Mathematics for Machine Learning

Vazgen Mikayelyan

August 22, 2020

FAST DISCOVERING THE FUTURE

Definite Integral

Theorem

Let $a < c < b$. *If* $f \in \mathcal{R}[a, c]$ *and* $f \in \mathcal{R}[c, b]$, *then* $f \in \mathcal{R}[a, b]$.

Definite Integral

Theorem

Let $a < c < b$. *If* $f \in \mathcal{R}[a, c]$ *and* $f \in \mathcal{R}[c, b]$, *then* $f \in \mathcal{R}[a, b]$.

Theorem

Let $a < c < b$. *If* $f \in \mathcal{R}[a, b]$ *then* $f \in \mathcal{R}[a, c]$ *and* $f \in \mathcal{R}[c, b]$, *moreover*

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

Definite Integral

Theorem

Let $a < c < b$. *If* $f \in \mathcal{R}[a, c]$ *and* $f \in \mathcal{R}[c, b]$, *then* $f \in \mathcal{R}[a, b]$.

Theorem

Let $a < c < b$. *If* $f \in \mathcal{R}[a, b]$ *then* $f \in \mathcal{R}[a, c]$ *and* $f \in \mathcal{R}[c, b]$, *moreover*

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

Theorem

If f *is monotone on* $[a, b]$ *then* $f \in \mathcal{R}[a, b]$.

Definite Integral

Theorem

If $f, g \in \mathcal{R}[a, b]$, *then*

1. $f \pm g \in \mathcal{R}[a, b]$ *and*

$$\int\limits_a^b (f(x) \pm g(x))\, dx = \int\limits_a^b f(x)\, dx \pm \int\limits_a^b g(x)\, dx,$$

Definite Integral

Theorem

If $f, g \in \mathcal{R}[a,b]$*, then*

1. $f \pm g \in \mathcal{R}[a,b]$ *and*

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

2. $cf \in \mathcal{R}[a,b]$, $c \in \mathbb{R}$ *and*

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

Definite Integral

Theorem

If $f, g \in \mathcal{R}[a,b]$, *then*

1. $f \pm g \in \mathcal{R}[a,b]$ *and*

$$\int\limits_a^b (f(x) \pm g(x))\, dx = \int\limits_a^b f(x)\, dx \pm \int\limits_a^b g(x)\, dx,$$

2. $cf \in \mathcal{R}[a,b]$, $c \in \mathbb{R}$ *and*

$$\int\limits_a^b cf(x)\, dx = c \int\limits_a^b f(x)\, dx.$$

3. $|f|, fg \in \mathcal{R}[a,b]$.

Definite Integral

Theorem

If $f, g \in \mathcal{R}[a,b]$ *and* $f(x) \geq g(x)$ *for all* $x \in [a,b]$, *then*

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

Definite Integral

Theorem

If $f, g \in \mathcal{R}[a,b]$ *and* $f(x) \geq g(x)$ *for all* $x \in [a,b]$, *then*

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

Theorem

If $f \in \mathcal{R}[a,b]$ *then*

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$$

Definite Integral

Let $f \in \mathcal{R}[a,b]$. Denote $F(x) = \int\limits_a^x f(t)\,dt$.

Definite Integral

Let $f \in \mathcal{R}[a,b]$. Denote $F(x) = \int\limits_a^x f(t)\,dt$.

Theorem

$F \in C[a,b]$.

Definite Integral

Let $f \in \mathcal{R}[a, b]$. Denote $F(x) = \int\limits_a^x f(t)\,dt$.

Theorem

$F \in C[a, b]$.

Theorem

If f is continuous at the point $x_0 \in [a, b]$, then F is differentiable and

$$F'(x_0) = f(x_0).$$

Definite Integral

Theorem

If $f \in C[a,b]$ and F is an antiderivative of f, then

$$\int_a^b f(t)\,dt = F(b) - F(a)$$

Definite Integral

Theorem

If $f \in C[a,b]$ *and* F *is an antiderivative of* f, *then*

$$\int_a^b f(t)\,dt = F(b) - F(a) = F(x)\Big|_{x=a}^{b}.$$

Definite Integral

Theorem

Let $f \in C[a,b]$, $\varphi : [\alpha,\beta] \to [a,b]$, $\varphi \in C^1[\alpha,\beta]$ *and* $\varphi(\alpha) = a$, $\varphi(\beta) = b$, *then*

