$w = (X^T.X + \alpha * I_n)^{-1}X^T t$

Nhung Đào Thị Hồng

October 2021

1 Posterior

Bayes theorem

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

$$\Leftrightarrow posterior = \frac{likelihood.prior}{evidence}$$

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

Vì $p(t|x,w)$ chưa biết phân phối nên giả sử nó phân phối chuẩn Ta có:

$$p(w|\alpha).p(t|x,w) = N(w|0,\alpha^{-1}I).\prod_{i=1}^{N} N(t_x|y(x_i,w),B^{-1})(*)$$

$$log(*) = log(\frac{1}{(2\pi)^{\frac{D}{2}.|\alpha^{-1}.I|}}.exp\left\{\frac{-1}{2}.w^T(\alpha^{-1}I)^{-1}.w\right\}.\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}.\beta^{-1}}.exp(\frac{-(t_i - y(x,w))^2.\beta}{2})$$

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

Mixímize: $log(*) <=> minimize(\sum_{i=1}^{N}(t_i - y(x_i,w))^2 - \lambda.w^T.w)$ Đặt $L = \sum_{i=1}^{N}(t_i - y(x_i,w))^2 + \lambda w^T w$

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

$$\frac{dL}{dw} = 2X^T.(Xw - t) + 2\lambda I_n w = 0$$

$$<=> X^T X w - X^T t + \lambda w = 0$$

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

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