

Limit and Calculus

沈威宇

March 17, 2025

Contents

Chapter 1	Limit and Calculus (極限與微積分)	1
1	Limit (極限)	1
I	A Limit for a Sequence	1
i	Glossary of Terms	1
ii	Definition of a Limit for a Sequence	2
iii	Infinite Limits	2
iv	Existence of a Limit	2
v	Convergence (收斂) and Divergence (發散)	2
vi	Limit Laws	2
vii	Special Sequences and Their Limits	3
viii	Squeeze (夾擠)/Sandwich (三明治) theorem	3
ix	Monotone Convergence Theorem (單調收斂定理)/Completeness of the Real Number (實數的完備性)	3
II	A Limit for a Series	4
i	Definition	4
ii	Convergence (收斂) and Divergence (發散)	4
iii	Limit Laws	4
III	A Limit for a Function	4
i	Limit at Finiteness	4
ii	Limit at Infinity	5
iii	Horizontal asymptote (水平漸近線)	5
iv	Slant asymptote (斜漸近線)	5
v	One-side Limits	5
vi	Infinite Limits	5
vii	Existence of Limits	6

viii	Limit Laws	6
ix	Squeeze/Sandwich theorem	6
x	Continuity (連續性)	7
xi	Theorem: Limits Involving Quotient Functions	7
2	Differentiation (微分).	7
I	Fréchet derivative (弗蘭歇導數)	7
II	Gateaux derivative (加托導數)	8
III	Taylor expansion (泰勒展開式)	8
3	Integration (積分)	8
I	Riemann integral (黎曼積分) and Darboux integral (達布積分)	8
II	Indicator function or characteristic function	10
III	Lebesgue integral (勒貝格積分).	10
4	Fundamental Theorem of Calculus (FTC) (微積分基本定理).	11
I	The First Theorem	11
II	The Second Theorem	12
5	Convention of multivariable (多變數/多變量/多元) calculus or multivariate calculus	12
I	Space convention	12
II	Operators notation convention	13
6	Multivariable differentiation	13
I	Notation convention	13
II	Gradient	13
III	Divergence	14
IV	Curl	14
V	Directional derivative	14
VI	Laplace operator	14
VII	Poisson's equation (卜瓦松/帕松/泊松方程)	14
VIII	Laplace's equation (拉普拉斯方程)	15
IX	Multi-index notation (多重指標記號).	15
X	High-order derivative	15
XI	Taylor expansion	15

7	Line integral (線積分)/Path integral (路徑積分)	16
I	Scalar field line/path integral	16
II	Vector field line/path integral	16
III	Conservative field (保守場).	16
8	Fundamental theorem of multivariable calculus (多變數微積分基本定理)	17
I	Gradient theorem (梯度定理)	17
II	Divergence theorem, Gauss's theorem, or Ostrogradsky's theorem (高斯散度定理)	17
III	Stokes' theorem (斯托克斯定理) or Kelvin–Stokes theorem	17
IV	Green's theorem (格林定理或綠定理)	17
V	Generalized Stokes theorem, Stokes–Cartan theorem, or fundamental theorem of multivariable calculus	18
9	Differential theorems	18
I	Intermediate Value Theorem, IVT (中間值定理)	18
II	Mean Value Theorem, MVT (均值定理)	18
III	Extreme Value Theorem, EVT (極值定理)	18
IV	L'Hôpital's rule (羅必達法則)/Bernoulli's rule	18
V	Critical point (臨界點)	18
VI	Relative extremum (相對極值), local extremum (局部極值), or extremum (極值)	18
VII	Absolute extremum (絕對極值/最值) or global extremum (全域極值)	19
VIII	Relative extreme theorem	19
IX	Concavity (凹性)	19
X	Concavity theorem	19
XI	Point of inflection or inflection point (反曲點/拐點).	19
XII	Inflection point theorem.	19
XIII	Lagrange multiplier (拉格朗日乘數)	19
i	Univariate form	20
ii	Multivariate form	20
10	Integral theorems	21
I	Integration by substitution (代換積分法), integration by change of variables (換元積分法), u-substitution (u-代換), reverse chain rule, substitution theorem (代換定理), change of variables theorem (換元定理), or transformation theorem (變換定理)	21
i	Univariate form	21

ii Multivariate form	22
iii Measure theory form	22
II Integration by parts (分部積分法) or partial integration (部分積分法)	23
i Theorem	23
ii Application	23
11 Common calculus results	23
I Polynomial function	23
II Exponential function	24
III Logarithmic function	24
IV Non-polynomial algebraic function	24
12 Numerical differentiation (數值微分)	25
I Newton's Method (牛頓法) or Newton-Raphson Method (牛頓-拉普森法)	25
13 Numerical integration (數值積分)	25
I The trapezoidal rule (梯形法)	25
II The Simpson's rule (辛普森法) or the Simpson's 1/3 rule	25

Chapter 1 Limit and Calculus (極限與微積分)

1 Limit (極限)

I A Limit for a Sequence

i Glossary of Terms

- Sequence (數列): A sequence is a function whose domain is an interval of integers, usually denoted as $\langle a_n \rangle$, $\{a_n\}$, or (a_n) , sometimes with domain as $\langle a_n \rangle_{n=1}^m$, $\{a_n\}_{n=1}^m$, or $(a_n)_{n=1}^m$, where the subscript n refers to the n th element of the sequence, that is, the function value when the independent variable is n .
- Finite sequence (有限數列): A finite sequence is a sequence with finite terms, e.g. $\langle a_n \rangle_{n=1}^m = \langle a_1, a_2, \dots, a_m \rangle$, $m \geq 1$ and m is finite.
- Infinite sequence (無限數列): An infinite sequence is a sequence with infinite terms, e.g. $\langle a_n \rangle_{n=1}^\infty = \langle a_1, a_2, \dots \rangle$. Unless otherwise specified, the sequences referred to below are infinite sequences.
- Monotone Increasing/Increasing/Non-Decreasing Sequence (單調遞增/遞增/非遞減數列): $\langle a_n \rangle$ is a monotone increasing/increasing/non-decreasing sequence if and only if $\forall n$ such that $\exists a_n, a_{n+1} : a_n \leq a_{n+1}$.
- Strictly Increasing Sequence (嚴格遞增數列): $\langle a_n \rangle$ is a strictly increasing sequence if and only if $\forall n$ such that $\exists a_n, a_{n+1} : a_n < a_{n+1}$.
- Monotone Decreasing/Decreasing/Non-Increasing Sequence (單調遞減/遞減/非遞增數列): $\langle a_n \rangle$ is a monotone decreasing/decreasing/non-increasing sequence if and only if $\forall n$ such that $\exists a_n, a_{n+1} : a_n \geq a_{n+1}$.
- Strictly Decreasing Sequence (嚴格遞減數列): $\langle a_n \rangle$ is a strictly decreasing sequence if and only if $\forall n$ such that $\exists a_n, a_{n+1} : a_n > a_{n+1}$.
- Monotone Sequence (單調數列): $\langle a_n \rangle$ is a monotone sequence if and only if $\langle a_n \rangle$ is either a increasing sequence or a decreasing sequence.
- Upper Bound (上界): M is an upper bound of $\langle a_n \rangle$ if and only if $\forall n$ such that $\exists a_n : a_n \leq M$. \exists upper bound of $\langle a_n \rangle \iff \langle a_n \rangle$ is bounded-above.
- Lower Bound (下界): m is a lower bound of $\langle a_n \rangle$ if and only if $\forall n$ such that $\exists a_n : a_n \geq m$. \exists lower bound of $\langle a_n \rangle \iff \langle a_n \rangle$ is bounded-below.

