Differentiability

Definitions

Definition 1. *$\boldsymbol{f}$: $\mathbb{R}^N \mapsto \mathbb{R}^M$ is differentiable at $\boldsymbol{x}_0$ if and only if there is a linear function $\boldsymbol{l}$: $\mathbb{R}^N \mapsto \mathbb{R}^M$ such that*

$$\lim_{\boldsymbol{x} \longrightarrow \boldsymbol{x}_0} \frac{\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0) - \boldsymbol{l}(\boldsymbol{x} - \boldsymbol{x}_0)\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0. \tag{1}$$

We denote $\boldsymbol{l}$ by $D f(\boldsymbol{x}_0)$, and call it the differential of $\boldsymbol{f}$ at $\boldsymbol{x}_0$.

Remark 2. The above is equivalent to

$$\lim_{\|\boldsymbol{x} - \boldsymbol{x}_0\| \longrightarrow 0} \frac{\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0) - \boldsymbol{l}(\boldsymbol{x} - \boldsymbol{x}_0)\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0. \tag{2}$$

Exercise 1. Prove that the definition is also equivalent to

$$\lim_{y \longrightarrow \boldsymbol{0}} \frac{\|\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{l}(\boldsymbol{y})\|}{\|\boldsymbol{y}\|} = 0. \tag{3}$$

Remark 3. The differentiability defined above, when we replace $\mathbb{R}^N$ be an abstract, possibly infinite dimensional space, is called "Frechét differentiability".

Example 4. Let $\boldsymbol{f}$: $\mathbb{R}^N \mapsto \mathbb{R}^M$ be linear. Then $\boldsymbol{f}$ is differentiable at every $\boldsymbol{x}_0 \in \mathbb{R}^N$, and $D\boldsymbol{f}(\boldsymbol{x}_0) = \boldsymbol{f}$.

Proof. Fix any $\boldsymbol{x}_0$. Take any $\boldsymbol{x} \in \mathbb{R}^N$. We have

$$\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0) - \boldsymbol{f}(\boldsymbol{x} - \boldsymbol{x}_0) = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0} \tag{4}$$

therefore

$$\lim_{\boldsymbol{x} \longrightarrow \boldsymbol{x}_0} \frac{\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0) - \boldsymbol{l}(\boldsymbol{x} - \boldsymbol{x}_0)\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \lim_{\boldsymbol{x} \longrightarrow \boldsymbol{x}_0} \frac{0}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0. \tag{5}$$

Thus ends the proof. □

Example 5. Let $f: \mathbb{R}^2 \mapsto \mathbb{R}$ be given by

$$f(x, y) := a\,x^2 + b\,x\,y + c\,y^2. \tag{6}$$

Prove that it is differentiable at $(1, 0)$ and find its differential there.

Proof. We need to find a linear transform $l(x, y)$ such that

$$\lim_{(x, y) \longrightarrow (0,0)} \frac{\|f(1 + x, y) - f(1, 0) - l(x, y)\|}{\|(x, y)\|} = 0. \tag{7}$$

By the representation theory of linear functions all we need to do is to find two numbers $l_1, l_2 \in \mathbb{R}$ such that

$$\lim_{(x, y) \longrightarrow (0,0)} \frac{\|f(1 + x, y) - f(1, 0) - (l_1\,x + l_2\,y)\|}{\|(x, y)\|} = 0. \tag{8}$$

Using the explicit formula for f, the above ratio reduces to

$$\frac{|[a\,(1+x)^2+b\,(1+x)\,y+c\,y^2]-a-(l_1\,x+l_2\,y)|}{(x^2+y^2)^{1/2}}. \tag{9}$$

Simplifying, we obtain

$$\frac{|2\,a\,x+a\,x^2+b\,y+b\,x\,y+c\,y^2-(l_1\,x+l_2\,y)|}{(x^2+y^2)^{1/2}}. \tag{10}$$

