Math 2010 Week 7

Last time:

Differentiability is defined in terms of linear approximation and error

<u>Ex</u> Show that a linear polynomial

$$f(\vec{x}) = c + b_1x_1 + \cdots + b_nx_n$$

is differentiable on $\mathbb{R}^n$ from definition.

<u>Rmk</u> ① $\frac{\partial f}{\partial x_i}(\vec{x}) = b_i \quad \forall \vec{x} \in \mathbb{R}^n$

② The linearization of $f(\vec{x})$ at any $a \in \mathbb{R}^n$ is

$$L(\vec{x}) = f(\vec{x})$$

<u>Thm</u> If $f, g : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ are differentiable at $\vec{a} \in \Omega$. Then

① $f(\vec{x}) \pm g(\vec{x})$, $cf(\vec{x})$, $f(\vec{x})g(\vec{x})$ are differentiable at $\vec{a}$

② $\dfrac{f(\vec{x})}{g(\vec{x})}$ is differentiable at $\vec{a}$ if $g(\vec{a}) \neq 0$.

③ (Special case of Chain Rule)
Let $h(x)$ be a one-variable function and is differentiable at $f(\vec{a})$
Then $h \circ f$ is differentiable at $\vec{a}$

$$\vec{a} \overset{f}{\longmapsto} f(\vec{a}) \overset{h}{\longmapsto} h \circ f(\vec{a})$$

<u>Rmk</u> We will discuss general case of chain rule later.

Proof of ①, ②, ③ are similar to those for one variable. (MATH 2050)

The results above give many examples of differentiable functions:

- constant functions $f(\vec{x}) = c$
- coordinate functions $f(\vec{x}) = x_i$
- Polynomials (Sum of products of x_i)

  eg. $4x^3y^2 + xy^2 - xyz + z^2$ (deg 5)
- Rational functions (Quotient of polynomials)

  eg. $\dfrac{x^3y + z}{x^2+y^2+z^2+1}$

- If $f(\vec{x})$ is differentiable, then the followings are differentiable:

$e^{f(\vec{x})}$, $\sin(f(\vec{x}))$, $\cos(f(\vec{x}))$

$\ln(f(\vec{x}))$ where $f(\vec{x}) > 0$

$|f(\vec{x})|$ where $f(\vec{x}) \neq 0$

$\sqrt{f(\vec{x})}$ where $f(\vec{x}) > 0$

$\ln|f(\vec{x})|$ where $f(\vec{x}) \neq 0$

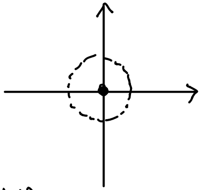
eg. $$\frac{e^{\sqrt{4+\sin(x^2+xy)}}}{\ln(1+\cos|x^2y|)}$$

Another way to check differentiability (Au 3.5.2)

Thm Let $\Omega \subseteq \mathbb{R}^n$ be open, f be C^1 on Ω
then f is differentiable on Ω

Rmk The assumption requires all $\frac{\partial f}{\partial x_i}$ exist on an open set, not just at a single point $\vec{a}$

eg. $f: \mathbb{R}^2 \to \mathbb{R}$



f_x, f_y exist and are continuous on a small open ball $B_\varepsilon(0,0)$

$\Rightarrow f$ is differentiable on $B_\varepsilon(0,0)$

$\Rightarrow f$ is differentiable at $(0,0)$

The theorem provides a simply way to verify differentiability if all $\frac{\partial f}{\partial x_i}$ can be easily showed to be continuous

eg $f(x,y,z) = xe^{x+y} - \log(x+z)$

Domain of f $= \{(x,y,z) \in \mathbb{R}^3 : x+z > 0\}$ is open

$$\frac{\partial f}{\partial x} = e^{x+y} + xe^{x+y} - \frac{1}{x+z}$$

$$\frac{\partial f}{\partial y} = xe^{x+y}$$

$$\frac{\partial f}{\partial z} = -\frac{1}{x+z}$$

All are continuous on domain of f

$\therefore f$ is C^1

$\Rightarrow f$ is differentiable

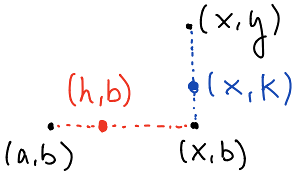
Pf of thm ($C^1 \Rightarrow$ differentiability on Ω)