- Supremum/Least Upper Bound (最小上界): M is the supremum/least upper bound of $\langle a_n \rangle$ if and only if \forall upper bound M of $\langle a_n \rangle$: $M_0 \leq M$, denoted as $\sup_n a_n$.
- Infimum/Greatest Lower Bound (最大下界): M is the infimum/greatest lower bound of $\langle a_n \rangle$ if and only if \forall lower bound m of $\langle a_n \rangle$: $m_0 \geq m$, denoted as $\inf_n a_n$.
- Bounded: A sequence is bounded if and only if it is bounded-above and bounded-below.
- Series (級數): The sum of the terms of a sequence.
- Finite Series (有限級數): The sum of the terms of a finite sequence.
- Infinite Series (無窮級數): The sum of the terms of an infinite sequence.

ii Definition of a Limit for a Sequence

For a sequence $\langle a_n \rangle$, the limit of $\langle a_n \rangle$ as n approaches ∞ is defined as follows:

$$\lim_{n \rightarrow \infty} a_n = L \equiv \forall \epsilon > 0 : \exists N \in \mathbb{N} \text{ such that } n \geq N \implies |a_n - L| < \epsilon.$$

In other words, as n becomes arbitrarily large, a_n gets arbitrarily close to L .

iii Infinite Limits

$$\lim_{n \rightarrow \infty} a_n = \infty \equiv \forall M > 0, \exists N \in \mathbb{N} \text{ such that } n > N \implies a_n > M.$$

$$\lim_{n \rightarrow \infty} a_n = -\infty \equiv \forall M < 0, \exists N \in \mathbb{N} \text{ such that } n > N \implies a_n < M.$$

iv Existence of a Limit

$$\exists \lim_{n \rightarrow \infty} a_n \iff \exists \text{ finite } L \text{ such that } \forall \epsilon > 0 : \exists N \in \mathbb{N} \text{ such that } n \geq N \implies |a_n - L| < \epsilon.$$

v Convergence (收斂) and Divergence (發散)

A sequence $\langle a_n \rangle$ converges to L if $\exists \lim_{n \rightarrow \infty} a_n \wedge \lim_{n \rightarrow \infty} a_n = L$. If no such L exists, we say the sequence does not converge, namely, the sequence diverges.

vi Limit Laws

Given sequences $\langle a_n \rangle$ and $\langle b_n \rangle$ which:

$$\exists \lim_{n \rightarrow \infty} a_n \wedge \exists \lim_{n \rightarrow \infty} b_n$$

Then:

$$\lim_{x \rightarrow a} \langle a_n \rangle + \langle b_n \rangle = \lim_{x \rightarrow a} \langle a_n \rangle + \lim_{x \rightarrow a} \langle b_n \rangle$$

$$\lim_{x \rightarrow a} \langle a_n \rangle - \langle b_n \rangle = \lim_{x \rightarrow a} \langle a_n \rangle - \lim_{x \rightarrow a} \langle b_n \rangle$$

$$\text{If } c \text{ is a constant : } \lim_{x \rightarrow a} c \langle a_n \rangle = c \lim_{x \rightarrow a} \langle a_n \rangle$$

$$\lim_{x \rightarrow a} \langle a_n \rangle \langle b_n \rangle = \lim_{x \rightarrow a} \langle a_n \rangle \lim_{x \rightarrow a} \langle b_n \rangle$$

$$\text{If } \langle b_n \rangle \neq 0 \wedge \lim_{x \rightarrow a} \langle b_n \rangle \neq 0 : \lim_{x \rightarrow a} \frac{\langle a_n \rangle}{\langle b_n \rangle} = \frac{\lim_{x \rightarrow a} \langle a_n \rangle}{\lim_{x \rightarrow a} \langle b_n \rangle}$$

vii Special Sequences and Their Limits

If c is a constant : $\lim_{n \rightarrow \infty} c = c$

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$$

$$\forall r \in (-1, 1) : \lim_{n \rightarrow \infty} r^n = 0$$

$$\exists \lim_{n \rightarrow \infty} r^n \iff r \in (-1, 1]$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

viii Squeeze (夾擠)/Sandwich (三明治) theorem

Given sequences $\langle a_n \rangle$, $\langle b_n \rangle$, and $\langle c_n \rangle$ which:

$$a_n \leq c_n \leq b_n$$

and

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = L,$$

then:

$$\lim_{n \rightarrow \infty} c_n = L.$$

ix Monotone Convergence Theorem (單調收斂定理)/Completeness of the Real Number (實數的完備性)

Statement.

(A) For a non-decreasing and bounded-above sequence of real numbers $\langle a_n \rangle_{n \in \mathbb{N}}$:

$$\lim_{n \rightarrow \infty} a_n = \sup_n a_n$$

(B) For a non-increasing and bounded-below sequence of real numbers $\langle a_n \rangle_{n \in \mathbb{N}}$:

$$\lim_{n \rightarrow \infty} a_n = \inf_n a_n$$

Proof. Let $\{a_n\}_{n \in \mathbb{N}}$ be the set of values of $\langle a_n \rangle$. By assumption, $\{a_n\}$ is non-empty and bounded-above by $\sup_n a_n$. Let $c = \sup_n a_n$.

$$\forall \epsilon > 0 : \exists N \text{ such that } c \geq a_N > c - \epsilon,$$

since otherwise $c - \epsilon$ is a strictly smaller upper bound of $\langle a_n \rangle$, contradicting the definition of the supremum.

Then since $\langle a_n \rangle$ is non decreasing, and c is an upper bound:

$$\forall \epsilon > 0 : \exists N \text{ such that } \forall n > N : |c - a_n| = c - a_n \leq c - a_N = |c - a_N| < \epsilon.$$

The proof of the (B) part is analogous or follows from (A) by considering $\langle -a_n \rangle_{n \in \mathbb{N}}$. □

Statement.

If $\langle a_n \rangle_{n \in \mathbb{N}}$ is a monotone sequence of real numbers, i.e., if $a_n \leq a_{n+1}$ for every $n \geq 1$ or $a_n \geq a_{n+1}$ for every $n \geq 1$, then this sequence has a finite limit if and only if the sequence is bounded.

Proof. "If"-direction: The proof follows directly from the proposition.

"Only If"-direction: By (ϵ, δ) -definition of limit, every sequence $\langle a_n \rangle_{n \in \mathbb{N}}$ with a finite limit L is necessarily bounded. □

II A Limit for a Series

i Definition

Let:

$$S_n = \sum_{i=1}^n a_i,$$

where a_i are terms of a sequence. The limit of S_n , denoted as $\lim_{n \rightarrow \infty} S_n$ or $\sum_{i=1}^{\infty} a_i$, is defined as the following:

$$\sum_{i=1}^{\infty} a_i = L \equiv \forall \epsilon > 0 : \exists N \in \mathbb{N} \text{ such that } n \geq N \implies |S_n - L| < \epsilon.$$

ii Convergence (收斂) and Divergence (發散)

A series S_n converges to L if $\exists \lim_{n \rightarrow \infty} S_n \wedge \lim_{n \rightarrow \infty} S_n = L$. If no such L exists, we say the series does not converge, namely, the series diverges.

iii Limit Laws

Given series $\sum_{i=1}^n a_i$ and $\sum_{i=1}^n b_i$ which:

$$\exists \lim_{n \rightarrow \infty} \sum_{i=1}^n a_i \wedge \exists \lim_{n \rightarrow \infty} \sum_{i=1}^n b_i$$

Then:

$$\sum_{i=1}^{\infty} a_i + \langle b_n \rangle = \sum_{i=1}^{\infty} a_i + \sum_{i=1}^{\infty} b_i$$

$$\sum_{i=1}^{\infty} a_i - \langle b_n \rangle = \sum_{i=1}^{\infty} a_i - \sum_{i=1}^{\infty} b_i$$

$$\text{If } c \text{ is a constant : } \sum_{i=1}^{\infty} c a_i = c \sum_{i=1}^{\infty} a_i$$

III A Limit for a Function

i Limit at Finitiy

Let I be an interval containing the point a . Let $f(x)$ be a function defined on I , except possibly at a itself. The limit of $f(x)$ as x approaches a is defined as follows:

$$\lim_{x \rightarrow a} f(x) = L \equiv \forall \epsilon > 0 : \exists \delta > 0 \text{ such that } 0 < |x - a| < \delta \implies |f(x) - L| < \epsilon.$$

In other words, as x becomes arbitrarily close to a , $f(x)$ gets arbitrarily close to L .

ii Limit at Infinity

- Let I be a left-bounded, right-unbounded interval with the point a being its endpoint on the left. Let $f(x)$ be a function defined on I . The limit of $f(x)$ as x approaches ∞ is defined as follows:

$$\lim_{x \rightarrow \infty} f(x) = L \equiv \forall \epsilon > 0 : \exists M > a \text{ such that } x > M \implies |f(x) - L| < \epsilon.$$

In other words, as x becomes arbitrarily close to ∞ , $f(x)$ gets arbitrarily close to L .

