DIFFERENTIABILITY

The real-valued case

Def: Let $\mathcal{U} \subset \mathbb{R}^n$ open, $f : \mathcal{U} \to \mathbb{R}$, $x \in \mathcal{U}$, $v \in \mathbb{R}^n$

The directional derivative of f at x along v is

$$\partial_v f(x) := \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \quad \text{whenever it exists}$$

Remark: since $\mathcal{U}$ is open, $x + tv \in \mathcal{U}$ for t "small enough".

Remark: if $\alpha \in \mathbb{R}$ then $\partial_{\alpha v} f(x) = \alpha \, \partial_v f(x)$.

Hence, if we know $\partial_v f(x)$ for some v, we know the directional derivatives for all the vectors with same direction.

Intuitively: by the above remark, we may assume that $\|v\| = 1$ and then $\partial_v f(x)$ is the instantaneous rate of change of f through x along the direction of v.

Ex: let $f(x,y) = x\cos(y) + y e^x$ and $v = (\cos\theta, \sin\theta)$

Then $\partial_v f(0,0) = \lim_{t \to 0} \frac{f((0,0) + t(\cos\theta, \sin\theta)) - f(0,0)}{t}$

$$= \lim_{t \to 0} \frac{f(t\cos\theta, t\sin\theta)}{t}$$

$$= \lim_{t \to 0} \cos\theta \cos(t\sin\theta) + \sin\theta \, e^{t\cos\theta}$$

$$= \cos(\theta) + \sin(\theta)$$

The highest rate of change through $(0,0)$ is along $\theta = \frac{\pi}{4}$ and the lowest at $\theta = \frac{5\pi}{4}$

Homework: plot the graph on MathSageCell.

Prop: $\mathcal{U} \subset \mathbb{R}^n$ open, $f, g : \mathcal{U} \to \mathbb{R}$, $x \in \mathcal{U}$, $v \in \mathbb{R}^n$, $c \in \mathbb{R}$

If $\partial_v f(x)$ and $\partial_v g(x)$ exist then $\partial_v (f+g)(x)$, $\partial_v (cf)(x)$, $\partial_v (fg)(x)$ exist and

1. $\partial_v (f+g)(x) = \partial_v f(x) + \partial_v g(x)$
2. $\partial_v (cf)(x) = c\, \partial_v f(x)$
3. $\partial_v (fg)(x) = g(x) \cdot \partial_v f(x) + f(x) \cdot \partial_v g(x)$ "Leibniz rule"

Δ ① & ②: obvious

③: $\lim\limits_{t \to 0} \dfrac{f(x+tv)\, g(x+tv) - f(x) g(x)}{t}$

$$= \lim_{t \to 0} \frac{f(x+tv)\, g(x+tv) - f(x)\, g(x+tv) + f(x) g(x+tv) - f(x) g(x)}{t}$$

$$= \lim_{t \to 0} g(x+tv) \frac{f(x+tv) - f(x)}{t} + f(x) \frac{g(x+tv) - g(x)}{t}$$

$$= g(x) \partial_v f(x) + f(x) \partial_v g(x)$$

Remark: $g(x+tv) = \dfrac{g(x+tv) - g(x)}{t} \cdot t + g(x) \xrightarrow[t \to 0]{} 0 + g(x)$ $\square$

Prop: $\mathcal{U} \subset \mathbb{R}^n$ open, $x \in \mathcal{U}$, $f : \mathcal{U} \to \mathbb{R}$, $h : I \to \mathbb{R}$, $v \in \mathbb{R}^n$

If $\partial_v f(x)$ exists and h differentiable at $f(x)$ then $\partial_v (h \circ f)(x)$ exists and $\partial_v (h \circ f)(x) = h'(f(x)) \cdot \partial_v f(x)$

Δ Later, but you can adapt the proof of the chain rule from MAT137 $\square$

Def: Let $U \subset \mathbb{R}^m$ open, $x \in U$, $i = 1, \dots, m$

We define the i-th partial derivative of f at x by

$$\frac{\partial f}{\partial x_i}(x) := \partial_{e_i} f(x) = \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t}$$

whenever it exists.

