posterior

Nguyen Thi Hoai Linh

October 2021

1 EX1:

Biến đổi lại posterior ra latex, từ p(w|D) $\Rightarrow w = (X^T X + \alpha * I)^{-1} X^T t$

Solution:

Bayes theorem:

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

$\Rightarrow posterior = \frac{likelihood*prior}{evidence}$

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

p(w|α) $= N(w, 0, \alpha^{-1}I)$

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

$\Rightarrow max \quad p(t|x,w)p(w|\alpha)$

Lấy Log:

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

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

$$= \frac{-\beta}{2} \sum_{i=1}^{N} (t - y(x_i, w))^2 - \frac{1}{2}\alpha w^T w \Rightarrow max$$

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

Let:

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

$\lambda > 0$

Loss function:

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

$\Rightarrow L = ||xw - t||_2^2 + \lambda ||w||_2^2$

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

$\Rightarrow w(x^T x + \lambda I) = x^T$

$\Rightarrow w(x^T x + \lambda I) = x^T t$

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