- Let I be a right-bounded, left-unbounded interval with the point a being its endpoint on the right. Let $f(x)$ be a function defined on I . The limit of $f(x)$ as x approaches $-\infty$ is defined as follows:

$$\lim_{x \rightarrow -\infty} f(x) = L \equiv \forall \epsilon > 0 : \exists M < a \text{ such that } x < M \implies |f(x) - L| < \epsilon.$$

In other words, as x becomes arbitrarily close to $-\infty$, $f(x)$ gets arbitrarily close to L .

iii Horizontal asymptote (水平漸近線)

$$(\lim_{x \rightarrow \infty} f(x) = L \vee \lim_{x \rightarrow -\infty} f(x) = L) \implies (y = L \text{ is a horizontal asymptote of } f(x))$$

iv Slant asymptote (斜漸近線)

$$(\lim_{x \rightarrow \infty} f(x) - (mx + b) = 0 \vee \lim_{x \rightarrow -\infty} f(x) - (mx + b) = 0) \implies ((y = mx + b) \text{ is a slant asymptote of } f(x))$$

v One-side Limits

- Right-hand Limit:** Let I be a left-open interval with the point a being its endpoint on the left. Let $f(x)$ be a function defined on I . The right-hand limit of $f(x)$ as x approaches a is defined as follows:

$$\lim_{x \rightarrow a^+} f(x) = L \equiv \forall \epsilon > 0 : \exists \delta > 0 \text{ such that } 0 < x - a < \delta \implies |f(x) - L| < \epsilon.$$

In other words, as x becomes arbitrarily close to a and is greater than a , $f(x)$ gets arbitrarily close to L .

- Left-hand Limit:** Let I be a right-open interval with the point a being its endpoint on the right. Let $f(x)$ be a function defined on I . The left-hand limit of $f(x)$ as x approaches a is defined as follows:

$$\lim_{x \rightarrow a^-} f(x) = L \equiv \forall \epsilon > 0 : \exists \delta > 0 \text{ such that } 0 < a - x < \delta \implies |f(x) - L| < \epsilon.$$

In other words, as x becomes arbitrarily close to a and is less than a , $f(x)$ gets arbitrarily close to L .

vi Infinite Limits

$$\lim_{x \rightarrow a} f(x) = \infty \equiv \forall M > 0, \exists \delta > 0 \text{ such that } 0 < |x - a| < \delta \implies f(x) > M.$$

$$\lim_{x \rightarrow a} f(x) = -\infty \equiv \forall M < 0, \exists \delta > 0 \text{ such that } 0 < |x - a| < \delta \implies f(x) < M.$$

$$\lim_{x \rightarrow \infty} f(x) = \infty \equiv \forall M > 0, \exists \delta > 0 \text{ such that } x > \delta \implies f(x) > M.$$

$$\lim_{x \rightarrow \infty} f(x) = -\infty \equiv \forall M > 0, \exists \delta > 0 \text{ such that } x > \delta \implies f(x) < M.$$

$$\lim_{x \rightarrow -\infty} f(x) = \infty \equiv \forall M > 0, \exists \delta < 0 \text{ such that } x < \delta \implies f(x) > M.$$

$$\lim_{x \rightarrow -\infty} f(x) = -\infty \equiv \forall M > 0, \exists \delta < 0 \text{ such that } x < \delta \implies f(x) < M.$$

$$a \in \mathbb{R}, \left(\left| \lim_{x \rightarrow a^+} f(x) \right| = \infty \vee \left| \lim_{x \rightarrow a^-} f(x) \right| = \infty \right) \implies x = a \text{ is the vertical asymptote (鉛直漸近線) of } f(x)$$

vii Existence of Limits

$$\exists \lim_{x \rightarrow a^+} f(x) \iff \exists \text{ finite } L \text{ such that } \forall \epsilon > 0 : \exists \delta > 0 \text{ such that } 0 < x - a < \delta \implies |f(x) - L| < \epsilon.$$

$$\exists \lim_{x \rightarrow a^-} f(x) \iff \exists \text{ finite } L \text{ such that } \forall \epsilon > 0 : \exists \delta > 0 \text{ such that } 0 < a - x < \delta \implies |f(x) - L| < \epsilon.$$

$$\exists \lim_{x \rightarrow a} f(x) \iff \exists \lim_{x \rightarrow a^+} f(x) \wedge \exists \lim_{x \rightarrow a^-} f(x) \wedge \lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x).$$

viii Limit Laws

Let I be an interval containing the point a . Let $f(x)$ and $g(x)$ be functions defined on I , except possibly at a itself which:

$$\exists \lim_{x \rightarrow a} f(x) \wedge \exists \lim_{x \rightarrow a} g(x).$$

Then:

$$\lim_{x \rightarrow a} f(x) + g(x) = \lim_{x \rightarrow a} f(x) + \lim_{x \rightarrow a} g(x)$$

$$\lim_{x \rightarrow a} f(x) - g(x) = \lim_{x \rightarrow a} f(x) - \lim_{x \rightarrow a} g(x)$$

$$\text{If } c \text{ is a constant : } \lim_{x \rightarrow a} c f(x) = c \lim_{x \rightarrow a} f(x)$$

$$\lim_{x \rightarrow a} f(x) g(x) = \lim_{x \rightarrow a} f(x) \lim_{x \rightarrow a} g(x)$$

$$\text{If } g(x) \neq 0 \wedge \lim_{x \rightarrow a} g(x) \neq 0 : \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} g(x)}$$

ix Squeeze/Sandwich theorem

Let I be an interval containing the point a . Let $f(x)$, $g(x)$, and $h(x)$ be functions defined on I , except possibly at a itself which:

$$f(x) \leq h(x) \leq g(x)$$

and

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = L,$$

then:

$$\lim_{x \rightarrow a} h(x) = L.$$

x Continuity (連續性)

- If $\exists \lim_{x \rightarrow a} f(x)$ and $\lim_{x \rightarrow a} f(x) = f(a)$, we say $f(x)$ is continuous (連續的) at $x = a$, or we say $f(x)$ is discontinuous (不連續的) at $x = a$.
- If $f(x)$ is continuous at all points in the open interval (a, b) , we say $f(x)$ is continuous on the open interval (a, b) .
- If $f(x)$ is continuous on the open interval (a, b) , and $\lim_{x \rightarrow a^+} f(x) = f(a)$, we say $f(x)$ is continuous on the left-close right-open interval $[a, b)$.
- If $f(x)$ is continuous on the open interval (a, b) , and $\lim_{x \rightarrow b^-} f(x) = f(b)$, we say $f(x)$ is continuous on the right-close left-open interval $(a, b]$.
- If $f(x)$ is continuous on the left-close right-open interval $[a, b)$ and the right-close left-open interval $(a, b]$, we say $f(x)$ is continuous on the open interval (a, b) .
- If $f(x)$ is continuous at all points in its domain, we say $f(x)$ is continuous function (連續函數).
- If both $f(x)$ and $g(x)$ are continuous at $x = a$, then $f(x) + g(x)$, $f(x) - g(x)$, and $f(x) \cdot g(x)$ are continuous at $x = a$.
- If both $f(x)$ and $g(x)$ are continuous at $x = a$ and $g(a) \neq 0$, then $\frac{f(x)}{g(x)}$ is continuous at $x = a$.
- If both $f(x)$ and $g(x)$ are continuous at $x = a$ and $g(a) \in \text{dom}(f)$, then $f \circ g(x)$ is continuous at $x = a$, namely, $\lim_{x \rightarrow a} ((f \circ g)(x)) = \lim_{x \rightarrow a} (f(g(x))) = f\left(\lim_{x \rightarrow a} g(x)\right) = f(g(a)) = f \circ g(a)$.

