Remark: $D_{\vec{u}}f(\vec{a})$ can be defined for any vector $\vec{v}$, not necessary $\|\vec{v}\|=1$ and could be $\vec{0}$, by the same definition

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

One can show that

$$D_{\vec{v}}f(\vec{a}) = \begin{cases} \|\vec{v}\|\, D_{\frac{\vec{v}}{\|\vec{v}\|}}f(\vec{a}), & \text{if } \vec{v}\neq\vec{0} \\ 0, & \text{if } \vec{v}=\vec{0} \end{cases}$$

and that

$$D_{\vec{v}}f(\vec{a}) = \vec{\nabla}f(\vec{a})\cdot\vec{v} \qquad \text{if } f \text{ is differentiable at } \vec{a}$$

(not true in general, if f is not differentiable)
eg $f(x,y)=\sqrt{|xy|}$ at $(0,0)$

Properties of Gradient

If $\begin{cases} \bullet\ f,g:\Omega\to\mathbb{R}\ (\Omega\subset\mathbb{R}^n, \text{open}) \text{ are differentiable,} \\ \bullet\ c \text{ is a constant,} \end{cases}$

then

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

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

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

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

(Pf: Easily from properties of partial derivatives)

Total Differential (of real-valued function)

$f: \Omega \to \mathbb{R}$ ($\Omega \subseteq \mathbb{R}^n$, open) differentiable at $\vec{a} \in \Omega$.

Then linearization at $\vec{a}$:

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

Usually denote: $\Delta f = f(\vec{x}) - f(\vec{a})$

$\Delta x_i = x_i - a_i$

Then $\Delta f \simeq \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a}) \Delta x_i$ $\left(\text{provided } \lim_{\vec{x} \to \vec{a}} \frac{|\mathcal{E}(\vec{x})|}{\|\vec{x} - \vec{a}\|} = 0\right)$

Classically, this approximation is presented as

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

(thinking: "$\Delta f \to df$" "$\Delta x_i \to dx_i$")

Def: Let
- $f: \Omega \to \mathbb{R}$, ($\Omega \subseteq \mathbb{R}^n$, open)
- $\vec{a} \in \Omega$

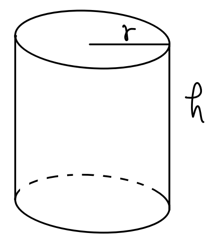
Suppose that f is <u>differentiable</u> on Ω. Then the <u>total differential</u> of f at $\vec{a}$ is defined to be the (formal) expression:

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

Remark: In the future, df and dx_i can be interpreted as a linear maps from $\mathbb{R}^n$ to $\mathbb{R}$.

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

(Volume of the 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 the change of V when (r,h) changes $(r,h) = (3, 12)$ to $(3+0.08, 12-0.3)$

Then let $dr = \Delta r = 0.08$

$dh = \Delta h = -0.3$

we have

$$\Delta V \approx dV = 2\pi r h \, dr + \pi r^2 dh \quad \left(\text{at } \begin{matrix} r=3, h=12 \\ dr, dh \text{ as above} \end{matrix}\right)$$

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

$$= 3.06\pi \approx 9.61.$$

Properties of Total Differential

If
- $f, g : \Omega \to \mathbb{R}$ ($\Omega \subseteq \mathbb{R}^n$, open) are differentiable, and
- $c \in \mathbb{R}$ is a constant.

Then

(1) $d(f \pm g) = df \pm dg$,

(2) $d(cf) = c\, df$

(3) $d(fg) = g\, df + f\, dg$

(4) $d\left(\frac{f}{g}\right) = \dfrac{g\, df - f\, dg}{g^2}$ provided $g \neq 0$

(Pf: Easily from properties of partial derivatives)

Summary (on differentiation of a real-valued function on $\mathbb{R}^n$)

$f: \mathbb{R}^n \to \mathbb{R}$

A. Types of differentiations (derivatives)

- Directional Derivative:

$$D_{\hat{u}} f(\vec{a}) = \lim_{t \to 0} \frac{f(\vec{a} + t\hat{u}) - f(\vec{a})}{t} \qquad (\|\hat{u}\| = 1)$$

- Partial derivatives:

$$\frac{\partial f}{\partial x_i}(\vec{a}) = D_{\vec{e}_i} f(\vec{a}), \qquad \vec{e}_i = (0, \cdots, \underset{\uparrow \text{ ith component}}{1}, \cdots 0)$$

- Gradient

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

- Total Differential

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

- Higher Derivatives

$$\frac{\partial^{k_1+\cdots+k_n} f}{\partial x_i^{k_1} \cdots \partial x_n^{k_n}}(\vec{a})$$

(provided f is C^k, $k = k_1 + \cdots + k_n$)

↑ all partial derivatives up to order k exist and continuous.

