

Question 1.

$$p(w|D) \propto \exp(-(X^T X + \alpha I)^{-1} X^T t)$$

My Answer 1.

$$p(w | x, t, \alpha, \beta) \propto p(t | x, w, \beta) p(w | \alpha)$$

$$\Rightarrow \log(p(w | x, t, \alpha, \beta)) \propto \log(p(t | x, w, \beta) p(w | \alpha))$$

we have:

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

$$p(t | x, w, \beta) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi\beta^{-1}}} \times e^{-\frac{(t - y(x_n, w))^2}{2\beta^{-1}}}$$

$$\log(p(t | x, w, \beta)) = -\frac{\beta}{2} \sum_{n=1}^N (t - y(x_n, w))^2 + \text{noise}$$

And:

$$p(w | \alpha) = N(w | 0, \alpha^{-1} I) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{(w-0)^T \Sigma^{-1} (w-0)}{2}}$$

$$\log(p(w | \alpha)) = -\frac{1}{2} w^T w + \text{noise}$$

So:

$$\log(p(w | x, t, \alpha, \beta)) \propto -\frac{\beta}{2} \sum_{n=1}^N (t - y(x_n, w))^2 + \frac{-1}{2} w^T w$$

we find that the maximum of the posterior is given by the minimum of:

$$\frac{\beta}{2} \sum_{n=1}^N (t - y(x_n, w))^2 + \frac{1}{2} w^T w$$

or we minimize:

$$Q = \|X\mathbf{w} - \mathbf{t}\|_2^2 + \lambda \mathbf{w}^T \mathbf{w}$$

$$\nabla Q_w = 2X^T(X\mathbf{w} - \mathbf{t}) + 2\lambda \mathbf{w}$$

$$\rightarrow \mathbf{w} = (X^T X + \lambda I)^{-1} X^T \mathbf{t}$$