha)$ on $[a,b]$. Define β and g on $[A,B]$ by*

$$\begin{aligned} \beta(y) &= \alpha(\varphi(y)); \\ g(y) &= f(\varphi(y)). \end{aligned}$$

Then, $g \in \mathcal{R}(\beta)$ and

$$\int_A^B g d\beta = \int_a^b f d\alpha.$$

7.3 Fundamental Theorem of Calculus

Theorem 78 *$f \in \mathcal{R}$ on $[a,b]$. For $a \leq x \leq b$, put $F(x) = \int_a^x f(t)dt$. Then, F is continuous on $[a,b]$. Furthermore, if f is continuous at $x_0 \in [a,b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.*

Theorem 79 (The Fundamental Theorem of Calculus) *$f \in \mathcal{R}$ on $[a,b]$. If there exists a differentiable function on $[a,b]$ such that $F' = f$, then*

$$\int_a^b f dx = F(b) - F(a).$$

Theorem 80 (Integration by Parts) *Suppose F and G are differentiable function on $[a,b]$, $F' = f \in \mathcal{R}$ and $G' = g \in \mathcal{R}$. Then,*

$$\int_a^b F(x)g(x)dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx.$$

8 Matrix Algebra

8.1 Sum, Product, Transpose

- $A : n \times m$ matrix
- $A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$
- $A_i^r = [a_{i1} \cdots a_{im}]$: i th raw of A
- $A_j^c = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}$: j th column of A
- $A = \begin{bmatrix} A_1^r \\ \vdots \\ A_n^r \end{bmatrix} = [A_1^c \cdots A_m^c]$
- $A : n \times m$, $B : n \times m$

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1m} + b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nm} + b_{nm} \end{bmatrix}$$

- $A : n \times m$, $B : m \times k$

$$\begin{aligned} AB &= \begin{bmatrix} A_1^r B_1^c & \cdots & A_1^r B_k^c \\ \vdots & \ddots & \vdots \\ A_n^r B_1^c & \cdots & A_n^r B_k^c \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^m a_{1i} b_{i1} & \cdots & \sum_{i=1}^m a_{1i} b_{ik} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{ni} b_{i1} & \cdots & \sum_{i=1}^m a_{ni} b_{ik} \end{bmatrix} \end{aligned}$$

  - **Note:** BA is not defined unless $n = k$, and even if BA is defined, BA may be different from AB.

Theorem 81

$$\begin{aligned} A + B &= B + A; \\ (A + B) + C &= A + (B + C); \\ (AB)C &= A(BC); \\ A(B + C) &= AB + AC. \end{aligned}$$

Definition 48 *For an $n \times m$ matrix A, the transpose of A is defined by*

$$A' = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ & \ddots & \\ \vdots & a_{ji} & \vdots \\ & & \ddots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}$$

Theorem 82

$$\begin{aligned} (A+B)' &= A' + B'; \\ (AB)' &= B'A'. \end{aligned}$$

8.2 Some Important Classes of Matrices

1. Square Matrix:

   $$\begin{aligned} A &: n \times n \\ A &= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \end{aligned}$$

   a_{ij} is called a diagonal element of A and a_{ij} for $i \neq j$ is called off diagonal element of A.

2. Symmetric Matrix:

   $$\begin{aligned} A &: n \times n \\ a_{ij} &= a_{ji} \text{ for all } i, j = 1, 2, ..., n \end{aligned}$$

   (a) Remark: A is symmetric $\Leftrightarrow A = A'$

3. Diagonal Matrix:

   $$\begin{aligned} A &: n \times n \\ A &= \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a_{nn} \end{bmatrix} \end{aligned}$$

   A diagonal matrix is a matrix such that all the off diagonal elements are 0.

4. Identity Matrix:

$$\begin{aligned} I_n &: n \times n \\ I_n &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

   An identity matrix is a diagonal matrix such that all of its diagonal elements are 1.