<u>In practice:</u> $f(x + te_i) = f(x_1, x_2, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_m)$

so $\frac{\partial f}{\partial x_i}(x) = g'(x_i)$ where $g(x_i) = f(x_1, \dots, x_i, \dots, x_m)$

where all the other variables are frozen

Ex: $f(x,y) = x^2 e^{xy}$

$\frac{\partial f}{\partial x}(x,y) = 2x e^{xy} + x^2 y e^{xy}$ $\qquad \frac{\partial f}{\partial x}(1,1) = 3e$

$\frac{\partial f}{\partial y}(x,y) = x^3 e^{xy}$ $\qquad \frac{\partial f}{\partial y}(1,1) = e$

Ex: $f(x,y) = |x|(1+y)$

$\frac{\partial f}{\partial x}(0,0)$ DNE : $\lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t \to 0} \frac{|t|}{t}$

$\frac{\partial f}{\partial y}(0,0) = 0$

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

$\frac{\partial f}{\partial x}(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$ exist

but $\partial_{(1,1)} f(0,0) = \lim_{t \to 0} \frac{f(t,t)}{t} = \lim_{t \to 0} \frac{1}{2t}$ DNE

⚠ So: the partial derivatives exist $\not\Rightarrow$ all the directional derivatives exist

Def: $U \subset \mathbb{R}^m$ open, $f: U \to \mathbb{R}$, $x \in U$, $\frac{\partial f}{\partial x_i}(x)$ exists $\forall i$

The **gradient of f at x** is

$$\nabla f(x) := \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \ldots, \frac{\partial f}{\partial x_m}(x)\right) \in \mathbb{R}^m$$

Ex: $f(x,y) = x\cos(y) + ye^x$

$$\nabla f(x,y) = \left(\cos(y) + ye^x, -x\sin(y) + e^x\right)$$

$$\nabla f(0,0) = (1,1)$$

Remark: We have already seen that $\underset{\nabla f(0,0)}{\overset{}{(1,1)}}$ $(\theta = \frac{\pi}{4})$ was the direction where f has a the highest rate of change through $(0,0)$.

We will see later that it is a general phenomenon

Def: $U \subset \mathbb{R}^m$ open, $f: U \to \mathbb{R}$, $x \in U$

We say that f is **differentiable** at x if there exists a linear function $d_x f: \mathbb{R}^m \to \mathbb{R}$ s.t.

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - d_x f(h)}{\|h\|} = 0$$

$d_x f(\lambda u + \mu v) = \lambda d_x f(u) + \mu d_x f(v)$

i.e. $f(x+h) = f(x) + d_x f(h) + E(h)$ where $\frac{E(h)}{\|h\|} \to 0$

"linear approximation of f at x"

Ex: ① $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, $f(x+h) = f(x) + 2xh + h^2$
so f is differentiable on $\mathbb{R}$ and $d_x f(h) = 2xh$

② $g: \mathbb{R}^m \to \mathbb{R}$, $g(x) = \|x\|^2$, $g(x+h) = \|x+h\|^2 = \|x\|^2 + 2(x \cdot h) + \|h\|^2$

so g is differentiable on $\mathbb{R}^m$ and $d_x g(h) = 2(x \cdot h)$

Theorem: If f is differentiable at x then its differential d_xf is unique

Δ Assume that f has 2 differentials at x: $\ell_1, \ell_2 : \mathbb{R}^m \longrightarrow \mathbb{R}$

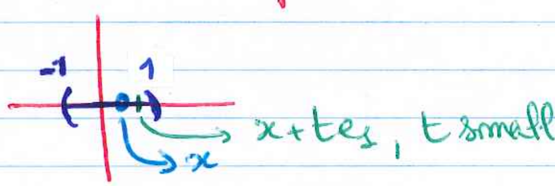
Let $\underset{\neq \vec{0}}{h} \in \mathbb{R}^m$. Since $\mathcal{U}$ is open, $x + th \in \mathcal{U}$ when $t > 0$ is small.

Then:
$$\ell_1(h) - \ell_2(h) = \frac{\ell_1(th) - \ell_2(th)}{t}$$
$$= -\frac{f(x+th) - f(x) - \ell_1(th)}{t\|h\|}\|h\| + \frac{f(x+th) - f(x) - \ell_2(th)}{t\|h\|}\|h\|$$
$$\xrightarrow[t \to 0^+]{} 0$$

Hence $\ell_1(h) = \ell_2(h)$ $\square$

⚠ What's why we want the domain to be open.
Otherwise the differential could not be unique.

