<u>Review</u> : 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

$= \begin{bmatrix} - \vec{a}_1 - \\ \vdots \\ - \vec{a}_m - \end{bmatrix}$ 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 $n \times 1$-matrix 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

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

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

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

In general: $(m\times n)$ times $(n\times k)$

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

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

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

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

<u>eg</u>:

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

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

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

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

<u>Differentiability of Vector-Valued Functions</u>

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

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

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

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

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

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,

- $\vec{f} : \Omega \to \mathbb{R}^m$ $\quad (\Omega \subset \mathbb{R}^n$, open$)$
- $\vec{f}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$ (in component form)
- $\vec{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \Omega$
- $\vec{x} - \vec{a} = \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}$

Def Jacobian Matrix 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)

Def Linearization 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 *differentiable* at $\vec{a} \in \Omega$,

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

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

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

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

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

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

Hence

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

Approximation:

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

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

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

i.e. $\quad \Delta\vec{f} \approx d\vec{f}$

(or $d\vec{f} = D\vec{f}(\vec{a})\, d\vec{x}$)

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

$$= \begin{pmatrix} (y+1)\ln x \\ x^2 - \sin y + 1 \end{pmatrix} = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} \quad \left(\text{Rewrite as column vector}\right)$$

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

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

Solu : (1) $D\vec{f}(x,y) = \begin{bmatrix} \frac{\partial}{\partial x}(y+1)\ln x & \frac{\partial}{\partial y}(y+1)\ln x \\ \frac{\partial}{\partial x}(x^2-\sin y+1) & \frac{\partial}{\partial y}(x^2-\sin y+1) \end{bmatrix} \left(= \begin{bmatrix} - \vec{\nabla} f_1 - \\ - \vec{\nabla} f_2 - \end{bmatrix} \right)$

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

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

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

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

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

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

$$= \begin{bmatrix} -0.1 \\ 1.7 \end{bmatrix} \quad \text{(check!)}$$

Chain Rule

Recall: 1-variable

$$\begin{cases} w = g(u) = 2u+1 \\ u = f(x) = x^2 \end{cases}$$

w can be regarded as a function x

$$w = g \circ f(x) = g(f(x)) = 2x^2+1$$

(Abuse of notation : $w = w(x)$ or $w = \underset{\overset{\shortparallel}{2x^2+1}}{g(x)}$)

Then $\dfrac{dw}{dx} = \dfrac{dw}{du}\dfrac{du}{dx}$ (usual way of writing)

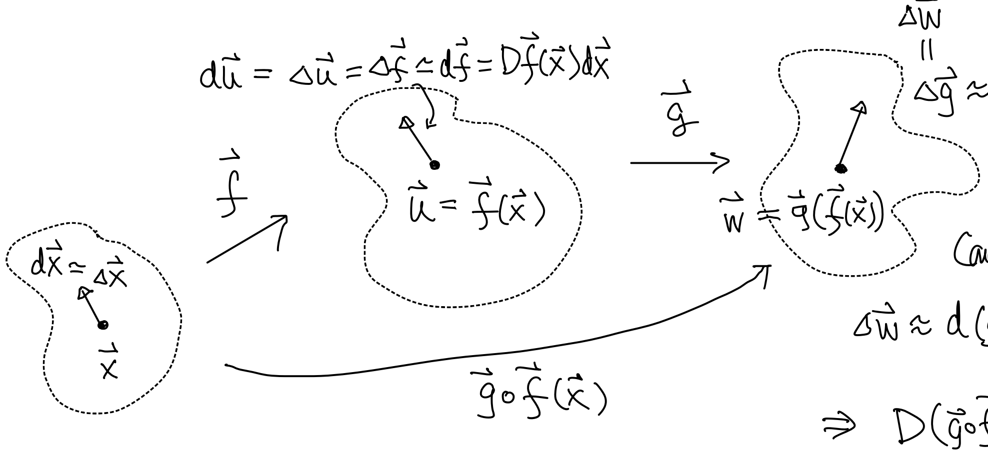
$$\left(\frac{dw}{dx}(x) = \frac{d(g\circ f)}{dx}(x) = \frac{dg}{du}(f(x)) \cdot \frac{df}{dx}(x) \right)$$

$$\frac{dw}{dx} = 2 \cdot 2x = 4x$$

<u>Caution</u>: Abuse of notation :