   (a) Remark: For any $n \times n$ matrix A,

$$\begin{aligned} I_n A &= A; \\ A I_n &= A. \end{aligned}$$

5. Lower- and Upper-Triangular Matrix:

$$\begin{aligned} A &: n \times n \\ A &= \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & & 0 \\ \vdots & & & \ddots & \vdots \\ a_{n1} & & \cdots & & a_{nn} \end{bmatrix} : \text{ lower triangular matrix} \\ A &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & & a_{nn} \end{bmatrix} : \text{ upper triangular matrix} \end{aligned}$$

   (a) Remark: If A is a lower triangular matrix, then A' is a upper triangular matrix.

6. Idempotent Matrix:

$$\begin{aligned} A &: n \times n \\ AA &= A \end{aligned}$$

8.3 Trace of an Matrix

- $A : n \times n$
- $trace(A) = a_{11} + a_{22} + \cdots + a_{nn}$
- Properties:

1. $A : n \times n$ and $B : n \times n \Rightarrow trace(AB) = trace(BA)$
2. $A : n \times n$ and $B : n \times n \Rightarrow trace(A + B) = trace(A) + trace(B)$
3. For $\lambda \in \mathbb{R}$,
$$trace(\lambda A) = \lambda trace(A)$$

8.4 Partitioned Matrix

- Example:
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ a_1 & a_2 \end{bmatrix}$$
$$\begin{aligned} A_1 &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ A_2 &= \begin{bmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{bmatrix} \\ a_1 &= \begin{bmatrix} a_{31} & a_{32} \end{bmatrix} \\ a_2 &= \begin{bmatrix} a_{33} & a_{34} \end{bmatrix} \end{aligned}$$
- In general
$$A = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nm} \end{bmatrix}$$
- Block diagonal matrix:
$$A = \begin{bmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_n \end{bmatrix}$$
- Remark:
  1. $A + B = \begin{bmatrix} A_{11} + B_{11} & \cdots & A_{1m} + B_{1m} \\ \vdots & \ddots & \vdots \\ A_{n1} + B_{n1} & \cdots & A_{nm} + B_{nm} \end{bmatrix}$
  2. $AB = \begin{bmatrix} \sum_{i=1}^{m} A_{1i}B_{i1} & \cdots & \sum_{i=1}^{m} A_{1i}B_{il} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} A_{ni}B_{i1} & \cdots & \sum_{i=1}^{m} A_{ni}B_{il} \end{bmatrix}$

8.5 Elementary Row Operations

$A : n \times n$

1. Multiplying a scalar for each element of a row of A:

$$A \xrightarrow{I} \begin{bmatrix} A_1^r \\ \vdots \\ \alpha A_i^r \\ \vdots \\ A_n^r \end{bmatrix} = E^I(\alpha; i)A$$

$$E^I(\alpha; i) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \alpha_{ii} & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

2. Interchanging two rows of A:

$$A \xrightarrow{II} \begin{bmatrix} A_1^r \\ \vdots \\ A_{i-1}^r \\ A_j^r \\ A_{i+1}^r \\ \vdots \\ A_{j-1}^r \\ A_i^r \\ A_{j+1}^r \\ \vdots \\ A_n^r \end{bmatrix} = E^{II}(i, j)A$$

$$E^{ii}(i, j) = \begin{bmatrix} 1 & \cdots & 0 & & \cdots & & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & & & \vdots & & \vdots \\ 0 & \cdots & 0_{ii} & & \cdots & & 1_{ij} & \cdots & 0 \\ & & & 1 & & & & & \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ & & & & & 1 & & & \\ 0 & \cdots & 1_{ji} & & \cdots & & 0_{jj} & \cdots & 0 \\ \vdots & & \vdots & & & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & & \cdots & & 0 & \cdots & 1 \end{bmatrix}$$

3. Adding a scalar multiple of one row to another:

$$A \overset{III}{\to} \begin{bmatrix} A_1^r \\ \vdots \\ A_{j-1}^r \\ \alpha A_i^r + A_j^r \\ A_{j+1}^r \\ \vdots \\ A_n^r \end{bmatrix} = E^{III}(\alpha; i, j)A$$

$$E^{III}(\alpha; i, j) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots \alpha_{ji} \cdots & 1_{ii} & & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

