$tr(a) = a$

$trA = trA^T$

$tr(A+B) = trA + trB$

$tr(aA) = a\,trA$

$$\nabla_A f(A) = \begin{bmatrix} \frac{\partial f}{\partial A_{11}} & \cdots & \frac{\partial f}{\partial A_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial A_{m1}} & \cdots & \frac{\partial f}{\partial A_{mn}} \end{bmatrix}$$

T1) $\nabla_A trAB = B^T$

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & & \\ a_{m1} & \cdots & & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & \ddots & & \vdots \\ \vdots & & & \\ b_{m1} & \cdots & & b_{mn} \end{bmatrix}$$

$\vec{a}_1, \ldots, \vec{a}_n$; $\vec{b}_1, \ldots, \vec{b}_n$

$$trAB = tr\begin{bmatrix} \vec{a}_1^T b_1 & \vec{a}_1 b_2 & \cdots & \vec{a}_1^T \vec{b}_n \\ \vec{a}_2^T b_1 & \ddots & & \vdots \\ \vdots & & & \\ \vec{a}_n^T b_1 & \cdots & & \vec{a}_n^T \vec{b}_n \end{bmatrix}$$

$$= \sum_{i=1}^{m} \vec{a}_{1i} \vec{b}_{i1} + \sum_{i=1}^{m} \vec{a}_{2i} \vec{b}_{i2} + \cdots + \sum_{i=1}^{m} \vec{a}_{ni} \vec{b}_{in}$$

$$\frac{\partial trAB}{\partial a_{ij}} = \frac{\partial}{\partial a_{ij}}\left[\sum_{i=1}^{n} \vec{a}_{1i} b_{i1} + \sum_{i=1}^{n} \vec{a}_{2i} \vec{b}_{i2} + \cdots + \sum_{i=1}^{n} \vec{a}_{ni} \vec{b}_{in}\right]$$

$$= b_{ji}$$

$$\therefore \nabla_A trAB = \begin{bmatrix} \frac{\partial}{\partial a_{11}} trAB & \cdots & \frac{\partial}{\partial a_{1n}} trAB \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial a_{m1}} trAB & \cdots & \frac{\partial}{\partial a_{mn}} trAB \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{m1} \\ b_{12} & \ddots & & \vdots \\ \vdots & & & \\ b_{1n} & \cdots & & b_{mn} \end{bmatrix}$$

$\nabla_A trAB = B^T$ #

T2) $\nabla_{A^T} f(A) = (\nabla_A f(A))^T$

$$\underbrace{\nabla_{A^T} f(A) = \begin{bmatrix} \frac{\partial f}{\partial A_{11}} & \cdots & \frac{\partial f}{\partial A_{m1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial A_{1n}} & \cdots & \frac{\partial f}{\partial A_{mn}} \end{bmatrix}}_{\nabla_{A^T} f(A)} = \underbrace{\begin{bmatrix} \frac{\partial f}{\partial A_{11}} & \cdots & \frac{\partial f}{\partial A_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial A_{m1}} & \cdots & \frac{\partial f}{\partial A_{mn}} \end{bmatrix}^T}_{(\nabla_A f(A))^T}$$

#

T3) $\nabla_A trABA^TC = CAB + C^TAB^T$

let $f(A) = AB$, $g(A) = tr(ABA^TC)$

then we declare

$H(f(A), A) = tr(f(A) \cdot A^T C)$

$$\therefore \nabla_A H(f(A), A) = \nabla_{f(A)} H(f(A), A) \cdot \nabla_A f(A) + \nabla_A H(f(A), A) \cdot \cancel{\nabla_A A}$$

$$= \cancel{\nabla_{f(A)} tr(f(A) \cdot A^T C)}^{(A^TC)^T} \cdot \cancel{\nabla_A AB}^{B^T} + \nabla_A tr(f(A) \cdot A^T C)$$

$$= C^TAB^T + \nabla_A tr(\underbrace{ABA^TC}_{f(A)\,A^T\,C})$$

$tr(ABC) = tr(CAB)$

$$= C^TAB^T + \nabla_A tr(CABA^T)$$

$$= C^TAB^T + \nabla_A tr(A(CAB)^T)^T$$

$$= C^TAB^T + CAB$$

#

Chompakorn Chaichot

593 122 98 21