Differential Calculus

Derivatives

Definition. ***One Variable Differentiability*** *A function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in \mathbb{R}$ if there exists an $m \in \mathbb{R}$ such that*

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

where $m = f'(a)$.

Remark.
The idea is that f is differentiable at a if it can be well-approximated by a linear function m,

$$f(a+h) = f(a) + mh + error(h)$$

such that the error go to zero faster than linearly in h.

$$\lim_{h \to 0} \frac{error(h)}{h} = 0$$

Also we can calculate derivative by evaluating

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

Note. Example of function continuous but not differentiable at 0

$$f(x) = \begin{cases} x \sin(\frac{1}{x}), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Example of differentiable function whose derivative is not continous

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Definition. ***Differentiability of vector valued function*** *A function $\gamma : \mathbb{R} \to \mathbb{R}^n$ is differentiable if at t_0,*

$$\begin{aligned} \gamma'(t_0) &= \lim_{h \to 0} \frac{\gamma(t_0+h) - \gamma(t_0)}{h} \\ &= \big(\lim_{h \to 0} \frac{\gamma_1(t_0+h) - \gamma_1(t_0)}{h}, \dots, \lim_{h \to 0} \frac{\gamma_2(t_0+h) - \gamma_2(t_0)}{h} \big) \end{aligned}$$

exists. γ is differentiable if all of its component functions are differentiable.

Proposition. ***Properties of vector valued function*** *Let $f, g : \mathbb{R} \to \mathbb{R}^n$ and $\varphi : \mathbb{R} \to \mathbb{R}$ be differentiable functions.*

1. $(\varphi f)' = \varphi' f + \varphi f'$
2. $(f \cdot g)' = f' \cdot g + f \cdot g'$
3. $(f \times g)' = f' \times g + f \times g'$ *(if $n = 3$)*

Definition. ***Multivariable differentiability*** *A function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $a \in \mathbb{R}^n$ if there exists $c \in \mathbb{R}^n$ such that*

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

where c if exists is called the ***gradient*** *of f, denoted as $\nabla f(a)$*

Theorem. *If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a then f is continous at a.*

Proof.

$$\begin{aligned}
\lim_{h \to 0} f(a+h) - f(a) &= \lim_{h \to 0} [f(a+h) - f(a) - \nabla f(a) \cdot h + \nabla f(a) \cdot h] \\
&= \lim_{h \to 0} f(a+h) - f(a) - \nabla f(a) \cdot h + \lim_{h \to 0} \nabla f(a) \cdot h \\
&= 0 + 0 = 0
\end{aligned}$$

□

Definition. *If $f : \mathbb{R}^n \to \mathbb{R}$. we define* ***partial derivatives*** *of f with respect to x_i at $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ as*

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a_1, \dots a_i + h, \dots a_n) - f(a_1, \dots, a_n)}{h}$$

That is $\frac{\partial f}{\partial x_i}$ is the one varirable derivative of $f(x_1, \dots, x_n)$ with respect to x_i where all other variables are held constant.

Theorem. *If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a then the partials of f exist at a and*

$$\nabla f(a) = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$$

Remark.
Example of function where **partials exist** but function **not differentiable**. This is reasonable bcause partials only measure differentiability in finitely many directions that the converse direction does not hold.

$$f(x,y) = \begin{cases} \frac{xy}{x^2 + y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

Proof. Function is not continous at $(x, y) = (0, 0)$ (prove this by taking a path and show limit is depends on the path) and therefore not differentiable. However partials exists at $(0, 0)$ by the limit definition.

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = 0$$

Note this could be explained by the fact that partials of f near zero is not continous

$$\frac{\partial f}{\partial x} = \frac{y^3 - x^2 y}{(x^2 + y^2)^2}$$

Partials does not exist as $(x, y) \to (0, 0)$ $\square$

Also example of function where **directional derivative exists** at every direction but function **not differentiable**.