Theorem 83 *After finite elementary row operations of A, we can obtain an upper triangular matrix.*

Theorem 84 *Any upper triangular matrix with no 0's on its diagonal can be changed into I_n after finite elementary row operations.*

8.6 Inverse Matrix

$A : n \times n$

Definition 49 *The inverse of A is defined to be an $n \times n$ matrix B such that*

$$AB = BA = I_n.$$

Theorem 85 *An $n \times n$ matrix A can have at most one inverse.*

- **Note:** Because of the theorem, the inverse is uniquely determined by A. So, we can denote the inverse by A^{-1}.

Definition 50 *If A has the inverse, it is called nonsingular or invertible.*

Theorem 86

$$\begin{aligned} \text{(i) } [\alpha A]^{-1} &= \frac{1}{\alpha} A^{-1}, \text{ where } \alpha \neq 0 \text{ is a scalar.} \\ \text{(ii) } (AB)^{-1} &= B^{-1}A^{-1} \end{aligned}$$

Theorem 87 *Any elementary matrix is invertible.*

- **Note:** The inverse of any elementary matrix is also an elementary matrix.

$$\begin{aligned} \left[E^{I}(\alpha; i)\right]^{-1} &= E^{I}(\frac{1}{\alpha}; i) \\ \left[E^{II}(i,j)\right]^{-1} &= E^{II}(j,i) \\ \left[E^{III}(\alpha; i, j)\right]^{-1} &= E^{III}(-\alpha; i, j) \end{aligned}$$

8.7 The Determinant

8.7.1 Definition

$A : n \times n$

Definition 51 *The determinant of A, $|A|$, is defined by*

(i) If A is 1×1, $|A| = a_{11}$.

(ii) If $n \geq 2$,

$$|A| = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \left|A_{1j}\right|,$$

where A_{ij} is the $(n-1) \times (n-1)$ matrix formed by deleting row i and column j from A.

- Example:

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ A_{11} &= \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \\ A_{23} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \\ |A| &= (-1)^2 a_{11} \left|A_{11}\right| + (-1)^3 a_{12} \left|A_{12}\right| + (-1)^4 \left|A_{13}\right| \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

Theorem 88

$$\begin{aligned} |A| &= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \left|A_{ik}\right| : \text{ \textit{expansion along row } } i \\ &= \sum_{h=1}^{n} (-1)^{h+j} a_{hj} \left|A_{hj}\right| : \text{ \textit{expansion along column } } h \end{aligned}$$

8.7.2 Properties of Determinants:

1. $|A| = |A'|$.
2. The determinant of a lower- (upper-) triangular matrix A is $|A| = a_{11}a_{22}\cdots a_{nn}$.
3. Given an $n \times n$ matrix A and a scalar α, construct a matrix B_i by

$$B_i = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ \alpha a_{i1} & \alpha a_{ii} & \alpha a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Then,

$$|B_i| = \alpha |A|.$$

4. $|\alpha A| = \alpha^n |A|$.
5. Given an $n \times n$ matrix A, construct an matrix B_{ij} by interchanging A's ith row and jth row:

$$Bij = E^{II}(i,j)A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \cdots & \vdots \\ a_{i1} & & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

Then,

$$|A| = -|B_{ij}|.$$

6. If two rows of A are equal,

$$|A| = 0.$$

7. Suppose two matrices A and B differ only in their ith row. Then,

$$C = \begin{pmatrix} A_1^r \\ \vdots \\ A_i^r + B_i^r \\ \vdots \\ A_n^r \end{pmatrix} \Rightarrow |C| = |A| + |B|.$$

8. $|AB| = |A|\,|B|$.

