數值分析 MA-3021

Chapter 1. Mathematical Preliminaries

§1.1 Review of Calculus

~~§1.2 Round-off Errors and Computer Arithmetic~~

~~§1.3 Algorithms and Convergence~~

§1.1 Review of Calculus - Limit of Functions and Sequences

Definition

Let I be a non-empty set in $\mathbb{R}$ (not necessary an interval), c be an accumulation point of I, and $f : I \to \mathbb{R}$ be a real-valued function. Then $\lim\limits_{x\to c} f(x) = L$ means for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big|f(x) - L\big| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta \text{ and } x \in X.$$

Definition

Let $\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of real (or complex) numbers and $x \in \mathbb{R}$ (or $\mathbb{C}$). Then $\lim\limits_{n\to\infty} x_n = x$ means for all $\varepsilon > 0$ there exists $N > 0$ such that

$$|x_n - x| < \varepsilon \quad \text{whenever} \quad n \geqslant N.$$

§1.1 Review of Calculus - Limit of Functions and Sequences

Definition

Let $I \subseteq \mathbb{R}$ be a non-empty set (not necessary an interval), c be an accumulation point of I, and $f : I \to \mathbb{R}$ be a real-valued function. Then $\lim_{x \to c} f(x) = L$ means for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left|f(x) - L\right| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta \text{ and } x \in X.$$

Definition

Let $\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of real (or complex) numbers and $x \in \mathbb{R}$ (or $\mathbb{C}$). Then $\lim_{n \to \infty} x_n = x$ means for all $\varepsilon > 0$ there exists $N > 0$ such that

$$|x_n - x| < \varepsilon \quad \text{whenever} \quad n \geqslant N.$$

§1.1 Review of Calculus - Limit of Functions and Sequences

Definition

Let $I \subseteq \mathbb{R}$ be ~~a~~ non-empty ~~set~~ (not necessary an interval), c be an accumulation point of I, and $f: I \to \mathbb{R}$ ~~be a real-valued function~~. Then $\lim_{x \to c} f(x) = L$ means for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| f(x) - L \right| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta \text{ and } x \in X.$$

Definition

Let $\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of real (or complex) numbers and $x \in \mathbb{R}$ (or $\mathbb{C}$). Then $\lim_{n \to \infty} x_n = x$ means for all $\varepsilon > 0$ there exists $N > 0$ such that

$$|x_n - x| < \varepsilon \quad \text{whenever} \quad n \geqslant N.$$

§1.1 Review of Calculus - Limit of Functions and Sequences

Definition

Let $\varnothing \neq I \subseteq \mathbb{R}$ (not necessary an interval), c be an accumulation point of I, and $f : I \to \mathbb{R}$. Then $\lim_{x \to c} f(x) = L$ means for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left|f(x) - L\right| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta \text{ and } x \in X.$$

Definition

Let $\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of real (or complex) numbers and $x \in \mathbb{R}$ (or $\mathbb{C}$). Then $\lim_{n \to \infty} x_n = x$ means for all $\varepsilon > 0$ there exists $N > 0$ such that

$$|x_n - x| < \varepsilon \quad \text{whenever} \quad n \geqslant N.$$

§1.1 Review of Calculus - Limit of Functions and Sequences

Theorem

Let $\emptyset \neq I \subseteq \mathbb{R}$, *c is an accumulation point of I, and f be a real-valued function defined on* $I - \{c\}$. *Then* $\lim_{x \to c} f(x) = L$ *if and only if*

every sequence $\{c_n\}_{n=1}^{\infty} \subseteq I - \{c\}$ *satisfying* $\lim_{n \to \infty} c_n = c$
also has the property that $\lim_{n \to \infty} f(c_n) = L$.

（一函數在 c 的極限為 L 如果「所有在 I 中取值不是 c 但收斂到 c 的數列其函數值所形成的數列都收斂到 L」）

Using the logic notation, $\lim_{x \to c} f(x) = L$ *if and only if*

$$\left(\forall \{c_n\}_{n=1}^{\infty} \subseteq I - \{c\}\right)\left(\lim_{n \to \infty} c_n = c \Rightarrow \lim_{n \to \infty} f(c_n) = L\right).$$

§1.1 Review of Calculus - Continuity of Functions

Definition

Let $\varnothing \neq I \subseteq \mathbb{R}$, $c \in I$, and $f : I \to \mathbb{R}$. Then f is said to be continuous at c if $\lim_{x \to c} f(x) = f(c)$. Using the ε-δ language, f is continuous at c if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big|f(x) - f(c)\big| < \varepsilon \quad \text{whenever} \quad |x - c| < \delta \text{ and } x \in I.$$

Theorem

Let $\varnothing \neq I \subseteq \mathbb{R}$, $c \in I$, and $f : I \to \mathbb{R}$. Then f is continuous at c if and only if

$$\lim_{n \to \infty} f(c_n) = f(c) \text{ as long as } \{c_n\}_{n=1}^{\infty} \subseteq I \text{ and } \lim_{n \to \infty} c_n = c.$$

（一函數 f 在 c 連續如果「所有在 I 中收斂到 c 的數列其函數值所形成的數列都收斂到 $f(c)$」）

§1.1 Review of Calculus - Continuity of Functions

Definition

Let $\varnothing \neq I \subseteq \mathbb{R}$. The collection of all continuous functions defined on I is denoted by $C(I)$.

