Matrix Differentiation

前面基本知识都一样

Convention3

$y = \psi(x)$ 　 y是m行的向量，x是n个元素的向量

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Propotion5

~~Convention 4~~

$y = Ax$ 　 $y: m\times 1$ 向量 　 $A: m\times n$ 矩阵 　 $x: n\times 1$ 向量

A与x无关

则 $\frac{\partial y}{\partial x} = A$.

证明：$y_i = A_{ij} x_j$

$$\frac{\partial y_i}{\partial x_j} = A_{ij} \longrightarrow \frac{\partial y}{\partial x} = A$$

$\cancel{\frac{\partial y_i}{\partial x_j}}$

Propotion6

$y = Ax$ 　 $x = x(z)$ 　 A与z无关

$$\frac{\partial y}{\partial z} = A\frac{\partial x}{\partial z}$$

Propotion7

① $\alpha = y^T A x$ （x与y无关 且A与xy无关）

② $y: m\times 1$

$A: m\times n$

$x: n\times 1$

$$\longrightarrow \begin{cases} \frac{\partial \alpha}{\partial x} = y^T A \\ \frac{\partial \alpha}{\partial y} = x^T A^T \end{cases}$$

证明：据Propotion5知 $\frac{\partial \alpha}{\partial x} = y^T A$.

令 $\alpha = w^T x$ 　 $w = A^T y$ 　 $w^T = y^T A$

由 $\alpha^T = \alpha = x^T A^T y$，同由P5得

$$\frac{\partial \alpha}{\partial y} = x^T A^T.$$

Propotion 8:

$\alpha = x^T A x$ 　 $x: n\times 1$ 　 $A: n\times n$ 　 A与x无关

$$\frac{\partial \alpha}{\partial x} = x^T(A + A^T)$$

证明：$\alpha = A_{ij} x_i x_j$ （张量记号）

$$\frac{\partial \alpha}{\partial x_k} = \sum_{j=1}^{n} A_{kj} x_j + \sum_{i=1}^{n} A_{ik} x_i = A_{kj} x_j + A_{ik} x_i$$

$$\frac{\partial \alpha}{\partial x} = x^T A^T + x^T A = x^T(A + A^T)$$

从而：$\frac{\partial \alpha}{\partial x_k \partial x_l} = A_{kl} + A_{lk}$

Propotion 9:

当A是对称阵

$$\frac{\partial \alpha}{\partial x} = 2x^T A$$

Propotion 10

① $\alpha = y^T x$，② $y: n\times 1$，$x: n\times 1$ 　 ③ $y = y(z)$，$x = x(z)$ 　 ④ $z: m\times 1$ 也可

↓

$$\frac{\partial \alpha}{\partial z} = x^T\frac{\partial y}{\partial z} + y^T\frac{\partial x}{\partial z}$$

证明：$\alpha = x_i y_i$

$$\frac{\partial \alpha}{\partial z_k} = \frac{\partial x_i}{\partial z_k} y_i + x_i \frac{\partial y_i}{\partial z_k} \qquad \frac{\partial \alpha}{\partial z} = x^T\frac{\partial y}{\partial z} + y^T\frac{\partial x}{\partial z}$$

Propotion 11:

① $\alpha = x^T x$ 　 ② $x = x(z)$ 　 ③ $x: n\times 1$，$z: m\times 1$

↓

$$\frac{\partial \alpha}{\partial z} = 2x^T\frac{\partial x}{\partial z}$$

Proportion 12:

$$\alpha = y^T A x$$

$$\frac{\partial \alpha}{\partial z} = x^T A^T \frac{\partial y}{\partial z} + y^T A \frac{\partial x}{\partial z}$$

P12 的条件为:

Propition 7 的条件加上 $x = x(z)$ $y = y(z)$ $z: 1 - num(vec(z))$

~~x=7~~

证明:

法一 $\alpha = A_{ij} y_i x_j \quad (i: 1-m \quad j: 1-n)$

$$\frac{\partial \alpha}{\partial z_k} = A_{ij} \frac{\partial y_i}{\partial z_k} x_j + A_{ij} y_i \frac{\partial x_j}{\partial z_k}$$

法二: 令

$\alpha = w^T x \quad w^T = y^T A$

由 proposition 10 得:

$$\frac{\partial \alpha}{\partial z} = x^T \frac{\partial w}{\partial z} + w^T \frac{\partial x}{\partial z}$$

$$\frac{\partial \alpha}{\partial z} = \frac{\partial \alpha}{\partial y} \frac{\partial y}{\partial z} + \frac{\partial \alpha}{\partial x} \frac{\partial x}{\partial z}$$

$$= x^T A^T \frac{\partial y}{\partial z} + y^T A \frac{\partial x}{\partial z}$$

↓

P7

Pro 13: (Pro 12 的特例)