9. If A is invertible,
$$\left|A^{-1}\right| = \frac{1}{|A|}.$$

8.7.3 Adjoint of a Matrix

Definition 52 *The adjoint of a matrix A is $n \times n$ matrix whose (i,j)th element is given by $(-1)^{i+j}\left|A_{ji}\right|$.*

- Example:

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \\ adjA &= \begin{bmatrix} (-1)^2 a_{22} & (-1)^3 a_{12} \\ (-1)^3 a_{21} & (-1)^4 a_{11} \end{bmatrix} \end{aligned}$$

Theorem 89 *For any $n \times n$ matrix A,*
$$A \cdot adjA = \begin{pmatrix} |A| & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |A| \end{pmatrix}.$$

Theorem 90 *If $|A| \neq 0$, then*
$$A^{-1} = \frac{1}{|A|} adjA.$$

Corollary 6 *A is invertible $\Leftrightarrow |A| \neq 0$.*

8.7.4 The determinant and the inverse of a partitioned matrix

$$\begin{aligned} A &= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \\ |A| &= |A_{11}|\,|A_{22}| \\ A^{-1} &= \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} A &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ |A| &= |A_{22}|\left|A_{11} - A_{12}A_{22}^{-1}A_{21}\right| \\ &= |A_{11}|\left|A_{22} - A_{21}A_{11}^{-1}A_{12}\right| \\ A^{-1} &= \begin{bmatrix} A_{11}^{-1}(I + A_{12}F_2A_{21}A_{11}^{-1}) & -A_{11}^{-1}A_{12}F_2 \\ -F_2A_{21}A_{11}^{-1} & F_2 \end{bmatrix} \\ \text{where } F_2 &= (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}. \end{aligned}$$

8.7.5 The determinant and the inverse of a Kronecker product

- Kronecker product of two matrices $A\,(n \times m)$ and $B\,(k \times l)$:

$$\underset{(nk \times ml)}{A \otimes B} = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix}$$

1. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
2. If A is $m \times m$ and B is $n \times n$, then

$$|A \otimes B| = |A|^n \, |B|^m \, .$$

3. $(A \otimes B)' = A' \otimes B'$.
4. $trace\,(A \otimes B) = trace\,(A)\, trace\,(B)\,.$

8.8 Systems of Linear Equations

$$\left\{ \begin{array}{c} a_{11}x_1 + \cdots + a_{1m}x_m = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m = b_n \end{array} \right. \qquad \Rightarrow Ax = b.$$

Theorem 91 (Cramer's Rule) *Let A be an $n \times n$ matrix. If the linear equation system $Ax = b$ has unique solution, the unique solution x is given by*

$$x_i = \frac{|B_i|}{|A|},$$

where B_i is obtained by replacing ith column of A by b:

$$B_i = \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix}.$$

9 Subspaces Attached to a Matrix

9.1 Vector Space

Definition 53 *A set V is a vector space if there exists a summation operation $(+)$ and a scalar multiplication, satisfying the following properties:*

1. $\forall u, v, w \in V$,
$$(u+v) = w = u + (v+w).$$
2. $\exists 0 \in V$ such that for any $u \in V$
$$0 + u = u + 0 = u.$$
3. $\forall u \in V, \exists u' \in V$ such that
$$u + u' = 0$$
where we denote $u' = -u$.
4. $\forall u, v \in V$,
$$u + v = v + u.$$
5. $\forall c \in \mathbb{R}, \forall u, v \in V$,
$$c(u+v) = cu + cv.$$
6. $\forall a, b \in \mathbb{R}, \forall u \in V$,
$$(a+b)u = au + bu.$$
7. $\forall a, b \in \mathbb{R}, \forall u \in V$,
$$(ab)u = a(bu).$$
8. $\forall u \in V$,
$$1u = u, \quad (1 \in \mathbb{R}).$$

Definition 54 *Let V be a vector space. A subset W of V is a subspace of V if it is a vector space.*

Theorem 92 *Let V be a vector space. $W \subset V$ is a subspace of V if and only if $\forall v, w \in V, \forall c \in \mathbb{R}$,*

$$\begin{aligned} \textit{(i) } v + w &\in W; \ \textit{and} \\ \textit{(ii) } cv &\in W. \end{aligned}$$

- Example:
  - $\mathbb{R}^n$ is a vector space.
  - Any plane (or line) in $\mathbb{R}^n$ including the origin is a subspace of $\mathbb{R}^n$.