B. Linear approximation

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

- $f(\vec{x}) = L(\vec{x}) + \mathcal{E}(\vec{x})$

error term

- f is differentiable at $\vec{a}$

$$\Leftrightarrow \quad \lim_{\vec{x} \to \vec{a}} \frac{|\mathcal{E}(\vec{x})|}{\|\vec{x} - \vec{a}\|} = 0$$

In this case, $df \simeq \Delta f$ (by identifying $dx_i = \Delta x_i$)

C. <u>Relations among various concepts</u>

- $C^\infty \Rightarrow \cdots \Rightarrow C^{k+1} \Rightarrow C^k \Rightarrow \cdots \Rightarrow C^1 \Rightarrow C^0$ (No reverse implication)

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

$$\Downarrow \quad \not\Uparrow$$

f is <u>differentiable</u> at $\vec{a}$

$$\swarrow \not\nearrow \qquad\qquad\qquad\qquad \searrow \not\nwarrow$$

$D_{\vec{u}} f(\vec{a})$ exists $\forall \vec{u} \in \mathbb{R}^n$, $\|\vec{u}\| = 1$ $\quad \not\Rightarrow \quad \not\Leftarrow \quad$ f is continuous at $\vec{a}$

$$\searrow \not\nwarrow \qquad\qquad\qquad\qquad \not\swarrow \not\nearrow$$

$\frac{\partial f}{\partial x_i}(\vec{a})$ exists, $\forall i = 1, \cdots, n$

<u>Counter examples</u>:

<u>eg1</u>: $f: \mathbb{R} \to \mathbb{R}$ (in MATH2050)

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

f is <u>differentiable</u> on $\mathbb{R}$ but

$f'(x)$ is <u>not continuous</u> at $x = 0$

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

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

Hence k-time differentiable $\not\Rightarrow C^k$

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

(For multivariable: $h(\vec{x}) = h(x_1, \cdots, x_n) = g(x_1)$.)

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

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

but f is <u>not continuous</u> at $(0,0)$ (Check!)

eg 3: $f(x,y) = |x+y|$ is continuous on $\mathbb{R}^2$ but

$f_x(0,0)$, $f_y(0,0)$ DNE. (Check!)

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

$f_x(0,0)$, $f_y(0,0)$ exist (in fact $=0$)

but $D_{\vec{u}} f(0,0)$ DNE for $\vec{u} \neq \pm \vec{e_1}, \pm \vec{e_2}$.

(Check!)

Review: Matrix Multiplication

Let $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ be an $m \times n$-matric

$$= \begin{bmatrix} - \vec{a}_1 - \\ \vdots \\ - \vec{a}_m - \end{bmatrix} \quad \text{where } \vec{a}_i = (a_{i1}, \cdots a_{in}) \in \mathbb{R}^n$$

If

$b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix}$ be a $n \times 1$-matric regarded as a column vector in $\mathbb{R}^n$,

then (matrix multiplication)

$$Ab = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} - \vec{a}_1 - \\ \vdots \\ - \vec{a}_m - \end{bmatrix} \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} b_1 + \cdots + a_{1n} b_n \\ \vdots \\ a_{m1} b_1 + \cdots + a_{mn} b_n \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \cdot \vec{b} \\ \vdots \\ \vec{a}_m \cdot \vec{b} \end{bmatrix}$$

(result $= m \times 1$- matrix $=$ column m-vector)

Similarly, for multiplication of $(1 \times n)$ & $(n \times k)$ matrices

$$[- \vec{a} -] \begin{bmatrix} | & & | \\ \vec{b}_1 & \cdots & \vec{b}_k \\ | & & | \end{bmatrix} \qquad (\underset{\substack{\uparrow \\ \text{row vector}}}{\vec{a}}, \underbrace{\vec{b}_1, \cdots, \vec{b}_k}_{\text{column vectors}} \in \mathbb{R}^n)$$

$$= [\vec{a} \cdot \vec{b}_1, \cdots, \vec{a} \cdot \vec{b}_k]$$

(result $= 1 \times k$-matrix $=$ row k-vector)