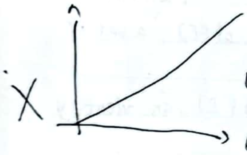
$$\alpha = x^T A x$$

$$\frac{\partial \alpha}{\partial z} = x^T (A + A^T) \frac{\partial x}{\partial z}$$

$$dA = dW = C_{ij}\varepsilon_j d\varepsilon_i \overset{6}{=} C_{11}\varepsilon_1 d\varepsilon_1 + C_{12}\varepsilon_2 d\varepsilon_1 \quad C_{21}\varepsilon_1 d\varepsilon_2 \quad \cancel{C_{12}\varepsilon} \quad C_{21}\varepsilon_1\varepsilon_2$$

$$+ C_{13}\varepsilon_3 d\varepsilon_1 + C_{21}\varepsilon_1 d\varepsilon_2 + C_{2\ldots}$$

$$\varepsilon_i \qquad \left(\tfrac{1}{2}C_{11}\varepsilon_1^2 \qquad C_{12}\varepsilon_2\varepsilon_1 \right) \qquad + C_{23}\varepsilon_3 d\varepsilon_2 + C_{31}\varepsilon_1 d\varepsilon_3 + C_{32}\varepsilon_2 d\varepsilon_3 + C_{33}\varepsilon_3 d\varepsilon_3$$



$$W = \tfrac{1}{2}C_{11}\varepsilon_1^2 + C_{12}\varepsilon_2\varepsilon_1 + C_{13}\varepsilon_3\varepsilon_1 + \cancel{C_{2\ldots}}\, C_{21}\varepsilon_{\ldots}$$

$$+ \tfrac{1}{2}C_{22}\varepsilon_2^2 + C_{23}\varepsilon_3\varepsilon_2 + \tfrac{1}{2}C_{33}\varepsilon_3^2 + C_{31}\varepsilon_{\ldots}$$

X $+ C_{32}\varepsilon_2\varepsilon_3$ (错误的，这样直接求是不行的)

$$W = \tfrac{1}{2}[\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4\ \varepsilon_5\ \varepsilon_6]\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{16} \\ C_{21} & C_{22} & \cdots & C_{26} \\ \vdots & & & \vdots \\ C_{61} & C_{62} & \cdots & C_{66} \end{bmatrix}\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_6 \end{bmatrix}$$

$$= \tfrac{1}{2}\varepsilon^T C \varepsilon$$

$$dW = \varepsilon^T C d\varepsilon$$

(共36项)

$$\frac{\partial W}{\partial \varepsilon} = \left[\frac{\partial\left(\frac{1}{2}C_{11}\varepsilon_1^2 + \frac{1}{2}C_{12}\varepsilon_1\varepsilon_2 + \frac{1}{2}C_{13}\varepsilon_1\varepsilon_3 + \frac{1}{2}C_{14}\varepsilon_1\varepsilon_4 + \frac{1}{2}C_{15}\varepsilon_1\varepsilon_5 + \frac{1}{2}C_{16}\varepsilon_1\varepsilon_6 + \frac{1}{2}C_{6\ldots} + \frac{1}{2}C_{22}\varepsilon_2^2 + \frac{1}{2}C_{23}\varepsilon_2\varepsilon_3 + \frac{1}{2}C_{24}\varepsilon_2\varepsilon_4 + \frac{1}{2}C_{25}\varepsilon_2\varepsilon_5 + \cdots + \frac{1}{2}C_{66}\varepsilon_6^2\right)}{\partial \varepsilon_1}, \right.$$

$$\left. \frac{\partial \frac{1}{2}C_{ij}\varepsilon_i\varepsilon_j}{\partial \varepsilon_2}, \frac{\partial \frac{1}{2}C_{ij}\varepsilon_i\varepsilon_j}{\partial \varepsilon_3}, \cdots\cdots, \frac{\partial \frac{1}{2}C_{ij}\varepsilon_i\varepsilon_j}{\partial \varepsilon_6} \right]$$

$$= \left[C_{11}\varepsilon_1 + \left(\tfrac{1}{2}C_{12}\varepsilon_2 + \tfrac{1}{2}C_{21}\varepsilon_2\right) + \left(\tfrac{1}{2}C_{13}\varepsilon_3 + \tfrac{1}{2}C_{31}\varepsilon_3\right) + \cdots \right.$$

$$\left. + \left(\tfrac{1}{2}C_{16}\varepsilon_6 + \tfrac{1}{2}C_{61}\varepsilon_6\right),\ C_{22}\varepsilon_2 + \left(\tfrac{1}{2}C_{21}\varepsilon_1 + \tfrac{1}{2}C_{12}\varepsilon_1\right) + \cdots + \left(\tfrac{1}{2}C_{26}\varepsilon_6 + \tfrac{1}{2}C_{62}\varepsilon_6\right), \cdots \right]$$ (共6项，事实上定 ~~$\varepsilon^T C$~~)

$$\boxed{\frac{\partial W}{\partial \varepsilon} = \frac{1}{2}\varepsilon^T (C + C^T)}$$ (未自知求)

由 $F'(v) = \frac{dF}{dv} = \frac{\partial F^{j}(v)}{\partial v^i} g_j \otimes g^i = \frac{\partial F}{\partial v^i} \otimes g^i$

$(v = v^i g_i)$ $dv = dv^i g_i$ 由 $dF = F'(v)\cdot dv$

между $dF = \frac{\partial F}{\partial v^i} \otimes g^i \cdot (dv^i g_i) = \frac{\partial F}{\partial v^i} dv^i$

$$dA = dW = \sigma_i d\varepsilon_i = C_{ij}\varepsilon_j d\varepsilon_i \longrightarrow W =$$

$$\sigma_i = C_{ij}\varepsilon_j \quad (i, j = 1, 2, \ldots, 6)$$

$$W = \tfrac{1}{2}C_{ij}\varepsilon_i\varepsilon_j$$

$$W = \tfrac{1}{2}\varepsilon^T C \varepsilon$$

年 月 日