Remark:

1. For simplicity, we also use $C[a, b]$ to denote $C([a, b])$, and use $C(a, b]$ to denote $C((a, b])$, and etc.
2. To be more precise, we use $C(I; J)$ to denote all continuous functions defined on I with codomain J. For example, we use $C(I; \mathbb{R})$ to denote all continuous real-valued function defined on I, and use $C(I; \mathbb{R}^3)$ to denote all continuous three vector-valued functions defined on I, and etc.

§1.1 Review of Calculus - Smoothness

Definition

Let I be a non-empty (open) interval in $\mathbb{R}$, $c \in I$, and $f : I \longrightarrow \mathbb{R}$.

1. If $\lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ exists, then we say f is differentiable at c and $f'(c) \equiv \lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ is the derivative of f at c.
2. If f is differentiable at each point in I, then we say f is differentiable on I.

Alternative definition: $f'(c) = \lim\limits_{h \to 0} \dfrac{f(c+h) - f(c)}{h}$.

Theorem

If f is differentiable at c, then f is continuous at c.

§1.1 Review of Calculus - Smoothness

Definition

Let $\emptyset \neq I \subseteq \mathbb{R}$ be an (open) interval , $c \in I$, and $f : I \to \mathbb{R}$.

1. If $\lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ exists, then we say f is differentiable at c and $f'(c) \equiv \lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ is the derivative of f at c.
2. If f is differentiable at each point in I, then we say f is differentiable on I.

Alternative definition: $f'(c) = \lim\limits_{h \to 0} \dfrac{f(c+h) - f(c)}{h}$.

Theorem

If f is differentiable at c, then f is continuous at c.

§1.1 Review of Calculus - Smoothness

Definition

Let $\varnothing \neq I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$. Function f is said to be continuously differentiable on I if

1. f is differentiable on I.
2. f' is continuous on I.

Function f is said to be k-times continuously differentiable on I if

1. $f, f', f'', \cdots, f^{(k)}$ exists on I.
2. $f^{(k)}$ is continuous on I.

The collection of all k-times continuously differentiable functions defined on I is denoted by $C^k(I)$, and the collection of all continuous functions defined on I that have derivatives of all order is denoted by $C^\infty(I)$.

§1.1 Review of Calculus - Mean Value Theorem

Theorem (Extreme Value Theorem)

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then there exists $c_1, c_2 \in [a, b]$ such that $f(c_1) \leqslant f(x) \leqslant f(c_2)$ for all $x \in [a, b]$.

Theorem (Fermat)

Let $f : (a, b) \to \mathbb{R}$ be differentiable. If $a < c < b$ and $f(c)$ is a local extreme value of f, then $f'(c) = 0$.

Extreme Value Theorem + Fermat's Theorem $\Rightarrow$

Theorem (Rolle)

Let $f : [a, b] \to \mathbb{R}$ be continuous. If f is differentiable on (a, b) and $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.

§1.1 Review of Calculus - Mean Value Theorem

Theorem (Mean Value Theorem)

Let $f : [a, b] \to \mathbb{R}$ *be continuous. If* f *is differentiable on* (a, b)*, then there exists* $c \in (a, b)$ *such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a} .$$

Theorem (Generalized Rolle's Theorem)

Let $f : [a, b] \to \mathbb{R}$ *be continuous. If* $f \in C^k((a, b))$ *and* f *has* $(k + 1)$ *distinct zeros in* $[a, b]$*, then there exists* $c \in (a, b)$ *such that*

$$f^{(k)}(c) = 0 .$$

§1.1 Review of Calculus - Intermediate Value Theorem

Theorem (Bolzano)

Let $f: [a, b] \to \mathbb{R}$ be continuous. If $f(a)f(b) < 0$, then there exists $c \in (a, b)$ such that $f(c) = 0$.

Theorem (Intermediate Value Theorem)

Let $f: [a, b] \to \mathbb{R}$ be continuous, and K is any number between $f(a)$ and $f(b)$; that is, $f(a) < K < f(b)$ or $f(b) < K < f(a)$, then there exists $c \in (a, b)$ such that $f(c) = K$.

Note: The Least-Upper-Bound Axiom + sign-preserving property $\Rightarrow$ Bolzano's Theorem $\Rightarrow$ Intermediate Value Theorem.

§1.1 Review of Calculus - Riemann integrals

Definition

A finite set $\mathcal{P} = \{x_0, x_1, \cdots, x_n\}$ is said to be a partition of the closed interval $[a, b]$ if $a = x_0 < x_1 < \cdots < x_n = b$. Such a partition $\mathcal{P}$ is usually denoted by $\{a = x_0 < x_1 < \cdots < x_n = b\}$. The norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$, is the number $\max\{x_i - x_{i-1} \mid 1 \leqslant i \leqslant n\}$; that is,

$$\|\mathcal{P}\| \equiv \max\left\{x_i - x_{i-1} \,\middle|\, 1 \leqslant i \leqslant n\right\}.$$

Let $f : [a, b] \to \mathbb{R}$ be a function. A Riemann sum of f for the partition $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ of $[a, b]$ is a sum which takes the form

$$\sum_{k=1}^{n} f(c_k)(x_k - x_{k-1}),$$

where $x_{k-1} \leqslant c_k \leqslant x_k$ for each $1 \leqslant k \leqslant n$.