$\dfrac{dw}{dx}$ is $\dfrac{d(g\circ f)}{dx}(x)$, $\dfrac{dw}{du}$ is $\dfrac{dg}{du}(f(x))$ & $\dfrac{du}{dx}$ is $\dfrac{df}{dx}(x)$

<u>General dimensions</u>:



$$\Delta\vec{w}$$
$$\shortparallel$$
$$\Delta\vec{g} \approx d\vec{g} = D\vec{g}(\vec{u})\,d\vec{u}$$
$$= D\vec{g}(\vec{f}(\vec{x}))D\vec{f}(\vec{x})d\vec{x}$$

Compare to

$$\Delta\vec{w} \approx d(\vec{g}\circ\vec{f}) = D(\vec{g}\circ\vec{f})(\vec{x})d\vec{x}$$

$$\Rightarrow D(\vec{g}\circ\vec{f})(\vec{x}) = D\vec{g}(\vec{f}(\vec{x}))D\vec{f}(\vec{x})$$

Thm (**Chain Rule**)

Let
- $\vec{f} : \Omega_1 \to \mathbb{R}^n$ $\quad (\Omega_1 \subseteq \mathbb{R}^k$, open$)$
- $\vec{g} : \Omega_2 \to \mathbb{R}^m$ $\quad (\Omega_2 \subseteq \mathbb{R}^n$, open$)$
- $\vec{f}(\Omega_1) \subset \Omega_2$,

If
- $\vec{f}$ <u>differentiable</u> at $\vec{a} \in \Omega_1 \subset \mathbb{R}^k$
- $\vec{g}$ <u>differentiable</u> at $\vec{b} = \vec{f}(\vec{a}) \in \Omega_2 \subset \mathbb{R}^n$

Then $\vec{g} \circ \vec{f}$ is <u>differentiable</u> at $\vec{a}$, and

$$D(\vec{g} \circ \vec{f})(\vec{a}) = D\vec{g}(\vec{f}(\vec{a}))\, D\vec{f}(\vec{a})$$

matrix multiplication

<u>eg1</u>: $\vec{f} : \mathbb{R} \to \mathbb{R}^2$,

$$\vec{f}(\theta) = (\cos\theta, \sin\theta) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$$\vec{g} : \mathbb{R}^2 \to \mathbb{R}^2$$

$$\vec{g}(u,v) = \begin{bmatrix} 2uv \\ u^2 - v^2 \end{bmatrix}$$

Find $D(\vec{g} \circ \vec{f})(\theta)$. $\quad \begin{pmatrix} u = \cos\theta \\ v = \sin\theta \end{pmatrix}$

<u>Soln</u>: <u>Method 1</u>: Find composition explicitly

$$\vec{g} \circ \vec{f}(\theta) = \vec{g}(\cos\theta, \sin\theta) = \begin{bmatrix} 2\cos\theta\sin\theta \\ \cos^2\theta - \sin^2\theta \end{bmatrix} = \begin{bmatrix} \sin 2\theta \\ \cos 2\theta \end{bmatrix}$$

$$D(\vec{g} \circ \vec{f})(\theta) = \begin{bmatrix} \frac{d}{d\theta} \sin 2\theta \\ \frac{d}{d\theta} \cos 2\theta \end{bmatrix} = \begin{bmatrix} 2\cos 2\theta \\ -2\sin 2\theta \end{bmatrix}$$

Method 2 Chain Rule

$$\vec{D}f(\theta) = \begin{bmatrix} f_1' \\ f_2' \end{bmatrix} = \begin{bmatrix} \frac{d}{d\theta}\cos\theta \\ \frac{d}{d\theta}\sin\theta \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$D\vec{g}(u,v) = \begin{bmatrix} -\vec{\nabla} g_1 - \\ -\vec{\nabla} g_2 - \end{bmatrix} = \begin{bmatrix} -\vec{\nabla}(2uv) - \\ -\vec{\nabla}(u^2-v^2) - \end{bmatrix} = \begin{bmatrix} 2v & 2u \\ 2u & -2v \end{bmatrix}$$