In general: (m×n) times (n×k)

$$AB = \begin{bmatrix} - & \vec{a}_1 & - \\ & \vdots & \\ - & \vec{a}_m & - \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{b}_1 & \cdots & \vec{b}_k \\ | & & | \end{bmatrix} \qquad \Big(\underbrace{\vec{a}_1, \cdots, \vec{a}_m}_{\text{row vectors}}, \underbrace{\vec{b}_1, \cdots, \vec{b}_k}_{\text{column vectors}} \in \mathbb{R}^n \Big)$$

$$= \begin{bmatrix} \vec{a}_1 \cdot \vec{b}_1 & \cdots & \vec{a}_1 \cdot \vec{b}_k \\ \vdots & & \vdots \\ \vec{a}_m \cdot \vec{b}_1 & \cdots & \vec{a}_m \cdot \vec{b}_k \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ A\vec{b}_1 & \cdots & A\vec{b}_k \\ | & & | \end{bmatrix} \qquad \left(= A \begin{bmatrix} | & & | \\ \vec{b}_1 & \cdots & \vec{b}_k \\ | & & | \end{bmatrix} \right)$$

$$= \begin{bmatrix} - & \vec{a}_1 B & - \\ & \vdots & \\ - & \vec{a}_m B & - \end{bmatrix} \qquad \left(= \begin{bmatrix} - & \vec{a}_1 & - \\ & \vdots & \\ - & \vec{a}_m & - \end{bmatrix} B \right)$$

eg:

$$\underset{A}{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}} \underset{B}{\begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}} = \begin{bmatrix} 21 & 24 & 27 \\ 47 & 54 & 61 \end{bmatrix} \quad \text{(check!)}$$

$$A \begin{bmatrix} 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 21 \\ 47 \end{bmatrix}, \quad A \begin{bmatrix} 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 24 \\ 54 \end{bmatrix}, \quad A \begin{bmatrix} 7 \\ 10 \end{bmatrix} = \begin{bmatrix} 27 \\ 61 \end{bmatrix}$$

$$[1, 2] B = [21, 24, 27]$$

$$[3, 4] B = [47, 54, 61]$$

Differentiability of Vector-Valued Functions

$$\vec{f} : \Omega \to \mathbb{R}^m \quad , (\Omega \subset \mathbb{R}^n, \text{ open})$$

$$\vec{f}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$$

Suppose $\dfrac{\partial f_i}{\partial x_j}(\vec{a})$ exists for each $i = 1, \cdots, m$ & $j = 1, \cdots, n$.

$$f_i(\vec{x}) = f_i(\vec{a}) + \vec{\nabla} f_i(\vec{a}) \bullet (\vec{x} - \vec{a}) + \varepsilon_i(\vec{x}) \quad \text{———} \; (*)_i$$

$\Big((1\times 1) \qquad (1\times 1) \qquad \underset{\text{row}}{\underset{\uparrow}{(1\times n)}} \cdot \underset{\text{column}}{\underset{\uparrow}{(n\times 1)}} \qquad (1\times 1) \text{ matrix} \Big)$

Put all $(*)_i$, we have

$$\begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix} = \underbrace{\begin{bmatrix} f_1(\vec{a}) \\ \vdots \\ f_m(\vec{a}) \end{bmatrix} + \underbrace{\begin{bmatrix} - \vec{\nabla} f_1(\vec{a}) - \\ \vdots \\ - \vec{\nabla} f_m(\vec{a}) - \end{bmatrix}}_{\substack{m\times n \text{ matrix} \\ \text{of } \left[\frac{\partial f_i}{\partial x_j}\right]_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}}}} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}}_{\vec{L}(\vec{x})} + \underbrace{\begin{bmatrix} \varepsilon_1(\vec{x}) \\ \vdots \\ \varepsilon_m(\vec{x}) \end{bmatrix}}_{\text{Errors}}$$

In the following definitions,

$$\left\{\begin{array}{l}
\bullet\ \vec{f} : \Omega \to \mathbb{R}^m \quad (\Omega \subset \mathbb{R}^n, \text{ open}) \\[2ex]
\bullet\ \vec{f}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix} \quad (\text{in component form}) \\[2ex]
\bullet\ \vec{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \Omega \\[2ex]
\bullet\ \vec{x} - \vec{a} = \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}
\end{array}\right.$$

<u>Def</u> <u>Jacobian Matrix</u> of $\vec{f}$ at $\vec{a}$ is defined to be

$$D\vec{f}(\vec{a}) = \begin{bmatrix} - \vec{\nabla} f_1(\vec{a}) - \\ \vdots \\ - \vec{\nabla} f_m(\vec{a}) - \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{bmatrix}$$

(a $m \times n$-matrix)