Ex: $\mathcal{U} = (-1, 1) \times \{0\}$



We can determine $d_xf(e_1)$:

But we have no condition for $d_xf(e_2)$ since $x + te_2 \notin \mathcal{U}$ $\forall t \neq 0$
so we can take whatever we want.

Prop: $\mathcal{U} \subset \mathbb{R}^m$ open, $f : \mathcal{U} \to \mathbb{R}$, $x \in \mathcal{U}$

If f is differentiable at x then f is continuous at x.

Δ $f(v) = f(x + (v - x)) = f(x) + d_xf(v - x) + E(v - x) \xrightarrow[v \to x]{} f(x) + 0 + 0$ $\square$

Prop: $f, g : \mathcal{U} \longrightarrow \mathbb{R}$ differentiable at $x \in \mathcal{U}$, $\mathcal{U} \in \mathbb{R}^m$ open, then

1. $d_x(f+g) = d_xf + d_xg$
2. $d_x(\lambda f) = \lambda\, d_xf$, $\lambda \in \mathbb{R}$
3. $d_x(fg) = g(x)\, d_xf + f(x)\, d_xg$
4. $d_x(1/f) = -1/f(x)^2 \cdot d_xf$ if $f(x) \neq 0$

Theorem: $\mathcal{U} \subset \mathbb{R}^n$ open, $f: \mathcal{U} \to \mathbb{R}$, $x \in \mathcal{U}$

If f is differentiable at x, then

① All the directional derivatives of f at x exist

② $\forall v \in \mathbb{R}^n$, $\partial_v f(x) = d_x f(v)$

③ $\forall h \in \mathbb{R}^n$, $d_x f(h) = \nabla f(x) \cdot h = \frac{\partial f(x)}{\partial x_1} h_1 + \cdots + \frac{\partial f}{\partial x_n}(x) \cdot h_n$

④ $\forall v \in \mathbb{R}^n$, $\partial_v f(x) = \nabla f(x) \cdot v$

(that's why physicists write $df = \sum_i \frac{\partial f}{\partial x_i} dx_i$)

Δ ①+②:

$$\frac{f(x+tv) - f(x)}{t} = \frac{f(x+tv) - f(x) - d_x f(tv) + d_x f(tv)}{t}$$

$$= \frac{f(x+tv) - f(x) - d_x f(tv)}{t\|v\|} \|v\| + d_x f(v)$$

$$\xrightarrow[t \to 0]{} 0 + d_x f(v)$$

So $\partial_v f(x)$ exists and $\partial_v f(x) = d_x f(v)$

③ $d_x f(e_i) = \partial_{e_i} f(x) = \frac{\partial f}{\partial x_i}(x)$

hence $d_x f(h_1, \ldots, h_n) = d_x f(\sum_i h_i e_i) = \sum_i h_i \, d_x f(e_i) = \sum_i h_i \frac{\partial f}{\partial x}(e_i) = \nabla f(x) \cdot h$

④ $\partial_v f(x) = d_x f(v) = \nabla f(x) \cdot v$ $\square$

Remark: if f is differentiable at x then $d_x f(h) = \nabla f(x) \cdot h$, there is no other possibility

Ex: $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases}$ is not differentiable at $(0,0)$ since $\partial_{(1,1)} f(0,0)$ DNE

⚠ The converse of the above theorem is false:

$f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases}$: all the directional derivatives exist at $(0,0)$ but f is not differentiable at $(0,0)$

- $\partial_v f(0,0) = \lim_{t \to 0} \frac{t^3 v_1^3}{t^3 v_1^2 + t^3 v_2^2} = \frac{v_1^3}{v_1^2 + v_2^2}$
- Assume by contradiction that f is differentiable at 0, then
$d_0 f(1,1) = d_0 f(1,0) + d_0 f(0,1) = \partial_{(1,0)} f(0,0) + \partial_{(0,1)} f(0,0) = 1 + 0 = 1$
$= \partial_{(1,1)} f(0,0) = 1/2$

Remark: If $f: \mathcal{U} \to \mathbb{R}$ is differentiable at x and $\nabla f(x) \neq \vec{0}$ then the direction of $\nabla f(x)$ is the direction of fastest increase at x and the magnitode of $\nabla f(x)$ is the instantaneous rate of change in that direction.