- Any element in a vector space is called a *vector*.

Definition 55 (Linear combination of vectors) *Given a vector space V, a set of scalars $\{\alpha_1, ..., \alpha_m\}$, and a set of vectors $\{v_1, ..., v_m\}$ in V,*

$$\alpha_1 v_1 + \cdots \alpha_m v_m \in V$$

is called a linear combination of $\{v_1, ..., v_m\}$.

Definition 56 *Given a vector space V and a set of vectors $\{v_1, ..., v_m\}$, define*

$$W \equiv \{x \in V : x = \alpha_1 v_1 + \cdots + \alpha_m v_m, \alpha_i \in \mathbb{R}\}.$$

Then, W is a subspace of V and called the subspace generated (spanned) by $v_1, ..., v_m$. We denote the subspace generated by $v_1, ..., v_m$ by $L(v_1, ..., v_n)$.

Definition 57 (Linear independence) *Let V be a vector space and $\{v_1, ..., v_m\}$ be a set of vectors in V. We say that $v_1, ..., v_n$ are linearly independent if*

$$\sum_{i=1}^{m} c_i v_i = 0(\in V) \Leftrightarrow c_1 = c_2 = \cdots = c_n = 0(\in \mathbb{R}).$$

$v_1, ..., v_m$ in V are said to be linearly dependent if they are not linearly independent. (i.e., $\exists (c_1, ..., c_m) \neq 0$ such that $\sum_{i=1}^{m} c_i v_i = 0$.)

- **Note:** For linearly dependent vectors, $\exists c_i \neq 0$ such that

  $$v_i = \sum_{j \neq i} -\frac{c_j}{c_i} v_j.$$

  That is, at least one vector can be represented by a linear combination of the others.

Definition 58 (Basis of a vector space) *Let V be a vector space and $\{v_1, ..., v_m\}$ be a set of vectors in V. If $L(v_1, ..., v_m) = V$ and $v_1, ..., v_m$ are linearly independent, then we call the set of vectors $\{v_1, ..., v_m\}$ a basis of V.*

Definition 59 *When a vector space V has a basis of m vectors, m is called the dimension of V, and we write $\dim V = m$.*

Theorem 93 *Let V be a vector space. If $\{v_1, ..., v_n\}$ generates V, and $\{v_1, ..., v_r\}$ is a maximal subset of $\{v_1, ..., v_n\}$ that is composed of linearly independent elements, then $\{v_1, ..., v_r\}$ is a basis of V.*

- **Note:** $\{v_1, ..., v_r\}$ is a maximal subset of $\{v_1, ..., v_n\}$ that is composed of linearly independent elements $\Leftrightarrow$

    1. $\{v_1, ..., v_r\}$ is a subset of $\{v_1, ..., v_n\}$;
    2. $v_1, ..., v_r$ are linearly independent; and
    3. There is no other subset of $\{v_1, ..., v_n\}$ that is composed of linearly independent vectors and larger than $\{v_1, ..., v_r\}$.

9.2 Row and Column Spaces of a Matrix

- $A : n \times m$ matrix

$$A = \begin{pmatrix} A_1^r \\ \vdots \\ A_n^r \end{pmatrix} = \begin{pmatrix} A_1^c & \cdots & A_m^c \end{pmatrix},$$

where

$$\begin{aligned} A_i^r &\in \mathbb{R}^m \text{ is called a row vector of } A; \text{ and} \\ A_j^c &\in \mathbb{R}^n \text{ is called a column vector of } A. \end{aligned}$$

- Given a matrix A, define $Row(A)$ be the subspace of $\mathbb{R}^m$ generated by vectors $\{A_1^r, ..., A_n^r\}$, and $Col(A)$ be the subspace of $\mathbb{R}^n$ generated by vectors $\{A_1^c, ..., A_m^c\}$.

Definition 60 *The row rank of a matrix A is the dimension of $Row(A)$, and the column rank of A is the dimension of $Col(A)$.*

- **Note:** By the previous theorem, the row rank is the maximum number of linearly independent row vectors, and the column rank is the maximum number of linearly independent column vectors.

