Week 5 Written Assignment

1. To show $\int_{\mathbb{R}^k} f(x)\,dx = 1$, since x run all over $\mathbb{R}^k$, $dx = d(x-\mu)$, we can replace $x-\mu$ by x. Σ is positive definite $\Leftrightarrow$ Σ^{-1} is positive definite, so we can replace Σ^{-1} by Σ. Therefore, it's sufficient to show

$$\sqrt{|\Sigma|}\int_{\mathbb{R}^k} e^{-\frac{1}{2}x^T\Sigma x}\,dx = \sqrt{(2\pi)^k} \quad \ldots (1)$$

Moreover, suppose Σ is symmetric, then $\exists\, Q \in M_k(\mathbb{R})$ s.t $\Sigma = Q^T Q$, and therefore $|Q| = |Q^T| = \sqrt{|\Sigma|}$. Let $y = Qx$,

$$\sqrt{|\Sigma|}\int_{\mathbb{R}^k} e^{-\frac{1}{2}x^T\Sigma x}\,dx = \sqrt{|\Sigma|}\cdot\int_{\mathbb{R}^k} e^{-\frac{1}{2}y^Ty}\frac{1}{\sqrt{|\Sigma|}}\,dy$$

$$= \int_{\mathbb{R}^k} e^{-\frac{1}{2}y^Ty}\,dy = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-\frac{1}{2}\sum_{i=1}^{k} y_i^2}\,dy_1\ldots dy_k$$

$$= \left(\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\,dy\right)^k = \sqrt{2\pi}^{\,k}$$

2(a). Let $\frac{\partial}{\partial A}\operatorname{trace}(AB) = C$, then

$$C_{ij} = \frac{\partial}{\partial a_{ij}}\sum_{k=1}^{n}\sum_{m=1}^{n} a_{km}b_{mk} = b_{ji}, \text{ hence } C = B^T$$

2.(b) $$\operatorname{trace}(xx^TA) = \sum_{i=1}^{n}\sum_{j=1}^{n}(x_ix_j)a_{ij}$$

$$= \sum_{i=1}^{n} x_i\left(\sum_{j=1}^{n} a_{ij}x_j\right) = x^TAx$$

2.(c) $L(\mu,\Sigma)=\prod_{m=1}^{n}\frac{1}{\sqrt{(2\pi)^{k}|\Sigma|}}e^{-\frac{1}{2}(x_m-\mu)^T\Sigma^{-1}(x_m-\mu)}$

$$l(\mu,\Sigma)=\ln\left(2\pi^{\frac{-nk}{2}}|\Sigma|^{\frac{-n}{2}}\right)+\sum_{m=1}^{n}-\frac{1}{2}(x_m-\mu)^T\Sigma^{-1}(x_m-\mu)$$

Solve $\frac{\partial l}{\partial \mu}=0$: let $\mu=(\mu_1,\mu_2\cdots\mu_k)$

$x_m=(x_{m1},x_{m2}\cdots x_{mk})$ for $m=1,2\ldots n$

$\Rightarrow \frac{\partial l}{\partial \mu_\alpha}=0$ for $\alpha=1,2\ldots k$

ij-element of Σ^{-1}

$$\Rightarrow \frac{\partial}{\partial\mu_\alpha}\sum_{m=1}^{n}\sum_{j=1}^{k}\sum_{i=1}^{k}\left[(x_{mj}-\mu_j)(x_{mi}-\mu_i)\Sigma^{-1}_{ij}\right]$$

$$=\sum_{m=1}^{n}\frac{\partial}{\partial\mu_\alpha}\left(2(x_{m\alpha}-\mu_\alpha)\sum_{\substack{j=1\\j\neq\alpha}}^{k}(x_{mj}-\mu_j)\Sigma^{-1}_{\alpha j}+(x_{m\alpha}-\mu_\alpha)^2\Sigma^{-1}_{\alpha\alpha}\right)$$

$$=\sum_{m=1}^{n}\left(-2\sum_{\substack{j=1\\j\neq\alpha}}^{k}(x_{mj}-\mu_j)\Sigma^{-1}_{\alpha j}+2\mu_\alpha\Sigma^{-1}_{\alpha\alpha}-2x_{m\alpha}\Sigma^{-1}_{\alpha\alpha}\right)$$

$$=\sum_{m=1}^{n}\left(2\sum_{j=1}^{k}\mu_j\Sigma^{-1}_{\alpha j}-2\sum_{j=1}^{k}x_{mj}\Sigma^{-1}_{\alpha j}\right)$$

$$= 2n \sum_{j=1}^{k} \mu_j \Sigma_{\alpha j}^{-1} - 2\sum_{j=1}^{k}\left(\sum_{m=1}^{n} x_{mj}\right)\Sigma_{\alpha j}^{-1} = 0$$

for $\alpha = 1, 2 \ldots k$

$$\Rightarrow \sum_{j=1}^{k}\left(\mu_j - \frac{1}{n}\sum_{m=1}^{n} x_{mj}\right)\Sigma_{\alpha j}^{-1} = 0 \quad \text{for } \alpha = 1, \ldots k$$

$$\Rightarrow \Sigma^{-1}\left(\mu - \frac{1}{n}\sum_{m=1}^{n} x_m\right) = 0 \quad \text{and } \Sigma^{-1} \text{ is inverteible}$$

$$\Rightarrow \mu - \frac{1}{n}\sum_{m=1}^{n} x_m = 0,$$

$$\hat{\mu} = \frac{1}{n}\sum_{m=1}^{n} x_m$$

Solve $\frac{\partial l}{\partial \Sigma} = 0$:

$$\frac{\partial l}{\partial \Sigma} = -\frac{n}{2}\frac{\partial}{\partial \Sigma}\ln|\Sigma| - \frac{1}{2}\sum_{m=1}^{n}\frac{\partial}{\partial \Sigma}\left((x_m-\mu)^T\Sigma^{-1}(x_m-\mu)\right)$$

$$\Rightarrow n\frac{|\Sigma|(\Sigma^{-1})^T}{|\Sigma|} - \sum_{m=1}^{n}(\Sigma^{-1})^T(x_m-\mu)(x_m-\mu)^T(\Sigma^{-1})^T = 0$$

$$\Rightarrow \Sigma = \frac{1}{n}\sum_{m=1}^{n}(x_m-\mu)(x_m-\mu)^T$$

$$\hat{\Sigma} = \frac{1}{n}\sum_{m=1}^{n}(x_m-\mu)(x_m-\mu)^T$$

3. 我發現第1題好像必須假設 Σ 是對稱的才有辦法證？如果只有 positive definite 很容易可以構造反例，因為 $(x-\mu)^T\Sigma(x-\mu) = (x-\mu)^T E(x-\mu)$，$E = \frac{\Sigma+\Sigma^T}{2}$，但是 $|\Sigma|$ 通常不等於 $|E|$，因此不假設 Σ 對稱的話 $\int_{\mathbb{R}^k} f\,dx$ 的值通常就會因選取不同的 Σ 而不同。從統計的角度來看，Σ 是共變異矩陣，Σ 要對稱好像也蠻符合直覺的