Q1 Define $R = \mathrm{diag}(r_1, r_2, \cdots, r_N)$

$$E_D(w) = \frac{1}{2}(t - \phi w)^T R (t - \phi w)$$

$$\frac{dE_D(w)}{dw} = 0 \qquad \phi^T R \phi w + w^T \phi^T R \phi = t^T R \phi + \phi^T R t$$

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

(1) $p(t|X, w, \beta) = \prod_{n=1}^{N} N(t_n | w^T \phi(x_n), \beta^{-1}) \quad \ln p(t|w,\beta) = \frac{N}{2}\ln\beta - \frac{N}{2}\ln 2\pi - \beta E_D'(w)$

$$E_D'(w) = \frac{1}{2}\sum_{n=1}^{N} \{t_n - w^T \phi(x_n)\}^2$$

Therefore, r_n can be regarded as β parameter particular to the datapoint (x_n, t_n)

(2) r_n can be regarded as an effective number of replicated observations of data point (x_n, t_n). If r_n is positive integer, it's more clear to be interpreted this way.

Q2 $\ln p(w, \beta | t) = \ln p(w, \beta) + \sum_{n=1}^{N} \ln p(t_n | w^T \phi(x_n), \beta^{-1})$

$$= \frac{M}{2}\ln\beta - \frac{1}{2}\ln|S_0| - \frac{\beta}{2}(w - m_0)^T S_0^{-1}(w - m_0) - b_0\beta + (a_0 - 1)\ln\beta$$

$$+ \frac{N}{2}\ln\beta - \frac{\beta}{2}\sum_{n=1}^{N}\{w^T\phi(x_n) - t_n\}^2 + C$$ where C is a constant

$p(w, \beta | t) = p(w|\beta, t)\, p(\beta|t)$

Consider the first dependence on w, we have

$$\ln p(w|\beta, t) = -\frac{\beta}{2} w^T [\phi^T\phi + S_0^{-1}] w + w^T [\beta S_0^{-1} m_0 + \beta \phi^T t] + C$$

Therefore, $p(w|\beta, t)$ is a Gaussian distribution with mean and covariance given by

$$m_N = S_N [S_0^{-1} m_0 + \phi^T t]$$

$$\beta S_N^{-1} = \beta (S_0^{-1} + \phi^T \phi)$$

$$S_N = (S_0^{-1} + \phi^T \phi)^{-1}$$

$$\ln p(\beta|t) = -\frac{\beta}{2} m_0^T S_0^{-1} m_0 + \frac{\beta}{2} m_N^T S_N^{-1} m_N + \frac{N}{2}\ln\beta - b_0\beta + (a_0 - 1)\ln\beta - \frac{\beta}{2}\sum_{n=1}^{N} t_n^2 + C$$

we recognize the log of Gamma distribution

$$\mathrm{Gamma}(\beta | a_N, b_N) = \frac{b_N^{a_N} \beta^{a_N - 1} e^{-b_N \beta}}{\Gamma(a_N)}$$

$$\log(\mathrm{Gamma}(\beta|a_N, b_N)) = a_N \log b_N + (a_N - 1)\log\beta - b_N\beta - \log\Gamma(a_N)$$

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

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

Question 3

$$\int \exp\{-E(w)\}\,dw = \int \exp\{-E(m_N) - \frac{1}{2}(w-m_N)^T A (w-m_N)\}\,dw$$

$$= \exp\{-E(m_N)\} \int \exp\{-\frac{1}{2}(w-m_N)^T A (w-m_N)\}\,dw$$

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

$$= \exp\{-E(m_N)\} (2\pi)^{\frac{M}{2}} |A|^{-\frac{1}{2}}$$

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

$$= \left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}} \left(\frac{\alpha}{2\pi}\right)^{\frac{M}{2}} \exp\{-E(m_N)\}\, 2\pi^{\frac{M}{2}} |A|^{-\frac{1}{2}}$$

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

Question 4

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

$$a = \frac{\sum_i X_i Y_i}{\sum_i X_i^2}$$

Question 5

$$\log p(y|\theta) = y\log\theta - \theta - \sum_{i=0}^{y} \log i$$

$$\therefore \sum_{i=1}^{n} (y_i \log\theta - \theta - \log y_i!)$$

$$= \sum_{i=1}^{n} (y_i\log\theta - \log y_i!) - n\theta$$

Question 6

$$\log f_X(x) = \alpha\log\lambda + (\alpha-1)\log x - \lambda x - \log\Gamma(\alpha)$$

$$g(\lambda) = n\alpha\log\lambda + (\alpha-1)\log\prod_{i=1}^{n} X_i - \lambda\sum_{i=1}^{n} X_i - n\log\Gamma(\alpha)$$

$$\frac{dg(\lambda)}{d\lambda} = \frac{n\alpha}{\lambda} - \sum_{i=1}^{n} X_i \qquad \hat{\lambda} = \frac{\alpha}{\frac{1}{n}\sum_{i=1}^{n} X_i}$$