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Definition. ***Continously differentiable*** *functions is in the collection of C^1 function on U,*

$$C^1(\mathbb{R}^n, \mathbb{R}) = \left\{ f : \mathbb{R}^n \to \mathbb{R} : \partial_i f \text{ exists and is continous for } i \in (1, \dots, n) \right\}$$

Theorem 0.1. ***C^1 functions are differentiable*** *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$, If $\partial_i f(x)$ all exists and are continous in an open neighborhood of a, then f is differentiable at a*

Remark. Example of function differentiable but not C^1

$$f(x, y) = \begin{cases} (x^2 + y^2) \sin(\dfrac{1}{\sqrt{x^2 + y^2}}), & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

We can see that that derivative not continous at 0.

Definition. *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$. If $u \in \mathbb{R}^n$ is a unit vector ($||u|| = 1$) then the* ***directional derivative*** *of f in the direction of u at a is*

$$\partial_u f(a) = \lim_{t \to 0} \frac{f(a + tu) - f(a)}{t} = \frac{d}{dt}|_{t=0} f(a + tu)$$

Theorem. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable at a, then for any unit vector u, $\partial_u f$ exists. Moreover,*

$$\partial_u f(a) = \nabla f(a) \cdot u$$

Remark. To ways to compute partial derivatives.

1. compute using limit definition
2. compute partials first and then $\partial_u f(a) = \nabla f(a) \cdot u$

Definition. ***Generalized differentiability*** *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$ if there exists an $m \times n$ matrix A such that*

$$\lim_{h \to 0} \frac{||f(a+h) - f(a) - Ah||_{\mathbb{R}^m}}{||h||_{\mathbb{R}^2}} = 0$$

Here $Df(a) = A$, the ***Jacobian Matrix***

Proposition. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is given by $f(x) = (f_1(x), \ldots, f_m(x))$, then f is differentiable if and only if each of the $f_i : \mathbb{R}^n \to \mathbb{R}$ is differentiable, that is*

$$Df(a) = \begin{bmatrix} \nabla f_1(a) \\ \nabla f_2(a) \\ \vdots \\ \nabla f_m(a) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Theorem. ***Chain Rule*** *Let $g : \mathbb{R}^k \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$. If g is differentiable at $a \in \mathbb{R}^k$ and f is differentiable at $g(a) \in \mathbb{R}^n$, then $f \circ g$ is differentiable at a, and*

$$D(f \circ g)(a) = Df(g(a))Dg(a)$$

Remark. Note that the gradient of a function $\mathbb{R}^n \to \mathbb{R}$ is a row vector and the derivative of a function $\mathbb{R} \to \mathbb{R}^n$ is a column vector.
special case 1, When $g : \mathbb{R} \to \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, so $f \circ g : \mathbb{R} \to \mathbb{R}$. Let $y = f(x)$ and let $(x_1, \ldots, x_n) = g(t) = (g_1(t), \ldots, g_n(t))$ so,

$$\frac{d}{dt}(f \circ g) = \frac{\partial y}{\partial x_1}\frac{\partial x_1}{\partial t} + \cdots + \frac{\partial y}{\partial x_n}\frac{\partial x_n}{\partial t}$$

special case 2, When $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}$ so that $f \circ g : \mathbb{R}^n \to \mathbb{R}$. if $y = f(x)$ and $x = g(t)$ then

$$\frac{\partial}{\partial t_i}(f \circ g)(x) = \frac{\partial y}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \cdots + \frac{\partial y}{\partial x_m}\frac{\partial x_m}{\partial t_i}$$

Another way of putting is

$$\partial_i (f \circ g)(a) = \nabla f(g(a)) Dg(a) = \nabla f(g(a)) \begin{pmatrix} \nabla g_i(a) \\ \vdots \\ \nabla g_m(a) \end{pmatrix} = \sum_{j=1}^{m} \partial_j f(g(a)) \cdot \partial_i g_j(a)$$

where $1 \leq j \leq m$ and $q \leq i \leq n$ and g_i is i-th component function of g
In summary we compute derivatives either with direct substitution or with the chain rule, where we compute jacobian matrix and compose them.