We prove it for 2-variable $f(x,y)$. Similar proof for more variables (Ex)

Suppose $(a,b) \in \Omega$ and $B_\delta(a,b) \subseteq \Omega$. For $(x,y) \in B_\delta(a,b)$,

$$f(x,y) - f(a,b) = f(x,y) - f(x,b) + f(x,b) - f(a,b) \qquad (MVT)$$

$$= f_y(x,k)(y-b) + f_x(h,b)(x-a) \text{ for some } k \text{ between } b,y \text{ and } h \text{ between } a,x$$



$$\left|\frac{\varepsilon(x,y)}{\|(x,y)-(a,b)\|}\right| = \left|\frac{f(x,y) - f(a,b) - f_x(a,b)(x-a) - f_y(a,b)(y-b)}{\sqrt{(x-a)^2+(y-b)^2}}\right|$$

$$= \left|\frac{[f_y(x,k) - f_y(a,b)](y-b) + [f_x(h,b) - f_x(a,b)](x-a)}{\sqrt{(x-a)^2+(y-b)^2}}\right|$$

$$\leq \left|\frac{[f_y(x,k) - f_y(a,b)](y-b)}{\sqrt{(x-a)^2+(y-b)^2}}\right| + \left|\frac{[f_x(h,b) - f_x(a,b)](x-a)}{\sqrt{(x-a)^2+(y-b)^2}}\right| \quad \text{by triangle inequality}$$

$$\leq |f_y(x,k) - f_y(a,b)| + |f_x(h,b) - f_x(a,b)|$$

Take $(x,y) \to (a,b)$, then $(x,k), (h,b) \to (a,b) \Rightarrow$ R.H.S $\to 0$ by continuity of f_x, f_y

By sandwich theorem, $\displaystyle\lim_{(x,y)\to(a,b)} \left|\frac{\varepsilon(x,y)}{\|(x,y)-(a,b)\|}\right| = 0 \Rightarrow f$ is differentiable at (a,b)

Gradient and Directional derivative

Defn Let $\Omega \subseteq \mathbb{R}^n$ be open, $\vec{a} \in \Omega$, $f : \Omega \to \mathbb{R}$. Define the gradient vector of f at $\vec{a}$ to be

$$\vec{\nabla} f(\vec{a}) = \left(\frac{\partial f}{\partial x_1}(\vec{a}), \cdots, \frac{\partial f}{\partial x_n}(\vec{a}) \right)$$

eg $f(x,y) = x^2 + 2xy$

$$\vec{\nabla} f(x,y) = (f_x, f_y) = (2x+2y, 2x)$$

$$\vec{\nabla} f(1,2) = (6, 2)$$

Rmk Using $\vec{\nabla} f$, linearization of f at $\vec{a}$ can be expressed as

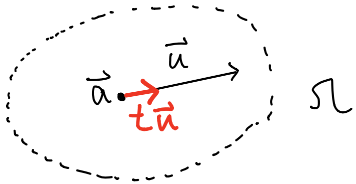
$$L(\vec{x}) = f(\vec{a}) + \sum \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i)$$

$$= f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot (\vec{x} - \vec{a})$$

Defn Let $\Omega \subseteq \mathbb{R}^n$ be open, $\vec{a} \in \Omega$, $f : \Omega \to \mathbb{R}$. Let $\vec{u} \in \mathbb{R}^n$ be an unit vector (i.e. $\|\vec{u}\| = 1$). Define the directional derivative of f in the direction of $\vec{u}$ at $\vec{a}$ to be

