Limit and Calculus

沈威宇

July 1, 2025

Contents

Chapter 1　Limit and Calculus (極限與微積分)

1　Limit (極限)

I　A Limit for a Sequence (數列)

i　Glossary of Terms

- **(Monotone) Increasing （（單調） 遞增)/Non-Decreasing (非遞減) Sequence**: $\langle a_n \rangle$ is a monotone increasing/increasing/non-decreasing sequence if and only if $\forall n$ s.t. $\exists a_n, a_{n+1}: a_n \leq a_{n+1}$.

- **Strictly Increasing (嚴格遞增) Sequence**: $\langle a_n \rangle$ is a strictly increasing sequence if and only if $\forall n$ s.t. $\exists a_n, a_{n+1}: a_n < a_{n+1}$.

- **(Monotone) Decreasing （（單調） 遞減)/Non-Increasing (非遞增) Sequence**: $\langle a_n \rangle$ is a monotone decreasing/decreasing/non-increasing sequence if and only if $\forall n$ s.t. $\exists a_n, a_{n+1}: a_n \geq a_{n+1}$.

- **Strictly Decreasing (嚴格遞減) Sequence**: $\langle a_n \rangle$ is a strictly decreasing sequence if and only if $\forall n$ s.t. $\exists a_n, a_{n+1}: a_n > a_{n+1}$.

- **Monotone (單調) Sequence**: $\langle a_n \rangle$ 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 $\forall n$ s.t. $\exists a_n: a_n \leq M$. $\exists$ upper bound of $\langle a_n \rangle \iff \langle a_n \rangle$ is bounded-above.

- **Lower Bound (下界)**: m is an lower bound of $\langle a_n \rangle$ if and only if $\forall n$ s.t. $\exists a_n: a_n \geq m$. $\exists$ 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.

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:

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

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

iii   Infinite Limits

$$\lim_{n \to \infty} a_n = \infty \equiv \forall M > 0, \exists N \in \mathbb{N} \text{ s.t. } n > N \implies a_n > M.$$

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

iv   Existence of a Limit

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

v   Convergence (收斂) and Divergence (發散)

A sequence $\langle a_n \rangle$ converges to L if $\exists \lim_{n \to \infty} a_n \wedge \lim_{n \to \infty} a_n = L$. 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:

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

Then:

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

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

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

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

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

vii   Special Sequences and Their Limits

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

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

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

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

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

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

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

then:

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

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

Statement.

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

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

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

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

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 $c = \sup_n a_n$.

$$\forall \epsilon > 0 : \exists N \text{ s.t. } c \geq a_N > c - \epsilon,$$

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:

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

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 (級數)

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 $\lim\limits_{n\to\infty} S_n$ or $\sum\limits_{i=1}^{\infty} a_i$, is defined as the following:

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

ii   Convergence (收斂) and Divergence (發散)

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

iii   Limit Laws

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

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

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 Finity

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:

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

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:

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

  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:

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

  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)

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

iv   Slant asymptote (斜漸近線)

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

v   One-side Limits

- **Right-hand Limit (右極限)**: 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:

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

  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 (左極限)**: 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:

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

  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

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

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

$$\lim_{x \to \infty} f(x) = \infty \equiv \forall M > 0, \exists \delta > 0 \text{ s.t. } x > \delta \implies f(x) > M.$$

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

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

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

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

vii Existence of Limits

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

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

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

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:

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

Then:

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

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

$$\lim_{x \to a} c f(x) = c \lim_{x \to a} f(x), \quad c \in \mathbb{R}$$

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

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim_{x \to a} f(x)}{\lim_{x \to a} g(x)}, \quad g(x) \neq 0 \wedge \lim_{x \to a} g(x) \neq 0$$

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

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

then:

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

- If $\exists \lim_{x \to a} f(x)$ and $\lim_{x \to a} f(x) = f(a)$, 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 $\lim_{x \to a^+} = f(a)$, 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 $\lim_{x \to b^-} = f(b)$, 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 $g(a) \leq 0$, then $\frac{f(x)}{g(x)}$ is continuous at $x = a$.