xi Theorem: Limits Involving Quotient Functions

Let a and b be real numbers, $A = \{f(x) \mid f(x) = x \vee \ln(f(x)) \in A \vee e^{f(x)} \in A\}$, and $f(x) \in A$. Then:

$$\lim_{x \rightarrow \infty} \frac{(f(x))^a}{(f(x))^b} = \infty \quad \text{if } a > b$$

$$\lim_{x \rightarrow \infty} \frac{af(x)}{bf(x)} = \frac{a}{b} \quad \text{if } b \neq 0$$

$$\lim_{x \rightarrow \infty} \frac{n^{af(x)}}{bf(x)} = \infty \quad \text{if } ab > 0 \wedge n > 1$$

$$\lim_{x \rightarrow \infty} \frac{af(x)}{b \log_n f(x)} = \infty \quad \text{if } ab > 0 \wedge n > 1$$

$$\lim_{x \rightarrow \infty} \frac{h(x)}{g(x)} = 0 \quad \text{if } \left| \Re \lim_{x \rightarrow \infty} \frac{g(x)}{h(x)} \right| + \left| i \Im \lim_{x \rightarrow \infty} \frac{g(x)}{h(x)} \right| = \infty$$

2 Differentiation (微分)

I Fréchet derivative (弗蘭歇導數)

Let V and W be normed vector spaces, and $U \subseteq V$ be an open subset of V . A function $f : U \rightarrow W$ is called Fréchet differentiable at $x \in U$ if there exists a bounded linear operator $A : V \rightarrow W$ such

that

$$\lim_{\|h\|_V \rightarrow 0} \frac{\|f(x+h) - f(x) - Ah\|_W}{\|h\|_V} = 0.$$

If there exists such an operator A , it is unique, so we write $Df(x) = A$ and call it the Fréchet derivative of f at x .

A function f that is Fréchet differentiable for any point of U is said to be C^1 if the function

$$Df : U \rightarrow B(V, W); x \mapsto Df(x)$$

is continuous.

II Gateaux derivative (加托導數)

Let V and W be locally convex topological vector spaces (LCTVSs), $U \subseteq V$ be an open subset of V , and a function $F : U \rightarrow W$. The Gateaux derivative $dF(u; \psi)$ of F at x in U in the direction $\psi \in V$ is defined to be

$$\begin{aligned} dF(u; \psi) &= \lim_{\tau \rightarrow 0} \frac{F(u + \tau\psi) - F(u)}{\tau} \\ &= \frac{d}{d\tau} F(u + \tau\psi) \Big|_{\tau=0} \end{aligned}$$

If the limit exists for all $\psi \in V$, then it is said that F is Gateaux differentiable at u .

III Taylor expansion (泰勒展開式)

Assume that $F : \mathbb{R} \rightarrow \mathbb{R}$ is an infinitely differentiable function, and its derivatives of every order exist on \mathbb{R} , then the Taylor expansion of F at a is

$$F(x) = \sum_{n \in \mathbb{N}_0} \frac{F^{(n)}(a)}{n!} (x - a)^n,$$

that is,

$$F(x) = \sum_{n=0}^k \frac{F^{(n)}(a)}{n!} (x - a)^n + \int_0^1 \frac{(1-t)^k}{k!} F^{(k+1)}(a + t(x-a)) (x-a)^{k+1} dt.$$

Also, the k th-order approximation of f near a is

$$F(x) \approx \sum_{n=0}^k \frac{F^{(n)}(a)}{n!} (x - a)^n,$$

and the first-order approximation near a is

$$F(x) \approx F(0) + F'(a)(x - a).$$

3 Integration (積分)

I Riemann integral (黎曼積分) and Darboux integral (達布積分)

1. Partition of an interval

A partition $P(x, n)$ of an interval $[a, b]$ is a finite sequence of numbers of the form

$$P(x, n) := \{x_i : a = x_0, b = x_n, \forall 1 \leq i < j \leq n : x_i < x_j\}_{i=0}^n.$$

Each $[x_i, x_{i+1}]$ is called a sub-interval of the partition. The mesh or norm of a partition is defined to be the length of the longest sub-interval, that is,

$$\max (x_{i+1} - x_i), \quad i \in [0, n-1].$$

A tagged partition $P(x, n, \xi)$ of an interval $[a, b]$ is a partition together with a choice of a sample point within each of all n sub-intervals, that is, numbers $\{\xi_i\}_{i=0}^{n-1}$ with $\xi_i \in [x_i, x_{i+1}]$ for each $i \in [0, n-1]$. The mesh of a tagged partition is the same as that of an ordinary partition.

Suppose that two tagged partitions $P(x, n, \xi)$ and $Q(y, m, \zeta)$ are both partitions of the interval $[a, b]$. We say that $Q(y, m, \zeta)$ is a refinement of $P(x, n, \xi)$ if for each integer $i \in [0, n]$, there exists an integer $r(i) \in [0, m]$ such that $x_i = y_{r(i)}$ and that $\forall i \in [0, n-1] : \exists j \in [r(i), r(i+1)]$ s.t. $\xi_i = \zeta_j$. That is, a tagged partition breaks up some of the sub-intervals and adds sample points where necessary, "refining" the accuracy of the partition.

We can turn the set of all tagged partitions into a directed set by saying that one tagged partition is greater than or equal to another if the former is a refinement of the latter.

2. Riemann sum

Let f be a real-valued function defined on the interval $[a, b]$. The Riemann sum of f with respect to the tagged partition $P(x, n, \xi)$ is defined to be

$$R(f, P) := \sum_{i=0}^{n-1} f(\xi_i) (x_{i+1} - x_i).$$

Each term in the sum is the product of the value of the function at a given point and the length of an interval. Consequently, each term represents the (signed) area of a rectangle with height $f(\xi_i)$ and width $x_{i+1} - x_i$. The Riemann sum is the (signed) area of all the rectangles.

3. Darboux sum

Lower and upper Darboux sums of f with respect to the partition $P(x, n)$ are two specific Riemann sums of which the tags are chosen to be the infimum and supremum (respectively) of f on each sub-interval:

$$L(f, P) := \sum_{i=0}^{n-1} \inf_{\xi \in [x_i, x_{i+1}]} f(\xi) (x_{i+1} - x_i),$$

$$U(f, P) := \sum_{i=0}^{n-1} \sup_{\xi \in [x_i, x_{i+1}]} f(\xi) (x_{i+1} - x_i).$$

4. Riemann integral

The Riemann integral of f exists and equals s if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for any tagged partition $P(x, n, \xi)$ whose mesh is less than δ ,

$$|R(f, P) - s| < \varepsilon.$$

5. Darboux integral

The Darboux integral of f exists and equals s if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for any partition P whose mesh is less than δ ,

$$|U(f, P) - s| < \varepsilon \wedge |L(f, P) - s| < \varepsilon.$$

6. Integrability

A function is Riemann-integrable if and only if it is Darboux-integrable.

II Indicator function or characteristic function

An indicator function or a characteristic function of a subset A of a set X is a function that maps elements of the subset to one, and all other elements to zero, often denoted as 1_A .

III Lebesgue integral (勒貝格積分)

Below, we will define the Lebesgue integral of measurable functions from a measure space (E, Σ, μ) into $\mathbb{R} \cup \{-\infty, \infty\}$.

1. Indicator functions

The integral of an indicator function of a measurable set S is defined to be

$$\int 1_S d\mu = \mu(S).$$

2. Simple functions

Simple functions are finite real linear combinations of indicator functions. A simple function s of the form

$$s := \sum_k a_k 1_{S_k},$$

where the coefficients a_k are real numbers and S_k are disjoint measurable sets, is called a measurable simple function. When the coefficients a_k are positive real numbers, s is called a non-negative measurable simple function. The integral of a non-negative measurable simple function $\sum_k a_k 1_{S_k}$ is defined to be

$$\int \left(\sum_k a_k 1_{S_k} \right) d\mu = \sum_k a_k \int 1_{S_k} d\mu = \sum_k a_k \mu(S_k).$$

whether this sum is finite or $+\infty$.