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t))\,\varphi'(t)\,dt$$

Definite Integral

Theorem

Let $f \in C\left[a,b\right]$, $\varphi : \left[\alpha,\beta\right] \to \left[a,b\right]$, $\varphi \in C^1\left[\alpha,\beta\right]$ *and* $\varphi\left(\alpha\right) = a$, $\varphi\left(\beta\right) = b$, *then*

$$\int\limits_a^b f\left(x\right)dx = \int\limits_\alpha^\beta f\left(\varphi\left(t\right)\right)\varphi'\left(t\right)dt$$

Theorem

If $f, g \in C^1\left[a,b\right]$, *then*

$$\int\limits_a^b f dg = fg|_a^b - \int\limits_a^b g df$$

Improper Integral

Definition

Let $f : [a, +\infty) \to \mathbb{R}$ *and* $f \in \mathcal{R}[a, A]$ *for all* $A > a$. *If there exists the limit*

$$\lim_{A \to +\infty} \int_a^A f(x)\,dx,$$

then this limit is called the improper integral of f *from* a *to* $+\infty$ *and it is denoted by*

$$\int_a^\infty f(x)\,dx.$$

Improper Integral

Definition

Let $f : [a, +\infty) \to \mathbb{R}$ *and* $f \in \mathcal{R}[a, A]$ *for all* $A > a$. *If there exists the limit*

$$\lim_{A \to +\infty} \int\limits_a^A f(x)\,dx,$$

then this limit is called the improper integral of f *from* a *to* $+\infty$ *and it is denoted by*

$$\int\limits_a^\infty f(x)\,dx.$$

In the same way we can define $\int\limits_{-\infty}^a f(x)\,dx$.

Improper Integral

Example

$$\int_0^{+\infty} \frac{dx}{1+x^2} =$$

Improper Integral

Example

$$\int_{0}^{+\infty} \frac{dx}{1+x^2} = \lim_{A \to +\infty} \operatorname{arctg} x \Big|_0^A =$$

Improper Integral

Example

$$\int_0^{+\infty} \frac{dx}{1+x^2} = \lim_{A \to +\infty} \operatorname{arctg} x \Big|_0^A = \lim_{A \to +\infty} \operatorname{arctg} A =$$

Example

$$\int\limits_{0}^{+\infty} \frac{dx}{1+x^2} = \lim_{A\to+\infty} \operatorname{arctg} x\Big|_0^A = \lim_{A\to+\infty} \operatorname{arctg} A = \frac{\pi}{2}.$$

Improper Integral

Proposition

If $\int\limits_{a}^{+\infty} f(x)\,dx$ *is convergent, then* $\lim\limits_{A\to\infty} \int\limits_{A}^{+\infty} f(x)\,dx = 0.$

Improper Integral

Proposition

If $\int\limits_a^{+\infty} f(x)\,dx$ *is convergent, then* $\lim\limits_{A\to\infty} \int\limits_A^{+\infty} f(x)\,dx = 0.$

Example

For the function

$$f(x) = \begin{cases} 1, x \in \mathbb{N}, \\ 0, x \notin \mathbb{N}, \end{cases}$$

the integral $\int\limits_0^{+\infty} f(x)\,dx$ *is convergent, but* $\lim\limits_{x\to+\infty} f(x) \neq 0.$

Improper Integral

Theorem

If $f(x) \geq 0$, *then the integral* $\int\limits_a^{+\infty} f(x)\,dx$ *is convergent if and only if the function* $F(A) = \int\limits_a^{A} f(x)\,dx$ *is bounded above. If it is not bounded above then* $\int\limits_a^{+\infty} f(x)\,dx = +\infty$.

Improper Integral

Theorem

Let $f, g : [a, +\infty) \to \mathbb{R}$ *and* $f, g \geq 0$. *If there exists a number* $A > a$ *such that* $f(x) \leq g(x)$ *for all* $x > A$, *then from the convergence of the integral* $\int\limits_a^{+\infty} g(x)\, dx$ *follows that* $\int\limits_a^{+\infty} f(x)\, dx$ *is convergent too.*

Improper Integral

Theorem

Let $f, g : [a, +\infty) \to \mathbb{R}$ *and* $f, g \geq 0$. *If there exists a number* $A > a$ *such that* $f(x) \leq g(x)$ *for all* $x > A$, *then from the convergence of the integral* $\int\limits_a^{+\infty} g(x)\,dx$ *follows that* $\int\limits_a^{+\infty} f(x)\,dx$ *is convergent too.*

