

Question 1

$$E_D(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N r_n (t_n - \mathbf{w}^T \phi(\mathbf{x}_n))^2$$

$$\frac{\partial E_D(\mathbf{w})}{\partial \mathbf{w}} = - \sum_{n=1}^N r_n \phi(\mathbf{x}_n)^T (t_n - \mathbf{w}^T \phi(\mathbf{x}_n)) = 0$$

It is solved that

$$\mathbf{w}^* = \left(\sum_{n=1}^N r_n \phi(\mathbf{x}_n) \phi(\mathbf{x}_n)^T \right)^{-1} \left(\sum_{n=1}^N r_n t_n \phi(\mathbf{x}_n) \right)$$

Data Dependent Noise Variance

The weighting factors r_n can be considered as inversely proportional to the variance of the noise associated with each data point. Higher weights imply lower noise variance and thus more reliable data points. This perspective is useful in scenarios where different data points have varying levels of reliability or precision.

Replicated Data Points

each weight r_n as representing the effective number of times that the data point n is replicated in the dataset. A higher weight for a data point means it's as if that data point appears multiple times, giving it more influence on the model.

Question 2

$$\text{Gam}(\beta|a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \beta^{a_0-1} \exp(-b_0 \beta)$$

$$p(\mathbf{w}, \beta | \mathbf{t}) \propto p(\mathbf{t} | \mathbf{X}, \mathbf{w}, \beta) p(\mathbf{w}, \beta)$$

$$p(\mathbf{t} | \mathbf{X}, \mathbf{w}, \beta) \propto \prod_{n=1}^N \beta^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (t_n - \mathbf{w}^T \Phi(x_n))^2 \right\}$$

$$p(\mathbf{w}, \beta) \propto \left(\frac{\beta}{|\mathbf{S}_0|} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\beta}{2} (\mathbf{w} - \mathbf{m}_0)^T \mathbf{S}_0 (\mathbf{w} - \mathbf{m}_0) \right\} b_0^{a_0} \beta^{a_0-1} \exp(-b_0 \beta)$$

Quadratic part

$$\sum_{n=1}^N \frac{\beta}{2} \mathbf{w}^T \phi(\mathbf{x}_n) \phi(\mathbf{x}_n)^T \mathbf{w} - \frac{\beta}{2} \mathbf{w}^T \mathbf{S}_0^{-1} \mathbf{w} = \frac{\beta}{2} \mathbf{w}^T \left(\sum_{n=1}^N \phi(\mathbf{x}_n) \phi(\mathbf{x}_n)^T + \mathbf{S}_0^{-1} \right) \mathbf{w}$$

Linear part

$$\beta \mathbf{m}_0^T \mathbf{S}_0^{-1} \mathbf{w} + \sum_{n=1}^N \beta t_n \phi(\mathbf{x}_n)^T \mathbf{w} = \beta \left(\mathbf{m}_0^T \mathbf{S}_0^{-1} + \sum_{n=1}^N t_n \phi(\mathbf{x}_n)^T \right) \mathbf{w}$$

Hence,

$$S_N = \left(\sum_{n=1}^N \phi(\mathbf{x}_n) \phi(\mathbf{x}_n)^T + S_0^{-1} \right)^{-1}$$

$$\mathbf{m}_N = S_N \left(S_0^{-1} \mathbf{m}_0 + \sum_{n=1}^N t_n \phi(\mathbf{x}_n) \right)$$

Constant part

$$\begin{aligned} p(\mathbf{w}, \beta | \mathbf{t}) &\propto -\frac{\beta}{2} \mathbf{m}_0^\top \mathbf{S}_0^{-1} \mathbf{m}_0 - b_0 \beta - \frac{\beta}{2} \sum_{n=1}^N t_n^2 \\ &= -\beta \left(\frac{1}{2} \mathbf{m}_0^\top \mathbf{S}_0^{-1} \mathbf{m}_0 + b_0 + \frac{1}{2} \sum_{n=1}^N t_n^2 \right) \\ &= -\beta \left(\frac{1}{2} \mathbf{m}_N^\top \mathbf{S}_N^{-1} \mathbf{m}_N + b_N \right) \\ \beta^{\frac{1}{2}} \beta^{a_0} \beta^{\frac{N}{2}} &= \beta^{\frac{N}{2} + a_0 - \frac{1}{2}} \\ a_N &= a_0 + \frac{N}{2} \\ b_N &= \frac{1}{2} \mathbf{m}_0^\top \mathbf{S}_0^{-1} \mathbf{m}_0 + b_0 + \frac{1}{2} \sum_{n=1}^N t_n^2 - \frac{1}{2} \mathbf{m}_N^\top \mathbf{S}_N^{-1} \mathbf{m}_N \end{aligned}$$

Question 3

$$\begin{aligned} E(\mathbf{w}) &= E(\mathbf{m}_N) + \frac{1}{2} (\mathbf{w} - \mathbf{m}_N)^\top \mathbf{A} (\mathbf{w} - \mathbf{m}_N) \\ \int \exp(-E(w)) dw &= \exp(-E(m_N)) \int \exp \left(-\frac{1}{2} (w - m_N)^\top A (w - m_N) \right) dw \\ &= \exp(-E(m_N)) (2\pi)^{\frac{M}{2}} |A|^{-\frac{1}{2}} \end{aligned}$$

Hence,

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

Question 4

$$\begin{aligned} L(a) &= \ln \prod_{i=1}^n p(Y_i | X_i, a) = -\frac{n}{2} \ln(2\pi) - n \ln \sigma - \sum_{i=1}^n \frac{1}{2\sigma^2} (Y_i - aX_i)^2 \\ \frac{\partial}{\partial a} L(a) &= \frac{1}{\sigma^2} \sum_{i=1}^n X_i (Y_i - aX_i) = 0 \end{aligned}$$

It is solved that

$$a_{\text{ML}} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}$$

Question 5

$$\begin{aligned} L(\theta) &= \ln \prod_{i=1}^n \frac{\theta^{y_i} e^{-\theta}}{y_i!} \\ &= \sum_{i=1}^n y_i \ln \theta - n\theta - \sum_{i=1}^n \ln(y_i!) \end{aligned}$$

Question 6

$$\begin{aligned} L(\lambda) &= \ln \prod_{i=1}^n \frac{1}{\Gamma(\alpha)} \lambda^\alpha X_i^{\alpha-1} e^{-\lambda X_i} \\ &= \ln \frac{1}{\Gamma(\alpha)^n} \lambda^{\alpha n} \left(\prod_{i=1}^n X_i^{\alpha-1} \right) e^{-\lambda \sum_{i=1}^n X_i} \\ \frac{\partial}{\partial \lambda} L(\lambda) &= \frac{\alpha n}{\lambda} - \sum_{i=1}^n X_i = 0 \end{aligned}$$

It is solved that

$$\lambda_{\text{ML}} = \frac{\alpha n}{\sum_{i=1}^n X_i}$$