§1.1 Review of Calculus - Riemann integrals

Conceptually, a function $f\colon [a,b] \to \mathbb{R}$ is integrable if

$$\lim_{\|\mathcal{P}\| \to 0} \sum_{k=1}^{n} f(c_k)(x_k - x_{k-1}) \quad \text{exists.}$$

The precise meaning of the limit above is the following

Definition

Let $f\colon [a,b] \to \mathbb{R}$ be a function. f is said to be Riemann integrable on $[a,b]$ if there exists a real number A such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is partition of $[a,b]$ satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums for the partition $\mathcal{P}$ belongs to the interval $(A-\varepsilon, A+\varepsilon)$. Such a number A (is unique and) is called the Riemann integral of f on $[a,b]$ and is denoted by $\displaystyle\int_{[a,b]} f(x)\,dx$.

§1.1 Review of Calculus - Riemann integrals

Definition

A set $A \subseteq \mathbb{R}$ is called a set of measure zero or is said to have measure zero if for every $\varepsilon > 0$ there exist intervals $I_1, I_2, \cdots, I_n, \cdots$ such that

1. $A \subseteq \bigcup_{k=1}^{\infty} I_k$, and
2. $\sum_{k=1}^{\infty} |I_k| < \varepsilon$, where $|I_k|$ denotes the length of I_k.

Theorem (Lebesgue)

*Let $f : [a, b] \to \mathbb{R}$ be a **bounded** function. Then f is Riemann integrable on $[a, b]$ if and only if the collection of discontinuities of f has measure zero.*

Therefore, if $f : [a, b] \to \mathbb{R}$ is continuous, then f is Riemann integrable on $[a, b]$.

§1.1 Review of Calculus - Weighted Mean Value Theorem for Integrals

Theorem

Let $f \in C([a,b])$, g is Riemann integrable on $[a,b]$ and does not change sign on $[a,b]$. There exists $c \in (a,b)$ such that

$$\int_a^b f(x)g(x)dx = f(c)\int_a^b g(x)dx\,.$$

Proof.

W.L.O.G., we assume that $g \geqslant 0$ on $[a,b]$ such that $\int_a^b g(x)\,dx > 0$. Since $f \in C([a,b])$, there exist $m = \min\limits_{x\in[a,b]} f(x)$ and $M = \max\limits_{x\in[a,b]} f(x)$. Then

$$\int_a^b mg(x)dx \leqslant \int_a^b f(x)g(x)dx \leqslant \int_a^b Mg(x)dx\,;$$

thus $m \leqslant \dfrac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \leqslant M$, and the assertion holds by the Intermediate Value Theorem. □

§1.1 Review of Calculus - Weighted Mean Value Theorem for Integrals

Remark:

1. When $g(x) \equiv 1$ on $[a, b]$, the weight MVT for integrals implies that $\int_a^b f(x)dx = f(c)(b-a)$. This is the original MVT for integrals.
2. The number $\frac{1}{b-a}\int_a^b f(x)dx$ is called the average value of f on $[a, b]$, and sometimes is denoted by $(f)_{[a,b]}$.

§1.1 Review of Calculus - Taylor's Theorem

Theorem (Taylor's Theorem for functions of one variable)

Let $f \in C^{m+1}([a,b])$ and $x_0 \in [a,b]$. Then for every $x \in [a,b]$, there exists $\xi(x)$ between x and x_0 such that

$$f(x) = P_m(x) + R_m(x)\,,$$

where the m-th Taylor polynomial $P_m(x)$ is given by

$$P_m(x) = \sum_{k=0}^{m} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$$

and the remainder (error) term $R_m(x)$ is given by

$$\begin{aligned} R_m(x) &= \frac{1}{m!}\int_{x_0}^{x} (x-t)^m f^{(m+1)}(t)\,dt && \text{(Integral form)} \\ &= \frac{1}{(m+1)!} f^{(m+1)}(\xi(x))(x-x_0)^{m+1} && \text{(Lagrange's form)} \end{aligned}$$

(the last "=" is by the weighted MVT for integrals)

§1.1 Review of Calculus - Taylor's Theorem

Remark: Assume that $f \in C^\infty([a, b])$.

- The series $\sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k$ is called the Taylor series of f at x_0. It is also called the Maclaurin series of f when $x_0 = 0$.
- If $R_m(x) \to 0$ as $m \to \infty$, then $P_m(x) \to f(x)$ as $m \to \infty$; i.e.,

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k.$$

Example

The Maclaurin series of the sine function is $\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$ and the Maclaurin series of the cosine function is $\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$. In fact,

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} \quad \text{and} \quad \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \quad \forall\, x \in \mathbb{R}.$$

§1.1 Review of Calculus - Taylor's Theorem

Example

Use the Taylor polynomial of $f(x) = \cos(x)$ at $x_0 = 0$ to estimate $\cos(0.01)$.

$f'(x) = -\sin(x),\ f''(x) = -\cos(x),\ f'''(x) = \sin(x),\ f^{(4)}(x) = \cos(x).$

$f(0) = 1,\ f'(0) = 0,\ f''(0) = -1,\ f'''(0) = 0,\ f^{(4)}(0) = 1.$

Case $m = 2$:

$\cos(x) = 1 - \frac{1}{2}x^2 + \frac{1}{6}x^3 \sin(\xi(x))$, where $\xi(x)$ is between 0 and x.

