Differentiability and C^1 Theorem

Key Definitions and Results

Definition 1. In this course **the only norm** we use is the "**two norm**". That is, for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ the notation $\|\cdot\|$ always means

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Definition 2. Let $F = (f_1, \ldots, f_m)$ be a vector-valued function defined near $a \in \mathbb{R}^n$. F is said to be **differentiable at a** if there is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\|x-a\|\to 0} \frac{\|F(x) - F(a) - T(x-a)\|}{\|x-a\|} = 0.$$

Such linear transformation T is unique and denoted by $DF(a)$. We say that $\boldsymbol{F'(a) := [\frac{\partial f_i}{\partial x_j}(a)]_{m\times n}}$ **exists** if all its first order partial derivatives at a exist.

Definition 3. We say that $F : \mathbb{R}^n \to \mathbb{R}^m$ is $\boldsymbol{C^1}$ **at $a \in \mathbb{R}^n$** if:

(i) Near a all first order partial derivatives exist$^{(*)}$.

(ii) All first order partial derivatives of F are continuous at a.

Remark 4 (Procedure to Check Differentiability and Nondifferentiability).

- **Step 1.** Compute all $\dfrac{\partial f_i}{\partial x_j}$'s of $F = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ at $a \in \mathbb{R}^n$.
- **Step 2.** Construct the Jacobian matrix

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

and compute the limit $L = \displaystyle\lim_{x\to a} \frac{\|F(x) - F(a) - F'(a)(x-a)\|}{\|x-a\|}$.

Case 1. If L exists and is zero, then F is differentiable at a by definition of differentiability.

Case 2. If L does not exist, or it exists but is nonzero, then F is not differentiable at a. Since if F is differentiable at a, its derivatives must make L vanish by definition.

$^{(*)}$ More precisely, there is an $r > 0$ such that all first order partial derivatives exist on $B(a, r)$.

Theorem 5 (Cauchy-Schwarz Inequality). Let $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$, we have

$$\left|\sum_{i=1}^n a_i b_i\right|^2 \le \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right).$$

Theorem 6. $\|\cdot\|$ satisfies the following properties:

(i) **Positivity:** $\|x\| \ge 0$ for all $x \in \mathbb{R}^n$ and $\|x\| = 0 \implies x = 0$.

(ii) **Scaling Property:** For every $a \in \mathbb{R}$ and $x \in \mathbb{R}^n$, $\|ax\| = |a| \cdot \|x\|$.

(iii) **Triangle Inequality:** For every $x, y \in \mathbb{R}^n$, $\|x + y\| \le \|x\| + \|y\|$.

Theorem 7. Let $F = (f_1, \ldots, f_m)$ be defined near $a \in \mathbb{R}^n$. If F is differentiable at a, then:

(i) All first order partial derivatives at a exist, i.e., $F'(a)$ exists.

(ii) The matrix of $DF(a)$ w.r.t. usual bases is $F'(a) := [\frac{\partial f_i}{\partial x_j}(a)]_{m\times n}$.

Theorem 8. Let $F = (f_1, \ldots, f_m)$ be defined near a, then F is differentiable at a $\iff$ all its coordinate functions are differentiable at a.

Theorem 9 (C^1). If $F : \mathbb{R}^n \to \mathbb{R}^m$ is C^1 at a, then F is differentiable at a.

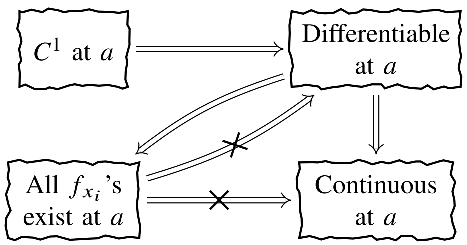
Theorem 10 (Chain Rule). Let $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^k$ be open balls and consider the composition $U \xrightarrow{G} V \xrightarrow{F} \mathbb{R}^n$. If G is differentiable at $a \in U$ and F is differentiable at $G(a)$, then $F \circ G$ is differentiable at a, moreover,

$$D(F \circ G)(a) = DF(G(a)) \circ DG(a),$$

or in matrix form (when domain and range are given the usual bases),

$$(F \circ G)'(a) = F'(G(a))G'(a).$$

Remark 11. We have the following diagram:



Example 1. For each of the following $f:\mathbb{R}^2\to\mathbb{R}$, find **(i)** $\lim_{(x,y)\to(0,0)} f(x,y)$, **(ii)** $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial y}(0,0)$. **(iii)** Determine if f_x and f_y are continuous at $(0,0)$.