- If both $f(x)$ and $g(x)$ are continuous at $x = a$ and $g(a) \in \operatorname{dom}(f)$, then $f \circ g(x)$ is continuous at $x = a$, namely, $\lim_{x \to a} ((f \circ g))(x)) = \lim_{x \to a} (f(g(x))) = f\left(\lim_{x \to a} g(x)\right) = f(g(a)) = f \circ g(a)$.

xi Theorem: Limits Involving Quotient Functions

Let a and b be real numbers and

$$f(x) \in \{f(x): U \subseteq \mathbb{R} \to \mathbb{R} \mid f(x) = x \vee \ln(f(x)) \in A \vee e^{f(x)} \in A\}.$$

Then:

$$\lim_{x \to \infty} \frac{(f(x))^a}{(f(x))^b} = \infty, \quad a > b$$

$$\lim_{x \to \infty} \frac{af(x)}{bf(x)} = \frac{a}{b}, \quad b \neq 0$$

$$\lim_{x \to \infty} \frac{n^{af(x)}}{bf(x)} = \infty, \quad a, b > 0 \wedge n > 1$$

$$\lim_{x \to \infty} \frac{n^{af(x)}}{bf(x)} = 0, \quad a, b > 0 \wedge 0 \leq n < 1$$

$$\lim_{x \to \infty} \frac{af(x)}{b \log_n f(x)} = \infty, \quad a, b > 0 \wedge n > 1$$

2 Differentiation (微分)

I Leibniz's notation (萊布尼茲符號) for differentiation

Suppose a dependent variable y represents a function f of an independent variable x, that is,

$$y = f(x).$$

Then the derivative of the function f can be written as

$$\frac{\mathrm{d}y}{\mathrm{d}x} \text{ or } \frac{\mathrm{d}}{\mathrm{d}x}y \text{ or } \frac{\mathrm{d}\,(f(x))}{\mathrm{d}x} \text{ or } \frac{\mathrm{d}}{\mathrm{d}x}\,(f(x))\,.$$

The nth derivative of the function f can be written as

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} \text{ or } \frac{\mathrm{d}^n}{\mathrm{d}x^n}y \text{ or } \frac{\mathrm{d}^n\,(f(x))}{\mathrm{d}x^n} \text{ or } \frac{\mathrm{d}^n}{\mathrm{d}x^n}\,(f(x))\,.$$

II   Lagrange's notation (拉格朗日符號) for differentiation

Suppose a dependent variable y represents a function f of an independent variable x, that is,

$$y = f(x).$$

Then the derivative of the function f can be written as

$$f'(x),$$

the second derivative of the function f can be written as

$$f''(x),$$

and so on.

The nth derivative of the function f can also be written as

$$f^{(n)}(x).$$

III   Newton's notation (牛頓符號), dot notation, or fluxions for differentiation

Suppose a dependent variable y represents a function f of an independent variable t, that is,

$$y = f(t),$$

where t usually represents time.

Then the derivative of the function f can be written as

$$\dot{y},$$

the second derivative of the function f can be written as

$$\ddot{y},$$

and so on.

IV 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 $f : U \to W$ is called Fréchet differentiable (弗蘭歇可微的) or differentiable at $x \in U$ if there exists a bounded linear operator $A : V \to W$ such that

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

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

And we call the function Df:

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

Fréchet derivative function (弗蘭歇導函數) or derivative function (導函數) of f.

The derivative function of the derivative function of f is called the second derivative function f; the derivative function of the k derivative function of f is called the $k + 1$th derivative function f.

V Smoothness (光滑性或平滑性)

A function f that has a kth derivative that is continuous in its domain is said to be of class C^k or be a C^k-function.

Generally, the term smooth function refers to a C^∞-function. However, it may also mean "sufficiently differentiable" for the problem under consideration.

VI 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 $F : U \to W$. The Gateaux derivative $dF(u; \psi)$ of F at x in U in the direction $\psi \in V$ is defined to be

$$dF(u; \psi) = \lim_{\tau \to 0} \frac{F(u + \tau\psi) - F(u)}{\tau}$$
$$= \frac{\mathrm{d}}{\mathrm{d}\tau} F(u + \tau\psi)\big|_{\tau=0}$$

If the limit exists for all $\psi \in V$, then it is said that F is Gateaux differentiable (加托可微的) at u.

VII Taylor series (泰勒級數) or Taylor expansion (泰勒展開)

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

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

that is,

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

Also, the kth-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

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

The Taylor series of F at 0 is called Maclaurin series (馬克勞林級數) or Maclaurin expansion (馬克勞林展開).

3   Integration (積分)

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

Two equivalent definitions of definite integral (定積分).

i   Partition of an interval

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

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

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,

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

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.

ii   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

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

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.

iii 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:

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

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

iv 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,

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

v Darboux integral (達布積分)

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,

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

vi Integrability

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

II Indicator function (指示函數或示性函數) or characteristic function (特徵函數)

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)

A definition of definite integral.