Now it is clear that we should take $l_1=2\,a$ and $l_2=b$. In the following we prove that

$$\lim_{(x,y)\longrightarrow(0,0)}\frac{|a\,x^2+b\,x\,y+c\,y^2|}{(x^2+y^2)^{1/2}}=0. \tag{11}$$

This is easy since we have

$$a\,x^2\leqslant|a|\,(x^2+y^2); \qquad b\,x\,y\leqslant\frac{|b|}{2}\,(x^2+y^2); \qquad c\,y^2\leqslant|c|\,(x^2+y^2) \tag{12}$$

which means

$$\frac{|a\,x^2+b\,x\,y+c\,y^2|}{(x^2+y^2)^{1/2}}\leqslant\left(|a|+|c|+\frac{|b|}{2}\right)(x^2+y^2)^{1/2}. \tag{13}$$

Now we clearly have

$$\lim_{(x,y)\longrightarrow(0,0)}\left(|a|+|c|+\frac{|b|}{2}\right)(x^2+y^2)^{1/2}=\lim_{(x,y)\longrightarrow(0,0)}0=0 \tag{14}$$

and then

$$\lim_{(x,y)\longrightarrow(0,0)}\frac{|a\,x^2+b\,x\,y+c\,y^2|}{(x^2+y^2)^{1/2}}=0 \tag{15}$$

follows from Squeeze Theorem.

Therefore f is differentiable at $(1,0)$ with derivative

$$D\,f(1,0)(x,y)=2\,a\,x+b\,y. \tag{16}$$

□

Remark 6. Note that

1. Checking differentiability by definition is surprisingly complicated for such a simple quadratic function;

2. If we simply fix y and treat $f(x,y)$ as a function of x along, its derivative at 1 would be $2\,a+b\,y$; On the other hand, fixing x and taking derivative of $f(x,y)$ as a function of y alone at 0 gives $b\,x$. Now evaluating them at $x=1$ and $y=0$, we obtain $2\,a$ and b. This is no coincidence!

Exercise 2. Prove that $f: \mathbb{R}^N \mapsto \mathbb{R}$ defined by

$$f(x_1,\ldots,x_N):=\sum_{i,j=1}^{N}a_{i\,j}\,x_i\,x_j \tag{17}$$

is differentiable at every $\boldsymbol{x}_0\in\mathbb{R}^N$ and calculate the derivative. What is the matrix representation of $D\,f(\boldsymbol{x}_0)$?

Theorem 7. *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be differentiable at $\boldsymbol{x}_0 \in \mathbb{R}^N$. Then $\boldsymbol{f}$ is continuous at $\boldsymbol{x}_0$.*

Exercise 3. Prove the above theorem.

Properties of the differential

Proposition 8. *If the differential exists, then it is unique.*

Exercise 4. Prove the above theorem.

Lemma 9. (Reduction to scalar functions) *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be written as $\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_N(\boldsymbol{x}) \end{pmatrix}$. Then $\boldsymbol{f}$ is differentiable at $\boldsymbol{x}_0 \in \mathbb{R}^N$ if and only if $f_1, \ldots, f_M$ are all differentiable at $\boldsymbol{x}_0$. Furthermore we have*

$$D\boldsymbol{f}(\boldsymbol{x}_0) = \begin{pmatrix} Df_1(\boldsymbol{x}_0) \\ \vdots \\ Df_N(\boldsymbol{x}_0) \end{pmatrix}. \tag{18}$$

Exercise 5. Prove the above theorem.

Theorem 10. (Arithmetics) *Let $a \in \mathbb{R}$. Let $f, g: \mathbb{R}^N \mapsto \mathbb{R}$ be differentiable at $\boldsymbol{x}_0 \in \mathbb{R}^N$. Then so are $f \pm g$, $a\, f$, $f\, g$ and f/g (as long as $g(\boldsymbol{x}_0) \neq 0$ in the last case), and furthermore*

$$D(f \pm g)(\boldsymbol{x}_0) = Df(\boldsymbol{x}_0) \pm Dg(\boldsymbol{x}_0); \qquad D(a\, f)(\boldsymbol{x}_0) = a\,(Df\,(\boldsymbol{x}_0)); \tag{19}$$

$$D(f\, g)(\boldsymbol{x}_0) = f(\boldsymbol{x}_0)\, Dg(\boldsymbol{x}_0) + g(\boldsymbol{x}_0)\, Df(\boldsymbol{x}_0); \tag{20}$$

$$D(f/g)(\boldsymbol{x}_0) = g(\boldsymbol{x}_0)^{-2}\, [g(\boldsymbol{x}_0)\, Df(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)\, Dg(\boldsymbol{x}_0)]. \tag{21}$$

Exercise 6. Prove the above lemma.

