Differentiation

Derivative

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

Total derivative

$f \in U \subset \mathbb{R}^n \to \mathbb{R}^m, \quad L \in \mathbb{R}^{m\times n}$

$$f'(a) := L \iff \lim_{a\to x} \frac{||f(x)-f(a)-L(x-a)||}{||x-a||} = 0$$

We then call L the **Jacobian matrix** and usually denote it by $f'(a)$.

In the one dimensional case no distinction has to be made between the total derivative and the derivative.

Notation

$f \in \mathcal{C}$	f is continuous in $\mathbb{R}^n$
$f \in \mathcal{C}(S)$	f is continuous in S
$f \in \mathfrak{C}$	f is differentiable on $\mathbb{R}^n$
$f \in \mathfrak{C}(S)$	f is differentiable on S
$f \in \mathfrak{C}_T$	f is total differentiable on $\mathbb{R}^n$
$f \in \mathfrak{C}_T(S)$	f is total differentiable on S
$f \in \mathfrak{C}_P$	f is partial differentiable on $\mathbb{R}^n$
$f \in \mathfrak{C}_P(S)$	f is partial differentiable on S
$f \in \mathcal{C}^1$	f is continuously differentiable on $\mathbb{R}^n$
$f \in \mathcal{C}^1(S)$	f is continuously differentiable on S

We use $\mathfrak{C}^n$ to denote an n-times differentiable function.

Note that $f \in \mathfrak{C}$ and $f \in \mathfrak{C}(S)$ is only used in the one dimensional case.

Elementary functions

$f(x)$	$\to$	$\frac{d}{dx}f(x)$	
c	$\to$	0	
cx	$\to$	c	
x^c	$\to$	cx^{c-1}	$c>0$
$\lvert x\rvert$	$\to$	$\text{sign}x$	$x\neq 0$
c^x	$\to$	$c^x \log c$	$c>0$
$\log_c x$	$\to$	$\frac{1}{x\log c}$	$c>0, c\neq 1$
$\frac{1}{x}$	$\to$	$\frac{-1}{x^2}$	$x\neq 0$
$\frac{1}{x^c}$	$\to$	$\frac{-c}{x^{c+1}}$	$x\neq 0$
$\sqrt{x}$	$\to$	$\frac{1}{2\sqrt{x}}$	$x>0$
e^x	$\to$	e^x	
$\ln\lvert x\rvert$	$\to$	$\frac{1}{x}$	
x^x	$\to$	$x^x(\log x+1)$	
$(f^g)'$	$\to$	$f^g(g'\ln f + \frac{g}{f}f')$	

Hyperbolic functions

$f(x)$	$\to$	$\frac{d}{dx}f(x)$
$\sinh x$	$\to$	$\cosh x$
$\cosh x$	$\to$	$\sinh x$
$\tanh x$	$\to$	$-\tanh^2 x + 1 = \frac{1}{\cosh^2 x}$
$\coth x$	$\to$	$-\coth^2 x + 1 = \frac{-1}{\sinh x}$
$\text{arsinh}x$	$\to$	$\frac{1}{\sqrt{x^2+1}}$
$\text{arcosh}x$	$\to$	$\frac{1}{\sqrt{x^2-1}}$
$\text{artanh}x$	$\to$	$\frac{1}{1-x^2}$
$\text{arcoth}x$	$\to$	$\frac{1}{1-x^2}$
$\text{csch}x = \frac{1}{\sinh x}$	$\to$	$\frac{-\cosh x}{\sinh^2 x}$
$\text{sech}x = \frac{1}{\cosh x}$	$\to$	$\frac{-\sinh x}{\cosh^2 x}$

