第三章作业

3.1. $\because \phi(x) = (f(x))^T g(x) \in R$，$f(x) \in R^{p\times1}$，$g(x) \in R^{p\times1}$

$$\therefore D_x\phi(x) = \frac{\partial \phi(x)}{\partial x^T} = \frac{\partial (f(x))^T g(x)}{\partial x^T}$$

$$= \left[\frac{\partial (f(x))^T g(x)}{\partial x^T}\right]\Bigg|_{g(x)\text{视作常数}} + \left[\frac{\partial (f(x))^T g(x)}{\partial x^T}\right]\Bigg|_{f(x)\text{视作常数}}$$

$\because$ 设 $x \in R^{n\times1}$

$\therefore$ 若使 $\left[\frac{\partial (f(x))^T g(x)}{\partial x^T}\right]\Big|_{g(x)\text{视作常数}}$ 结果为 $1\times n$ 的行向量

则需 $(1\times p)\times(p\times n)$ 的向量乘矩阵形式

$$\left[\frac{\partial (f(x))^T g(x)}{\partial x^T}\right]\Bigg|_{g(x)\text{视作常数}} = (g(x))^T \cdot D_x f(x)，\text{其中 } D_x f(x) = \frac{\partial \begin{bmatrix} f(x)_1 \\ \vdots \\ f(x)_p \end{bmatrix}}{\partial [x_1, \cdots, x_n]} \in R^{p\times n}$$

$\therefore$ 同理可得

$$\left[\frac{\partial (f(x))^T g(x)}{\partial x^T}\right]\Bigg|_{f(x)\text{视作常数}} = (f(x))^T D_x g(x)$$

$\therefore$ 综上，$D_x\phi(x) = (g(x))^T D_x f(x) + (f(x))^T D_x g(x)$.

3.11 $f(x) = a^T x$，a, x 是 $m\times1$ 的列向量

$$\therefore H[f(x)] = H[a^T x] = \begin{bmatrix} \frac{\partial a^T x}{\partial x_1 x_1} & \frac{\partial a^T x}{\partial x_1 \partial x_2} & \cdots & \frac{\partial a^T x}{\partial x_1 \partial x_m} \\ \vdots & & & \vdots \\ \frac{\partial a^T x}{\partial x_m \partial x_1} & \frac{\partial a^T x}{\partial x_m \partial x_2} & \cdots & \frac{\partial a^T x}{\partial x_m \partial x_2} \end{bmatrix}$$

$\because a^T x = \sum_{i=1}^{m} a_i x_i$

$\therefore$ 对任何 i, j，有 $\frac{\partial \sum_{i=1}^{m} a_i x_i}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left[\frac{\partial \sum_{i=1}^{m} a_i x_i}{\partial x_j}\right] = \frac{\partial a_j}{\partial x_i} = 0$

$\therefore H[a^T x] = 0_{m\times m}$

$\because$ 若 $f(x) = x^T A x \in R$，则 $\frac{\partial f(x)}{\partial x \partial x^T} = \frac{\partial}{\partial x}\left[\frac{\partial f(x)}{\partial x^T}\right] = \frac{\partial [x^T(A^T + A)]}{\partial x} = (A + A^T)$

$\therefore H[x^T A x] = A + A^T$

3.19. (1) 设矩阵函数 $F(X)=AXB$，$A\in p\times m$，$X\in m\times n$，$B\in n\times q$.

$\because \frac{\partial [F(X)]_{kl}}{\partial x_{ij}} = \frac{\partial [\sum_{u=1}^{m}\sum_{v=1}^{n} a_{ku} x_{uv} b_{vl}]}{\partial x_{ij}} = b_{jl} a_{ki}$

$\therefore D(AXB) = \frac{\partial AXB}{\partial X^T} = B^T \otimes A.$

(2) 设矩阵函数为 $F(X) = AX^{-1}B$.

$\therefore dF(x) = A\, d(X^{-1}) B$

$= -A\cdot X^{-1}(dX)X^{-1}\cdot B$

$\because d(\mathrm{vec}F(X)) = A\, d(\mathrm{vec}X)$，A为 Jacobian 矩阵

$\therefore d(\mathrm{vec}(AX^{-1}B)) = -[(X^{-1}\cdot B)^T \otimes (A\cdot X^{-1})]\, d\mathrm{vec}X)$

$\therefore A = -[(X^{-1}B)^T \otimes (AX^{-1})]$

3.29. (1) $D_{z_1} f(z) = \frac{\partial f(z)}{\partial z_1^T} = \frac{\partial (z_1^H A z_2)}{\partial z_1^T} = 0_{n\times 1}$ ($\because \frac{\partial z_1^H}{\partial z_1^T} = 0_{n\times 1}$)

$D_{z_1^*} f(z) = \frac{\partial f(z)}{\partial z_1^H} = \frac{\partial A z_2}{\partial z_1^H} = z_2^T A^T.$

(2) 同理 $D_{z_2} f(z) = \frac{\partial f(z)}{\partial z_2^T} = \frac{\partial (z_1^H A z_2)}{\partial z_2^T} = z_1^H A$

$D_{z_2^*} f(z) = \frac{\partial f(z)}{\partial z_2^H} = 0_{n\times 1}$

(3). $\because z$ 为 $2n\times 1$ 列向量，$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}_{2n\times 1}$

$\therefore D_z f(x) = \frac{\partial f(z)}{\partial z^T} = \left[\frac{\partial f(z)}{\partial z_1^T} \quad \frac{\partial f(z)}{\partial z_2^T} \right]$

$= [0_{1\times n},\ z_1^H A]_{1\times 2n}$

$D_{z^*} f(x) = \frac{\partial f(z)}{\partial z^H} = [z_2^T A^T,\ 0_{1\times n}]_{1\times 2n}$