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

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

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

Solution. (a) (i) If we let $y=kx$, then for $x\neq 0$ we have

$$f(x,kx)=\frac{k}{1+k^2},$$

therefore different paths produce different limits, thus $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

(ii) Since $f(x,0)=f(0,y)=0$ for $x,y\in\mathbb{R}$, we have

$$f_x(0,0)=\frac{d}{dx}f(x,0)\bigg|_{x=0}=\frac{d}{dx}0\bigg|_{x=0}=0$$

and

$$f_y(0,0)=\frac{d}{dy}f(0,y)\bigg|_{y=0}=\frac{d}{dy}0\bigg|_{y=0}=0.$$

(iii) For $(x,y)\neq(0,0)$, we have

$$f_x=\frac{y^3-x^2y}{(x^2+y^2)^2}\quad\text{and}\quad f_y=\frac{x^3-xy^2}{(x^2+y^2)^2},$$

and thus $f_x(0,y)=1/y\to\infty$ as $y\to 0$ and $f_y(x,0)=1/x\to\infty$ as $x\to 0$, so f_x,f_y are both not continuous at $(0,0)$.

(b) (i) Since for $(x,y)\neq(0,0)$ we have

$$|f(x,y)|=\left|\frac{x^2}{x^2+y^2}\right|\cdot|y|\le|y|\le\sqrt{x^2+y^2}=\|(x,y)\|,$$

thus $\lim_{(x,y)\to(0,0)} f(x,y)=0$.

(ii) We still have $f(x,0)=0$ and $f(0,y)=0$ for each $x,y\in\mathbb{R}$, therefore

$$f_x(0,0)=f_y(0,0)=0$$

for the same reason as in (a) (ii).

(iii) For $(x,y)\neq(0,0)$ we have

$$f_x=\frac{2xy^3}{(x^2+y^2)^2}\quad\text{and}\quad f_y=\frac{x^4-x^2y^2}{(x^2+y^2)^2}$$

and thus by polar coordinate $(x,y)=(r\cos\theta,r\sin\theta)$, where $\theta=\theta(r)$, we have

$$f_x=2\cos\theta\sin^3\theta\quad\text{and}\quad f_y=\cos^4\theta-\cos^2\theta\sin^2\theta$$

thus different choices of $\theta(r)$ (i.e., different paths) will product different limits, f_x,f_y are not continuous at $(0,0)$.

(c) (i) By polar coordinate we have $f=r^3\cos^2\theta\sin^2\theta$, therefore $f\to 0$ as $(x,y)\to(0,0)$.

(ii) We still have $f(x,0)=f(0,y)=0$ for each $x,y\in\mathbb{R}$, therefore

$$f_x(0,0)=f_y(0,0)=0$$

as in part (a) (ii).

(iii) For $(x,y)\neq(0,0)$ we have

$$f_x=\frac{x^3y^2+2xy^4}{(x^2+y^2)^{3/2}}\quad\text{and}\quad f_y=\frac{y^3x^2+2yx^4}{(x^2+y^2)^{3/2}}$$

both $f_x,f_y\to 0$ as $(x,y)\to(0,0)$, this is obvious by using polar coordinate system

Example 2. For each of the following $f:\mathbb{R}^2\to\mathbb{R}$, determine f_x and f_y at every $(a,b)\in\mathbb{R}^2$, where they exist.

(a) $f(x,y)=\cos|xy|$.

(b) $f(x,y)=x\sin|y|$.