Definition. *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable at a, then we define the **Jacobian (determinant)** of f to be $detDf(a)$*

Definition. *Some properties of multivariate differentiable function*

1. *If f is a constant function ($\exists y \in \mathbb{R}^m, f(x) = y$ for all $x \in \mathbb{R}^n$) then $Df(a) = T_o$ where $T_o = \vec{0}$*

2. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $Df(a) = f$, i.e. the derivative is itself.*

   Proof. Since f differentible, error approach 0 as $h \to 0$

   $$0 = error(h) = f(a+h) - f(a) - Ah = f(a) + f(h) - f(a) - Ah \Rightarrow f(h) = A(h)$$

   Meaning that the linear map $Df = A = f$ □

   As an example, If $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + y = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} x & y \end{smallmatrix}\right)$, i.e. f is linear, then $Df(a) = s$

   Proof.
   $$\lim_{h \to 0} \frac{f(a+h) - f(a) - \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} h_1 & h_2 \end{smallmatrix}\right)}{||h||} = 0$$

   □

3. *$f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a if and only if f_i, the i-th component function, is differentiable at a*

4. *$f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto xy$, then $Df(a) : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto a_2 x + a_1 y$*

   Theorem. *Let $f, g : \mathbb{R}^n \to \mathbb{R}$, then*

1. ***Sum Rule:***
   $$D(f + g)(a) = Df(a) + Dg(a)$$

2. ***Product Rule:***

$$D(f \cdot g)(a) = f(a)Dg(a) + g(a)Df(a)$$

Proof. Let s represent the summation and

$$\begin{aligned} D(f+g)(a) &= D(s \circ (f,g))(a) \\ &\overset{\text{chain rule}}{=} Ds(f(a),g(a)) \circ D(f,g)(a) \\ &= Ds(f(a),g(a)) \circ (Df(a), Dg(a)) = Df(a) + Dg(a) \end{aligned}$$

□

3. ***Quotient Rule:***

$$D(\frac{f}{g})(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{g(a)^2} \text{ if } g(a) \neq 0$$

Theorem. ***Mean Value Theorem for One Variable*** *In one variable, if $f : [a,b] \to \mathbb{R}$ is continous on $[a,b]$ and differentiable on (a,b), then there exists $c \in (a,b)$ such that*

$$f(b) - f(a) = f'(c)(b-a)$$

Corollary. *A short list of propositions*

1. *There is a point such that the tangent line has the same slope as the secant between $(a, f(a))$ and $(b, f(b))$*
2. *If $f : [a,b] \to \mathbb{R}$ is differentiable with bounded derivative, say $|f'(x)| \leq M$ for all $x, y \in [a,b]$, then $|f(y) - f(x)| \leq M|y-x|$*
3. *If $f'(x) \equiv 0$ for all $x \in [a,b]$ then f is the constant function on $[a,b]$*
4. *If $f'(x) > 0$ for all $x \in [a,b]$ then f is an increasing (and hence injective) function*

Theorem. ***Mean Value Theorem for Multivariate Functions*** *Let $U \subseteq \mathbb{R}^n$ and let $a, b \in U$ be such that the straight line connecting them lives entirely within U. More precisely, the curve $\gamma : [0,1] \to \mathbb{R}^n$ given by $\gamma(t) = (1-t)a + tb$ satisfies $\gamma(t) \in U$ for all $t \in [0,1]$. If $f : U \to \mathbb{R}$ is a function such that $f \circ \gamma$ is continous on $[0,1]$ and differentiable on $(0,1)$, then there exists a $t_0 \in (0,1)$ such that $c = \gamma(t_0)$ and*

$$f(b) - f(a) = \nabla f(c) \cdot (b-a)$$

Corollary. *If $U \subseteq \mathbb{R}^n$ is convex and $f : U \to \mathbb{R}$ is a differentiable function such that $|\nabla f(x)| \leq M$ for all $x \in U$, then for every $a, b \in U$, we have*

$$|f(b) - f(a)| \leq M|b-a|$$

Corollary. *If $U \subseteq \mathbb{R}^n$ is convex and $f : U \to \mathbb{R}$ is a differentiable function such that $\nabla f(x) = 0$ for all $x \in U$, then f is a constant function on U*