If B is a measurable subset of E and $s := \sum_k a_k 1_{S_k}$ is a non-negative measurable simple function, one defines

$$\int_B s d\mu = \int 1_B s d\mu = \sum_k a_k \mu(S_k \cap B).$$

3. Non-negative measurable functions

Let f be a non-negative measurable function on some measurable subset B of E . We define

$$\int_B f \, d\mu = \sup \left\{ \int_B s \, d\mu : \forall x \in B : 0 \leq s(x) \leq f(x) \wedge s \text{ is a measurable simple function} \right\}.$$

4. Signed functions

Let f be a measurable function from a measure set E into $\mathbb{R} \cup \{-\infty, \infty\}$. We define

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) > 0, \\ 0 & \text{otherwise,} \end{cases}$$
$$f^-(x) = \begin{cases} -f(x) & \text{if } f(x) < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that both f^+ and f^- are non-negative and that

$$f = f^+ - f^-, \quad |f| = f^+ + f^-.$$

We say that the Lebesgue integral of f exists, if

$$\min \left(\int f^+ \, d\mu, \int f^- \, d\mu \right) < \infty.$$

In this case we define

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

If

$$\int |f| \, d\mu < \infty,$$

we say that f is Lebesgue integrable.

4 Fundamental Theorem of Calculus (FTC) (微積分基本定理)

I The First Theorem

Let $F(x)$ be a differentiable function defined on $[a, b]$. Then:

$$\int_a^b F'(x) \, dx = F(b) - F(a).$$

Proof.

By the definition of the Riemann integral:

$$\int_a^b F'(x) \, dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n F'(x_i^*) \Delta x_i,$$

where $\{x_i^*\}$ are sample points in the subintervals of a partition $P = \{x_0, x_1, \dots, x_n\}$ of $[a, b]$, and $\Delta x_i = x_i - x_{i-1}$. By the Mean Value Theorem for derivatives, since $F(x)$ is differentiable, there exists

an adequately refined partition $P = \{x_0, x_1, \dots, x_n\}$ such that on each subinterval $[x_{i-1}, x_i]$ there exists a point $x_i^* \in [x_{i-1}, x_i]$ such that:

$$F'(x_i^*) \cdot \Delta x_i = F(x_i) - F(x_{i-1}).$$

Thus, the Riemann sum becomes:

$$\sum_{i=1}^n F'(x_i^*) \Delta x_i = \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = F(b) - F(a).$$

□

II The Second Theorem

Let $f(x)$ be a continuous function defined on $[a, b]$. Then:

$$\frac{d}{dx} \int_a^x f(t) dt = f(x).$$

Proof.

$$\frac{d}{dx} \left(\int_a^x f(t) dt \right) = \lim_{h \rightarrow 0} \frac{\int_a^{x+h} f(t) dt - \int_a^x f(t) dt}{h}$$

Using the additivity property of integrals:

$$\int_a^{x+h} f(t) dt - \int_a^x f(t) dt = \int_x^{x+h} f(t) dt$$

By the Mean Value Theorem for integrals, since $f(t)$ is continuous on $[x, x+h]$, there exists a point $c \in [x, x+h]$ such that:

$$\int_x^{x+h} f(t) dt = f(c) \cdot h.$$

Substituting into the difference quotient:

$$\frac{\int_x^{x+h} f(t) dt}{h} = f(c).$$

$$\lim_{h \rightarrow 0} \frac{\int_x^{x+h} f(t) dt}{h} = f(x).$$

□

5 Convention of multivariable (多變數/多變量/多元) calculus or multivariate calculus

I Space convention

- The domain of the functions or maps below are subsets of a Euclidean vector space. If not otherwise specified, the coordinates are the Cartesian coordinates, the norms are the Euclidean norms, and the measures are the Lebesgue measures.
- $\mathbf{0}$ or 0 refers to the zero tensor (零張量) in the interested Euclidean tensor space V , that is, it satisfies

$$\forall \mathbf{v} \in V : \mathbf{v} + \mathbf{0} = \mathbf{v}.$$

II Operators notation convention

- Dot product (點積) operator: \cdot
- Cross product (叉積) operator: \times
- Gradient (梯度) operator: ∇
- Divergence (散度) operator: $\nabla \cdot$
- Curl (旋度) operator: $\nabla \times$
- Directional derivative (方向導數) operator: $\cdot \nabla$
- Laplace (拉普拉斯) operator: ∇^2 或 Δ
- Line/Path integral operator: \int
- Surface integral operator: \iint
- Volume integration operator: \iiint
- Closed line integral operator: \oint
- Closed surface Integral Operator: \oiint
- $\int \mathbf{F} \cdot d\mathbf{S}$ is used as a shorthand for $\int (\mathbf{F} \cdot \hat{\mathbf{n}}) dS$, where $\hat{\mathbf{n}}$ is the outward pointing unit normal at almost each point on S .

6 Multivariable differentiation

I Notation convention

- Unit vector (單位向量): \mathbf{e}_i is the unit vector in the i th direction, i.e., a vector with zero norm.
- Independent variable vector: $\mathbf{x} = (x_1, x_2, \dots, x_n)$
- Vector fields: $\mathbf{F}(\mathbf{x}) = \sum_{i=1}^n F_i(\mathbf{x})\mathbf{e}_i$ 、 \mathbf{G}
- Scalar fields: $A(\mathbf{x})$ 、 $B(\mathbf{x})$
- Tensor fields: $f(\mathbf{x})$ 、 $g(\mathbf{x})$
- Three-dimensional tensor space field: $\mathbf{T}(\mathbf{x})$
- The i -th component of the map f : f_i

II Gradient

$$\nabla f = \left(\left(\frac{\partial f}{\partial x_1} \right)^T \quad \left(\frac{\partial f}{\partial x_2} \right)^T \quad \dots \quad \left(\frac{\partial f}{\partial x_n} \right)^T \right)$$

The gradient of a scalar field is a vector field, the gradient of a vector field is a second-order tensor (matrix) field, and the gradient of a k -order tensor field is a $k + 1$ -order tensor field.

In particular, the gradient of a scalar field A is

$$\nabla A = \sum_{i=1}^n \frac{\partial A}{\partial x_i} \mathbf{e}_i.$$

And the gradient of a vector field \mathbf{F} is also called the Jacobian matrix (雅可比矩陣) of it and also denoted as $\mathbf{J}(\mathbf{F})$, $J(\mathbf{F})$, or $J_{\mathbf{F}}$, of which the (i, j) th entry is

$$\mathbf{J}_{ij} = \frac{\partial F_i}{\partial x_j}.$$

The determinant $\det(J_{\mathbf{F}})$ of a Jacobian matrix is called a Jacobian determinant, or Jacobian for short.

III Divergence

$$\nabla \cdot \mathbf{f} = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}$$

The divergence of a vector field is a scalar field, the gradient of a second-order tensor (i.e. matrix) field is a vector field, and the divergence of a $k + 1$ -order tensor field is a k -order tensor field.

IV Curl

The curl is only defined on three-dimensional vector field.

$$\nabla \times \mathbf{T} = \begin{pmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ T_1 & T_2 & T_3 \end{pmatrix}$$

The curl of a three-dimensional vector field is a three-dimensional vector field.

V Directional derivative

$$(\mathbf{f} \cdot \nabla) \mathbf{g} = \sum_{i=1}^n f_i \frac{\partial \mathbf{g}}{\partial x_i}$$

VI Laplace operator

$$\nabla^2 f = \nabla \cdot (\nabla f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$$

The Laplace operator applied to a tensor field is a tensor field of the same order and same dimension (but not necessarily the same field).

VII Poisson's equation (卜瓦松/帕松/泊松方程)

$$\nabla^2 A = B(\mathbf{x})$$

VIII Laplace's equation (拉普拉斯方程)

$$\nabla^2 A = 0$$

A real function A with real independent variables that is second-order differentiable for all independent variables is called a harmonic function if A satisfies Laplace's equation.

IX Multi-index notation (多重指標記號)

Multiindex α is a convenient notation for partial derivatives and polynomial expansions in multiple variables. Suppose there are n variables x_1, x_2, \dots, x_n , then a multiindex is a vector of n non-negative integers:

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n), \quad \text{where } \alpha_i \in \mathbb{N}_0.$$

Define:

- Norm $\|\alpha\|$:

$$\|\alpha\| = \alpha_1 + \alpha_2 + \dots + \alpha_n.$$

- Factorial $\alpha!$:

$$\alpha! = \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n!.$$

- Power \mathbf{x}^α : If $\mathbf{x} = (x_1, x_2, \dots, x_n)$, then

$$\mathbf{x}^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

- High-order mixed partial derivatives $D^\alpha f$:

$$D^\alpha f = \frac{\partial^{\|\alpha\|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}.$$

X High-order derivative

The k th order derivative of $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, denoted as $\mathbf{F}^{(k)}(\mathbf{x})$ or $D^k \mathbf{F}(\mathbf{x})$, is a $(\mathbb{R}^n)^k \rightarrow \mathbb{R}$ function, where $(\mathbb{R}^n)^k$ is a Cartesian product of k copies of \mathbb{R}^n vector, that is,

$$D^k \mathbf{F}(\mathbf{x}) = \sum_{\|\alpha\|=k} (D^\alpha \mathbf{F}(\mathbf{a})).$$

In particular, the first-order derivative of \mathbf{F} is the gradient of it.