Theorem

Let $f, g : [a, +\infty) \to \mathbb{R}$ *and* $f, g \geq 0$. *If* $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = K$, *where* $0 \leq K < +\infty$, *then from the convergence of the integral* $\int\limits_a^{+\infty} g(x)\,dx$ *follows that* $\int\limits_a^{+\infty} f(x)\,dx$ *is convergent too.*

Improper Integral

Example

a) $\int\limits_{0}^{1} \frac{1}{\sqrt{x}} dx =$

Improper Integral

Example

a) $\int\limits_0^1 \frac{1}{\sqrt{x}} dx = 2\sqrt{x}\Big|_0^1 =$

Improper Integral

Example

a) $$\int_0^1 \frac{1}{\sqrt{x}} dx = 2\sqrt{x}\Big|_0^1 = 2.$$

Improper Integral

Example

a) $\int\limits_0^1 \frac{1}{\sqrt{x}} dx = 2\sqrt{x}\Big|_0^1 = 2.$

b) *Prove that the gamma function* $\Gamma(\alpha)$, *defined as*

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

satisfies the relation

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1), \alpha > 1.$$

Taylor Series

Taylor Series

Definition

The **Taylor polynomial** *of degree* n *of* $f : \mathbb{R} \to \mathbb{R}$ *at* x_0 *is defined as*

$$T_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

where $f^{(k)}(x_0)$ *is the* k*th derivative of* f *at* x_0 *(which we assume exists) and* $\frac{f^{(k)}(x_0)}{k!}$ *are the coefficients of the polynomial.*

Taylor Series

Definition

The **Taylor polynomial** *of degree* n *of* $f : \mathbb{R} \to \mathbb{R}$ *at* x_0 *is defined as*

$$T_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

where $f^{(k)}(x_0)$ *is the* k*th derivative of* f *at* x_0 *(which we assume exists) and* $\frac{f^{(k)}(x_0)}{k!}$ *are the coefficients of the polynomial.*

Definition

For a smooth function $f \in C^{\infty}, f : \mathbb{R} \to \mathbb{R}$ *the* **Taylor series** *of* f *at* x_0 *is defined as*

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

Remark

In general, a Taylor polynomial of degree n is an approximation of a function, which does not need to be a polynomial. The Taylor polynomial is similar to f in a neighborhood around x_0.

Taylor Series

Remark

In general, a Taylor polynomial of degree n is an approximation of a function, which does not need to be a polynomial. The Taylor polynomial is similar to f in a neighborhood around x_0.

Example

Find the Taylor polynomial $T_5(x)$ of $f(x) = x^3$ at point $x_0 = 2$. Verify that $T_5(x) = f(x)$.

Taylor Series

Remark

In general, a Taylor polynomial of degree n is an approximation of a function, which does not need to be a polynomial. The Taylor polynomial is similar to f in a neighborhood around x_0.

Example

Find the Taylor polynomial $T_5(x)$ of $f(x) = x^3$ at point $x_0 = 2$. Verify that $T_5(x) = f(x)$.

Remark

Taylor polynomial of degree n is an exact representation of a polynomial f of degree $k \leq n$.

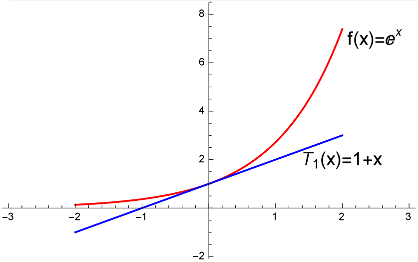
Taylor Series

Example

Find the Taylor series of the function $f(x) = e^x$.

Example

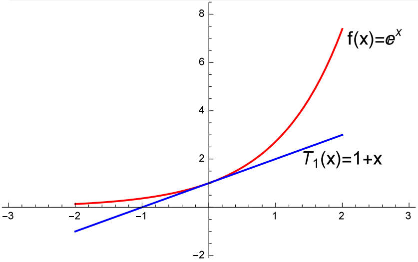
Find the Taylor series of the function $f(x) = e^x$.

