Vector Calculus

Tuan Nguyen

Ngày 13 tháng 1 năm 2021

Overview

Taylor Polynomial

Differentiation Rules

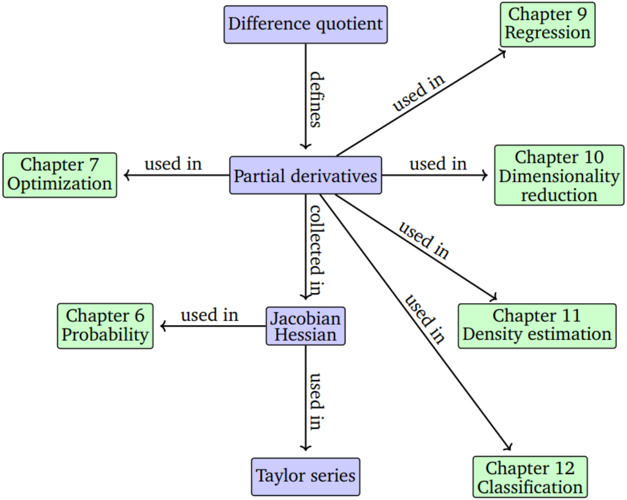
Partial Differentiation and Gradients

Rules of Partial Differentiation

Gradients of Vector-Valued Functions

Gradients of Matrices

Higher-Order Derivatives



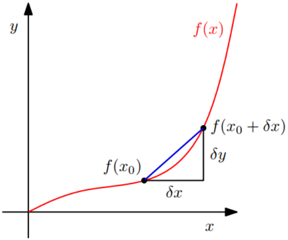
Hình 1: The use of calculus

Differentiation of Univariate Functions

Univariate function y = f(x), x, y ∈ R, we define:

Difference Quotient

$$\frac{\delta y}{\delta x} := \frac{f(x + \delta x) - f(x)}{\delta x}$$



The difference quotient can also be considered the average slope of f between x and x + δx if we assume f to be a linear function

The derivative of f at x is defined as the limit

$$\frac{df}{dx} := \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

The derivative of f points in the direction of steepest ascent of f.

Exercise: Use the definition to calculate the derivative of the function:

$$f(x) = x^2$$

Taylor Polynomial

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

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

Proof?

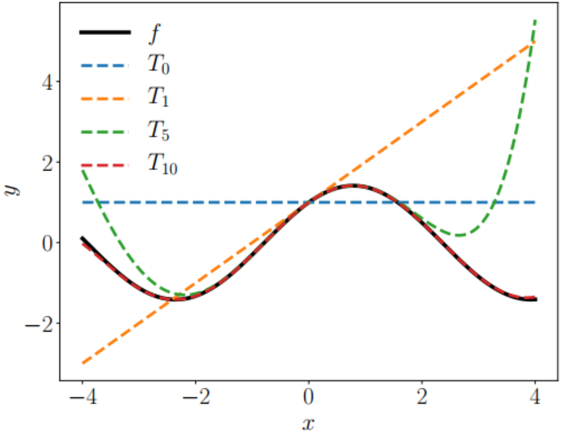
For a smooth function f: $\mathrm{R} \to \mathrm{R}$, the Taylor series of f at x_0 is defined as

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

Write Taylor series for a function:

$$f(x) = \sin(x) + \cos(x)$$

Taylor Polynomial (cont.)



Hình 2: Taylor series approximation

Basic derivative:

- (c)' $= 0$, c is a constant
- $(a^n)' = na^{n-1}$
- $(e^x)' = e^x$
- $(log(x))' = \frac{1}{x}$

Differentiation Rules

- Product rule: $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$
- Quotient rule: $(\frac{f(x)}{g(x))})' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$
- Sum rule: $(f(x) + g(x))' = f'(x) + g'(x)$
- Chain rule: $g(f(x))' = g'(f(x))f'(x)$

Differentiation Rules (cont.)

Calculate derivative of following functions:

- $\sigma(x) = \frac{1}{1+e^{-x}}$
- $f(x) = (3x + 2)^2$
- $f(x) = -(y \log z + (1 - y) \log(1 - z)), z = \sigma(x)$

Partial Differentiation and Gradients

For a function f : R $\rightarrow$ R; $x \rightarrow f(x)$; $x \in$ R^n of n variables $x_1, x_2, ..., x_n$, we define the partial derivatives as

$$\frac{\partial f}{\partial x_1} = \lim_{h \to 0} \frac{f(x_1 + h, x_2, ..., x_n) - f(x_1, x_2, ..., x_n)}{h}$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = \lim_{h \to 0} \frac{f(x_1, x_2, ..., x_n + h) - f(x_1, x_2, ..., x_n)}{h}$$

And

$$\frac{df}{dx} = [\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_1}] \in \mathrm{R}^{1 \times n}$$

is called the gradient of f or the Jacobian.

Exercise: Calculate the gradient of function
$f(x_1, x_2) = x_1^2 x_2 + x_1 x_2^3 \in$ R

Rules of Partial Differentiation

Here are the general product rule, sum rule, and chain rule:

- $\frac{\partial}{\partial x}(f(x)g(x)) = \frac{\partial f(x)}{\partial x}g(x) + \frac{\partial g(x)}{\partial x}f(x)$
- $\frac{\partial}{\partial x}(f(x) + g(x)) = \frac{\partial f(x)}{\partial x} + \frac{\partial g(x)}{\partial x}$
- $\frac{\partial}{\partial x}(g(f(x)) = \frac{\partial f}{\partial g}\frac{\partial g}{\partial x}$

Given $f(x_1, x_2) = x_1^2 + 2x_2$, where $x_1 = sin(t)$ and $x_2 = cos(t)$, calculate $\frac{df}{dt}$

Gradients of Vector-Valued Functions

For a function f : $R^n \to R^m$ and a vector $x = [x_1, \ldots, x_n] \in R^n$, the corresponding vector of function values is given as

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \in R^m$$

$f_i : R^n \to R$ that map onto R $\Rightarrow \frac{df_1(x)}{dx} \in R^{1\times n} \Rightarrow \frac{df(x)}{dx} \in R^{m\times n}$

$$\frac{df}{dx} = \begin{bmatrix} \frac{df_1(x)}{dx_1} & \cdots & \frac{df_1(x)}{dx_n} \\ \vdots & \vdots & \vdots \\ \frac{df_1(x)}{dx_1} & \cdots & \frac{df_1(x)}{dx_n} \end{bmatrix}$$

It is called Jacobian.

Given f(x) = Ax, f(x)$\in R^m$; A$\in R^{m\times n}$; $x \in R^n$. Calculate: $\dfrac{df}{dx}$

Calculate the gradient of following functions:

- f(x) $= a^T x$
- f(x) $=$ Ax
- f(x) $= x^T Ax$
- f(x) $= \|Ax - b\|_2^2$
- f(x) $= a^T Xb$

Higher-Order Derivatives

Consider a function f : $R^2 \to R$ of two variables x; y. We use the following notation for higher-order partial derivatives:

- $\frac{\partial^2 f}{\partial x^2}$ is the second partial derivative of f with respect to x.
- $\frac{\partial^n f}{\partial x^n}$ is the n^{th} partial derivative of f with respect to x.
- $\frac{\partial^2 f}{\partial x \partial y}$ is the partial derivative obtained by first partial differentiating with respect to x and then with respect to y.

The **Hessian** is the collection of all second-order partial derivatives.

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$