$\cos(0.01) = 0.99995 + 0.1\bar{6} \times 10^{-6} \sin(\xi)$, where $0 < \xi < 0.01$.

$$\begin{aligned}|\cos(0.01) - 0.99995| &\leqslant 0.1\bar{6} \times 10^{-6} |\sin(\xi)| \leqslant 0.1\bar{6} \times 10^{-6} \times 0.01 \\ &= 0.1\bar{6} \times 10^{-8},\end{aligned}$$

where we use the fact $|\sin(x)| \leqslant |x|$ for all $x \in \mathbb{R}$.

Case $m = 3$:

$\cos(x) = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 \cos(\tilde{\xi}(x))$, where $\tilde{\xi}(x)$ is between 0 and x.

$$|\cos(0.01) - 0.99995| \leqslant \frac{1}{24}(0.01)^4 \times 1 \leqslant 4.2 \times 10^{-10}.$$

§1.1 Review of Calculus - Taylor's Theorem

Example (continued)

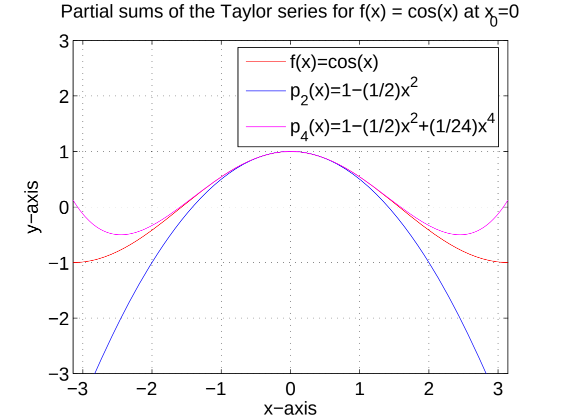
$$\begin{aligned}\int_0^{0.1} \cos(x)dx &= \int_0^{0.1} (1 - \frac{1}{2}x^2)dx + \int_0^{0.1} \frac{1}{24}x^4 \cos(\tilde{\xi}(x))dx \\ &= (x - \frac{1}{6}x^3)\Big|_0^{0.1} + \int_0^{0.1} \frac{1}{24}x^4 \cos(\tilde{\xi}(x))dx \\ &= 0.0998\bar{3} + \int_0^{0.1} \frac{1}{24}x^4 \cos(\tilde{\xi}(x))dx.\end{aligned}$$

$$\begin{aligned}\left|\int_0^{0.1} \cos(x)dx - 0.0998\bar{3}\right| &\leqslant \frac{1}{24}\int_0^{0.1} x^4|\cos(\tilde{\xi}(x))|dx \\ &\leqslant \frac{1}{24}\int_0^{0.1} x^4 dx = 8.\bar{3} \times 10^{-8}.\end{aligned}$$

True value is 0.099833416647, actual error for this approximation is 8.3314×10^{-8}.

§1.1 Review of Calculus - Taylor's Theorem

Partial sums of the Taylor series for $f(x) = \cos(x)$ at $x_0 = 0$



Note: A Taylor series converges rapidly near the point of expansion and slowly (or not at all) at more remote points.

§1.1 Review of Calculus - Taylor's Theorem

Taylor's Theorem for functions of multiple variables:

Let $U \subseteq \mathbb{R}^n$ be open, and $\boldsymbol{a} \in U$. Suppose that $\boldsymbol{x} \in U$ is such that $\overline{\boldsymbol{ax}} \subseteq U$; that is, $\boldsymbol{x}$ satisfies that the point

$$p(t) = (1-t)\boldsymbol{a} + t\boldsymbol{x} \in U \quad \text{whenever} \quad t \in [0,1].$$

For a function $f : U \to \mathbb{R}$, define a function h by

$$h(t) = f(p(t)) = f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})).$$

If $h \in C^{m+1}([0,1])$, then Taylor's theorem for functions of one variable implies that

$$h(1) = h(0) + h'(0) + \frac{1}{2!}h''(0) + \cdots + \frac{1}{m!}h^{(m)}(0) + R_m,$$

where the remainder R_m, in Lagrange's form, is given by

$$R_m = \frac{1}{(m+1)!}h^{(m+1)}(s)$$

for some $s \in [0,1]$.

§1.1 Review of Calculus - Taylor's Theorem

Questions:

1. When is $h \in C^{m+1}([0,1])$?
2. What is $h^{(k)}(t)$ for general $k \in \mathbb{N}$?

Definition (Multi-index)

An n-dimensional multi-index is a vector $\alpha = (\alpha_1, \cdots, \alpha_n)$ of non-negative integers, Given an n-dimensional multi-index $\alpha = (\alpha_1, \cdots, \alpha_n)$, $|\alpha|$ and $\alpha!$ are defined by

$$|\alpha| = \sum_{k=1}^{n} \alpha_k \quad \text{and} \quad \alpha! = \prod_{k=1}^{n} \alpha_k! \,.$$

The differential operator $D_{\boldsymbol{x}}^{\alpha}$ is defined by

$$D_{\boldsymbol{x}}^{\alpha} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \,.$$

When the variable is specified, we simply use D^{α} to denote $D_{\boldsymbol{x}}^{\alpha}$.

