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

$= \begin{bmatrix} - \vec{a}_1 - \\ \vdots \\ - \vec{a}_m - \end{bmatrix}$ 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 $(n\times k)$

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

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

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

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

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

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

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

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

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

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

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,

- $\vec{f} : \Omega \to \mathbb{R}^m$ ($\Omega \subset \mathbb{R}^n$, 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

<u>Def</u>: $\vec{f}$ 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

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

satisfies

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

<u>Remarks</u> (1) $\left[D\vec{f}(\vec{a})\right]_{ij}$ (ij-entry of $D\vec{f}(\vec{a})$)

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

(2)

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

(3) If f is real-valued ($m = 1$), then

$$Df(\vec{a}) = \vec{\nabla} f(\vec{a}) \quad ((1\times n)\text{-matrix})$$

(4) $\|\vec{E}(\vec{x})\|$ & $\|\vec{x} - \vec{a}\|$ 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:

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

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

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

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

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

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

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

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

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

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

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

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