

Machine learning

Huong Bui

October 6, 2021

$$\mathbf{1} \quad \mathbf{p}(\mathbf{w} \mid D) \rightarrow w(X^T X + \lambda I) - X^T t$$

Solve:

$$p(w \mid x, t) = \frac{p(t \mid x, w)p(w \mid \alpha)}{p(D)}$$

Purpose : $P(w \mid \alpha)$ is normal distribution. We have :

$$p(w \mid \alpha) = \mathcal{N}(w \mid 0, \alpha^{-1} I) = \left(\frac{\alpha}{2\pi}\right)^{\frac{(M+1)}{2}} \exp\left(-\frac{\alpha}{2} w^T w\right)$$

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

$p(D)$ is constant So max posterior $\Rightarrow \max p(t \mid x, w)p(w \mid \alpha)$

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

$$\leftrightarrow \log(p(t \mid x, w)) + \log(p(w \mid \alpha))$$

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

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

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

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

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

$$Let : \frac{\alpha}{\beta} = \lambda$$

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

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

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

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

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