温兆和 102055014432 数学基础作业7

1、解：$f(A)=\frac{1}{2}(Ax+b-y)^T(Ax+b-y)$

$=\frac{1}{2}(x^TA^TAx+x^TA^T(b-y)+(b-y)^TAx+b^T(b-y)+y^T(y-b))$

故 $\frac{\partial f}{\partial A}=\frac{1}{2}\frac{\partial}{\partial A}(x^TA^TAx+x^TA^T(b-y)+(b-y)^TAx)$

$=\frac{1}{2}\frac{\partial}{\partial A}(x^TA^TAx+2(b-y)^TAx)$

$=\frac{1}{2}(2Axx^T+2(b-y)x^T)$

$=(Ax+b-y)x^T$

$\frac{\partial f}{\partial x}=\frac{1}{2}\frac{\partial}{\partial x}(x^TA^TAx+x^TA^T(b-y)+(b-y)^TAx)$

$=\frac{1}{2}(2A^TAx+2A^T(b-y))$

$=A^T(Ax+b-y)$

2、解：由于 $x^TAx=x_1\sum_{i=1}^{n}x_ia_{i1}+\cdots+x_n\sum_{i=1}^{n}x_ia_{in}$

$=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_ja_{ij}$

故 $\frac{\partial x^TAx}{\partial x_k}=\sum_{i=1}^{n}x_i(a_{ik}+a_{ki})$

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

而 $\frac{\partial x^TAx}{\partial a_{ij}}=x_ix_j$

故 $\frac{\partial x^TAx}{\partial A}=xx^T$

3、解：由于 $d(w^{-1})=-w^{-1}dww^{-1}$

故 ~~原式~~ $=d\,Tr(w^{-1})=Tr(d(w^{-1}))$

$=Tr(-w^{-1}dww^{-1})$

$=Tr(-(w^{-1})^2dw)$

故原式 $=[-(w^{-1})^2]^T$.

4. 证：(1). $\log(q) = \log(f(z)) = \log\left(\frac{\exp(z)}{\sum_i (\exp(z))_i}\right)$

故：$(\log(q))_i = \log \frac{e^{z_i}}{e^{z_1} + \cdots + e^{z_n}}$

$= z_i - \log(\sum_i e^{z_i})$

$J = -(p_1 z_1 - p_1 \log(\sum_i e^{z_i}) + \cdots + p_n z_n - p_n \log(\sum_i e^{z_i}))$

$\boxed{= -(\sum_{i=1}^{n} p_i z_i - \log(\sum_i e^{z_i}))}$

故 $\frac{\partial J}{\partial z_i} = -\left(p_i - \frac{e^{z_i}}{\sum_i e^{z_i}}\right) = -(p_i - q_i)$

故 $\frac{\partial J}{\partial z} = q - p.$

(2) 此时：$J = -\left(\sum_{i=1}^{n} p_i (w_{i1} x_1 + \cdots + w_{im} x_m) - \log(\sum_i e^{z_i})\right)$

其中：$e^{z_i} = e^{w_{i1} x_1 + \cdots + w_{im} x_m}$.

$\frac{\partial J}{\partial w_{ij}} = -\left(p_i x_j - \frac{x_j e^{z_i}}{\sum e^{z_i}}\right) = -x_j (p_i - q_i)$

故 $\frac{\partial J}{\partial W} = (q-p)x^T$ 成立。

5. 解：(1). $\frac{\partial L}{\partial \mu} = -\frac{1}{2} \frac{\partial}{\partial \mu} \sum_{t=1}^{N} (\mu^T \Sigma^{-1} \mu - x_t^T \Sigma^{-1} \mu - \mu^T \Sigma^{-1} x_t)$

$= -\frac{N}{2} \frac{\partial}{\partial \mu} \mu^T \Sigma^{-1} \mu + \frac{\partial}{\partial \mu} (\sum_t x_t)^T \Sigma^{-1} \mu$.

$= -N \Sigma^{-1} + \Sigma^{-1} (\sum_t x_t)^T$

(2). 令 $A = -\frac{N}{2} \frac{\partial}{\partial \Sigma} \ln|\Sigma|$.

则 $A = -\frac{N}{2|\Sigma|} \frac{\partial}{\partial \Sigma} |\Sigma|$

$= -\frac{N}{2|\Sigma|} |\Sigma| \Sigma^{-1}$.

$= -\frac{N}{2}(\cancel{\Sigma^{-1}}\Sigma^{-T}$

令 $B = -\frac{1}{2}\sum_t \frac{\partial}{\partial\Sigma}(x_t-\mu)^T\Sigma^{-1}(x_t-\mu)$.

则 $B = -\frac{1}{2}\sum_t \frac{\partial}{\partial\Sigma}(x_t-\mu)^T\Sigma^{-1}(x_t-\mu)$

$= \frac{1}{2}\sum_t \Sigma^{-T}(x_t-\mu)(x_t-\mu)^T\Sigma^{-T}$.

故 $\frac{\partial L}{\partial\Sigma} = A+B$. 再将 $\mu = \frac{1}{N}\sum_t x_t$ 代入得:

$\frac{\partial L}{\partial\Sigma} = -\frac{N}{2}\Sigma^{-T} + \frac{1}{2}\sum_t \Sigma^{-T}(x_t-\frac{1}{N}\sum_t x_t)(x_t-\frac{1}{N}\sum_t x_t)^T\Sigma^{-T}$.

当 $\frac{\partial L}{\partial\Sigma}=0$:

$\frac{N}{2}\Sigma^{-T} = \frac{1}{2}\sum_t \Sigma^{-T}(x_t-\mu)(x_t-\mu)^T\Sigma^{-T}$.

$N = (\sum_t (x_t-\mu)(x_t-\mu)^T)\Sigma^{-T}$

$\Sigma^T = \frac{1}{N}\sum_t (x_t-\mu)(x_t-\mu)^T$

故 $\Sigma = \frac{1}{N}\sum_t (x_t-\mu)(x_t-\mu)^T$

$= \frac{1}{N}\sum_t (x_t-\frac{1}{N}\sum_t x_t)(x_t-\frac{1}{N}\sum_t x_t)^T$.

6.解: $d|X^k| = Tr(X^{k*}dX^k)$

$= Tr(X^{k*}\cdot kX^{k-1}dX)$

故 $= Tr(kX^{k*}X^kX^{-1}dX)$

$= Tr(k|X^k|X^{-1}dX)$.

故 $\frac{\partial|X^k|}{\partial X} = k|X^k|X^{-T}$.

7.解: $d\,Tr(AXBX^Tc)$

$= Tr(d(AXBX^Tc))$

$= Tr(A\,dX\,BX^Tc + AXB\,dX^Tc)$

$= Tr(A\,dX\,BX^Tc + c^T dX\,B^TX^TA^T)$

$$= Tr\left((BX^T CA + B^T X^T A^T C^T) dX\right)$$

故原式 $= (BX^T CA + B^T X^T A^T C^T)^T$

$$= A^T C^T X B^T + CAXB.$$