$$D_{\vec{u}} f(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$$

= the rate of change of f in the direction of $\vec{u}$ at the point $\vec{a}$



Rmk Let $e_i = (0, \cdots, 0, 1, 0, \cdots 0) \in \mathbb{R}^n$

i-th term

Then $D_{e_i} f(\vec{a}) = \frac{\partial f}{\partial x_i}(\vec{a})$

eg $e_2 = (0,1) \in \mathbb{R}^2$

$$D_{e_2} f(a,b) = \lim_{t \to 0} \frac{f((a,b)+te_2) - f(a,b)}{t}$$

$$= \lim_{t \to 0} \frac{f(a,b+t) - f(a,b)}{t}$$

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

Thm Suppose f is differentiable at $\vec{a}$

Let $\vec{u} \in \mathbb{R}^n$ be a unit vector. Then

$$D_{\vec{u}} f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \vec{u}$$

Recall that if $\vec{v} \neq \vec{0} \in \mathbb{R}^n$, then

the direction of $\vec{v}$ is defined to be

the unit vector $\frac{\vec{v}}{\|\vec{v}\|}$

eg Let $f(x,y) = \sin^{-1}\left(\frac{x}{y}\right)$

Find the rate of change of f at $(1,\sqrt{2})$

in the direction of $\vec{v} = (1,-1)$

Sol Let $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$

Recall: $(\sin^{-1} z)' = \frac{1}{\sqrt{1-z^2}}$

$$\frac{\partial f}{\partial x} = \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}} \cdot \frac{1}{y} \qquad \frac{\partial f}{\partial y} = \frac{1}{\sqrt{1-\left(\frac{x}{y}\right)^2}} \cdot \frac{-x}{y^2}$$

Note f, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ are continuous near $(1,\sqrt{2})$

$\Rightarrow f$ is C^1 near $(1,\sqrt{2})$

$\Rightarrow f$ is differentiable at $(1,\sqrt{2})$

$\therefore$ Answer $= D_{\vec{u}} f(1,\sqrt{2})$

$= \vec{\nabla} f(1,\sqrt{2}) \cdot \vec{u}$

$$= \left(\frac{\partial f}{\partial x}(1,\sqrt{2}), \frac{\partial f}{\partial y}(1,\sqrt{2}) \right) \cdot \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

$$= \left(1, -\frac{1}{\sqrt{2}} \right) \cdot \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

$$= \frac{1}{\sqrt{2}} + \frac{1}{2}$$

<u>Pf</u> (differentiable $\Rightarrow D_{\vec{u}} f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \vec{u}$)

Let $L(\vec{x})$ be linearization of $f(\vec{x})$ at $\vec{a}$

$$f(\vec{x}) = L(\vec{x}) + \varepsilon(\vec{x})$$

$$= f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot (\vec{x} - \vec{a}) + \varepsilon(\vec{x})$$

Put $\vec{x} = \vec{a} + t\vec{u}$:

$$f(\vec{a} + t\vec{u}) = f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot (t\vec{u}) + \varepsilon(\vec{a} + t\vec{u})$$

$$D_{\vec{u}} f(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$$

$$= \lim_{t \to 0} \frac{\vec{\nabla} f(\vec{a}) \cdot (t\vec{u}) + \varepsilon(\vec{a} + t\vec{u})}{t}$$

$$= \vec{\nabla} f(\vec{a}) \cdot \vec{u} + \lim_{t \to 0} \frac{\varepsilon(\vec{a} + t\vec{u})}{t}$$

Differentiability of f at $\vec{a}$

$$\Rightarrow \lim_{t \to 0} \frac{|\varepsilon(a + t\vec{u})|}{\|\vec{a} + t\vec{u} - \vec{a}\|} = 0$$

$$\Rightarrow \lim_{t \to 0} \left| \frac{\varepsilon(a + t\vec{u})}{t} \right| = 0$$

By Sandwich theorem,

$$D_{\vec{u}} f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \vec{u} + 0 = \nabla f(\vec{a}) \cdot \vec{u}$$

Geometric Meanings of $\vec{\nabla} f(\vec{a})$

If f is differentiable at $\vec{a}$, $\|\vec{u}\| = 1$,

$$D_{\vec{u}} f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \vec{u}$$

By Cauchy-Schwarz

$$|\vec{\nabla} f(\vec{a}) \cdot \vec{u}| \leq \|\vec{\nabla} f(\vec{a})\| \|\vec{u}\| = \|\vec{\nabla} f(\vec{a})\|$$

