Basic knowledge ①

matrix derivation

For variable is a vector

1. constant.

$f(x)$, $x = [x_1, x_2, \cdots x_n]^T$

$$\frac{\partial f(x)}{\partial x} = \left[\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right]^T$$

$$\frac{\partial c}{\partial x} = 0_{n\times 1}$$

prove:

$$\frac{\partial c}{\partial x} = \begin{bmatrix} \frac{\partial c}{\partial x_1} \\ \vdots \\ \frac{\partial c}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0_{n\times 1}$$

2. Linear Rule.

$$\frac{\partial [c_1 f(x) + c_2 g(x)]}{\partial x} = c_1 \frac{\partial f(x)}{\partial x} + c_2 \frac{\partial g(x)}{\partial x}$$, here c_1, c_2 are constants.

prove:

$$\frac{\partial [c_1 f(x) + c_2 g(x)]}{\partial x} = \begin{bmatrix} \frac{\partial}{\partial x_1}[c_1 f + c_2 g] \\ \vdots \\ \frac{\partial}{\partial x_n}[c_1 f + c_2 g] \end{bmatrix} = \begin{bmatrix} c_1 \frac{\partial f}{\partial x_1} + c_2 \frac{\partial g}{\partial x_1} \\ \vdots \\ c_1 \frac{\partial f}{\partial x_n} + c_2 \frac{\partial g}{\partial x_n} \end{bmatrix}$$

$$= c_1 \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} + c_2 \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \vdots \\ \frac{\partial g}{\partial x_n} \end{bmatrix} = c_1 \frac{\partial f(x)}{\partial x} + c_2 \frac{\partial g(x)}{\partial x}$$

//

3. multiply Rule.

$$\frac{d[f(x)g(x)]}{dx} = \frac{df(x)}{dx} \cdot g(x) + f(x) \cdot \frac{dg(x)}{dx}$$

Prove:

$$\frac{d[f(x)g(x)]}{dx} = \begin{bmatrix} \frac{d(fg)}{dx_1} \\ \vdots \\ \frac{d(fg)}{dx_n} \end{bmatrix} = \begin{bmatrix} \frac{df}{dx_1} \cdot g + f \cdot \frac{dg}{dx_1} \\ \vdots \\ \frac{df}{dx_n} \cdot g + f \cdot \frac{dg}{dx_n} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{df}{dx_1} \\ \vdots \\ \frac{df}{dx_n} \end{bmatrix} \cdot g + f \cdot \begin{bmatrix} \frac{dg}{dx_1} \\ \vdots \\ \frac{dg}{dx_n} \end{bmatrix} = \frac{df(x)}{dx} \cdot g(x) + f(x) \cdot \frac{dg(x)}{dx}.$$

4. Division rule.

$$\frac{d\left[\frac{f(x)}{g(x)}\right]}{dx} = \frac{\frac{df(x)}{dx} \cdot g(x) - \frac{dg(x)}{dx} \cdot f(x)}{g(x)^2}, \text{ here } g(x) \neq 0.$$

Prove:

$$\frac{d\left[\frac{f(x)}{g(x)}\right]}{dx} = \begin{bmatrix} \frac{d\left[\frac{f}{g}\right]}{dx_1} \\ \vdots \\ \frac{d\left[\frac{f}{g}\right]}{dx_n} \end{bmatrix} = \begin{bmatrix} \frac{1}{g^2}\left(\frac{df}{dx_1} \cdot g - \frac{dg}{dx_1} \cdot f\right) \\ \vdots \\ \frac{1}{g^2}\left(\frac{df}{dx_n} \cdot g - \frac{dg}{dx_n} \cdot f\right) \end{bmatrix}$$

$$= \frac{1}{g^2}\left(\begin{bmatrix} \frac{df}{dx_1} \\ \vdots \\ \frac{df}{dx_n} \end{bmatrix} \cdot g - \begin{bmatrix} \frac{dg}{dx_1} \\ \vdots \\ \frac{dg}{dx_n} \end{bmatrix} \cdot f\right)$$

$$= \frac{1}{g^2}\left(\frac{df(x)}{dx} \cdot g(x) - \frac{dg(x)}{dx} \cdot f(x)\right)$$

5. Some equations.

① $\frac{\partial(x^Ta)}{\partial x} = \frac{\partial(a^Tx)}{\partial x} = a$, here a is constant vector.

$a = [a_1 \cdots a_n]^T$.

prove:

$$\frac{\partial(x^Ta)}{\partial x} = \frac{\partial(a^Tx)}{\partial x}$$

$$= \begin{bmatrix} \frac{\partial(a_1x_1 + \cdots a_nx_n)}{\partial x_1} \\ \vdots \\ \frac{\partial(a_1x_1 + \cdots a_nx_n)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = a$$

② $\frac{\partial(x^Tx)}{\partial x} = 2x$

prove: $\frac{\partial(x^Tx)}{\partial x} = \begin{bmatrix} \frac{\partial(x_1^2 + x_2^2 + \cdots x_n^2)}{\partial x_1} \\ \vdots \\ \frac{\partial(x_1^2 + \cdots x_n^2)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ \vdots \\ 2x_n \end{bmatrix} = 2x$

③ $\frac{d(x^TAx)}{dx} = Ax + A^Tx$. here $A_{n\times n}$ is constant matrix .