By Chain Rule

$$D(\vec{g}\circ\vec{f})(\theta) = D\vec{g}(\vec{f}(\theta))\, D\vec{f}(\theta)$$

$$= D\vec{g}(\cos\theta, \sin\theta)\, D\vec{f}(\theta)$$

$$= \begin{bmatrix} 2\sin\theta & 2\cos\theta \\ 2\cos\theta & -2\sin\theta \end{bmatrix} \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} -2\sin^2\theta + 2\cos^2\theta \\ -2\sin\theta\cos\theta - 2\sin\theta\cos\theta \end{bmatrix} = \begin{bmatrix} 2\cos 2\theta \\ -2\sin 2\theta \end{bmatrix} \quad \divideontimes$$

eg2 (Abuse of notations)

$$f(x,y) = (x^2, 3xy, x+y^2) \quad (= \vec{f})$$

$$g(u,v,w) = \frac{uw}{v}$$

Consider $g\circ f$:

$$\begin{matrix} x \\ y \end{matrix} \xmapsto{f} \begin{matrix} (f_1 =)\ u \\ (f_2 =)\ v \\ (f_3 =)\ w \end{matrix} \xmapsto{g} g$$

Find $\frac{\partial g}{\partial x}(1,1)$ (regard g as a function of x,y)

Solu: (g is real-valued, $Dg = \vec{\nabla} g$)

$$Dg = \vec{\nabla} g = \left[\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}, \frac{\partial g}{\partial w} \right] = \left[\frac{w}{v}, -\frac{uw}{v^2}, \frac{u}{v} \right]$$

At $(1,1)$, $\begin{cases} u = x^2 = 1 \\ v = 3xy = 3 \\ w = x + y^2 = 2 \end{cases} \Rightarrow f(1,1) = (1,3,2)$

$$Dg(f(1,1)) = Dg(1,3,2) = \left[\frac{2}{3}, -\frac{2}{9}, \frac{1}{3} \right]$$

$$Df = \begin{bmatrix} \vec{\nabla} f_1 \\ \vec{\nabla} f_2 \\ \vec{\nabla} f_3 \end{bmatrix} = \begin{bmatrix} \vec{\nabla} x^2 \\ \vec{\nabla}(3xy) \\ \vec{\nabla}(x+y^2) \end{bmatrix} = \begin{bmatrix} 2x & 0 \\ 3y & 3x \\ 1 & 2y \end{bmatrix}$$

$$Df(1,1) = \begin{bmatrix} 2 & 0 \\ 3 & 3 \\ 1 & 2 \end{bmatrix}$$

Hence Chain rule $\Rightarrow$ $D(g \circ f)(1,1) = \left[\frac{2}{3}, -\frac{2}{9}, \frac{1}{3} \right] \begin{bmatrix} 2 & 0 \\ 3 & 3 \\ 1 & 2 \end{bmatrix}$

$= [1, 0]$

$\left[\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right]$

$\therefore \frac{\partial g}{\partial x}(1,1) = 1$.

<u>Remark</u>: We should just calculate the 1^{st} column

$$\begin{bmatrix} \frac{2}{3}, & -\frac{2}{9}, & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & * \\ 3 & * \\ 1 & * \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial g}{\partial u}, & \frac{\partial g}{\partial v}, & \frac{\partial g}{\partial w} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial f_1}{\partial x} & * \\ \frac{\partial f_2}{\partial x} & * \\ \frac{\partial f_3}{\partial x} & * \end{bmatrix}$$

$$\begin{pmatrix} u = f_1 \\ v = f_2 \\ w = f_3 \end{pmatrix}$$

$$\frac{\partial g}{\partial x} = \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial x}$$

In general, for 2-variables $\xrightarrow{\vec{f}}$ 3-variables $\xrightarrow{g}$ real-valued

i.e $\quad k=2, \quad n=3, \quad m=1$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \xrightarrow{\vec{f}} \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \\ f_3(x_1, x_2) \end{pmatrix} \xrightarrow{g} g(f_1, f_2, f_3)$$

we usually use classical notation,

(x, y) for (x_1, x_2),

$u = f_1(x,y)$, $v = f_2(x,y)$, $w = f_3(x,y)$, and

$g = g(u, v, w)$

$(x, y) \longmapsto (u, v, w) \longmapsto g$ can be regarded as function of (x,y) : $g = g(x,y)$