Trigonometric functions

$f(x)$	$\to$	$\frac{d}{dx}f(x)$
$\sin x$	$\to$	$\cos x$
$\cos x$	$\to$	$-\sin x$
$\tan x$	$\to$	$\tan^2 x + 1 = \frac{1}{\cos^2 x}$
$\cot x$	$\to$	$-\cot^2 x - 1 = \frac{-1}{\sin^2 x}$
$\arcsin x$	$\to$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos x$	$\to$	$\frac{-1}{\sqrt{1-x^2}}$
$\arctan x$	$\to$	$\frac{1}{1+x^2}$
$\text{arccot}x$	$\to$	$\frac{-1}{1+x^2}$
$\csc x = \frac{1}{\sin x}$	$\to$	$\frac{-\cos x}{\sin^2 x}$
$\sec x = \frac{1}{\cos x}$	$\to$	$\frac{\sin x}{\cos^2 x}$

Single variable rules

$(cf)'$	$=$	cf'	Linearity 1
$(f\pm g)'$	$=$	$f'\pm g'$	Linearity 2
$(fg)'$	$=$	$f'g+fg'$	Product rule
$\left(\frac{f}{g}\right)'$	$=$	$\frac{f'g-fg'}{g^2}$	Quotient rule
$\left(\frac{1}{f}\right)'$	$=$	$\frac{-f'}{f^2}$	Reciprocal rule
$(f\circ g)'$	$=$	$(f'\circ g)\cdot g'$	Chain rule
$(f^{-1})'$	$=$	$\frac{1}{f'\circ f^{-1}}$	Inverse function rule

Mean value theorem

$f \in \mathcal{C}([a,b]) \wedge f \in \mathfrak{C}((a,b)) \implies$
$\exists c \in (a,b) \ : \ f'(c) = \frac{f(b)-f(a)}{b-a}$

Logarithmic derivative

$$(\ln f)' = \frac{f'}{f}$$

Real-valued functions

$f : D \subset \mathbb{R}^n \to \mathbb{R}$ where D is open.
$\boldsymbol{x} : I \subset \mathbb{R} \to \mathbb{R}^n$

Partial derivative

$\frac{\partial f}{\partial x} = \partial_x f$ — all variables except x are considered to be constants

If all partial derivates exist we call the function **partial differentiable**.

$f \in \mathfrak{C}_P(D) \not\Longrightarrow f \in \mathcal{C}^1(D)$
$f \in \mathfrak{C}_T(D) \Longleftarrow f \in \mathcal{C}^1(D)$

$f \in \mathfrak{C}_T(D) \implies f' = [\partial_{x1} f \ \partial_{x2} f \ \dots]$

Schwarz's theorem

$$f \in \mathcal{C}^2(D) \implies \frac{\partial^2 f}{\partial xy} = \frac{\partial^2 f}{\partial yx}$$

Chain rule

$f \in \mathcal{C}^1(D), \ \boldsymbol{x} \in \mathfrak{C}(I) \implies$

$$\frac{d}{dt} f(\boldsymbol{x}(t)) = f'(\boldsymbol{x}(t)) \cdot \dot{\boldsymbol{x}}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x}(t)) \cdot \dot{x}_i(t)$$

Mean value theorem

$f \in \mathcal{C}^1(D) \implies$
$\exists \theta \in (0,1) \ : \ f(\boldsymbol{x}) - f(\boldsymbol{a}) = f'(\boldsymbol{a} + \theta(\boldsymbol{x}-\boldsymbol{a}))(\boldsymbol{x}-\boldsymbol{a})$

Directional derivative

$$\partial_{\boldsymbol{v}} f = \langle \nabla f, \boldsymbol{v} \rangle$$

Vector-valued functions

$\boldsymbol{f} : B \subset \mathbb{R}^l \to \mathbb{R}^m$
$\boldsymbol{g} : D \subset \mathbb{R}^n \to \mathbb{R}^l$

Jacobian matrix

If $\boldsymbol{f} \in \mathcal{C}^1(B)$

$$\boldsymbol{f}' = \begin{bmatrix} \partial_{x_1} f_1 & \dots & \partial_{x_n} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m & \dots & \partial_{x_n} f_m \end{bmatrix}$$

Chain rule

$\boldsymbol{f} \in \mathcal{C}^1(B), \boldsymbol{g} \in \mathfrak{C}_P(D) \implies$
$(\boldsymbol{f}\circ\boldsymbol{g})' = (\boldsymbol{f}'\circ\boldsymbol{g})\cdot\boldsymbol{g}'$

from that

$$\partial_t f(x(t), y(t)) = \partial_x f \cdot \partial_t x + \partial_y f \cdot \partial_t y$$