Exercise 7. Prove the corresponding results for $\boldsymbol{f}, \boldsymbol{g}: \mathbb{R}^N \mapsto \mathbb{R}^M$.

Remark 11. Note that thanks to Lemma 9, we haven't lost any generality here.

Theorem 12. (Chain rule) *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be differentiable at $\boldsymbol{x}_0 \in \mathbb{R}^N$ and $\boldsymbol{g}: \mathbb{R}^M \mapsto \mathbb{R}^K$ be differentiable at $\boldsymbol{y}_0 := \boldsymbol{f}(\boldsymbol{x}_0)$. Then the composite function $\boldsymbol{g} \circ \boldsymbol{f}$: $\mathbb{R}^N \mapsto \mathbb{R}^K$ is differentiable at $\boldsymbol{x}_0$ with derivative*

$$D\,(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}_0) = (D\boldsymbol{g}(\boldsymbol{y}_0)) \circ (D\boldsymbol{f}(\boldsymbol{x}_0)) \tag{22}$$

Proof. We need to show

$$\lim_{\boldsymbol{x} \longrightarrow \boldsymbol{x}_0} \frac{\|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) - \boldsymbol{g}(\boldsymbol{y}_0) - [D\boldsymbol{g}(\boldsymbol{y}_0) \circ D\boldsymbol{f}(\boldsymbol{x}_0)](\boldsymbol{x} - \boldsymbol{x}_0)\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0. \tag{23}$$

To make the presentation easier to understand, we denote the two linear transformations $D\boldsymbol{g}(\boldsymbol{y}_0)$ and $D\boldsymbol{f}(\boldsymbol{x}_0)$ by L_g, L_f. Now writing

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) - \boldsymbol{g}(\boldsymbol{y}_0) - L_g(L_f(\boldsymbol{x} - \boldsymbol{x}_0)) \;&=\; \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) - \boldsymbol{g}(\boldsymbol{y}_0) - L_g(\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y}_0) \\ &\quad + L_g(\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y}_0 - L_f(\boldsymbol{x} - \boldsymbol{x}_0)). \end{aligned} \tag{24}$$

By triangle inequality and property of limit all we need to show are

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\frac{\|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}))-\boldsymbol{g}(\boldsymbol{y}_0)-L_g(\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}=0 \tag{25}$$

and

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\frac{\|L_g(\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0-L_f(\boldsymbol{x}-\boldsymbol{x}_0))\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}=0. \tag{26}$$

- The first limit.

  We define the function

$$H(\boldsymbol{x}):=\begin{cases}0 & \boldsymbol{f}(\boldsymbol{x})=\boldsymbol{y}_0:=\boldsymbol{f}(\boldsymbol{x}_0)\\ \dfrac{\|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}))-\boldsymbol{g}(\boldsymbol{y}_0)-L_g(\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} & \boldsymbol{f}(\boldsymbol{x})\neq\boldsymbol{y}_0:=\boldsymbol{f}(\boldsymbol{x}_0)\end{cases}. \tag{27}$$

  Then one can write

$$H(\boldsymbol{x})=\frac{\|\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}))-\boldsymbol{g}(\boldsymbol{y}_0)-L_g(\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0)\|}{\|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0\|}\cdot\frac{\|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \tag{28}$$

  and prove $\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}H(\boldsymbol{x})=0$. (See exercise below).

- The second limit.

