Homework 5: Regularized Linear Regression

Trần Minh Khoa

October, 2021

$$posterior = \frac{likelihood \times prior}{evidence}$$

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

$$p(w|\alpha) = N(w|0,\alpha^{-1}.I)$$

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

We are trying to maximize the posterior to find w

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

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

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

$$= \sum_{i=1}^{N} log(N|y(x_i,w),\beta^{-1}) + log(N(w|0,\alpha^{-1}.I))$$

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

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

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

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

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

$$\Rightarrow \frac{\partial L}{\partial w} = 2X^T(Xw - t) + 2\lambda w = 0$$

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

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