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

i Indicator functions

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

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

1. 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 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

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

ii Non-negative measurable functions

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

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

iii 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.

IV Antiderivative (反導函數), inverse derivative, primitive function, primitive integral or indefinite integral (不定積分)

An antiderivative of a continuous function f is a differentiable function F whose derivative is equal to the original function f, that is,

$$F' = f.$$

Suppose f is a function of an independent variable x, then its antiderivative F is denoted as

$$F = \int f \, dx.$$

The process of solving for antiderivatives is called antidifferentiation (反微分) or indefinite integration (不定積分).

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:

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

where $\{x_i^*\}$ are sample points in the subintervals of a partition $P = \{x_0, x_1, \ldots, x_n\}$ 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 $P = \{x_0, x_1, \ldots, x_n\}$ 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:

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

$\square$

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.

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

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:

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

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

$\square$

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

I Space convention

- The domain of the funcitons 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.

- **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: $\oiint$

- $\int \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ is used as a shorthand for $\int (\mathbf{F} \cdot \hat{\mathbf{n}}) \, \mathrm{d}S$, where $\hat{n}$ 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 ith direction, i.e., a vector with zero norm.

- Independent variable vector: $\mathbf{x} = (x_1, x_2, \ldots, x_n)$

- 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

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

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

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

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

III Divergence

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

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.

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

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

V Directional derivative

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

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 $x_1, x_2, \ldots, x_n$, then a multiindex is a vector of n non-negative integers:

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

Define:

- Norm $\|\alpha\|$:
$$\|\alpha\| = \alpha_1 + \alpha_2 + \cdots + \alpha_n.$$

- Factorial $\alpha!$:
$$\alpha! = \alpha_1! \cdot \alpha_2! \cdot \cdots \cdot \alpha_n!.$$

- Power $\mathbf{x}^\alpha$: If $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, 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 kth order derivative of $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$, denoted as $\mathbf{F}^{(k)}(\mathbf{x})$ or $D^k \mathbf{F}(\mathbf{x})$, is a $(\mathbb{R}^n)^k \to \mathbb{R}$ function, where $(\mathbb{R}^n)^k$ is a Cartesian product of k copies of $\mathbb{R}^n$ vector, that is,

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

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

XI Taylor expansion

Assume that $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ 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,

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

Also, the kth-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 $A: U \subseteq \mathbb{R}^n \to \mathbb{R}$ and the path $C \in U$, the line integral of A is:

$$\int_C A \, \mathrm{d}s = \int_a^b A(\mathbf{r}(t)) \| \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{r}(t) \| \, \mathrm{d}t,$$

where $\mathbf{r}: [a, b] \to C$ 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 $\mathbf{F}: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and the path $C \in U$, the line integral of $\mathbf{F}$ is:

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

where $\mathbf{r}: [a, b] \to C$ 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

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

are the same.

This implies

- For any closed path C,

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

- 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,

$$\iiint_V \left(\nabla \cdot \mathbf{F} \right) dV = \oiint_S \mathbf{F} \cdot d\mathbf{S}$$

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 $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ 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,

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

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 $\mathrm{d}\omega$ over the whole of Ω, i.e.,

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

9 Differential theorems

I Distributive property (分配性)

Differentiation is distributive over addition and subtraction.

II Product rule (乘法定則)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(f(x)g(x)\right) = f'(x)g(x) + f(x)g'(x).$$

III Quotient rule (除法定則)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}.$$

IV Chain rule (連鎖律)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\left(f \circ g\right)(x)\right) = f'\left(g(x)\right)g'(x).$$

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

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

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

VI Mean Value Theorem, MVT (均值定理)

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

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

VII Extreme Value Theorem, EVT (極值定理)

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

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

VIII   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 $\lim\limits_{x \to a} f(x) = \lim\limits_{x \to a} g(x) = 0$ or ∞ or $-\infty$, and $\exists \lim\limits_{x \to a} \frac{f'(x)}{g'(x)}$, then:

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

IX   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.

X   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.

XI   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.

XII   Relative extreme theorem

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

XIII   Concavity (凹性)

Let $f : J \subseteq \mathbb{R} \to \mathbb{R}$ 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.

XIV Concavity theorem

Let $f : J \subseteq \mathbb{R} \to \mathbb{R}$ 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.

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

Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ 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.

XVI Inflection point theorem

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

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

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:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x} = 0$$

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 $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be the independent variable vector and $\mathbf{0}$ be the zero vector. Now we have $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^c$. 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 to λ such that:

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

$\square$

10　Integral theorems

I　Distributive property (分配性)

Integration is distributive over addition and subtraction.