9.2.1 Row Rank

Definition 61 *A row of a matrix is said to have k leading zeros if the first k elements of the row are all zeros and the $(k+1)$st elements of the row is not zero:*

$$A_i^r = (\underbrace{0 \cdots 0}_{k} a_{i,k+1} \cdots a_m).$$

Definition 62 *A matrix is in row echelon form if each row has more leading zeros than the row preceding it:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \\ 0 & a_{22} & \cdots & \\ 0 & 0 & a_{33} & \cdots \end{pmatrix}.$$

- Similar to square matrices, we can obtain a row echelon form from any matrix by using three elementary operations.

Theorem 94 *Let A_r be any row echelon form obtained from A by using elementary operations. Then,*

$$\dim\left[Row(A)\right] = \dim A_r.$$

- **Note:** To prove this, we need to use the following lemma:

Lemma 95 *Let $v_1, ..., v_m$ be a collection of vectors in $\mathbb{R}^n$. For some $j > 1$, let $w = c_1 v_1 + c_j v_j$ with $c_1 \neq 0$. Then,*

$$L(v_1, ..., v_m) = L(w_1, v_2, ..., v_m).$$

- This lemma implies that elementary operations do not change the dimension of $Row(A)$.

$$\Downarrow$$

- Therefore, for the row rank of a matrix A,we just need to know $\dim A_r$.

Lemma 96 *Let $v_1, ..., v_m \in V$ be nonzero vectors such that for each i, v_{i+1} has more leading zeros than v_i. Then, $v_1, ..., v_m$ are linearly independent.*

Theorem 97 *The number of nonzero rows of a row echelon form of a matrix A is the row rank of A.*

9.2.2 Column Rank

Definition 63 *A pivot of a matrix in row echelon form is an element which is the first non zero element in its row:*

- Example:

$$A = \begin{pmatrix} \boxed{1} & 8 & 7 & 3 \\ 0 & \boxed{2} & 9 & 5 \\ 0 & 0 & 0 & \boxed{4} \end{pmatrix}.$$

  Given A, a_{11}, a_{22}, and a_{34} are pivots.

Definition 64 *A_j^c is a basic column of A if the corresponding column of a row echelon form contains a pivot.*

- Example:

$$A = \begin{pmatrix} 4 & 8 & 1 & 9 \\ -8 & -16 & 1 & -15 \end{pmatrix}$$

$$\downarrow \text{ by elementary operations}$$

$$A_r = \begin{pmatrix} 4 & 8 & 1 & 9 \\ 0 & 0 & 3 & 3 \end{pmatrix}.$$

Then, $A_1^c = \begin{pmatrix} 4 \\ -8 \end{pmatrix}$ and $A_3^c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ are basic columns of A.

Theorem 98 *The basic columns of A form a basis for $Col(A)$.*

Theorem 99 *For any $n \times m$ matrix A,*

$$\textit{row rank of } A = \textit{column rank of } A,$$

and define

$$rank(A) \equiv \textit{row rank of } A.$$

9.3 Nullspace and Affine Subspace

9.3.1 Nullspace

$A : n \times m$ matrix

Definition 65 *$N(A)$ is the null space (or kernel) of a matrix A and defined by*

$$N(A) \equiv \{x \in \mathbb{R}^m : Ax = 0\}.$$

- $N(A)$ is a set of solutions to a linear equation system $Ax = 0$.

Theorem 100 *$N(A)$ is a subspace of $\mathbb{R}^m$.*

- **Note:** Since $N(A)$ always includes 0, $N(A)$ is not an empty set for any matrix A.

9.3.2 Affine Subspace

- **Note:** A set of solutions to $Ax = b$ is not a subspace (if $b \neq 0$). But it looks very "similar" to the set of solutions to $Ax = 0$. (Actually, it is just a linear transformation of $N(A)$.)