Mean value theorem

$\boldsymbol{f} \in \mathcal{C}^1(B) \implies$
$\exists \theta_i \in (0,1) \ : \ f_i(\boldsymbol{x}) - f_i(\boldsymbol{a}) = f_i'(\boldsymbol{a} + \theta_i(\boldsymbol{x}-\boldsymbol{a}))(\boldsymbol{x}-\boldsymbol{a})$

Inverse function rule

$\boldsymbol{f} : D \subset \mathbb{R}^n \to \mathbb{R}^n$
D open $\wedge$ $\boldsymbol{f} \in \mathcal{C}^1$ $\wedge$ $\exists \boldsymbol{f}^{-1}$
$\implies (\boldsymbol{f}^{-1})'(\boldsymbol{y}) = [\boldsymbol{f}'(\boldsymbol{f}^{-1}(\boldsymbol{y}))]^{-1}$

Gradient, Curl, Divergence

$$\nabla f = \begin{bmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{bmatrix} = (f')^T$$

$$\text{curl}\boldsymbol{f} = \nabla \times \boldsymbol{f}$$

$$\text{div}\boldsymbol{f} = \langle \nabla, \boldsymbol{f} \rangle$$

Hessian matrix

$$\mathcal{H}_f = \begin{bmatrix} \frac{\partial^2 \boldsymbol{f}}{\partial x_1^2} & \dots & \frac{\partial^2 \boldsymbol{f}}{\partial x_n x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \boldsymbol{f}}{\partial x_1 x_n} & \dots & \frac{\partial^2 \boldsymbol{f}}{\partial x_n^2} \end{bmatrix}$$

in the two dimensional case

$$\mathcal{H}_f = \begin{bmatrix} \partial_{xx} \boldsymbol{f} & \partial_{yx} \boldsymbol{f} \\ \partial_{xy} \boldsymbol{f} & \partial_{yy} \boldsymbol{f} \end{bmatrix}$$

Applications of Differentiation

L'Hôpital's rule

$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \quad \vee$

$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \pm\infty \quad \Longrightarrow$

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

Monotonicity

$f \in \mathcal{C}^1$

$f' \geq 0 \iff$ monotonically increasing

$f' \leq 0 \iff$ monotonically decreasing

$f' > 0 \Longrightarrow$ strictly increasing

$f' < 0 \Longrightarrow$ strictly decreasing

Newton's method

A zero (root) of f can be computed by (if x_0 is choosen near a zero)

$f'(x_n) \neq 0 \Longrightarrow$

$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

Newton's method (multivariate)

$\boldsymbol{f}'(\boldsymbol{x}_n)(\boldsymbol{x}_{n+1} - \boldsymbol{x}_n) = -\boldsymbol{f}(\boldsymbol{x}_n)$

Implicit function theorem

$\boldsymbol{f} : D \subset \mathbb{R}^{n+m} \to \mathbb{R}^m,\ (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$

$\boldsymbol{f} \in \mathcal{C}^1 \ \wedge\ D$ open

$\boldsymbol{f}' = [\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \ \ \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}]$

If

$$\boldsymbol{f}(\boldsymbol{a}, \boldsymbol{b}) = 0 \ \wedge\ \det \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \neq 0$$

Then there exists an open neighbourhood $U \subset \mathbb{R}^{n+m}$ of $(\boldsymbol{a}, \boldsymbol{b})$ such that there exits a function $\boldsymbol{g} : \mathbb{R} \to \mathbb{R}^m$ which describes the solutions of $\boldsymbol{f}(\boldsymbol{a}, \boldsymbol{b}) = 0$ in U.

