Calculus

沈威宇

December 20, 2024

Contents

Chapter 1 Calculus (微積分)

1 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 $f : U \to W$ is called Fréchet 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 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

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

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

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

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

2 Differential theorems

I Taylor expansion (泰勒展開式)

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 expansion of F at a is

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

that is,

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

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

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

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

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

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

III 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

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

IV 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

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

V 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:

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

VI 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.

VII 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.

VIII 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.

IX 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.

X 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.

XI 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.

XII 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.

XIII Inflection point theorem

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

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

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

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 $P(x,\, n,\, t)$ 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 $\{t_i\}_{i=0}^{n-1}$ with $t_i \in [x_i,\, x_{i+1}]$ 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 partitions $P(x,\, n,\, t)$ and $Q(y,\, m,\, s)$ are both partitions of the interval $[a,\, b]$. We say that $Q(y,\, m,\, s)$ is a refinement of $P(x,\, n,\, t)$ 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. $t_i = s_j$. 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 $P(x,\, n,\, t)$ is defined to be

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

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 $f(t_i)$ 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:

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

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 $P(x, n, t)$ 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

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

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 positive real numbers, s is called a non-negative measurable simple function. The integral of a non-negative measurable simple function $\sum\limits_k a_k 1_{S_k}$ is defined to be

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

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

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

3. Non-negative measurable functions

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

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

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

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

In this case we define

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

If

$$\int |f| \,\mathrm{d}\mu < \infty,$$

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:

$$\int_a^b F'(x) \,\mathrm{d}x = F(b) - F(a).$$

Proof.
By the definition of the Riemann integral:

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

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

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_a^x f(t)\,\mathrm{d}t = f(x).$$

Proof.

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

Using the additivity property of integrals:

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

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:

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

Substituting into the difference quotient:

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

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

□

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

I Space convention

- The domain of the 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.
- $\mathbf{0}$ 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: $\displaystyle\int$
- Surface integral operator: $\displaystyle\iint$
- Volume integration operator: $\displaystyle\iiint$
- Closed line integral operator: $\displaystyle\oint$
- Closed surface Integral Operator: $\displaystyle\oiint$
- $\displaystyle\int \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ is used as a shorthand for $\displaystyle\int (\mathbf{F} \cdot \hat{\mathbf{n}})\, \mathrm{d}S$, 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: $\mathbf{x} = (x_1, x_2, \dots, x_n)$
- 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

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

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

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

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.

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

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 $x_1, x_2, \dots, x_n$, 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 $\mathbf{x}^\alpha$: 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 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,

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

Also, the kth-order approximation of $\mathbf{F}$ near $\mathbf{a}$ is

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

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,
$$\oint_C f(\mathbf{x}) \cdot \mathrm{d}\mathbf{x} = 0.$$
- 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)

If $\varphi : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is a differentiable function and γ a differentiable curve in U which starts at a point p and ends at a point q, then,

$$\int_\gamma \nabla\varphi(\mathbf{r}) \cdot \mathrm{d}\mathbf{r} = \varphi\left(\mathbf{q}\right) - \varphi\left(\mathbf{p}\right).$$

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,

$$\iiint_V (\nabla \cdot \mathbf{F})\ \mathrm{d}V = \oiint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$$

Divergence theorem is a special case of generalized Stokes theorem.

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

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,

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

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,

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

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 over the whole of Ω, i.e.,

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

9 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 For single variables

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:

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

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

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

ii For multiple variables

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:

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

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

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

iii In measure theory

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

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

II 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

10 Common calculus results

I Polynomial function

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

Proof.

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

□

II Exponential function

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

Proof.

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

□

III Logarithmic function

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

Proof.

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

□

IV Irrational algebraic function

$$\int\sqrt{a^2-x^2}\ \mathrm{d}x = \frac{1}{2}\Bigl(x\sqrt{a^2-x^2}+\sin^{-1}\frac{x}{a}\Bigr)+C,\quad a>x$$

Proof.
Let:

$$x = a\sin\theta,\quad \text{so}\quad \mathrm{d}x = a\cos\theta\ \mathrm{d}\theta\,.$$

Also, note:

$$\sqrt{a^2-x^2}=\sqrt{a^2-a^2\sin^2\theta}=\sqrt{a^2(1-\sin^2\theta)}=a\cos\theta.$$

Substitute back:

$$\int\sqrt{a^2-x^2}\ \mathrm{d}x=\int a\cos\theta\cdot a\cos\theta\ \mathrm{d}\theta=a^2\int\cos^2\theta\ \mathrm{d}\theta\,.$$

Use the trigonometric identity:

$$\cos^2\theta=\frac{1+\cos(2\theta)}{2}.$$

Thus:

$$a^2\int\cos^2\theta\ \mathrm{d}\theta=a^2\int\frac{1+\cos(2\theta)}{2}\ \mathrm{d}\theta\,.$$

Split the integral:

$$a^2\int\cos^2\theta\ \mathrm{d}\theta=\frac{a^2}{2}\Bigl(\int 1\ \mathrm{d}\theta+\int\cos(2\theta)\ \mathrm{d}\theta\Bigr)$$
$$=\frac{a^2}{2}\theta+\frac{a^2}{4}\sin(2\theta)+C.$$

Substitute back using:

$$\sin\theta=\frac{x}{a},\quad \theta=\sin^{-1}\frac{x}{a},\quad\text{and}\quad \sin(2\theta)=2\sin\theta\cos\theta=2\frac{x}{a}\cdot\frac{\sqrt{a^2-x^2}}{a}.$$

Thus:

$$\frac{a^2}{4}\sin(2\theta)=\frac{a^2}{4}\cdot 2\frac{x\sqrt{a^2-x^2}}{a^2}=\frac{x\sqrt{a^2-x^2}}{2}.$$

and:

$$\frac{a^2}{2}\theta=\frac{a^2}{2}\sin^{-1}\frac{x}{a}$$

So:

$$\int\sqrt{a^2-x^2}\ \mathrm{d}x=\frac{1}{2}\Bigl(x\sqrt{a^2-x^2}+a^2\sin^{-1}\frac{x}{a}\Bigr)+C.$$

□

11 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.

12 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

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

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 $\dfrac{n}{2} \in \mathbb{N}$.
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

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