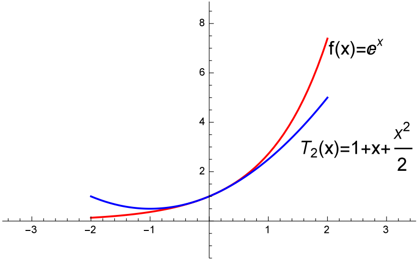


Taylor Series

Example

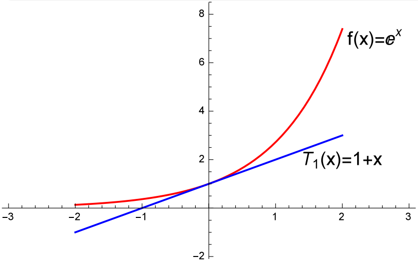
Find the Taylor series of the function $f(x) = e^x$.

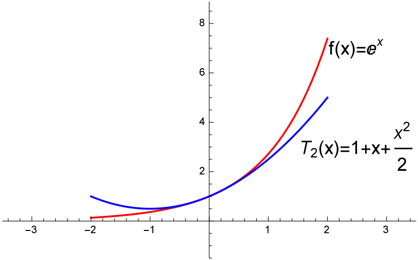


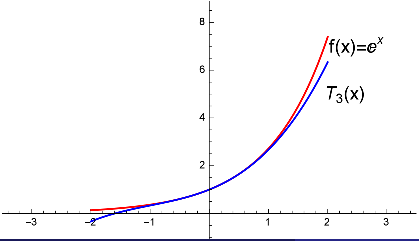


Example

Find the Taylor series of the function $f(x) = e^x$.







Taylor's theorem

If $|f^{(n+1)}(x)| \leq M$ for $x \in (x_0 - h, x_0 + h)$, then

$$f(x) = T_n(x) + r_n(x), \text{ where } |r_n(x)| \leq \frac{M}{n!}|x - x_0|^{n+1},$$

for $x \in (x_0 - h, x_0 + h)$.

Taylor Series

Theorem

1. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \; x \in \mathbb{R},$

Taylor Series

Theorem

1. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},\ x \in \mathbb{R},$
2. $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1},\ x \in \mathbb{R},$

Taylor Series

Theorem

1. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \; x \in \mathbb{R},$
2. $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \; x \in \mathbb{R},$
3. $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \; x \in \mathbb{R},$

Taylor Series

Theorem

1. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $x \in \mathbb{R}$,
2. $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$, $x \in \mathbb{R}$,
3. $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$, $x \in \mathbb{R}$,
4. $\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$, $x \in (-1, 1)$,

Taylor Series

Theorem

1. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $x \in \mathbb{R}$,
2. $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$, $x \in \mathbb{R}$,
3. $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$, $x \in \mathbb{R}$,
4. $\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$, $x \in (-1, 1)$,
5. $(1+x)^\alpha = \sum_{n=0}^{\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n$, $x \in (-1, 1)$.

Partial Differentiation and Gradients

Partial Differentiation and Gradients

Definition

For a function $f : \mathbb{R}^n \to \mathbb{R}$ *(of* n *variables* $x_1, \ldots, x_n$*)*

Definition

For a function $f : \mathbb{R}^n \to \mathbb{R}$ (of n variables $x_1, \ldots, x_n$) we define the **partial derivatives** *as*

$$\begin{aligned}
f'_{x_1}(\mathbf{x}) = \frac{\partial f}{\partial x_1}(\mathbf{x}) &= \lim_{h \to 0} \frac{f(x_1 + h, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{h} \\
&\vdots \\
f'_{x_n}(\mathbf{x}) = \frac{\partial f}{\partial x_n}(\mathbf{x}) &= \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_n + h) - f(x_1, x_2, \ldots, x_n)}{h}.
\end{aligned}$$

Partial Differentiation and Gradients

Definition

For a function $f : \mathbb{R}^n \to \mathbb{R}$ (of n variables $x_1, \dots, x_n$) we define the **partial derivatives** *as*

$$f'_{x_1}(\mathbf{x}) = \frac{\partial f}{\partial x_1}(\mathbf{x}) \quad = \lim_{h \to 0} \frac{f(x_1 + h, x_2, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{h}$$

$$\vdots$$

$$f'_{x_n}(\mathbf{x}) = \frac{\partial f}{\partial x_n}(\mathbf{x}) \quad = \lim_{h \to 0} \frac{f(x_1, x_2, \dots, x_n + h) - f(x_1, x_2, \dots, x_n)}{h}.$$

The row vector

$$\nabla f = \frac{df}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} & \frac{\partial f(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{1 \times n},$$

is called the **gradient** *of f or the Jacobian.*