  Due to differentiability of $\boldsymbol{f}$ at $\boldsymbol{x}_0$ we have

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\frac{\|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0-L_f(\boldsymbol{x}-\boldsymbol{x}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}=0 \tag{29}$$

  which means the function:

$$\boldsymbol{F}(\boldsymbol{x}):=\frac{\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0-L_f(\boldsymbol{x}-\boldsymbol{x}_0)}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \tag{30}$$

  satisfies

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{0}. \tag{31}$$

  Now taking advantage of the continuity of linear functions, we have

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\frac{\|L_g(\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{y}_0-L_f(\boldsymbol{x}-\boldsymbol{x}_0))\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}=\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\|L_g(\boldsymbol{F}(\boldsymbol{x}))\|=\|L_g(\boldsymbol{0})\|=0. \tag{32}$$

Thus ends the proof. □

Exercise 8. Let $\boldsymbol{l}:\mathbb{R}^N\mapsto\mathbb{R}^M$ be a linear function. Prove that

$$\sup_{\boldsymbol{x}\in\mathbb{R}^N,\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{l}(\boldsymbol{x})\|}{\|\boldsymbol{x}\|}<+\infty. \tag{33}$$

(Hint: Use the matrix representation of $\boldsymbol{l}$).

Exercise 9. Let $\boldsymbol{f}:\mathbb{R}^N\mapsto\mathbb{R}^M$ be differentiable at $\boldsymbol{x}_0\in\mathbb{R}^N$. Prove that

$$\limsup_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\frac{\|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}<+\infty. \tag{34}$$

Remark 13. If A is the matrix representation of $D\boldsymbol{g}(\boldsymbol{y}_0)$ and B is the matrix representation of $D\boldsymbol{f}(\boldsymbol{x}_0)$, then the matrix representation of $D(\boldsymbol{g}\circ\boldsymbol{f})(\boldsymbol{x}_0)$ is $A\,B$.

Exercise 10. Let $u(x,t)=f(x-t)$ for differentiable functions $f\colon\mathbb{R}\mapsto\mathbb{R}$. Prove that u is differentiable, and furthermore u satisfies the following partial differential equation:

$$\frac{\partial u}{\partial x}+\frac{\partial u}{\partial t}=0. \tag{35}$$

Exercise 11. (Change of variables) Let $f(x,y)$ be differentiable. Define

$$u(r,\theta):=f(r\cos\theta, r\sin\theta). \tag{36}$$

Prove

$$\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2=\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2. \tag{37}$$

Question 14. *Critique the following statement:*

> *A function $\boldsymbol{f}\colon\mathbb{R}^N\mapsto\mathbb{R}^M$ is differentiable at $\boldsymbol{x}_0\in\mathbb{R}^N$ if and only if there is a linear function $\boldsymbol{l}\colon\mathbb{R}^N\mapsto\mathbb{R}^M$ such that for all other linear functions $\boldsymbol{L}\colon\mathbb{R}^N\mapsto\mathbb{R}^M$,*
>
> $$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{x}_0}\frac{\|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x}_0)-\boldsymbol{l}(\boldsymbol{x}-\boldsymbol{x}_0)\|}{\|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x}_0)-\boldsymbol{L}(\boldsymbol{x}-\boldsymbol{x}_0)\|}=0. \tag{38}$$

If you think it is true, provide a proof; If you think it is false, construct a counter-example.

Question 15. (Euler's Theorem) *$f\colon\mathbb{R}^N\mapsto\mathbb{R}$ is said to be homogeneous of degree m if*

$$f(t\,\boldsymbol{x})=t^m f(\boldsymbol{x}). \tag{39}$$

a) *Give examples of $f\colon\mathbb{R}^3\mapsto\mathbb{R}$ that are homogeneous of degrees $1,3,-1$.*

b) *Assume f is differentiable. Then it satisfies the following partial differential equation:*

$$x_1\frac{\partial f}{\partial x_1}+\cdots+x_N\frac{\partial f}{\partial x_N}=m\,f. \tag{40}$$

c) *Prove that if f is differentiable and satisfies the above equation, then f must be homogeneous.*