Also, if $\vec{\nabla} f(\vec{a}) \neq \vec{0}$, then

$$-\|\vec{\nabla} f(\vec{a})\| \leq \vec{\nabla} f(\vec{a}) \cdot \vec{u} \leq \|\vec{\nabla} f(\vec{a})\|$$

"=" $\Leftrightarrow$ $\vec{\nabla} f(\vec{a}) = k\vec{u}$ for some $k < 0$

"=" $\Leftrightarrow$ $\vec{\nabla} f(\vec{a}) = k\vec{u}$ for some $k > 0$

> At $\vec{a}$, f increases (decreases) most rapidly in the direction of $\vec{\nabla} f(\vec{a})$ $(-\vec{\nabla} f(\vec{a}))$ at a rate of $\|\vec{\nabla} f(\vec{a})\|$

Properties of Gradient

If $f, g : \Omega \underset{\text{open}}{\subseteq} \mathbb{R}^n \longrightarrow \mathbb{R}$ are differentiable

c is a constant. Then

① $\vec{\nabla}(f+g) = \vec{\nabla} f + \vec{\nabla} g$

$\quad \vec{\nabla}(cf) = c\vec{\nabla} f$

② $\vec{\nabla}(fg) = g\vec{\nabla} f + f\vec{\nabla} g$

③ $\vec{\nabla}\left(\frac{f}{g}\right) = \dfrac{g\vec{\nabla} f - f\vec{\nabla} g}{g^2}$ if $g \neq 0$

Pf Follow easily from partial differentiations.

Rmk In definition of $D_{\vec{u}} f(\vec{a})$,

$\vec{u}$ is assumed to be a unit vector

It can also be generalized to $D_{\vec{v}} f(\vec{a})$

for any $\vec{v}$ (any length).

In that case

$$D_{\vec{v}} f(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a}+t\vec{v}) - f(\vec{a})}{t}$$

and $D_{\vec{v}} f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \vec{v}$

Note

$$D_{\vec{v}} f = \begin{cases} \|\vec{v}\| D_{\vec{u}} f & \text{if } \vec{v} \neq \vec{0},\ \vec{u} = \dfrac{\vec{v}}{\|\vec{v}\|} \\ 0 & \text{if } \vec{v} = \vec{0} \end{cases}$$

Total Differential (of a real-valued function)

Let $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\vec{a} \in \Omega$

Consider linearization at $\vec{a}$:

$$f(\vec{x}) = f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i) + \varepsilon(\vec{x})$$

Denote $\Delta f = f(\vec{x}) - f(\vec{a})$, $\Delta x_i = x_i - a_i$

Then $\Delta f \approx \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a}) \Delta x_i$

The approximation is good up to 1st order since

$$\lim_{\vec{x} \to \vec{a}} \frac{\varepsilon(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0$$

1st order

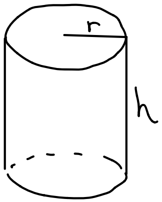
Classically, this 1st order approximated change is denoted by

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})\, dx_i$$

and is called the total differential of f at $\vec{a}$

Rmk In more advanced level, df and dx_i can be interpreted as linear maps from $\mathbb{R}^n$ to $\mathbb{R}$.

eg Let $V(r,h) = \pi r^2 h$



V = volume of cylinder

V is $C^1 \Rightarrow$ differentiable

$$dV = \frac{\partial V}{\partial r} dr + \frac{\partial V}{\partial h} dh$$

$$= 2\pi r h \, dr + \pi r^2 dh$$

For application:

Suppose we want to approximate change of V when (r,h) changes from $(r,h) = (3,12)$ to $(3+0.08, 12-0.3)$

Let $dr = \Delta r = 0.08$,

$dh = \Delta h = -0.3$

Then $\Delta V \approx dV$ ← approximated change

actual change

$$= 2\pi r h \, dr + \pi r^2 dh$$

$$= 2\pi(3)(12)(0.08) + \pi(3)^2(-0.3)$$

$$= 3.06\pi \approx 9.61$$

Properties of total differential

If $f, g : \underset{\text{open}}{\Omega \subseteq} \mathbb{R}^n \to \mathbb{R}$ are differentiable, c is a constant. Then

① $d(f+g) = df + dg$

$d(cf) = c\,df$

② $d(fg) = g\,df + f\,dg$

③ $d\left(\frac{f}{g}\right) = \frac{g\,df - f\,dg}{g^2}$

Pf Follow easily from partial differentiations.