XI Taylor expansion

Assume that $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an infinitely differentiable function, and its partial derivatives of every order exist on \mathbb{R}^n , then the Taylor expansion of \mathbf{F} at \mathbf{a} is

$$\mathbf{F}(\mathbf{x}) = \sum_{\|\alpha\| \in \mathbb{N}_0} \frac{D^\alpha \mathbf{F}(\mathbf{a})}{\alpha!} (\mathbf{x} - \mathbf{a})^\alpha,$$

that is,

$$\mathbf{F}(\mathbf{x}) = \sum_{\|\alpha\| \leq k} \frac{D^\alpha \mathbf{F}(\mathbf{a})}{\alpha!} (\mathbf{x} - \mathbf{a})^\alpha + \int_0^1 \frac{(1-t)^k}{k!} D^{k+1} \mathbf{F}(\mathbf{a} + t(\mathbf{x} - \mathbf{a})) (\mathbf{x} - \mathbf{a})^{k+1} dt.$$

Also, the k th-order approximation of \mathbf{F} near \mathbf{a} is

$$\mathbf{F}(\mathbf{x}) \approx \sum_{\|\alpha\| \leq k} \frac{D^\alpha \mathbf{F}(\mathbf{a})}{\alpha!} (\mathbf{x} - \mathbf{a})^\alpha,$$

and the first-order approximation near \mathbf{a} is

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{a}) + \nabla \mathbf{F}(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}).$$

7 Line integral (線積分)/Path integral (路徑積分)

I Scalar field line/path integral

For a scalar field $A : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and the path $C \in U$, the line integral of A is:

$$\int_C A ds = \int_a^b A(\mathbf{r}(t)) \left\| \frac{d}{dt} \mathbf{r}(t) \right\| dt,$$

where $\mathbf{r} : [a, b] \rightarrow C$ is a one-to-one parametric function with $\mathbf{r}(a)$ and $\mathbf{r}(b)$ being the two endpoints of the path C .

A is called the integral function, C is called the integral path, and the result of the line integration does not depend on the parametric function r .

II Vector field line/path integral

For a scalar field $\mathbf{F} : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ and the path $C \in U$, the line integral of \mathbf{F} is:

$$\int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \frac{d}{dt} \mathbf{r}(t) dt$$

where $\mathbf{r} : [a, b] \rightarrow C$ is a one-to-one parametric function with $\mathbf{r}(a)$ and $\mathbf{r}(b)$ being the two endpoints of the path C .

\mathbf{F} is called the integral function, C is called the integral path, and the result of the line integration does not depend on the parametric function \mathbf{r} .

III Conservative field (保守場)

A field f whose domain is a subset U of a Euclidean tensor space is called a conservative field if for all paths C between point a and b , the integral

$$\int_C f(\mathbf{x}) \cdot d\mathbf{x}$$

are the same.

This implies

- For any closed path C ,

$$\int_C f(\mathbf{x}) \cdot d\mathbf{x} = 0.$$

- If $\dim(U) = 3$, then for any subset of U where f is smooth,

$$\nabla \times f = 0.$$

8 Fundamental theorem of multivariable calculus (多變數微積分基本定理)

I Gradient theorem (梯度定理)

Suppose r is a oriented differentiable curve that starts at a point \mathbf{p} and ends at a point \mathbf{q} . If \mathbf{F} is a differentiable tensor field defined on a neighborhood of \mathbf{F} , then,

$$\int_r (\nabla \mathbf{F}) \cdot d\mathbf{r} = \mathbf{F}(\mathbf{q}) - \mathbf{F}(\mathbf{p}).$$

Gradient theorem is a special case of generalized Stokes theorem.

II Divergence theorem, Gauss's theorem, or Ostrogradsky's theorem (高斯散度定理)

Suppose $V \subseteq \mathbb{R}^n$ is compact and has a piecewise smooth boundary S (also indicated with $\partial V = S$). The closed, measurable set ∂V is oriented by outward-pointing normals. If F is a continuously differentiable vector field defined on a neighborhood of V , then,

$$\iiint_V (\nabla \cdot \mathbf{F}) dV = \iint_S \mathbf{F} \cdot d\mathbf{S}$$

Divergence theorem is a special case of generalized Stokes theorem.

III Stokes' theorem (斯托克斯定理) or Kelvin–Stokes theorem

Let S be a positively oriented, piecewise smooth surface in \mathbb{R}^3 with boundary $\partial S \equiv L$. If a vector field $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is defined and has continuous first order partial derivatives in a region containing S , then,

$$\iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \oint_L \mathbf{F} \cdot d\mathbf{L}$$

Stokes' theorem is a special case of generalized Stokes theorem.

IV Green's theorem (格林定理或綠定理)

Let S be a positively oriented, piecewise smooth surface in \mathbb{R}^2 with boundary $\partial S \equiv L$. If scalar function $P, (x, y) Q(x, y)$ are defined and has continuous first order partial derivatives in a region containing S , then,

$$\oint_L (Pdx + Qdy) = \iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$$

where the path of integration along C is counterclockwise.

Green's theorem is a special case of Stokes' theorem.

V Generalized Stokes theorem, Stokes–Cartan theorem, or fundamental theorem of multivariable calculus

The generalized Stokes theorem says that the integral of a differential form ω over the boundary $\partial\Omega$ of some orientable manifold Ω is equal to the integral of its exterior derivative $d\omega$ over the whole of Ω , i.e.,

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega$$

9 Differential theorems

I Intermediate Value Theorem, IVT (中間值定理)

Let $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function, then

$$\forall [a, b] \subseteq I \text{ s.t. } f(a) \neq f(b) : k \in (f(a), f(b)) \implies (\exists c \in (a, b) \text{ s.t. } f(c) = k)$$

II Mean Value Theorem, MVT (均值定理)

Let $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function, and f is differentiable on an interval $(a, b) \subseteq I$, then

$$\exists c \in (a, b) \text{ s.t. } f'(c) = \frac{f(b) - f(a)}{b - a}$$

III Extreme Value Theorem, EVT (極值定理)

Let function $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be continuous on an interval $[a, b]$, then

$$\exists x_{\max}, x_{\min} \in [a, b] \text{ s.t. } (\forall x \in [a, b] : f(x_{\max}) \geq f(x) \geq f(x_{\min}))$$

IV L'Hôpital's rule (羅必達法則)/Bernoulli's rule

Let I be an interval containing the point a . Let $f(x)$ and $g(x)$ be functions defined on I , except possibly at a itself. Let $f(x)$ and $g(x)$ be differentiable at all points except a in I . If $\lim_{x \rightarrow a} f(x) =$

$\lim_{x \rightarrow a} g(x) = 0$ or ∞ or $-\infty$, and $\exists \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$, then:

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

V Critical point (臨界點)

Let f be a real function and c be a point in D_f , if $f'(c) = 0$ or f' does not exist at c , then c is a critical point of f .

VI Relative extremum (相對極值), local extremum (局部極值), or extremum (極值)

Let f be a real function and c be a point in D_f , a relative maximum or maximum $f(c)$ of f is said to occur at c if there exists an open interval I where $c \in I \subseteq D_f$ such that $\forall x \in I : f(c) \geq f(x)$.

Let f be a real function and c be a point in D_f , a relative minimum or minimum $f(c)$ of f is said to occur at c if there exists an open interval I where $c \in I \subseteq D_f$ such that $\forall x \in I : f(c) \leq f(x)$.
The relative maximum and relative minimum are collectively called the relative extreme.

