Q1

$$E_D(w) = \frac{1}{2} tr\{(XW-T)^T(XW-T)\} = \frac{1}{2} tr\{(W^TX^T - T^T)(XW-T)\}$$

$$= \frac{1}{2} tr\{(W^TX^TXW - T^TXW - W^TX^TT + T^TT)\}$$

$$\frac{dE_D(w)}{dw} = \frac{1}{2}\left\{\frac{d}{dw} tr(W^TX^TXW) - \frac{d}{dw} tr(T^TXW) - \frac{d}{dw} W^TX^TT\right\}$$

$$= \frac{1}{2}\{X^TXW + X^TXW - X^TT - X^TT\} = X^TXW - X^TT = 0$$

$$W = (X^TX)^{-1}X^TT$$

4.8

$$\mathrm{p}(\mathrm{x}|C_1) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\}$$

$$\mathrm{p}(\mathrm{x}|C_2) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-\mu_2)^T\Sigma^{-1}(x-\mu_2)\right\}$$

$$\mathrm{a} = ln \frac{\exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)p(C_1)}{\exp\left(-\frac{1}{2}(x-\mu_2)^T\Sigma^{-1}(x-\mu_2)\right)p(C_2)}$$

$$= -\frac{1}{2}(x^T\Sigma^{-1}x - x^T\Sigma^{-1}\mu_1 - {\mu_1}^T\Sigma^{-1}x + {\mu_1}^T\Sigma^{-1}\mu_1 - x^T\Sigma^{-1}x$$

$$+ x^T\Sigma^{-1}\mu_2 + {\mu_2}^T\Sigma^{-1}x - {\mu_2}^T\Sigma^{-1}\mu_2) + \ln\frac{p(C_1)}{p(C_2)}$$

$$= (\mu_1-\mu_2)^T\Sigma^{-1}x - \frac{1}{2}({\mu_1}^T\Sigma^{-1}\mu_1 - {\mu_2}^T\Sigma^{-1}\mu_2) + \ln\frac{p(C_1)}{p(C_2)}$$

$$= \boldsymbol{w^T x} + w_0$$

4.9

Likelihood : $\mathrm{p}(\{\phi_n, t_n\}|\{\pi_k\}) = \prod_{n=1}^{N}\prod_{k=1}^{K}\{p(\phi_n|C_k)\pi_k\}^{t_{nk}}$

⇨ Taking log : $\ln \mathrm{p}(\{\phi_n, t_n\}|\{\pi_k\}) = \sum_{n=1}^{N}\sum_{k=1}^{K} t_{nk}\{\ln p(\phi_n|C_k) + \ln \pi_k\}$

Since $\sum_k \pi_k$ must equal 1, we try and maximize the following :

$$\ln \mathrm{p}(\{\phi_n, t_n\}|\{\pi_k\}) + \lambda(\sum_{k=1}^{K}\pi_k - 1)$$

We then set the derivative with respect to π_k to 0:

⇨ $\sum_{n=1}^{N}\frac{t_{nk}}{\pi_k} + \lambda = 0 \quad -\pi_k\lambda = \sum_{n=1}^{N} t_{nk} = N_k \quad \lambda = -\mathrm{N} \quad \pi_k = \frac{N_k}{N}$