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

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

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

$$\int f \circ g(x) \cdot g'(x) \, \mathrm{d}x = \int f(u) \, \mathrm{d}u.$$

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 $g'(\xi_i) \approx \dfrac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}$, 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 $\int_a^b f(g(x))g'(x) \, \mathrm{d}x$. Similarly, as the partition Q of $[g(a), g(b)]$ gets finer, the Riemann sum T_Q converges to $\int_{g(a)}^{g(b)} f(u) \, \mathrm{d}u$. $\qquad\square$

ii Multivariate form

Let $\mathbf{T} : I \subseteq \mathbb{R}^n \to \mathbb{R}^n$ 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 $f : K \supseteq \mathbf{T}(D) \to \mathbb{R}$ be integrable on $\mathbf{T}(D)$. Then:

$$\int_{\mathbf{T}(D)} f(x_1\,x_2\,\dots\,x_n) \cdot \mathrm{d}x_1\,\mathrm{d}x_2\,\dots\,\mathrm{d}x_n = \int_D f(u_1\,u_2\,\dots\,u_n)\left|\det\left(\nabla\mathbf{T}\right)\right|\,\mathrm{d}u_1\,\mathrm{d}u_2\,\dots\,\mathrm{d}u_n,$$

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

$$\int f(x_1\,x_2\,\dots\,x_n) \cdot \mathrm{d}x_1\,\mathrm{d}x_2\,\dots\,\mathrm{d}x_n = \int f(u_1\,u_2\,\dots\,u_n)\left|\det\left(\nabla\mathbf{T}\right)\right|\,\mathrm{d}u_1\,\mathrm{d}u_2\,\dots\,\mathrm{d}u_n.$$

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 $\phi : X \to Y$ 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 $f : Y \to \mathbb{R}$, the function $(f \circ \phi) \cdot w$ is Lebesgue integrable on X, and for every open subset U of X

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

Furthermore, there exists some Borel measurable function g such that

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

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

i Theorem

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

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 integrals

I Rational function

$$\int \frac{f'(x)}{f(x)} \, dx = \ln|f(x)| + C.$$

$$\int \frac{1}{x^2 + a^2} \, dx = \frac{1}{a} \arctan \frac{x}{a} + C.$$

$$\int \frac{1}{x^2 - a^2} \, dx = \frac{1}{2a} \ln \frac{x - a}{x + a} + C.$$

$$\int \frac{1}{a^2 - x^2} \, dx = \frac{1}{2a} \ln \frac{a + x}{a + x} + C.$$

$$\int \frac{1}{ax + b} \, dx = \frac{1}{a} \ln|ax + b| + C.$$

$$\int (ax + b)^n \, dx = \frac{(ax + b)^{n-1}}{a(n + 1)} + C, \quad n \neq -1.$$

$$\int \frac{x}{ax + b} \, dx = \frac{x}{a} - \frac{b}{a^2} \ln|ax + b| + C.$$

$$\int \frac{x}{(ax + b)^2} \, dx = \frac{b}{a^2(ax + b)} + \frac{1}{a^2} \ln|ax + b| + C.$$

$$\int x(ax + b)^n \, dx = \frac{a(n + 1)x - b}{a^2(n + 1)(n + 2)}(ax + b)^{n+1} + C, \quad n \notin \{-1, -2\}.$$

II Irrational function

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

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

$$\int \sqrt{a^2 - x^2} \, dx = \frac{x}{2}\sqrt{a^2 - x^2} + \frac{a^2}{2} \arcsin \frac{x}{a} + C, \quad |a| \geq |x|.$$

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

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

The error is defined as

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

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

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

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

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

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

The error is defined as

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

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

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

14 Differential Equation (微分方程)

I Ordinary Differential Equations (常微分方程)

i Characteristic Equation Method for Solving Linear Ordinary Differential Equations with Constant Coefficients (解常係數線性常微分方程的特徵方程法)

Equation to be solved:

$$\sum_{i=0}^n a_i \frac{\mathrm{d}^i y}{\mathrm{d}x^i} = f(x)$$

where a_i are constants.

1. **Take the homogeneous equation:**

$$\sum_{i=0}^{n} a_i \frac{\mathrm{d}^i y_h}{\mathrm{d}x^i} = 0$$

2. **Take the characteristic equation:**

Assume the homogeneous solution is in the form of $y_h = e^{\lambda x}$. By substituting it into the homogeneous equation, we get the characteristic equation:

$$\sum_{i=0}^{n} a_i \lambda^i = 0$$

3. **Find characteristic roots:**

Solving the above characteristic equation yields m distinct complex roots, where the ith is λ_i (the order does not matter).