§1.1 Review of Calculus - Taylor's Theorem

Example

$\alpha = (1,5,3)$ is a three-dimensional multi-index satisfying

$$|\alpha| = 9 \quad \text{and} \quad \alpha! = 5! \cdot 3! = 720\,.$$

Example

Suppose that f is a function of three variables x_1, x_2, x_3. Then

$$D^{(1,5,3)} f(x_1, x_2, x_3) = \frac{\partial^9 f}{\partial x_1 \partial x_2^5 \partial x_3^3}(x_1, x_2, x_3)\,.$$

The chain rule for functions of multiple variables:

$$\frac{d}{dt} f\big(x_1(t), \cdots, x_n(t)\big) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\big(x_1(t), \cdots, x_n(t)\big) x_j'(t)\,.$$

Therefore, if $x_j(t) = a_j + t(x_j - a_j)$ for all $1 \leqslant j \leqslant n$,

$$\frac{d}{dt} f\big(x_1(t), \cdots, x_n(t)\big) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\big(x_1(t), \cdots, x_n(t)\big)(x_j - a_j)\,.$$

§1.1 Review of Calculus - Taylor's Theorem

Example

$\alpha = (1,5,3)$ is a three-dimensional multi-index satisfying

$$|\alpha| = 9 \quad \text{and} \quad \alpha! = 5! \cdot 3! = 720 \,.$$

Example

Suppose that f is a function of three variables x_1, x_2, x_3. Then

$$D^{(1,5,3)} f(x_1,x_2,x_3) = \frac{\partial^9 f}{\partial x_1 \partial x_2^5 \partial x_3^3}(x_1,x_2,x_3)\,.$$

The chain rule for functions of multiple variables:

$$\frac{d}{dt} f\big(x_1(t),\cdots,x_n(t)\big) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\big(x_1(t),\cdots,x_n(t)\big) x_j'(t)\,.$$

Therefore, if $x_j(t) = a_j + t(x_j - a_j)$ for all $1 \leqslant j \leqslant n$,

$$\frac{d}{dt} f\big(x_1(t),\cdots,x_n(t)\big) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\big(x_1(t),\cdots,x_n(t)\big)(x_j - a_j)\,.$$

§1.1 Review of Calculus - Taylor's Theorem

Questions:

1. When is $h \in C^{m+1}([0,1])$?
2. What is $h^{(k)}(t)$ for general $k \in \mathbb{N}$?

Answers:

1. The mixed partial derivatives $D^{\alpha} f$ is continuous in an open set containing $\overline{\boldsymbol{ax}}$ for all n-dimensional multi-index α satisfying $|\alpha| \leqslant m+1$.
2. By the chain rule for function of multiple variables,

$$h^{(k)}(t) = \sum_{|\alpha|=k} \frac{|\alpha|!}{\alpha!} (D^{\alpha} f)(p(t))(\boldsymbol{x} - \boldsymbol{a})^{\alpha},$$

where

$$(\boldsymbol{x}-\boldsymbol{a})^{\alpha} = \prod_{k=1}^{n} (x_k - a_k)^{\alpha_k} = (x_1 - a_1)^{\alpha_1} (x_2 - a_2)^{\alpha_2} \cdots (x_n - a_n)^{\alpha_n}.$$

§1.1 Review of Calculus - Taylor's Theorem

Therefore, if $D^\alpha f$ is continuous in an open set containing $\overline{\boldsymbol{ax}}$ for all n-dimensional multi-index α satisfying $|\alpha| \leqslant m+1$, we have

$$\begin{aligned} f(\boldsymbol{x}) &= h(1) = h(0) + h'(0) + \frac{1}{2!}h''(0) + \cdots + \frac{1}{m!}h^{(m)}(0) + R_m \\ &= \sum_{k=0}^{m} \frac{1}{k!} h^{(k)}(0) + R_m \\ &= \sum_{k=0}^{m} \frac{1}{k!} \sum_{|\alpha|=k} \frac{|\alpha|!}{\alpha!} (D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + R_m \,, \end{aligned}$$

where

$$R_m = \frac{1}{(m+1)!} \sum_{|\alpha|=m+1} \frac{|\alpha|!}{\alpha!} (D^\alpha f)(p(s))(\boldsymbol{x}-\boldsymbol{a})^\alpha$$

for some $s \in [0,1]$.

§1.1 Review of Calculus - Taylor's Theorem

Therefore, if $D^\alpha f$ is continuous in an open set containing $\overline{\boldsymbol{ax}}$ for all n-dimensional multi-index α satisfying $|\alpha| \leqslant m+1$, we have

$$\begin{aligned} f(\boldsymbol{x}) &= h(1) = h(0) + h'(0) + \frac{1}{2!}h''(0) + \cdots + \frac{1}{m!}h^{(m)}(0) + R_m \\ &= \sum_{k=0}^{m} \frac{1}{k!} h^{(k)}(0) + R_m \\ &= \sum_{k=0}^{m} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + R_m\,, \end{aligned}$$

where

$$R_m = \sum_{|\alpha|=m+1} \frac{1}{\alpha!}(D^\alpha f)(p(s))(\boldsymbol{x}-\boldsymbol{a})^\alpha$$

for some $s \in [0,1]$.