Abuse of notations

Then the Chain rule is

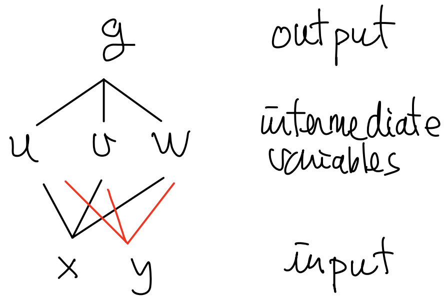
$$\boxed{\begin{aligned} \frac{\partial g}{\partial x} &= \frac{\partial g}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\cdot\frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}\cdot\frac{\partial w}{\partial x} \\ \frac{\partial g}{\partial y} &= \frac{\partial g}{\partial u}\cdot\frac{\partial u}{\partial y} + \frac{\partial g}{\partial v}\cdot\frac{\partial v}{\partial y} + \frac{\partial g}{\partial w}\cdot\frac{\partial w}{\partial y} \end{aligned}}$$

$$\begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} = D(g\circ\vec{f}) = Dg(\vec{f})\, D\vec{f} = \begin{bmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} & \frac{\partial g}{\partial w} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix}$$

(1×2 matric) $\qquad$ (1×3)(3×2) matric

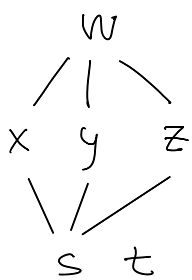
(Similarly for other low dimensional situations)

<u>Remark</u>: Branch Diagram (in Textbook)



<u>eg 3</u> $\quad w(x,y,z) = \sqrt{x^2+y^2+z^2}$

$$\begin{cases} x = 3e^t \sin s \\ y = 3e^t \cos s \\ z = 4e^t \end{cases}$$



Find $\frac{\partial w}{\partial s}$ at $s=t=0$.

Soln: $\dfrac{\partial W}{\partial S} = \dfrac{\partial W}{\partial x}\dfrac{\partial x}{\partial S} + \dfrac{\partial W}{\partial y}\dfrac{\partial y}{\partial S} + \dfrac{\partial W}{\partial z}\dfrac{\partial z}{\partial S}$

$$= \left(\frac{\partial}{\partial x}\sqrt{x^2+y^2+z^2}\right)\cdot\frac{\partial}{\partial s}(3e^t\sin s) + \left(\frac{\partial}{\partial y}\sqrt{x^2+y^2+z^2}\right)\frac{\partial}{\partial s}(3e^t\cos s)$$

$$+\left(\frac{\partial}{\partial z}\sqrt{x^2+y^2+z^2}\right)\frac{\partial}{\partial s}(4e^t)$$

$$= \frac{x}{\sqrt{x^2+y^2+z^2}}\,3e^t\cos s + \frac{y}{\sqrt{x^2+y^2+z^2}}(-3e^t\sin s) + 0$$

Put $s=t=0$ ($\Rightarrow$ $x=3e^t\sin s=0$, $y=3e^t\cos s=3$, $z=4$)

$$\frac{\partial W}{\partial S}(0,0) = 0+0 = 0 \quad . \quad ※$$

eg4. John is walking with position at time t given by

$$\begin{cases} x(t) = t^3+1 \\ y(t) = 2t^2 \end{cases}$$

Altitude is $H(x,y) = x^2-y^2+100$

(1) Is John going up/down at $t=1$?

(2) Which direction should he go instead at $t=1$ to go down most quickly?

Soln: (1) Find $\left.\dfrac{dH}{dt}\right|_{t=1}$ $\qquad t \begin{smallmatrix} < x > \\ < y > \end{smallmatrix} H$

Chain rule: $\dfrac{dH}{dt} = \dfrac{\partial H}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial H}{\partial y}\dfrac{dy}{dt}$

$$= 2x\cdot 3t^2 - 2y\cdot 4t$$

At $t=1$, $(x,y)=(2,2)$,

$$\therefore \left.\frac{dH}{dt}\right|_{t=1} = (2\cdot 2)\cdot(3\cdot 1^2) - 2(2)\cdot(4\cdot 1) = -4 < 0$$

$\therefore$ John is <u>going down</u> at $t=1$.

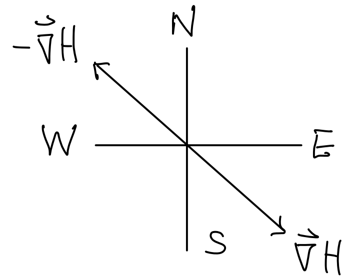
(2) go down most quickly in the direction $-\vec{\nabla}H$
(by the geometric interpretation of $\vec{\nabla}H$)

$\vec{\nabla}H = (2x, -2y)$ at $(x,y)=(2,2)$ (when $t=1$)