$A_{n\times n} = (a_{ij})_{i=1,j=1}^{n,n}$

Prove.

$$[x_1 \cdots x_n]\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \vdots & & & & \\ a_{n1} & a_{n2} & \cdots & & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= [a_{11}x_1 + a_{21}x_2 + \cdots a_{n1}x_n,\ a_{12}x_1 + a_{22}x_2 + \cdots a_{n2}x_n,\ \cdots,\ a_{1n}x_1 + a_{2n}x_2 + \cdots a_{nn}x_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \left\{\begin{matrix} a_{11}x_1x_1 + a_{21}x_1x_2 + a_{31}x_1x_3 + \cdots a_{n1}x_1x_n + \\ a_{12}x_1x_2 + a_{22}x_2x_2 + a_{32}x_2x_3 + \cdots a_{n2}x_2x_n + \\ a_{13}x_1x_3 + a_{23}x_2x_3 + a_{33}x_3x_3 + \cdots a_{n3}x_3x_n + \cdots \end{matrix}\right) \to G(x)$$

$$\frac{d(x^TAx)}{dx} = \begin{bmatrix} \frac{\partial}{\partial x_1}G(x) \\ \vdots \\ \frac{\partial}{\partial x_n}G(x) \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{11}x_1 + a_{21}x_2 + a_{31}x_3 + \cdots a_{n1}x_n + a_{12}x_2 + a_{13}x_3 + \cdots a_{1n}x_n \\ a_{12}x_1 + a_{22}x_2 + a_{22}x_2 + a_{32}x_3 + \cdots a_{n2}x_n + a_{21}x_1 + a_{23}x_3 + \cdots a_{2n}x_n \\ \vdots \\ a_{1n}x_1 + a_{2n}x_2 + a_{3n}x_3 + a_{nn}x_n + a_{nn}x_n + a_{n1}x_1 + a_{n2}x_2 + \cdots a_{n,n-1}x_{n-1} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + \cdots a_{n1}x_n \\ a_{12}x_1 + \cdots a_{n2}x_n \\ \vdots \\ a_{1n}x_1 + \cdots a_{nn}x_n \end{bmatrix} + \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots a_{nn}x_n \end{bmatrix}$$

$$A^TX \quad + \quad AX$$

//

④ $\frac{\partial(a^T x x^T b)}{\partial x} = ab^T x + ba^T x$, here a, b are constant vector. $a = (a_1 \cdots a_n)^T$, $b = (b_1 \cdots b_n)^T$.

prove:

m1:

$$\frac{d(a^T x x^T b)}{dx} = \begin{bmatrix} \frac{d}{dx_1}[(a_1x_1 + \cdots a_nx_n)(b_1x_1 + \cdots b_nx_n)] \\ \frac{d}{dx_2}[(a_1x_1 + \cdots a_nx_n)(b_1x_1 + \cdots b_nx_n)] \\ \vdots \\ \frac{d}{dx_n}[(a_1x_1 + \cdots a_nx_n)(b_1x_1 + \cdots b_nx_n)] \end{bmatrix}$$

$$= \begin{bmatrix} a_1(b_1x_1 + \cdots b_nx_n) + b_1(a_1x_1 + \cdots a_nx_n) \\ a_2(b_1x_1 + \cdots b_nx_n) + b_2(a_1x_1 + \cdots a_nx_n) \\ \vdots \\ a_n(b_1x_1 + \cdots b_nx_n) + b_n(a_1x_1 + \cdots a_nx_n) \end{bmatrix}$$

$$= \begin{bmatrix} a_1(b_1x_1 + \cdots b_nx_n) \\ \vdots \\ a_n(b_1x_1 + \cdots b_nx_n) \end{bmatrix} + \begin{bmatrix} b_1(a_1x_1 + \cdots a_nx_n) \\ \vdots \\ b_n(a_1x_1 + \cdots a_nx_n) \end{bmatrix}$$

$$= \begin{bmatrix} a_1 b^T x \\ \vdots \\ a_n b^T x \end{bmatrix} + \begin{bmatrix} b_1 a^T x \\ \vdots \\ b_n a^T x \end{bmatrix}$$

$$= ab^T x + a^T b x$$

m2: for $a^T x = x^T a$, $x^T b = b^T x$.

so $\frac{d(a^T x x^T b)}{dx} = \frac{d(x^T a b^T x)}{dx}$

Here ab^T is a $n \times n$ constant matrix, so: (from ③)

$$\frac{d(a^T x x^T b)}{dx} = \frac{d(x^T a b^T x)}{dx} = ab^T x + a^T b x$$