Linear approximation (tangent) at a

$f \in \mathcal{C}^1 \Longrightarrow$

$f(x) \approx f(a) + f'(a) \cdot (x - a)$

Approximation by a polynomial of degree n at a

$f \in \mathcal{C}^{n+1}(I) \wedge x, a \in I \Longrightarrow$

$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} + R_{n+1}(x, a)$

$R_{n+1}(x, a) = \frac{1}{n!} \int_a^x (x - t)^n f^{(n+1)}(t)\ dt$ Remainder (Lagrange form)

$R_{n+1}(x, a) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - a)^{n-1}$ Remainder (Cauchy form)

$\xi \in (x, a) \vee \xi \in (a, x)$

We often don't know the remainder explicitely; However we can bound it by $\mathcal{O}(x^{n+1})$.

Taylor series (of f at a)

$f \in \mathcal{C}^\infty \Longrightarrow$

$T(x, a, f) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x - a)^k$

Taylor's theorem

$\lim_{n\to\infty} R_n(x) = 0 \iff$

$f(x) = T(x, a, f)$

In this case the function at x is equal to its taylor series.

Multivariate approximation

$T_1(\boldsymbol{x}, \boldsymbol{a}) = f(\boldsymbol{a}) + f'(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})$

$T_2(\boldsymbol{x}, \boldsymbol{a}) = f(\boldsymbol{a}) + f'(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}) + (\boldsymbol{x} - \boldsymbol{a})^T \mathcal{H}_f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})$

Multivariate Taylor series

$(Df(a))\,(\cdot) := \langle \nabla f, (\cdot) \rangle$

$$T(\boldsymbol{x}, \boldsymbol{a}, f) = \sum_{n=0}^{\infty} \frac{1}{n!} D^n f(\boldsymbol{a})\,(\boldsymbol{x} - \boldsymbol{a}, \ldots, \boldsymbol{x} - \boldsymbol{a})$$

The Taylor series is generalized componentwise to vector-valued functions.

Single variable Extrema

$f : [a, b] \to \mathbb{R}$

$\exists r \forall x\ :\ |x - c| < r \Longrightarrow f(x) < f(c)$ local minima

We call a, b and all points $c \in [a, b]$ where f isn't differentiable or where $f'(x) = 0$ a **critical point**.

Fermat's theorem

c extrema $\Longrightarrow c$ critical point

First derivative test

If f' changes sign at c, f has an extremum at c. If the sign changes from $+$ to $-$ the function has a maximum at c.

Second derivative test

$f''(c) < 0 \Longrightarrow f$ has a maxima at c

$f''(c) > 0 \Longrightarrow f$ has a minima at c

$f''(c) = 0 \Longrightarrow$ the test is inconclusive

Existence theorem

$f \in \mathcal{C}^1([a, b]) \ \wedge\ f \in \mathfrak{C}((a, b))$

$\Longrightarrow f$ has a maxima and minima on $[a, b]$

Multivariable Extrema

$f : D \subset \mathbb{R}^n \to \mathbb{R}$

$f \in \mathfrak{C}_P \ \wedge\ D$ open

$\exists r \forall \boldsymbol{x}\ :\ ||\boldsymbol{x} - \boldsymbol{c}|| < r \Longrightarrow f(\boldsymbol{x}) < f(\boldsymbol{c})$ local minima

f has a **critical** point at c iff $\nabla f(\boldsymbol{c}) = \boldsymbol{0}$.

Second derivative test

$\mathcal{H}_f(c)$ positive definite $\Longrightarrow$ minima

$\mathcal{H}_f(c)$ negative definite $\Longrightarrow$ maxima

$\mathcal{H}_f(c)$ indefinit $\Longrightarrow$ no extremum

Definiteness

$$\boldsymbol{H}_1 = a_{11} \quad \boldsymbol{H}_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \ldots \quad \boldsymbol{H}_n = \boldsymbol{A}$$

$\boldsymbol{A}$ positive definite $\Longleftarrow \forall i\ :\ \det \boldsymbol{H}_i > 0$

$\boldsymbol{A}$ negative definite $\Longleftarrow \forall i\ :\ \operatorname{sgn}(\det \boldsymbol{H}_i) = (-1)^i$

Extrema with constraints

$f : \mathbb{R}^n \to \mathbb{R}$

$g_k : \mathbb{R}^n \to \mathbb{R}$

$1 < k < m \ \wedge\ f \in \mathfrak{C}_P \ \wedge\ D$ open

The maxima of f under the constraints of $g_1, g_2, \cdots = 0$.