Solution. (a) As cos is an even function, we have $f=\cos(xy)$, thus direct differentiation yields

$$f_x(a,b)=-b\sin(ab) \quad\text{and}\quad f_y(a,b)=-a\sin(ab).$$

(b) For every $(a,b)\in\mathbb{R}^2$ we have

$$f_x(a,b)=\frac{d}{dx}f(x,b)\bigg|_{x=a}=\frac{d}{dx}(x\sin|b|)\bigg|_{x=a}=\sin|b|$$

Now we compute $f_y(a,b)$.

Case 1. If $b>0$, then (x,y) near (a,b) implies $y>0$ implies $f(x,y)=x\sin y$, thus direct differentiation gives $f_y(a,b)=a\cos b$.

Case 2. If $b<0$, then similarly we have (x,y) near (a,b) implies $y<0$ implies $f=-x\sin y$, and also direct differentiation gives $f_y(a,b)=-a\cos b$.

Case 3. If $b=0$, then since

$$f_y(a,0)=\frac{d}{dy}f(a,y)\bigg|_{y=0}=\frac{d}{dy}a\sin|y|\bigg|_{y=0}=a\lim_{h\to0}\frac{\sin|h|}{h},$$

the limit exists only when $a=0$, therefore we conclude

$$f_y(0,0)=0.$$

Example 3. Let $\alpha>\frac{1}{2}$ and $f:\mathbb{R}^2\to\mathbb{R}$ be defined by

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

Prove that f is differentiable at $(0,0)$.

Solution. We need to find a linear map T such that

$$\frac{f(\vec{x})-f(\vec{0})-T\vec{x}}{\|\vec{x}\|}\to0 \quad\text{as } \vec{x}\to\vec{0}.$$

By the definition of differentiability the only candidate of such T is the linear map induced by the Jacobian matrix

$$f'(\vec{0})=\nabla f(\vec{0})=\begin{bmatrix}f_x(\vec{0}) & f_y(\vec{0})\end{bmatrix}.$$

Step 1. Since $f(x,0)=f(0,y)=0$ for every $x,y\in\mathbb{R}$, by direct computation we have

$$f_x(0,0)=\frac{d}{dx}f(x,0)\bigg|_{x=0}=\frac{d}{dx}0\bigg|_{x=0}=0.$$

And similarly $f_y(0,0)=0$.

Step 2. Since $f(0,0)=0$ and $T=f'(0,0)=\begin{bmatrix}0 & 0\end{bmatrix}$, we have

$$\lim_{\|(x,y)\|\to0}\frac{f(x,y)-f(0,0)-\begin{bmatrix}0 & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}}{\|(x,y)\|}=\lim_{\|(x,y)\|\to0}\frac{|xy|^\alpha\ln(x^2+y^2)}{\sqrt{x^2+y^2}}$$

By polar coordinate the latter expression becomes $2r^{2\alpha-1}\ln r|\cos\theta|^\alpha|\sin\theta|^\alpha$. As $2\alpha>1$ by the given condition, thus

$$\lim_{\|(x,y)\|\to0}\frac{f(x,y)-f(0,0)-\begin{bmatrix}0 & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}}{\|(x,y)\|}=0.$$

Exercise 1. Determine with proof if

$$f(x,y):=\begin{cases}\dfrac{\ln(1+x^4+y^4)}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)\neq(0,0)\end{cases}$$

is differentiable at $(0,0)$.

Exercise 2. Let $F:\mathbb{R}^n\to\mathbb{R}^n$ be a linear map defined by $F(x)=Ax$, for some matrix A. Show that F is differentiable on $\mathbb{R}^n$, moreover, $F'(x)=A$, for each $x\in\mathbb{R}^n$.

Example 4. Let $\alpha < 3/2$, prove that

$$f(x,y) = \begin{cases} \dfrac{x^4+y^4}{(x^2+y^2)^\alpha}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0) \end{cases}$$

is differentiable on $\mathbb{R}^2$ **by using C^1 theorem**.