VII Absolute extremum (絕對極值/最值) or global extremum (全域極值)

Let f be a real function and c be a point in D_f , an absolute maximum (絕對極大值/最大值) $f(c)$ of f is said to occur at c if $\forall x \in D_f : f(c) \geq f(x)$.

Let f be a real function and c be a point in D_f , an absolute minimum (絕對極小值/最小值) $f(c)$ of f is said to occur at c if $\forall x \in D_f : f(c) \leq f(x)$.

The absolute maximum and absolute minimum are collectively called the absolute extreme.

VIII Relative extreme theorem

A point in I where a relative extreme of $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ occurs must be a critical point.

IX Concavity (凹性)

Let $f : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be differentiable on the open interval $I \subseteq J$. If f' is strictly increasing on I , the graph of f is said to concave upward on I ; if f' is strictly decreasing on I , the graph of f is said to concave downward on I ; if f' is a constant on I , the graph of f is said to be neither upward nor downward (or both upward and downward, or undefined, in some contexts) on I .

X Concavity theorem

Let $f : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be differentiable on the open interval $I \subseteq J$. If $\forall x \in I : f'' > 0$, then the graph of f is concave upward on I ; if $\forall x \in I : f'' < 0$, then the graph of f is concave downward on I .

XI Point of inflection or inflection point (反曲點/拐點)

Let $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function and be differentiable on an open interval $J \subseteq I$, and let $a < b < c \in J$. If the graph of f is concave upward on interval (a, b) and concave downward on interval (b, c) , or concave downward on interval (a, b) and concave upward on interval (b, c) , then $(b, f(b))$ is called an inflection point of the graph of f .

XII Inflection point theorem

If $(c, f(c))$ is an inflection point of f , then $f''(c) = 0$ or f'' does not exist at c .

XIII Lagrange multiplier (拉格朗日乘數)

The Lagrange multiplier method is a method for finding the extreme value of a differentiable function under differential constraints. The solution obtained by the Lagrange multiplier method will contain all the extreme points of the original problem, but it is not guaranteed that every critical point obtained by the Lagrange multiplier method is the critical point of the original problem.

i Univariate form

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$. We want to find the extreme points of $f(x)$ under the constraint $g(x) = 0$.

First, construct the Lagrangian function $\mathcal{L}(x, \lambda)$:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda \cdot g(x),$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier.

Find the derivative of \mathcal{L} and set it to zero:

$$\frac{d\mathcal{L}}{dx} = 0$$

Solve the equation to find x and λ .

Statement: The solution to x obtained by the Lagrange multiplier method above will contain all the extreme points of the original problem.

ii Multivariate form

Let $\mathbf{x} = (x_1, x_2, \dots, x_n)$ be the independent variable vector and $\mathbf{0}$ be the zero vector. Now we have $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}^c$. We want to find the extreme points of $f(\mathbf{x})$ under the constraint $g(\mathbf{x}) = \mathbf{0}$.

First, construct the Lagrangian function $\mathcal{L}(\mathbf{x}, \lambda)$:

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda \cdot g(\mathbf{x}),$$

where $\lambda \in \mathbb{R}^c$ is the Lagrange multiplier vector.

Find the gradient of \mathcal{L} and set it to zero:

$$\nabla \mathcal{L} = \mathbf{0}$$

Solve the equation to find \mathbf{x} and λ .

Statement: The solution to \mathbf{x} obtained by the Lagrange multiplier method above will contain all the extreme points of the original problem.

Proof.

Consider an extreme point \mathbf{x}^* in the original problem. Since it is an extreme point under the constraint, it must satisfy the constraint:

$$g(\mathbf{x}^*) = \mathbf{0}.$$

Any feasible point \mathbf{x} near \mathbf{x}^* must satisfy the constraint. We can represent \mathbf{x} as:

$$\mathbf{x} = \mathbf{x}^* + \delta \mathbf{x},$$

where $\delta \mathbf{x}$ is a differential change perpendicular to the manifold defined by $g(\mathbf{x})$ in the kernel of $\nabla g(\mathbf{x}^*)$, that is,

$$g(\mathbf{x}^* + \delta \mathbf{x}) = \mathbf{0}.$$

Find the first-order Taylor expansion approximation of f at \mathbf{x}^* :

$$f(\mathbf{x}^* + \delta \mathbf{x}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*) \cdot \delta \mathbf{x} + O(\|\delta \mathbf{x}\|^2)$$

Find the first-order Taylor expansion approximation of g at \mathbf{x}^* :

$$g(\mathbf{x}^* + \delta\mathbf{x}) \approx g(\mathbf{x}^*) + \nabla g(\mathbf{x}^*) \cdot \delta\mathbf{x} + O(\|\delta\mathbf{x}\|^2)$$

Since \mathbf{x}^* is an extreme point, for any feasible $\delta\mathbf{x}$ we must have:

$$\nabla f(\mathbf{x}^*) \cdot \delta\mathbf{x} = 0.$$

Since $g(\mathbf{x}^*) = \mathbf{0}$, we have

$$\nabla g(\mathbf{x}^*) \cdot \delta\mathbf{x} = O(\|\delta\mathbf{x}\|^2).$$

Because $\delta\mathbf{x} \in \ker(\nabla g(\mathbf{x}^*))$, $\nabla f(\mathbf{x}^*)$ can be expressed as a linear combination of $\nabla g(\mathbf{x}^*)$. This means that there exists a vector λ such that:

$$\nabla \mathcal{L} = \nabla (f(\mathbf{x}) - \lambda \cdot g(\mathbf{x})) = \mathbf{0}$$

□

10 Integral theorems

I Integration by substitution (代換積分法), integration by change of variables (換元積分法), u-substitution (u-代換), reverse chain rule, substitution theroem (代換定理), change of variables theorem (換元定理), or transformation theorem (變換定理)

i Univariate form

Let $g : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be injective and differentiable on $[a, b] \subseteq I$, with g' being integrable on $[a, b]$, and $f : K \supseteq g([a, b]) \rightarrow \mathbb{R}$ be integrable on $g([a, b])$. Then:

$$\int_a^b f \circ g(x) \cdot g'(x) dx = \int_{g(a)}^{g(b)} f(u) du,$$

and for $K = g([a, b])$:

$$\int f \circ g(x) \cdot g'(x) dx = \int f(u) du.$$

Proof.

Consider the interval $[a, b]$ partitioned as

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\}.$$

For each subinterval $[x_{i-1}, x_i]$, let $\xi_i \in [x_{i-1}, x_i]$ be a sample point. The Riemann sum for the left-hand side of the integral is:

$$S_P = \sum_{i=1}^n f(g(\xi_i))g'(\xi_i)(x_i - x_{i-1}).$$

Since g is injective and differentiable, it is either strictly increasing or strictly decreasing on $[a, b]$. Assume g is strictly increasing (the proof for g strictly decreasing follows similarly).

Under this assumption, g maps $[a, b]$ to $[g(a), g(b)]$ (with $g(a) < g(b)$). Let $[g(a), g(b)]$ be partitioned as

$$Q = \{g(a) = u_0 < u_1 < \cdots < u_m = g(b)\},$$

where $u_i = g(x_i)$. For g increasing, $(u_i - u_{i-1}) = g(x_i) - g(x_{i-1})$.

The Riemann sum for the right-hand side of the integral is:

$$T_Q = \sum_{i=1}^n f(u_i)(u_i - u_{i-1}).$$

Since $u_i = g(x_i)$ and $g'(\xi_i) \approx \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}$, we rewrite:

$$g'(\xi_i)(x_i - x_{i-1}) \approx g(x_i) - g(x_{i-1}) = u_i - u_{i-1}.$$

Thus, the left-hand Riemann sum S_P becomes:

$$S_P = \sum_{i=1}^n f(g(\xi_i))(u_i - u_{i-1}),$$

which matches the structure of the right-hand Riemann sum T_Q if we let $\xi_i = g^{-1}(u_i)$.