Summary: Differentiating a real-valued function $f(\vec{x}) = f(x_1, \cdots, x_n)$ at $\vec{a} \in \mathbb{R}^n$

Different types of derivatives

- Directional derivative: $D_{\vec{u}} f(\vec{a}) = \lim_{t \to 0} \dfrac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t}$ for $\|\vec{u}\| = 1$
- Partial derivative: $\dfrac{\partial f}{\partial x_i}(\vec{a}) = D_{e_i} f(\vec{a})$ $e_i = (0, \cdots, 0, 1, 0 \cdots 0)$
- Gradient: $\vec{\nabla} f(\vec{a}) = \left(\dfrac{\partial f}{\partial x_1}(\vec{a}), \cdots, \dfrac{\partial f}{\partial x_n}(\vec{a}) \right)$
- Total differential: $df = \sum_{i=1}^{n} \dfrac{\partial f}{\partial x_i}(\vec{a})\, dx_i$
- Higher derivatives: eg. $\dfrac{\partial^2 f}{\partial x_1 \partial x_2} = f_{x_2 x_1}$

  f is C^k means f and all its partial derivatives up to order k exists and are continuous

Linear approximation of $f(\vec{x})$ near $\vec{a}$

- $L(\vec{x}) = f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot (\vec{x} - \vec{a})$
- $f(\vec{x}) = L(\vec{x}) + \varepsilon(\vec{x})$
- f is differentiable at $\vec{a}$ if $\lim_{\vec{x} \to \vec{a}} \dfrac{\varepsilon(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0 \Rightarrow df \approx \Delta f$

Relations among derivatives

① $C^\infty \Rightarrow \cdots \Rightarrow C^{k+1} \Rightarrow C^k \Rightarrow \cdots \Rightarrow C^1 \Rightarrow C^0$

② f is C^1 on an open set containing $\vec{a}$

$\Downarrow$

f is differentiable at $\vec{a}$

$\swarrow$ $D_{\vec{u}} f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \vec{u}$ $\searrow$

$D_{\vec{u}} f(\vec{a})$ exists for any unit vector $\vec{u} \in \mathbb{R}^n$ $\not\Rightarrow$ f is continuous at $\vec{a}$

$D_{\vec{u}} f(\vec{a})$ exists for any unit vector $\vec{u} \in \mathbb{R}^n$ $\not\Leftarrow$ f is continuous at $\vec{a}$

$\Downarrow$

$\dfrac{\partial f}{\partial x_i}(\vec{a})$ exists for $i = 1, \cdots, n$ $\not\Leftrightarrow$ f is continuous at $\vec{a}$

③ All the $\Rightarrow$ in the reverse direction are false. See next page for counter examples.

Verify the following (counter-) examples :

eg 1 $f: \mathbb{R} \to \mathbb{R}$

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

f is differentiable on $\mathbb{R}$ but

$f'(x)$ is not continuous at $x = 0$.

Similarly,

$g(x) = x^{2k-2} f(x)$ is k-time differentiable

but $g^{(k)}(x)$ is not continuous at $x = 0$.

Hence, k-time differentiable $\not\Longrightarrow$ C^k

In particular, $C^{k-1} \not\Longrightarrow C^k$

For multivariable, let $h(\vec{x}) = g(x_1)$

eg 2

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

$D_{\vec{u}} f(0,0)$ exists for any unit vector $\vec{u} \in \mathbb{R}^2$

but f is not continuous at $(0,0)$

eg 3 $f(x,y) = |x+y|$

f is continuous on $\mathbb{R}^2$ but $f_x(0,0)$, $f_y(0,0)$ DNE

eg 4 $f(x,y) = \sqrt{|xy|}$

$f_x(0,0)$, $f_y(0,0)$ exist

but $D_{\vec{u}} f(0,0)$ DNE for $\vec{u} \neq \pm \vec{e}_1, \pm \vec{e}_2$.