Δ Let $v \in \mathbb{R}^m$ be a unit vector then

$$\partial_v f(x) = d_x f(v) = \nabla f(x) \cdot v = |v| \cdot |\nabla f(x)| \cdot \cos\theta = |\nabla f(x)| \cdot \cos\theta$$

the max is when $\cos\theta = 1$, ie $v = \dfrac{\nabla f(x)}{|\nabla f(x)|}$ and then $\partial_v f(x) = |\nabla f(x)|$ $\square$

Compare with the examples about $f(x,y) = x\cos(y) + ye^x$ above.

Theorem: $\mathcal{U} \subset \mathbb{R}^m$ open, $f: \mathcal{U} \longrightarrow \mathbb{R}$, $x \in \mathcal{U}$ at least in an open ball around x

If the partial derivatives of f exist on $\mathcal{U}$ and are continuous at x then f is differentiable at x

Δ ~~Idea~~: we apply the MVT to $t \mapsto f(t, x_2, \dots, x_m)$ on $[x_1, x_1+h_1]$

then $\exists\, \theta_1$ s.t. $f(x_1+h_1, x_2, \dots, x_m) - f(x_1, \dots, x_m) = h_1 \dfrac{\partial f}{\partial x_1}(x_1+\theta_1 h_1, x_2, \dots, x_m)$

By MVT to $t \mapsto f(x_1+h_1, t, x_3, \dots, x_m)$ on $[x_2, x_2+h_2]$

$\exists\, \theta_2$ s.t. $f(x_1+h_1, x_2+h_2, x_3, \dots, x_m) - f(x_1+h_1, x_2, \dots, x_m) = h_2 \dfrac{\partial f}{\partial x_2}(x_1+h_1, x_2+\theta_2 h_2, \dots)$

And so on for $x_3, \dots, x_m$.

Then: $f(x+h) - f(x) - \sum_i h_i \dfrac{\partial f}{\partial x_i}(x) = \sum_i h_i \overbrace{\left(\dfrac{\partial f}{\partial x_i}(a_i) - \dfrac{\partial f}{\partial x_i}(x) \right)}^{\xrightarrow[h\to 0]{} 0 \text{ by continuity of } \frac{\partial f}{\partial x_i}}$

where $a_i = (x_1+h_1, \dots, x_i + h_i\theta_i, x_{i+1}, \dots, x_m) \xrightarrow[h\to 0]{} x$ $\square$

The converse is false: $f(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$

is differentiable on $\mathbb{R}$ but f' is not continuous at 0

Summary: $f : \mathcal{U} \to \mathbb{R}$, $\mathcal{U} \subset \mathbb{R}^n$ open

Name	Nature	Notation
Directional derivative at $x \in \mathcal{U}$ along $v \in \mathbb{R}^n$	Real number	$\partial_v f(x)$
Partial derivative at x	Real number	$\frac{\partial f}{\partial x_i}(x)$
Gradient at x	Vector of $\mathbb{R}^n$	$\nabla f(x)$
Differential at x	Linear function $\mathbb{R}^n \to \mathbb{R}$	$d_x f(h)$

Partial derivatives exist on $\mathcal{U}$ and are continuous at x $\Rightarrow$ Differentiable at x (converse false: ⓐ)

Differentiable at x $\Rightarrow$ Continuous at x (converse false: ⓑ)

Differentiable at x $\Rightarrow$ Directional derivatives exist at x and (converse false: ⓒ)

① $\partial_v f(x) = d_x f(v)$

② $d_x f(h) = \nabla f(x) \cdot h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)\, h_i$

③ $\partial_v f(x) = \nabla f(x) \cdot v$

useful to prove that a function is not differentiable. See ⓒ

Counter-examples:

ⓐ $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

ⓑ $f(x) = |x|$

ⓒ $f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{otherwise} \end{cases}$ on $\mathbb{R}^2$

Partial derivatives exist $\not\Rightarrow$ Directional derivatives exist

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