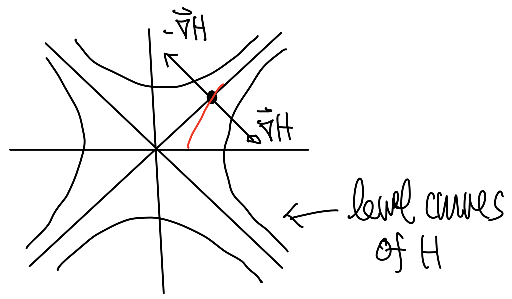
$= (4,-4)$

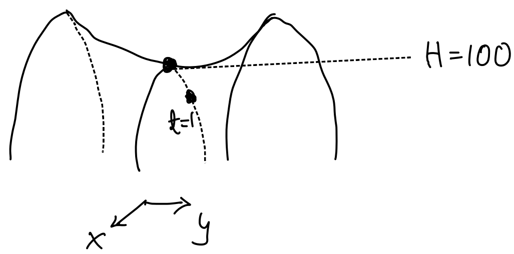
$\therefore$ H decreases most rapidly in the <u>direction of</u>

$-\vec{\nabla}H(2,2) = (-4,4)$



i.e. John should go NW
(North-west)





Idea of Pf of Chain Rule

$$\left\{\begin{array}{l}
\bullet\ \vec{f}: \Omega_1 \to \mathbb{R}^n \quad (\Omega_1 \subseteq \mathbb{R}^k, \text{ open}) \\
\bullet\ \vec{g}: \Omega_2 \to \mathbb{R}^m \quad (\Omega_2 \subseteq \mathbb{R}^n, \text{ open}) \\
\bullet\ \vec{f}(\Omega_1) \subset \Omega_2, \\
\bullet\ \vec{f} \text{ differentiable at } \vec{a} \in \Omega_1 \subset \mathbb{R}^k \\
\bullet\ \vec{g} \text{ differentiable at } \vec{b} = \vec{f}(\vec{a}) \in \Omega_2 \subset \mathbb{R}^n
\end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{ll}
\vec{f}(\vec{x}) - \vec{f}(\vec{a}) = D\vec{f}(\vec{a})(\vec{x}-\vec{a}) + \vec{\varepsilon}_{\vec{f}}(\vec{x}) & \text{———} (1) \\
\vec{g}(\vec{y}) - \vec{g}(\vec{b}) = D\vec{g}(\vec{b})(\vec{y}-\vec{b}) + \vec{\varepsilon}_{\vec{g}}(\vec{y}) & \text{———} (2)
\end{array}\right.$$

Put $\vec{y} = f(\vec{x})$ (and $\vec{b} = \vec{f}(\vec{a})$) in (2), we have

$$\vec{g}(\vec{f}(\vec{x})) - \vec{g}(\vec{b}) = D\vec{g}(\vec{b})(\vec{f}(\vec{x}) - \vec{f}(\vec{a})) + \vec{\varepsilon}_{\vec{g}}(\vec{f}(\vec{x}))$$

$$\begin{aligned}
(\text{by}(1)) \qquad &= D\vec{g}(\vec{b})\left(D\vec{f}(\vec{a})(\vec{x}-\vec{a}) + \vec{\varepsilon}_{\vec{f}}(\vec{x})\right) + \vec{\varepsilon}_{\vec{g}}(\vec{f}(\vec{x})) \\
&= D\vec{g}(\vec{b})D\vec{f}(\vec{a})(\vec{x}-\vec{a}) \\
&\quad + \left(D\vec{g}(\vec{b})\vec{\varepsilon}_{\vec{f}}(\vec{x}) + \vec{\varepsilon}_{\vec{g}}(\vec{f}(\vec{x}))\right)
\end{aligned}$$

(expecting $\vec{\varepsilon}_{\vec{g}\circ\vec{f}}(\vec{x}) = D\vec{g}(\vec{b})\vec{\varepsilon}_{\vec{f}}(\vec{x}) + \vec{\varepsilon}_{\vec{g}}(\vec{f}(\vec{x}))$)

So one need to show that

$$\lim_{\vec{x}\to\vec{a}} \frac{\| D\vec{g}(\vec{b})\vec{\varepsilon}_{\vec{f}}(\vec{x}) + \vec{\varepsilon}_{\vec{g}}(\vec{f}(\vec{x})) \|}{\|\vec{x}-\vec{a}\|} = 0$$

(Proof: Omitted
see MATH 2050
for 1-variable case)

in order to prove that $D(\vec{g}\circ\vec{f})(\vec{x}) = D\vec{g}(\vec{b})D\vec{f}(\vec{a})$. #