Solution. By direct computation for every $(x,y) \neq (0,0)$,

$$f_x = -2\alpha \frac{x(x^4+y^4)}{(x^2+y^2)^{\alpha+1}} + \frac{4x^3}{(x^2+y^2)^\alpha}$$

and

$$f_x = -2\alpha \frac{y(x^4+y^4)}{(x^2+y^2)^{\alpha+1}} + \frac{4y^3}{(x^2+y^2)^\alpha},$$

therefore f is C^1 at every point $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$. It remains to check f is C^1 at $(0,0)$, for this, let's compute $f_x(0,0)$ and $f_y(0,0)$. Since $f(x,0) = x^{4-2\alpha}$, $f(0,y) = y^{4-2\alpha}$ and $2\alpha < 3$ implies $\boxed{4-2\alpha > 1}$, thus we have

$$f_x(0,0) = \frac{d}{dx} x^{4-2\alpha}\bigg|_{x=0} = 0 \quad \text{and} \quad f_y(0,0) = \frac{d}{dy} y^{4-2\alpha}\bigg|_{y=0} = 0.$$

It remains to check $f_x, f_y \to 0$ as $(x,y) \to (0,0)$, this follows easily by using polar coordinate.

In the following exercise we slightly loosen the hypothesis in C^1 theorem.

Exercise 3. Let $f(x,y)$ be defined near (x_0,y_0). Suppose:

(a) $f_x(x_0,y_0)$ exists;

(b) f_y exists near (x_0,y_0) and is continuous at (x_0,y_0).

Show that $f(x,y)$ is differentiable at (x_0,y_0). Try to extend this exercise to three or more variables.

Example 5. Let $F : \mathbb{R}^2 \to \mathbb{R}^3$ and $G : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$F(x,y) = (x,y,x^2y) \quad \text{and} \quad G(s,t) = (s+t, s^2-t^2).$$

What is $(F \circ G)'(2,1)$? What is $\big((F \circ G)(\underbrace{(0,1)+t(2,0)}_{:=\gamma(t)})\big)'(1)$?

Solution. By chain rule we have

$$\begin{aligned}
(F \circ G)'(2,1) &= F'(G(2,1))G'(2,1) \\
&= \begin{bmatrix} \text{——} \nabla x \text{——} \\ \text{——} \nabla y \text{——} \\ \text{——} \nabla x^2y \text{——} \end{bmatrix}(3,3) \begin{bmatrix} \text{——} \nabla(s+t) \text{——} \\ \text{——} \nabla(s^2-t^2) \text{——} \end{bmatrix}(2,1) \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2xy & x^2 \end{bmatrix}(3,3) \begin{bmatrix} 1 & 1 \\ 2s & -2t \end{bmatrix}(2,1) \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 18 & 9 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & -2 \\ 54 & 0 \end{bmatrix}.
\end{aligned}$$

Also we have

$$\begin{aligned}
(F \circ G \circ \gamma)'(1) &= (F \circ G)'((\gamma(1))\gamma'(1) \\
&= (F \circ G)'(2,1) \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & -2 \\ 54 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 108 \end{bmatrix}.
\end{aligned}$$

Recall that in page 137 of the lecture notes of Math3033 we have defined the norm of an $m \times n$ matrix $A = [a_{ij}]_{m\times n}$ by $\|A\| = \sum_{i,j} a_{ij}^2$ and shown the fact that $\|Ax\| \leq \|A\| \cdot \|x\|$.

Exercise 4 ("Mean Value Theorem"). Suppose $F : U \to \mathbb{R}^m$ is a function on some convex domain $U \subseteq \mathbb{R}^n$. Let $z \in \mathbb{R}^m$, show that for every $x, y \in U$, there is a point $c \in L := \{tx + (1-t)y : t \in (0,1)\}$ such that

$$z \cdot (F(x) - F(y)) = z \cdot F'(c)(x-y),$$

hence show that there is $\tilde{c} \in L$ such that

$$\|F(x) - F(y)\| \leq \|F'(\tilde{c})\| \cdot \|x-y\|.$$