4. **Construct homogeneous solution:**

Let the root λ_i be a repeated root with multiplicity of k_i. The corresponding homogeneous solution y_h is:

$$y_h = \sum_{i=1}^{m} e^{\Re(\lambda_i)x} \sum_{j=1}^{k_i} x^{j-1} \left(C_{ij} \cos\left(\Im \lambda_i x\right) + D_{ij} \sin\left(\Im \lambda_i x\right) \right)$$

where C_{ij} and D_{ij} are constants. If λ_i is a real number, then $\left(C_{ij} \cos\left(\Im r_i x\right) + D_{ij} \sin\left(\Im r_i x\right) \right)$ degenerates to C_{ij}. If $f(x) = 0$, that is, the original equation is a linear homogeneous ordinary differential equation with constant coefficients, then the homogeneous solution is the general solution.

5. **Heuristic method for constructing particular solutions:**

The heuristic method is to assume a particular solution y_p that satisfies the original non-homogeneous equation. The form of the particular solution depends on the inhomogeneous term $f(x)$ on the right-hand side. Some $f(x)$ are difficult to handle using the heuristic method.

If $f(x)$ has k different terms, we can find each particular solutions y_{pi} separately, and the total particular solution $y_p = \sum_{i=1}^{k} y_{pi}$. The particular solution of a term can be assumed to be the result of the term after a finite number of translations (i.e., replacing x with $x - h$, where h is a constant, or adding a constant to the expression) and/or stretches (i.e., replacing x with ax, where a is a constant, or multiplying the expression by a constant). However, if a particular solution is assumed to be a polynomial of a term in the homogeneous solution, the particular solution assumption must be multiplied by $a(x - h)$, where a and h are constants. If the particular solutions of different terms of $f(x)$ have a term of the same form, we can retain any one of them and skip the others. Common terms of $f(x)$ that can be handled by the heuristic method include:

- linear combinations of polynomials, exponential functions, and sine/cosine functions,

- logarithmic function, and

- tangent function.

6. Construct general solution:

The general solution y is:

$$y = y_h + y_p$$

Example:

Consider the second-order non-homogeneous linear ordinary differential equation:

$$y'' - 2y' + 3y = \ln(x) + \tan(x)$$

The homogeneous equation is:
$$y'' - 2y' + 3y = 0$$

The characteristic equation is:
$$\lambda^2 - 2\lambda + 3 = 0$$

The characteristic roots are:
$$\lambda = \frac{2 \pm \sqrt{4 - 12}}{2} = 1 \pm \sqrt{2}i$$

The homogeneous solution is:

$$y_h = e^x(C \cos\left(\sqrt{2}x\right) + D \sin\left(\sqrt{2}x\right))$$

Assume the particular solution for $\ln(x)$ is $y_{p1} = A \ln(x) + B$.

$$y'_{p1} = \frac{A}{x}$$

$$y''_{p1} = -\frac{A}{x^2}$$

Substitute into the original equation:

$$-\frac{A}{x^2} - 2\left(\frac{A}{x}\right) + 3(A \ln(x) + B) = \ln(x)$$

$$-\frac{A}{x^2} - \frac{2A}{x} + 3A \ln(x) + 3B = \ln(x)$$

$$A = \frac{1}{3}$$

$$-\frac{A}{x^2} - \frac{2A}{x} + 3B = 0$$

$$B = 0$$

$$y_{p1} = \frac{1}{3} \ln(x)$$

Assume the particular solution for $\ln(x)$ is $y_{p2} = C \tan(x)$.

$$y'_{p2} = C \sec^2(x)$$

29

$$y''_{p2} = 2C \sec^2(x) \tan(x)$$

Substitute into the original equation:

$$2C \sec^2(x) \tan(x) - 2(C \sec^2(x)) + 3(C \tan(x)) = \tan(x)$$

$$2C \sec^2(x) \tan(x) - 2C \sec^2(x) + 3C \tan(x) = \tan(x)$$

$$(2C - 2C + 3C) \tan(x) = \tan(x)$$

$$C = \frac{1}{3}$$

$$y_{p2} = \frac{1}{3} \tan(x)$$

Construct the general solution:

$$y = y_h + y_{p1} + y_{p2} = e^x\left(C \cos\left(\sqrt{2}x\right) + D \sin\left(\sqrt{2}x\right)\right) + \frac{1}{3} \ln(x) + \frac{1}{3} \tan(x)$$