§1.1 Review of Calculus - Taylor's Theorem

Therefore, if $D^\alpha f$ is continuous in an open set containing $\overline{\boldsymbol{ax}}$ for all n-dimensional multi-index α satisfying $|\alpha| \leqslant m+1$, we have

$$
\begin{aligned}
f(\boldsymbol{x}) &= h(1) = h(0) + h'(0) + \frac{1}{2!}h''(0) + \cdots + \frac{1}{m!}h^{(m)}(0) + R_m \\
&= \sum_{k=0}^{m} \frac{1}{k!} h^{(k)}(0) + R_m \\
&= \sum_{k=0}^{m} \sum_{|\alpha|=k} \frac{1}{\alpha!} (D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + R_m ,
\end{aligned}
$$

where

$$
R_m = \sum_{|\alpha|=m+1} \frac{1}{\alpha!} (D^\alpha f)(p(s))(\boldsymbol{x}-\boldsymbol{a})^\alpha
$$

for some $s \in [0,1]$.

§1.1 Review of Calculus - Taylor's Theorem

Therefore, if $D^\alpha f$ is continuous in an open set containing $\overline{\boldsymbol{ax}}$ for all n-dimensional multi-index α satisfying $|\alpha| \leqslant m+1$, we have

$$
\begin{aligned}
f(\boldsymbol{x}) = h(1) &= h(0) + h'(0) + \frac{1}{2!}h''(0) + \cdots + \frac{1}{m!}h^{(m)}(0) + R_m \\
&= \sum_{k=0}^{m} \frac{1}{k!} h^{(k)}(0) + R_n \\
&= \sum_{k=0}^{m} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + R_m \,,
\end{aligned}
$$

where

$$
R_m = \sum_{|\alpha|=m+1} \frac{1}{\alpha!}(D^\alpha f)(\ \xi\ )(\boldsymbol{x}-\boldsymbol{a})^\alpha
$$

for some $\boldsymbol{\xi} \in \overline{\boldsymbol{ax}}$.

§1.1 Review of Calculus - Taylor's Theorem

Theorem (Taylor's Theorem for functions of multiple variables)

Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$, and $\boldsymbol{a}, \boldsymbol{x} \in U$ be such that $\overline{\boldsymbol{ax}} \subseteq U$. If $D^\alpha f$ is continuous in an open set containing $\overline{\boldsymbol{ax}}$ for all n-dimensional multi-index α satisfying $|\alpha| \leqslant m+1$, then there exists $\boldsymbol{\xi} \in \overline{\boldsymbol{ax}}$ such that

$$f(\boldsymbol{x}) = P_m(\boldsymbol{x}) + R_m(\boldsymbol{x}),$$

where the m-th Taylor polynomial $P_m(x)$ is given by

$$P_m(\boldsymbol{x}) = \sum_{k=0}^{m} \sum_{|\alpha|=k} \frac{1}{\alpha!} (D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$$

and the remainder (error) term $R_m(\boldsymbol{x})$ is given by

$$R_m(\boldsymbol{x}) = \sum_{|\alpha|=m+1} \frac{1}{\alpha!} (D^\alpha f)(\boldsymbol{\xi})(\boldsymbol{x}-\boldsymbol{a})^\alpha.$$

§1.1 Review of Calculus - Taylor's Theorem

Example

The second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is $P_2(x) = \sum_{k=0}^{2} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$; thus

$$P_2(\boldsymbol{x}) = \sum_{|\alpha|=0} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=1} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=2} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$$

$$= f(\boldsymbol{a}) + \frac{1}{(1,0)!}(D^{(1,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(1,0)} + \frac{1}{(0,1)!}(D^{(0,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,1)} + \frac{1}{(2,0)!}(D^{(2,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(2,0)} + \frac{1}{(1,1)!}(D^{(1,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(1,1)} + \frac{1}{(0,2)!}(D^{(0,2)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,2)}.$$

§1.1 Review of Calculus - Taylor's Theorem

Example

The second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is $P_2(x) = \sum_{k=0}^{2} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$; thus

$$
\begin{aligned}
P_2(\boldsymbol{x}) &= \sum_{|\alpha|=0} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=1} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&\quad + \sum_{|\alpha|=2} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&= f(\boldsymbol{a}) + \frac{1}{(1,0)!}(D^{(1,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(1,0)} + \frac{1}{(0,1)!}(D^{(0,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,1)} \\
&\quad + \frac{1}{(2,0)!}(D^{(2,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(2,0)} + \frac{1}{(1,1)!}(D^{(1,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(1,1)} \\
&\quad + \frac{1}{(0,2)!}(D^{(0,2)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,2)} .
\end{aligned}
$$

§1.1 Review of Calculus - Taylor's Theorem

Example

The second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is $P_2(x) = \sum_{k=0}^{2} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$; thus

$$
\begin{aligned}
P_2(\boldsymbol{x}) &= \sum_{|\alpha|=0} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=1} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&\quad + \sum_{|\alpha|=2} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&= f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{1}{(0,1)!}(D^{(0,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,1)} \\
&\quad + \frac{1}{(2,0)!}(D^{(2,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(2,0)} + \frac{1}{(1,1)!}(D^{(1,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(1,1)} \\
&\quad + \frac{1}{(0,2)!}(D^{(0,2)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,2)} .
\end{aligned}
$$

§1.1 Review of Calculus - Taylor's Theorem

Example

The second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is $P_2(x) = \sum_{k=0}^{2} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$; thus

$$
\begin{aligned}
P_2(\boldsymbol{x}) &= \sum_{|\alpha|=0} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=1} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&\quad + \sum_{|\alpha|=2} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&= f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{1}{(0,1)!}(D^{(0,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,1)} \\
&\quad + \frac{1}{(2,0)!}(D^{(2,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(2,0)} + \frac{1}{(1,1)!}(D^{(1,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(1,1)} \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a})(x_2 - a_2)^2 \quad .
\end{aligned}
$$

