# Homework 4

Thanh Hue

October 2021

## 1 Proof: $p(w|D)->w = (X^TX+\alpha I)^{-1}X^Tt$

We have:

$$p(w|x,t,\alpha,\beta) = \frac{p(t|x,w,\beta)p(w|\alpha)}{p(x,t,\alpha,\beta)}$$

Suppose $p(w|\alpha)$ is a normal distribution. So:

$$p(w|\alpha) = N(w|0,\alpha^{-1}I) = (\frac{\alpha}{2\pi})^{\frac{M+1}{2}} exp(-\frac{\alpha}{2}w^Tw)$$

$$p(t|x,w) = \prod_{i=1}^{N} p(t_i|x_i,w) = \prod_{i=1}^{N} N(t_i|y(x_i,w),\beta^{-1})$$

To maximize the posterior to find w, we have to maximize $p(t|x,w,\beta)p(w|\alpha)$.

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

$$\rightarrow log(p(t|(x,w))+log(p(w|\alpha))$$

$$= \sum_{i=1}^{N} log(N(y(x_i,w),\beta^{-1}+log(p(w|\alpha^{-1}I)$$

$$= \sum_{i=1}^{N} log(\frac{1}{\beta^{-1}\sqrt{2\pi}}e^{\frac{-(t_i-y(x_i,w))^2\beta}{2}}) + log(\frac{1}{\sqrt{(2\pi)^D} \mid \alpha^{-1}I \mid}e^{\frac{-1}{2}w^T(\alpha^{-1}I)^{-1}w})$$

$$= \frac{-\beta}{2}\sum_{i=1}^{N}(t_i-y(x_i,w))^2 + \frac{-1}{2}\alpha w^Tw$$

(because $(\alpha^{-1}I)^{-1} = \alpha I_D = \alpha$)

$$\Rightarrow min: \sum_{i=1}^{N}(t_i-y(x_i,w))^2 + \frac{\alpha}{\beta}w^Tw$$

$$Let\ \frac{\alpha}{\beta} = \lambda\ (\lambda > 0)$$

$$L = \sum_{i=1}^{N} (t_i - y(x_i, w))^2 + \lambda w^T w$$

$$\Leftrightarrow L = \|Xw - t\|_2^2 + \lambda \|w\|_2^2$$

$$\frac{\partial L}{\partial w} = 0$$

$$\Leftrightarrow 2X^T(Xw - t) + 2\lambda w = 0$$

$$\Leftrightarrow w(X^T X + \lambda I) - X^T t = 0$$

$$\Leftrightarrow w(X^T X + \lambda I) = X^T t$$

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