As the partition P of $[a, b]$ gets finer, the Riemann sum S_P converges to $\int_a^b f(g(x))g'(x) dx$. Similarly, as the partition Q of $[g(a), g(b)]$ gets finer, the Riemann sum T_Q converges to $\int_{g(a)}^{g(b)} f(u) du$. \square

ii Multivariate form

Let $\mathbf{T} : I \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ be injective and differentiable on $D \subseteq I$, with all elements of its gradient (i.e. Jacobian matrix) $\nabla \mathbf{T}$ being continuous on D , and $f : K \supseteq \mathbf{T}(D) \rightarrow \mathbb{R}$ be integrable on $\mathbf{T}(D)$. Then:

$$\int_{\mathbf{T}(D)} f(x_1 x_2 \dots x_n) \cdot dx_1 dx_2 \dots dx_n = \int_D f(u_1 u_2 \dots u_n) |\det(\nabla \mathbf{T})| du_1 du_2 \dots du_n,$$

and for $K = \mathbf{T}(D)$:

$$\int f(x_1 x_2 \dots x_n) \cdot dx_1 dx_2 \dots dx_n = \int f(u_1 u_2 \dots u_n) |\det(\nabla \mathbf{T})| du_1 du_2 \dots du_n.$$

iii Measure theory form

Let X be a locally compact Hausdorff space equipped with a finite Radon measure μ , and let Y be a σ -compact Hausdorff space with a σ -finite Radon measure ρ . Let $\phi : X \rightarrow Y$ be an absolutely continuous function, (which implies that $\mu(E) = 0 \implies \rho(\phi(E)) = 0$). Then there exists a real-valued Borel measurable function w on X such that for every Lebesgue integrable function $f : Y \rightarrow \mathbb{R}$, the function $(f \circ \phi) \cdot w$ is Lebesgue integrable on X , and for every open subset U of X

$$\int_{\phi(U)} f(y) d\rho(y) = \int_U (f \circ \phi)(x) \cdot w(x) d\mu(x).$$

Furthermore, there exists some Borel measurable function g such that

$$w(x) = (g \circ \phi)(x).$$

II Integration by parts (分部積分法) or partial integration (部分積分法)

i Theorem

$$\frac{d}{dx} \prod_{i=1}^n f_i(x) = \sum_{j=1}^n \left(\frac{df_j(x)}{dx} \frac{\prod_{i \neq j}^n f_i(x)}{f_j(x)} \right)$$

ii Application

Integration by parts is a heuristic rather than a purely mechanical process for solving integrals; given a single function to integrate, the typical strategy is to carefully separate this single function into a product of two functions such that the residual integral from the integration by parts formula is easier to evaluate than the single function.

The DETAIL rule or the LIATE rule is a rule of thumb for integration by parts. It involves choosing as u the function that comes first in the following list:

- L: Logarithmic function
- I: Inverse trigonometric function
- A: Algebraic function (such as polynomial function)
- T: Trigonometric function
- E: Exponential function

11 Common calculus results

I Polynomial function

$$\frac{d}{dx} x^n = n \cdot x^{n-1}$$

Proof.

$$\begin{aligned} \frac{d}{dx} x^n &= \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^n - x^n}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\left(\sum_{k=0}^n \binom{n}{k} x^{n-k} (\Delta x)^k \right) - x^n}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\sum_{k=1}^n \binom{n}{k} x^{n-k} (\Delta x)^k}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \sum_{k=1}^n \binom{n}{k} x^{n-k} (\Delta x)^{k-1} \\ &= nx^{n-1} \end{aligned}$$

□

II Exponential function

$$\frac{d}{dx}e^x = e^x$$

Proof.

$$\begin{aligned}\frac{d}{dx}e^x &= \lim_{\Delta x \rightarrow 0} \frac{e^{x+\Delta x} - e^x}{\Delta x} \\&= \lim_{\Delta x \rightarrow 0} \frac{e^x \cdot e^{\Delta x} - e^x}{\Delta x} \\&= e^x \cdot \lim_{\Delta x \rightarrow 0} \frac{e^{\Delta x} - 1}{\Delta x} \\&= e^x \cdot \lim_{\Delta x \rightarrow 0} \frac{\lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{(\Delta x)^i}{i!} - 1}{\Delta x} \\&= e^x \cdot \lim_{\Delta x \rightarrow 0} \frac{\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{(\Delta x)^i}{i!}}{\Delta x} \\&= e^x \cdot \lim_{\Delta x \rightarrow 0} \sum_{i=1}^n \frac{(\Delta x)^{i-1}}{i!} \\&= e^x\end{aligned}$$

□

III Logarithmic function

$$\frac{d}{dx} \ln(x) = \frac{1}{x}$$

Proof.

$$\begin{aligned}\frac{d}{dx} \ln(x) &= \lim_{\Delta x \rightarrow 0} \frac{\ln(x + \Delta x) - \ln(x)}{\Delta x} \\&= \lim_{\Delta x \rightarrow 0} \frac{\ln\left(\frac{x+\Delta x}{x}\right)}{\Delta x} \\&= \lim_{\Delta x \rightarrow 0} \frac{\ln\left(1 + \frac{\Delta x}{x}\right)}{\Delta x} \\&= \lim_{\frac{\Delta x}{x} \rightarrow 0} \frac{\ln\left(1 + \frac{\Delta x}{x}\right)}{\frac{\Delta x}{x}} \cdot \frac{1}{x} \\&= \frac{1}{x}\end{aligned}$$

□

IV Non-polynomial algebraic function

$$\int \sqrt{a^2 - x^2} \, dx = \frac{1}{2} \left(x \sqrt{a^2 - x^2} + \arcsin \frac{x}{a} \right) + C, \quad a > x.$$

$$\int (a^2 + x^2)^{-1} \, dx = \frac{1}{a} \arctan \left(\frac{x}{a} \right) + C.$$

12 Numerical differentiation (數值微分)

I Newton's Method (牛頓法) or Newton-Raphson Method (牛頓-拉普森法)

Newton's method, also known as Newton-Raphson method, is an iterative technique used to approximate the roots of a real-valued function. Given a function $f(x)$ and an initial guess x_0 close to a root, Newton's method refines this guess by repeatedly applying the formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, n \in \mathbb{N}$$

where:

- x_n is the current approximation,
- $f(x_n)$ is the value of the function at x_n ,
- $f'(x_n)$ is the derivative of $f(x)$ evaluated at x_n .

13 Numerical integration (數值積分)

I The trapezoidal rule (梯形法)

Let f be a continuous real-valued function on $[a, b]$, the trapezoidal rule gives the approximation

$$\int_a^b f(x) dx \approx \frac{b-a}{2n} \left(2 \left(\sum_{i=0}^n f\left(a + \frac{i}{n}(b-a)\right) \right) - f(a) - f(b) \right).$$

The error is defined as

$$E_n = \int_a^b f(x) dx - \frac{b-a}{2n} \left(2 \left(\sum_{i=0}^n f\left(a + \frac{i}{n}(b-a)\right) \right) - f(a) - f(b) \right).$$

When $\frac{d^2 f(x)}{dx^2}$ is continuous on $[a, b]$, the error satisfies that

$$|E_n| \leq \frac{(b-a)^3}{12n^2} \max_{a \leq x \leq b} \left(\left| \frac{d^2 f(x)}{dx^2} \right| \right).$$

II The Simpson's rule (辛普森法) or the Simpson's 1/3 rule

Let f be a continuous real-valued function on $[a, b]$, the Simpson's rule gives the approximation

$$\int_a^b f(x) dx \approx \frac{b-a}{3n} \left(2 \left(\sum_{i=0}^n f\left(a + \frac{i}{n}(b-a)\right) \right) + 2 \left(\sum_{i=1}^{\frac{n}{2}} f\left(a + \frac{2i-1}{n}(b-a)\right) \right) - f(a) - f(b) \right),$$

where $\frac{n}{2} \in \mathbb{N}$.

The error is defined as

$$E_n = \int_a^b f(x) dx - \frac{b-a}{3n} \left(2 \left(\sum_{i=0}^n f\left(a + \frac{i}{n}(b-a)\right) \right) + 2 \left(\sum_{i=1}^{\frac{n}{2}} f\left(a + \frac{2i-1}{n}(b-a)\right) \right) - f(a) - f(b) \right).$$

When $\frac{d^4 f(x)}{dx^4}$ is continuous on $[a, b]$, the error satisfies that

$$|E_n| \leq \frac{(b-a)^5}{180n^4} \max_{a \leq x \leq b} \left(\left| \frac{d^4 f(x)}{dx^4} \right| \right).$$