§1.1 Review of Calculus - Taylor's Theorem

Example

The second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is $P_2(x) = \sum_{k=0}^{2} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$; thus

$$
\begin{aligned}
P_2(\boldsymbol{x}) &= \sum_{|\alpha|=0} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=1} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&\quad + \sum_{|\alpha|=2} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha \\
&= f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{1}{(0,1)!}(D^{(0,1)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(0,1)} \\
&\quad + \frac{1}{(2,0)!}(D^{(2,0)} f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^{(2,0)} + \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2) \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a})(x_2 - a_2)^2 \quad .
\end{aligned}
$$

§1.1 Review of Calculus - Taylor's Theorem

Example

The second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is $P_2(x) = \sum_{k=0}^{2} \sum_{|\alpha|=k} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$; thus

$$P_2(\boldsymbol{x}) = \sum_{|\alpha|=0} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=1} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha + \sum_{|\alpha|=2} \frac{1}{\alpha!}(D^\alpha f)(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})^\alpha$$

$$= f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(\boldsymbol{a})(x_2 - a_2) + \frac{1}{2}\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a})(x_1 - a_1)^2 + \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2) + \frac{1}{2}\frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a})(x_2 - a_2)^2 .$$

§1.1 Review of Calculus - Taylor's Theorem

Example (Cont.)

Therefore, the second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is

$$P_2(\boldsymbol{x}) = f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(\boldsymbol{a})(x_2 - a_2) + \frac{1}{2}\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a})(x_1 - a_1)^2 + \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2) + \frac{1}{2}\frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a})(x_2 - a_2)^2 .$$

Similarly, the third Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is

$$P_3(\boldsymbol{x}) = P_2(\boldsymbol{x}) + \frac{1}{3!}\left[\frac{\partial^3 f}{\partial x_1^3}(\boldsymbol{a})(x_1 - a_1)^3 + 3\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\boldsymbol{a})(x_1 - a_1)^2(x_2 - a_2) + 3\frac{\partial^3 f}{\partial x_1 \partial x_2^2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2)^2 + \frac{\partial^3 f}{\partial x_2^3}(\boldsymbol{a})(x_2 - a_2)^3\right].$$

§1.1 Review of Calculus - Taylor's Theorem

Example (Cont.)

Therefore, the second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is

$$P_2(\boldsymbol{x}) = f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(\boldsymbol{a})(x_2 - a_2) + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a})(x_1 - a_1)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2) + \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a})(x_2 - a_2)^2\right].$$

Similarly, the third Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is

$$P_3(\boldsymbol{x}) = P_2(\boldsymbol{x}) + \frac{1}{3!}\left[\frac{\partial^3 f}{\partial x_1^3}(\boldsymbol{a})(x_1 - a_1)^3 + 3\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\boldsymbol{a})(x_1 - a_1)^2(x_2 - a_2) + 3\frac{\partial^3 f}{\partial x_1 \partial x_2^2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2)^2 + \frac{\partial^3 f}{\partial x_2^3}(\boldsymbol{a})(x_2 - a_2)^3\right].$$

§1.1 Review of Calculus - Taylor's Theorem

Example (Cont.)

Therefore, the second Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is

$$P_2(\boldsymbol{x}) = f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(\boldsymbol{a})(x_2 - a_2) + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{a})(x_1 - a_1)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2) + \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{a})(x_2 - a_2)^2\right].$$

Similarly, the third Taylor polynomial of $f = f(\boldsymbol{x}) = f(x_1, x_2)$ about $\boldsymbol{a} = (a_1, a_2)$ is

$$P_3(\boldsymbol{x}) = P_2(\boldsymbol{x}) + \frac{1}{3!}\left[\frac{\partial^3 f}{\partial x_1^3}(\boldsymbol{a})(x_1 - a_1)^3 + 3\frac{\partial^3 f}{\partial x_1^2 \partial x_2}(\boldsymbol{a})(x_1 - a_1)^2(x_2 - a_2) + 3\frac{\partial^3 f}{\partial x_1 \partial x_2^2}(\boldsymbol{a})(x_1 - a_1)(x_2 - a_2)^2 + \frac{\partial^3 f}{\partial x_2^3}(\boldsymbol{a})(x_2 - a_2)^3\right].$$

§1.1 Review of Calculus - Big O notation for sequences

Big $\mathcal{O}$ notation is used to describe the limiting behavior of a function when the argument tends towards a particular value or infinity.

Definition

Suppose that $\lim_{x \longrightarrow a} G(x) = 0$ and $\lim_{x \longrightarrow a} F(x) = L$. If there exists $K > 0$ and $\delta > 0$ such that $|F(x) - L| \leqslant K|G(x)|$ for all $0 < |x - a| < \delta$, then we say that $F(x)$ converges to L with rate of convergence $\mathcal{O}(G(x))$ and write $F(x) = L + \mathcal{O}(G(x))$ as $x \to a$.

Definition

Suppose that $\lim_{n \longrightarrow \infty} \beta_n = 0$ and $\lim_{n \longrightarrow \infty} \alpha_n = \alpha$. If there exists $K > 0$ and $n_0 \in \mathbb{N}$ such that $|\alpha_n - \alpha| \leqslant K|\beta_n|$ for all $n \geqslant n_0$, then we say that $\{\alpha_n\}$ converges to α with rate of convergence $\mathcal{O}(\beta_n)$ and write $\alpha_n = \alpha + \mathcal{O}(\beta_n)$.

