Exercise 1

$$p(t|w) = \frac{p(t|w)p(w)}{p(t)}$$

To maxmize posterior, we maximize $p(t|w)p(w)$

$$With\ p(w) \sim N\left(0,\ \alpha^{-1}I\right),\ p(t|w) \sim N\left(t|y(x,w),\ \beta^{-1}\right)$$

$$p(t|w)p(w)\ max$$

$$\rightarrow\ N\left(0,\ \alpha^{-1}I\right)\prod N\left(t_i|y(x_i,\ w),\ \beta^{-1}\right)\ max$$

$$log\ p(t|w)p(w) = \sum log\left(N\left(t_i|y(x_i,w),\mathrm{\beta}^{-1}\right)\right) + N\left(0,\ \alpha^{-1}I\right)$$

$$= -\frac{\beta}{2}\sum\left[y(x_i,w) - t_i\right]^2 - \frac{\alpha}{2}w^Tw + const$$

$$max\ log\ p(t|x,w,\mathrm{\beta}) = max\ -\frac{\beta}{2}\sum\left[y(x_i,w) - t_i\right]^2 - \frac{\alpha}{2}w^Tw$$

$$= min\ \frac{\beta}{2}\sum\left[y(x_i,w) - t_i\right]^2 + \frac{\alpha}{2}w^Tw$$

$$= min\ ||Xw - t||^2 + \lambda w^Tw\ with\ \lambda = \frac{\alpha}{\beta} \tag{1}$$

To minimize (1), suppose

$$\begin{aligned} S &= ||Xw - t||^2 + \lambda w^Tw \\ &= (Xw - t)^T(Xw - t) + \lambda w^Tw \\ &= t^Tt - (Xw)^Tt - t^T(Xw) + (Xw)^T(Xw) + \lambda w^Tw \\ &= t^Tt - 2(Xw)^Tt + (Xw)^T(Xw) + \lambda w^Tw \\ &= t^Tt - 2w^TX^Tt + w^TX^TXw + \lambda w^Tw \end{aligned}$$

To minimize S, w must satisfy

$$\frac{\partial S}{\partial w} = -2X^Tt + 2X^TXw + 2\lambda Iw = 0$$

$$\leftrightarrow\ 2\left(X^TX + 2\lambda I\right)w = 2X^Tt$$

$$\leftrightarrow\ w = \left(X^TX + 2\lambda I\right)^{-1}X^Tt$$