Q1. Since $E_D(w) = \frac{1}{2}\sum_{n=1}^{N} r_n (t_n - w^T\phi(x_n))^2$

then $\frac{\partial}{\partial w} E_D(w) = \sum_{n=1}^{N} r_n (t_n - w^T\phi(x_n))\,\phi(x_n)^T$

Let $\frac{\partial}{\partial w} E_D(w) = 0$, we get

$$\sum_{n=1}^{N} r_n t_n \phi(x_n)^T = w^T \sum_{n=1}^{N} r_n \cdot \phi(x_n)\cdot\phi(x_n)^T$$

$$\Rightarrow \sum_{n=1}^{N} t_n' \phi'(x_n)^T = w^T \sum_{n=1}^{N} \phi'(x_n)\,\phi'(x_n)^T$$

where $t_n' = \sqrt{r_n}\cdot t_n$, $\phi'(x_n) = \sqrt{r_n}\cdot\phi(x_n)$

then we get solution:

$$w^* = (\Phi^T\Phi)^{-1}\Phi^T t$$

where $t = [\sqrt{r_1}t_1, \sqrt{r_2}t_2, \cdots, \sqrt{r_n}t_n]^T$

Φ is an $N\times M$ design matrix. $\Phi_{ij} = \sqrt{r_i}\phi_j(x_i)$

(i) If we substitute β^{-1} by $r_n\beta^{-1}$ in the summation term, the equation will become the expression above.

(ii) r_n can be viewed as the effective number of observation (x_n, t_n)

Q2. According to Bayesian theorem, we have

$$p(w,\beta|t) \propto p(t|X, w, \beta)\cdot p(w,\beta)$$

Since $p(t|w,\beta) = \prod_{n=1}^{N} N(t_n|w^T\phi(x_n), \beta^{-1})$

$$\propto \prod_{n=1}^{N} \beta^{\frac{1}{2}} e^{-\frac{\beta}{2}(t_n - w^T\phi(x_n))^2}$$

$$p(w,\beta) = N(w|m_0, \beta^{-1}S_0)\cdot Gam(\beta|a_0, b_0)$$

$$\propto \left(\frac{\beta}{|S_0|}\right)^2 e^{-\frac{1}{2}(w-m_0)^T\cdot\beta S_0^{-1}(w-m_0)}\, b_0^{a_0}\beta^{a_0-1}e^{-b_0\beta}$$

then ① quadratic term $= -\frac{\beta}{2}w^T S_0^{-1}w + \sum_{n=1}^{N} -\frac{\beta}{2}w^T\phi(x_n)\phi(x_n)^T w$

$$= -\frac{\beta}{2}w^T\left[S_0^{-1} + \sum_{n=1}^{N}\phi(x_n)\cdot\phi(x_n)^T\right]w$$

$$\Rightarrow S_N = \left[S_0^{-1} + \sum_{n=1}^{N}\phi(x_n)\cdot\phi(x_n)^T\right]^{-1}$$

② linear term $= \beta m_0^T S_0^{-1} w + \sum_{n=1}^{N}\beta t_n\phi(x_n)^T w$

$$= \beta\left[m_0^T S_0^{-1} + \sum_{n=1}^{N} t_n\phi(x_n)^T\right]w$$

$$\Rightarrow m_N = S_N\left[S_0^{-1}m_0 + \sum_{n=1}^{N} t_n \phi(x_n)\right]$$

③ constant term $= \left(-\frac{\beta}{2} m_0^T S_0^{-1} m_0 - b_0\beta\right) - \frac{\beta}{2}\sum_{n=1}^{N} t_n^2$

$$= -\beta\left[\frac{1}{2}m_0^T S_0^{-1} m_0 + b_0 + \frac{1}{2}\sum_{n=1}^{N} t_n^2\right]$$

$$\Rightarrow b_N = \frac{1}{2}m_0^T S_0^{-1} m_0 + b_0 + \frac{1}{2}\sum_{n=1}^{N} t_n^2 - \frac{1}{2}m_N^T S_N^{-1} m_N$$

④ exponential term $= (2 + a_0 - 1) + \frac{N}{2}$

$$= 2 + a_N - 1$$

$$\Rightarrow a_N = a_0 + \frac{N}{2}$$

Q3 Since $\int \frac{1}{(2\pi)^{M/2}} \frac{1}{|A|^{1/2}} e^{-\frac{1}{2}(w-m_N)^T A (w-m_N)}\, dw = 1$

then $\int e^{-\frac{1}{2}(w-m_N)^T A (w-m_N)}\, dw = (2\pi)^{\frac{M}{2}} \cdot |A|^{\frac{1}{2}}$

Since $E(m_N)$ doesn't depend on w, we get

$$\int e^{-E(w)} dw = e^{-E(m_N)} \cdot (2\pi)^{\frac{M}{2}} \cdot |A|^{-\frac{1}{2}}$$

then from the textbook, we have

$$p(t|\alpha,\beta) = \left(\frac{\beta}{2\pi}\right)^{N/2} \left(\frac{\alpha}{2\pi}\right)^{M/2} \int e^{-E(w)} dw$$

$$\Rightarrow \ln p(t|\alpha,\beta) = \frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - E(m_N) - \frac{1}{2}\ln|A| - \frac{N}{2}\ln(2\pi)$$

Q4.

$$F(a) = \frac{1}{2}\sum_i (Y_i - aX_i)^2$$

$$\frac{\partial}{\partial a}F(a) = \sum aX_i^2 - X_iY_i = 0$$

$$\Rightarrow a = \frac{\sum X_iY_i}{\sum X_i^2}$$

Q5

$$L(\theta|y_1, y_2, \cdots, y_n) = \frac{\theta^{y_1}e^{-\theta}}{y_1!} \cdot \frac{\theta^{y_2}e^{-\theta}}{y_2!} \cdots \frac{\theta^{y_n}e^{-\theta}}{y_n!}$$

$$= e^{-n\theta} \cdot \frac{\theta^{\sum y_i}}{\prod_{i=1}^{n} y_i!}$$

$$\ln L(\theta|y_1, y_2, \cdots, y_n) = -n\theta + \sum_{i=1}^{n} y_i \cdot \ln\theta - \sum_{i=1}^{n}\ln y_i$$

Let $\frac{\partial}{\partial\theta}\ln L = -n + \frac{\sum_{i=1}^{n} y_i}{\theta} = 0$

$$\Rightarrow \theta = \frac{\sum_{i=1}^{n} y_i}{n}$$

Q6. $L(X|\alpha,\lambda) = \frac{1}{(\Gamma(\alpha))^n} \cdot \lambda^{\alpha n} \cdot \prod_{i=1}^{n} x_i^{\alpha-1} \cdot e^{-\lambda x_i}$

$\ln L(X|\alpha,\lambda) = \alpha n \cdot \ln\lambda - n \cdot \ln(\Gamma(\alpha)) + \sum_{i=1}^{n}\left[(\alpha-1)\cdot\ln x_i - \lambda x_i\right]$

Let $\frac{\partial}{\partial\lambda}\ln L = \frac{\alpha n}{\lambda} - \sum_{i=1}^{N} x_i = 0$

$\Rightarrow \quad \lambda = \frac{\alpha n}{\sum_{i=1}^{n} x_i}$