§1.1 Review of Calculus - Big O notation for sequences

Big $\mathcal{O}$ notation is used to describe the limiting behavior of a function when the argument tends towards a particular value or infinity.

Definition

Suppose that $\lim_{x\to a} G(x) = 0$ and $\lim_{x\to a} F(x) = L$. If there exists $K > 0$ and $\delta > 0$ such that $|F(x) - L| \leqslant K|G(x)|$ for all $0 < |x-a| < \delta$, then we say that $F(x)$ converges to L with rate of convergence $\mathcal{O}(G(x))$ and write $F(x) = L + \mathcal{O}(G(x))$ as $x \to a$.

Definition

Suppose that $\lim_{n\to\infty} \beta_n = 0$ and $\lim_{n\to\infty} \alpha_n = \alpha$. If there exists $K > 0$ and $n_0 \in \mathbb{N}$ such that $|\alpha_n - \alpha| \leqslant K|\beta_n|$ for all $n \geqslant n_0$, then we say that $\{\alpha_n\}$ converges to α with rate of convergence $\mathcal{O}(\beta_n)$ and write $\alpha_n = \alpha + \mathcal{O}(\beta_n)$.

§1.1 Review of Calculus - Big O notation for functions

Definition (Alternative)

One writes

$$f(x) = \mathcal{O}(g(x)) \quad \text{as } x \to a$$

provided that

$$\limsup_{x \to a} \frac{|f(x)|}{|g(x)|} < \infty \,.$$

Example

By Taylor's theorem,

$$\cos(h) = 1 - \frac{1}{2}h^2 + \frac{1}{24}h^4 \cos(\xi(h))$$

for some $\xi(h)$ between 0 and h. Then

$$\left|\cos(h) + \frac{1}{2}h^2 - 1\right| = \left|\frac{1}{24}\cos(\xi(h))\right| h^4 \leqslant \frac{1}{24}h^4 \quad \forall\, h\,;$$

thus $\cos(h) + \frac{1}{2}h^2 = 1 + \mathcal{O}(h^4)$.

§1.1 Review of Calculus - Big O notation for functions

Definition (Alternative)

One writes

$$f(x) = \mathcal{O}(g(x)) \quad \text{as } x \to a$$

provided that

$$\limsup_{x \to a} \frac{|f(x)|}{|g(x)|} < \infty .$$

Example

By Taylor's theorem,

$$\cos(h) = 1 - \frac{1}{2}h^2 + \frac{1}{24}h^4 \cos(\xi(h))$$

for some $\xi(h)$ between 0 and h. Then

$$\left|\cos(h) + \frac{1}{2}h^2 - 1\right| = \left|\frac{1}{24}\cos(\xi(h))\right| h^4 \leqslant \frac{1}{24}h^4 \quad \forall\, h;$$

thus $\cos(h) + \frac{1}{2}h^2 = 1 + \mathcal{O}(h^4)$.

§1.1 Review of Calculus - Big O notation for sequences

Example

Let $\alpha_n = 1 + \frac{n+1}{n^2}$. Then $\lim_{n\to\infty} \alpha_n = \alpha = 1$.
If $\beta_n = \frac{1}{n}$, then $\lim_{n\to\infty} \beta_n = 0$ and

$$|\alpha_n - 1| = \frac{n+1}{n^2} \leqslant \frac{n+n}{n^2} = 2\frac{1}{n} = 2|\beta_n - 0|.$$

Therefore, $\alpha_n = 1 + \mathcal{O}\left(\frac{1}{n}\right)$.

Example

Let $\alpha_n = 2 + \frac{n+3}{n^3}$. Then $\lim_{n\to\infty} \alpha_n = \alpha = 2$.
If $\beta_n = \frac{1}{n^2}$, then $\lim_{n\to\infty} \beta_n = 0$ and

$$|\alpha_n - 2| = \frac{n+3}{n^3} \leqslant \frac{n+3n}{n^3} = 4\frac{1}{n^2} = 4|\beta_n - 0|.$$

Therefore, $\alpha_n = 2 + \mathcal{O}\left(\frac{1}{n^2}\right)$.

§1.1 Review of Calculus - Big O notation for sequences

Example

Let $\alpha_n = 1 + \frac{n+1}{n^2}$. Then $\lim_{n\to\infty} \alpha_n = \alpha = 1$.
If $\beta_n = \frac{1}{n}$, then $\lim_{n\to\infty} \beta_n = 0$ and

$$|\alpha_n - 1| = \frac{n+1}{n^2} \leqslant \frac{n+n}{n^2} = 2\frac{1}{n} = 2|\beta_n - 0|.$$

Therefore, $\alpha_n = 1 + \mathcal{O}\left(\frac{1}{n}\right)$.

Example

Let $\alpha_n = 2 + \frac{n+3}{n^3}$. Then $\lim_{n\to\infty} \alpha_n = \alpha = 2$.
If $\beta_n = \frac{1}{n^2}$, then $\lim_{n\to\infty} \beta_n = 0$ and

$$|\alpha_n - 2| = \frac{n+3}{n^3} \leqslant \frac{n+3n}{n^3} = 4\frac{1}{n^2} = 4|\beta_n - 0|.$$

Therefore, $\alpha_n = 2 + \mathcal{O}\left(\frac{1}{n^2}\right)$.