First Method

$[g(x, y) = 0 \iff y = h(x)]$

$\Longrightarrow f(x, h(x))' = 0$

Second method

$\{(x, g(x, y)) | g(x, y) = 0\} = \{(x(t), y(t)), t \in \mathbb{R})$

$\Longrightarrow f(x(t), y(t))' = 0$

Lagrange multipliers

$$L(x_1, \ldots, x_n, \lambda_1 \ldots \lambda_m) = f + \sum_{i=1}^{m} \lambda_i g_i$$

Then the extrema are given by

$\nabla L = \boldsymbol{0}$

Zeroes of the gradient

$\nabla(uv) = u \cdot \nabla v + v \cdot \nabla u$

u, v linearly independent $\Longrightarrow u = 0 \ \wedge\ v = 0$

u, v linearly dependent $\Longrightarrow$ equation with fewer variables

Linear dependence in the 2D case is determined by $\det[\nabla v\ \nabla u] = 0$.

Elementary taylor series

$\frac{1}{1-x}$	$=$	$\sum_{n\geq 0} x^n$	$\lvert x\rvert < 1$
$\frac{x^m}{1-x}$	$=$	$\sum_{n\geq m} x^n$	$\lvert x\rvert < 1,\ m \in \mathbb{N}_{\geq 0}$
$\frac{x}{(1-x)^2}$	$=$	$\sum_{n\geq 1} n x^n$	$\lvert x\rvert < 1$
$(1+x)^z$	$=$	$\sum_{n\geq 0} \binom{z}{n} \cdot x^n$	$\lvert x\rvert < 1,\ z \in \mathbb{C}$
$\sqrt{1+x}$	$=$	$\sum_{n\geq 0} \frac{(-1)^n (2n)!}{(1-2n) n!^2 4^n} x^n$	
e^x	$=$	$\sum_{n\geq 0} \frac{x^n}{n!}$	
$\log(1-x)$	$=$	$-\sum_{n\geq 1} \frac{x^n}{n}$	$\lvert x\rvert < 1$

Hyperbolic/Trigonometric taylor series

$\sin x$	$=$	$\sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$	
$\cos x$	$=$	$\sum_{n\geq 0} \frac{(-1)^n}{(2n)!} x^{2n}$	
$\tan x$	$=$	$\sum_{n\geq 1} \frac{B_{2n}(-4)^n(1-4)^n}{(2n)!} x^{2n-1}$	$\lvert x\rvert < \frac{\pi}{2}$
$\arcsin x$	$=$	$\sum_{n\geq 0} \frac{(2n)!}{4^n n!^2 (2n+1)} x^{2n+1}$	$\lvert x\rvert \leq 1$
$\arccos x$	$=$	$\frac{\pi}{2} - \arcsin x$	
$\arctan x$	$=$	$\sum_{n\geq 0} \frac{(-1)^n}{(2n+1)} x^{2n+1}$	$\lvert x\rvert \leq 1$
$\sec x$	$=$	$\sum_{n\geq 0} \frac{(-1)^n E_{2n}}{(2n)!} x^{2n}$	
$W_0(x)$	$=$	$\sum_{n\geq 1} \frac{(-n)^{n-1}}{n!} x^n$	$\lvert x\rvert < \frac{1}{e}$

The taylor series of *sinh*/*cosh*/*tanh* can be formed by removing $(-1)^n$ from the formulas of *sin*/*cos*/*tan*; The taylor series holds for $\mathbb{R}$, $\mathbb{R}$, and $|x| < \pi/2$ respectively. The formulas of *arcsinh*/*arccosh*/*arctanh* can be formed by adding $(-1)^n$ to the formulas for *arcsin*/*arccos*/*arctan*; They hold for $|x| < 1$.

B_n are the Bernoulli numbers. E_n are the Euler numbers.