Definition 66 *If a subset of* $\mathbb{R}^m$ *has the form* $c + W$ *for some subspace* W *of* $\mathbb{R}^m$ *and for some vector* c *in* $\mathbb{R}^m$*. Then, we call the subset an affine subspace of* $\mathbb{R}^m$*.*

Theorem 101 *Let* A *be an* $n \times m$ *matrix. Then,* $\{x \in \mathbb{R}^m : Ax = b\}$ *be an affine subspace of* $\mathbb{R}^m$*.*

Definition 67 *The dimension of an affine subspace* $c + W$ *is the dimension of* W*.*

9.4 Fundamental Theorem of Linear Algebra

Theorem 102 (Fundamental Theorem of Linear Algebra) *Let* A *be an* $n \times m$ *matrix. Then,*

$$\dim(A) = m - rank(A).$$

- **Note:** This theorem characterizes a set of solutions to a linear equation system $Ax = b$. If the set of solutions to $Ax = b$ is nonempty, then its dimension is the number of variables m minus the rank of A.

10 Quadratic Forms

10.1 Quadratic Forms and Definiteness

Definition 68 *Let A be an $n \times n$ matrix. The quadratic form associated with A is a function from $\mathbb{R}^n$ to $\mathbb{R}$ of the form*

$$g_A(x) = x'Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j.$$

- Example:

$$\begin{aligned} g_A(x) &= x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 6x_2x_3 \\ &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 3 \\ -1 & 3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned}$$

Definition 69 *An $n \times n$ matrix A is said to be*

1. positive definite if $x'Ax > 0$ for any $x \in \mathbb{R}^n$ with $x \neq 0$;
2. positive semidefinite if $x'Ax \geq 0$ for any $x \in \mathbb{R}^n$;
3. negative definite if $x'Ax < 0$ for any $x \in \mathbb{R}^n$ with $x \neq 0$; and
4. negative semidefinite if $x'Ax \leq 0$ for any $x \in \mathbb{R}^n$.

10.2 Identifying Definiteness and Semidefiniteness

Definition 70 *Let A_k be a submatrix of an $n \times n$ matrix A obtained when only the first k rows and columns are retained:*

$$A_k = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}.$$

Then, A_k is called the k-th naturally ordered principal minor of A.

Theorem 103 *Let A be an $n \times n$ matrix. Then,*

1. A is negative definite $\Leftrightarrow$

$$(-1)^k |A_k| > 0 \text{ for any } k = 1, ..., n.$$

2. A is positive definite $\Leftrightarrow$

$$|A_k| > 0 \text{ for all } k = 1, ..., n.$$

Definition 71 *π is a permutation if it is a one-to-one function from $\{1, ..., n\}$ onto $\{1, ..., n\}$.*

- Given an $n \times n$ matrix and a permutation $\pi = (\pi_1, ..., \pi_n)$, define

$$A^\pi = \begin{bmatrix} a_{\pi_1\pi_1} & \cdots & a_{\pi_1\pi_n} \\ \vdots & a_{\pi_i\pi_j} & \vdots \\ a_{\pi_n\pi_1} & \cdots & a_{\pi_n\pi_n} \end{bmatrix}.$$

  A^π is another $n \times n$ matrix obtained by applying the permutation π to both the rows and the columns of A.

- Example: $\pi = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 \end{pmatrix}$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\downarrow$$

$$\begin{aligned} A &= \begin{bmatrix} a_{22} & a_{23} & a_{21} \\ a_{32} & a_{33} & a_{31} \\ a_{12} & a_{13} & a_{11} \end{bmatrix} \\ &= \begin{bmatrix} 5 & 6 & 4 \\ 8 & 9 & 7 \\ 2 & 3 & 1 \end{bmatrix}. \end{aligned}$$

Theorem 104 *Let A be an $n \times n$ matrix. Then,*

1. A is positive semidefinite $\Leftrightarrow$

$$|A_k^\pi| \geq 0 \text{ for any } k = 1, ..., n, \text{ and for all possible permutations } \pi.$$

2. A is negative semidefinite $\Leftrightarrow$

$$(-1)^k |A_k^\pi| \geq 0 \text{ for all } k = 1, ..., n, \text{ and for all possible permutations } \pi.$$