Assignment 5

1.

$$\int_{\mathbb{R}^k} f(x)\,dx = \int_{\mathbb{R}^k} \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)dx \quad \text{, let } y=\sqrt{\Sigma^{-1}}(x-\mu)$$

$$dy = \sqrt{\Sigma^{-1}}\,dx$$

$$= \int_{\mathbb{R}^k} \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} y^T y\right) \cdot \frac{dy}{\sqrt{\Sigma^{-1}}}$$

$$= \int_{\mathbb{R}^k} \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}\sum_{i=1}^{k} y_i^2} \cdot \frac{dy}{\sqrt{\Sigma^{-1}}}$$

$$= \frac{1}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} e^{-\frac{1}{2}\sum_{i=1}^{k} y_i^2}\,dy$$

$$= \frac{1}{(2\pi)^{k/2}} \prod_{i=1}^{k} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y_i^2}\,dy_i \quad \text{, note that } \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz = \sqrt{2\pi},\ z\in\mathbb{R}^1$$

$$= \frac{1}{(2\pi)^{k/2}} \cdot (\sqrt{2\pi})^k$$

$$= 1.$$

Suppose $I=\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz$, Then

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-\frac{z_1^2}{2}}dz_1\right)\left(\int_{-\infty}^{\infty} e^{-\frac{z_2^2}{2}}dz_2\right)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(z_1^2+z_2^2)/2}\,dz_1\,dz_2$$

$$= \int_{0}^{\infty}\int_{0}^{2\pi} e^{-r^2/2}\,r\,d\theta\,dr \quad \text{, by polar coordinate}$$

$$= 2\pi\left(-e^{-\frac{r^2}{2}}\right)\Big|_0^{\infty} = 2\pi.$$

Thus, $I=\sqrt{2\pi}$ □

2.

(a) $tr(AB) = \sum_{i,j=1}^{n} a_{ij} b_{ji}$

$$\Rightarrow \frac{\partial}{\partial A} tr(AB) = \begin{bmatrix} \frac{\partial}{\partial a_{11}}tr(AB) & \frac{\partial}{\partial a_{12}}tr(AB) & \cdots \\ \frac{\partial}{\partial a_{21}}tr(AB) & \ddots & \\ & & \frac{\partial}{\partial a_{nn}}tr(AB) \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{21} & \cdots \\ b_{12} & \ddots & \\ \vdots & & b_{nn} \end{bmatrix} = B^T$$

since $\frac{\partial}{\partial a_{pq}} \sum_{i,j=1}^{n} a_{ij} b_{ji} = b_{qp}$ for specific p,q.

(b) $tr((xx^T)A) = tr(A(xx^T))$ by the property of trace

$$= \sum_{i,j=1}^{n} a_{ij} x_j x_i$$

$$= \sum_{i,j=1}^{n} x_i a_{ij} x_j$$

$$= x^T A x$$

(c) $f(x_i) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$

take log $\Rightarrow$ $L = \frac{-1}{2}\sum_{i=1}^{n}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) - \frac{n}{2}\ln(2\pi)^n - \frac{n}{2}\ln|\Sigma|$

$\overset{(b)}{=} \frac{-1}{2}\sum_{i=1}^{n} tr((x_i-\mu)(x_i-\mu)^T\Sigma^{-1}) - \frac{n}{2}\ln(2\pi)^n + \frac{n}{2}\ln|\Sigma^{-1}|$ by $|A| = \frac{1}{|A^{-1}|}$

A: invertible matrix

To find $\bar{\mu}, \bar{\Sigma}$ s.t. $L(\bar{\mu},\bar{\Sigma}) = \min L(\mu,\Sigma)$, we have

$\frac{\partial}{\partial\mu}L = \sum_{i=1}^{n}(x_i-\mu)\Sigma^{-1} = 0 \Rightarrow \bar{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and

$\frac{\partial}{\partial\Sigma^{-1}}L = \frac{\partial}{\partial\Sigma^{-1}}\frac{-1}{2}\sum_{i=1}^{n} tr((x_i-\mu)(x_i-\mu)^T\Sigma^{-1}) + \frac{n}{2}\Sigma$

$\overset{(a)}{=} \frac{-1}{2}\sum_{i=1}^{n}(x_i-\mu)^T(x_i-\mu) + \frac{n}{2}\Sigma = 0 \Rightarrow \bar{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)(x_i-\mu)^T.$

Thus, MLE gives $\bar{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{\mu})(x_i-\bar{\mu})^T$ $\square$

3.

In GDA, we know that $\Sigma_k = \frac{1}{n_k}\sum_{i:y^{(i)}=k}(x^{(i)}-\mu_k)(x^{(i)}-\mu_k)^T$.

Note that $rank(AA^T) = rank(A)$. Then for $(x^{(i)}-\mu_k)$, $rank(\Sigma_k) = rank(x^{(i)}-\mu_k) \le n_k - 1$ since $\sum_{i=1}^{n_k}(x^{(i)}-\mu_k) = 0$ implies that $x^{(i)}-\mu_k$ are dependent.

So if $n_k < n$, then $rank(\Sigma_k) < n$ implies that Σ_k isn't invertible, and we need Σ_k is invertible. So how to solve it? Is doing regularization to Σ_k the solution?