1. proof $\int_{R^k} f(x)\, dx = 1$

令 $y = x - u$, $\int_{R^k} f(x)\, dx = \int_{R^k} \frac{1}{(2\pi)^k |\Sigma|} \exp\left(-\frac{1}{2} y^T \Sigma^{-1} y\right) dy$

其中, $\Sigma = Q \Lambda Q$, $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_k)$, $|\Sigma| = \prod_{i=1}^{k} \lambda_i$

令 $z = Q^T y \Rightarrow y^T \Sigma^{-1} y = z^T \Lambda^{-1} z = \sum_{i=1}^{k} \frac{z_i^2}{\lambda_i}$

$$\Rightarrow \int_{R^k} \frac{1}{\sqrt{(2\pi)^k \prod_i \lambda_i}} \exp\left(-\frac{1}{2} \sum_{i=1}^{k} \frac{z_i^2}{\lambda_i}\right) dz$$

$$= \int_{\mathbb{R}^k} \left( \frac{1}{\sqrt{2\pi\lambda_i}} \int_{-\infty}^{\infty} \exp\left(-\frac{z_i^2}{2\lambda_i}\right) dz_i \right)$$

$$\because \int_{-\infty}^{\infty} \exp\left(-\frac{z^2}{2\lambda}\right) dz = \sqrt{2\pi\lambda}$$

$$\therefore \int_{\mathbb{R}^k} f(x)\, dx = 1$$

2.

(a) $$\frac{\partial}{\partial A} tr(AB) = B^T$$

$$tr(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} B_{ij}$$

$$\frac{\partial}{\partial A_{kl}} tr(AB) = \sum_{ij} \frac{\partial A_{ij}}{\partial A_{kl}} B_{ji} = B_{lk}$$

$$\Rightarrow \frac{\partial}{\partial A} tr(AB) = B^T$$

(b)

$$\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$$

$$x^T A x = \mathrm{tr}(x^T A x) = \mathrm{tr}(A x x^T) = \mathrm{tr}(x x^T A)$$

(c)

1.

$$\frac{\partial l}{\partial \mu} = -\frac{1}{2} \sum_{i=1}^{N} \left(-2 \Sigma^{-1}(x_i - \mu)\right) = \Sigma^{-1} \sum_{i=1}^{N} (x_i - \mu)$$

$$\Sigma^{-1} \sum_{i=1}^{N} (x_i - \mu) = 0 \Rightarrow \sum_{i=1}^{N} x_i - N\mu = 0$$

$$\Rightarrow \hat{\mu}_{ML} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

2.

$$S = \sum_{i=1}^{N} (x_i - \hat{u})(x_i - \hat{u})^T$$

$$\sum_{i=1}^{N} (x_i - u)^T \Sigma^{-1} (x_i - u) = tr(S\Sigma^{-1})$$

$$\frac{\partial \ell}{\partial w} = \frac{N}{2}\Sigma - \frac{1}{2}S$$

令 $\frac{N}{2}\Sigma - \frac{1}{2}S = 0 \Rightarrow N\Sigma = S$

$$\therefore \Sigma_{mL} = \frac{1}{N}S = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{u})(x_i - \hat{u})^T$$