In the following definitions,

- $\vec{f}: \Omega \to \mathbb{R}^m$ ($\Omega \subset \mathbb{R}^n$, 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 <u>differentiable</u> at $\vec{a} \in \Omega$,

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

Remarks

(1) $[D\vec{f}(\vec{a})]_{ij}$ (ij-entry of $D\vec{f}(\bar{a})$)

$$= \frac{\partial f_i}{\partial x_j}(\vec{a})$$

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

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

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

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

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:

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

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

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

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

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

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

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

(Rewrite as column vector)

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

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

<u>Solu</u>: (1) $D\vec{f}(x,y) = \begin{bmatrix} \frac{\partial}{\partial x}[(y+1)\ln x] & \frac{\partial}{\partial y}[(y+1)\ln x] \\ \frac{\partial}{\partial x}[x^2-\sin y+1] & \frac{\partial}{\partial y}[x^2-\sin y+1] \end{bmatrix} = \begin{bmatrix} - \vec{\nabla} f_1 - \\ - \vec{\nabla} f_2 - \end{bmatrix}$

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

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

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

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

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

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

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0.9-1 \\ 0.1 \end{bmatrix}$$

$$= \begin{bmatrix} -0.1 \\ 1.7 \end{bmatrix}$$ (Check!)

※

<u>Chain Rule</u>

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 $w = \underset{\displaystyle 2x^2+1}{\underset{\shortparallel}{g(x)}}$)

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

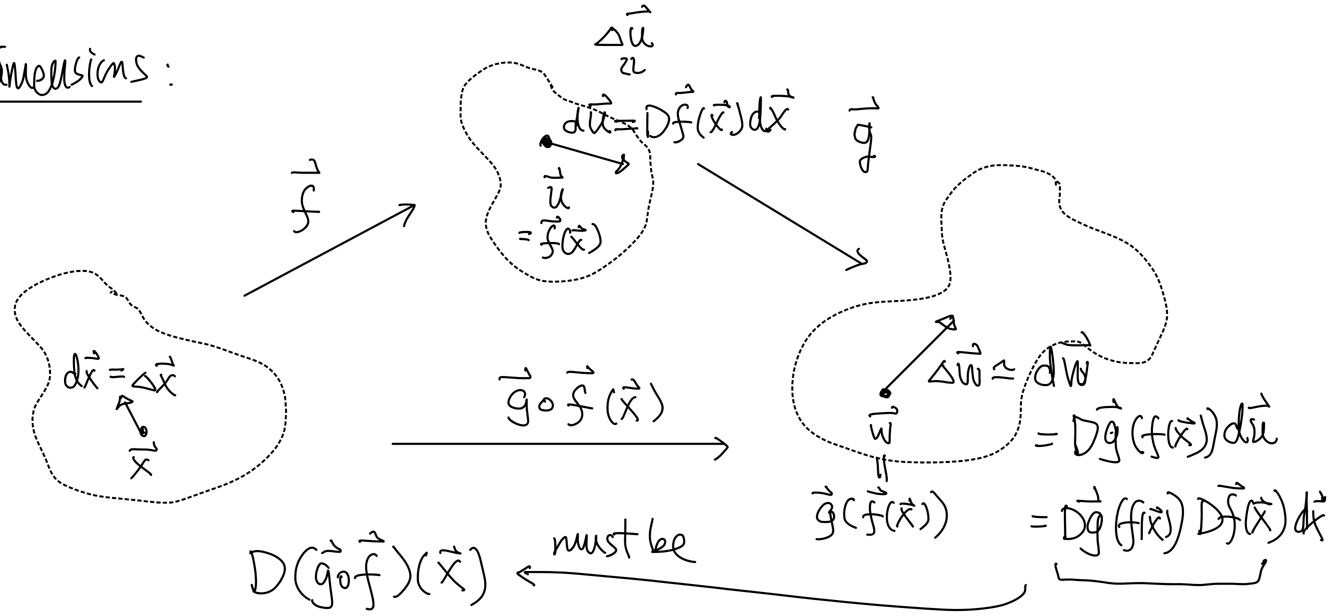
$$\left(\frac{dw}{dx}(x) = \frac{d(g\circ f)}{dx}(x) = \frac{dg}{du}(f(x)) \cdot \frac{df}{dx}(x) \right)$$

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

<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>:



Hence

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

Then $\vec{g} \circ \vec{f}$ is differentiable 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