<u>Def</u> <u>Linearization</u> of $\vec{f}$ at $\vec{a}$ is defined to be

$$\vec{L}(\vec{x}) = \vec{f}(\vec{a}) + D\vec{f}(\vec{a})(\vec{x} - \vec{a})$$

matrix multiplication

Def: $\vec{f}$ is said to be <u>differentiable</u> at $\vec{a} \in \Omega$,

if
- $\frac{\partial f_i}{\partial x_j}(\vec{a})$ exists $\forall\, i=1,\cdots,m$ & $j=1,\cdots,n$
- Error term of the linear approximation
$$\vec{\varepsilon}(\vec{x}) = \vec{f}(\vec{x}) - \vec{L}(\vec{x})$$
satisfies
$$\lim_{\vec{x}\to\vec{a}} \frac{\|\vec{\varepsilon}(\vec{x})\|}{\|\vec{x}-\vec{a}\|} = 0 .$$

<u>Remarks</u>

(1) $[D\vec{f}(\vec{a})]_{ij}$ (ij-entry of $D\vec{f}(\vec{a})$)
$$= \frac{\partial f_i}{\partial x_j}(\vec{a})$$

(2)
$$\underset{\substack{\uparrow\\ \text{column}\\ m\text{-vector}\\ \uparrow\\ m\times 1}}{\vec{f}(\vec{x})} = \underset{\substack{\uparrow\\ \text{column}\\ m\text{-vector}\\ \uparrow\\ m\times 1}}{\vec{f}(\vec{a})} + \underbrace{\underset{\substack{\uparrow\\ m\times n\\ \text{matrix}}}{D\vec{f}(\vec{a})}\,\underset{\substack{\uparrow\\ \text{column}\\ n\text{-vector}}}{(\vec{x}-\vec{a})}}_{\substack{\uparrow\\ (m\times n)\cdot(n\times 1)}} + \underset{\substack{\uparrow\\ \text{column}\\ m\text{-vector}\\ \uparrow\\ m\times 1}}{\vec{\varepsilon}(\vec{x})} \quad (\text{matrix})$$

(3) If f is real-valued ($m=1$), then
$$Df(\vec{a}) = \vec{\nabla} f(\vec{a}) \quad ((1\times n)\text{-matrix})$$

(4) $\|\vec{\varepsilon}(\vec{x})\|$ & $\|\vec{x}-\vec{a}\|$ are length in $\mathbb{R}^m$ & $\mathbb{R}^n$ respectively.

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

Hence

$$\boxed{\vec{f} \text{ is differentiable at } \vec{a} \iff f_i \text{ is differentiable at } \vec{a},\ \forall i=1,\cdots,m}$$

<u>Approximation:</u>

$$\vec{f}(\vec{x}) \simeq \vec{L}(\vec{x}) = \vec{f}(\vec{a}) + D\vec{f}(\vec{a})(\vec{x}-\vec{a})$$

$$\Rightarrow \quad \underbrace{\vec{f}(\vec{x})-\vec{f}(\vec{a})}_{\Delta\vec{f} = \text{change of } \vec{f}} \simeq \underbrace{D\vec{f}(\vec{a})}_{\text{Jacobian matrix}} \underbrace{(\vec{x}-\vec{a})}_{\Delta\vec{x} = \text{change in } \vec{x}}$$

<u>Notation</u>: $d\vec{f} = D\vec{f}(\vec{a})(\vec{x}-\vec{a})$ approximated change of f (total differential)

i.e. $\Delta\vec{f} \simeq d\vec{f}$

eg: $\vec{f}(x,y) = ((y+1)\ln x,\ x^2-\sin y+1)$

$$= \begin{pmatrix} (y+1)\ln x \\ x^2-\sin y+1 \end{pmatrix} = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix}$$

(Rewrite as column vector)

(1) Find $D\vec{f}(1,0)$

(2) Approximate $\vec{f}(0.9, 0.1)$

<u>Solu</u>:

$$D\vec{f}(x,y) = \begin{bmatrix} - & \vec{\nabla} f_1 & - \\ - & \vec{\nabla} f_2 & - \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{y+1}{x} & \ln x \\ 2x & -\cos y \end{bmatrix}$$

$$\therefore\ D\vec{f}(1,0) = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$$

(2) $$\vec{L}(x,y) = \vec{f}(1,0) + D\vec{f}(1,0) \cdot \begin{bmatrix} x-1 \\ y-0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix}$$

$$\vec{f}(0.9, 0.1) \simeq \vec{L}(0.9, 0.1)$$

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0.9-1 \\ 0.1 \end{bmatrix}}_{d\vec{f}}$$

$\Delta\vec{x} = d\vec{x}$

$$= \begin{bmatrix} -0.1 \\ 1.7 \end{bmatrix}$$