Linear Regression

ThienCao

September 2021

I.Bien doi lai bai tren lop:

$$t = y(x,w) + noise = N(y(x,w), \beta^{-1}$$

$$\Longrightarrow p(t \mid x,w,\beta) = N(t \mid y(x,w), \beta^{-1}$$

$$p(t \mid x,w,\beta) = \prod_{n=1}^{N} N(t_n \mid y + (x_n,w), \beta^{-1})$$

$$\longrightarrow logp(t \mid x,w,\beta) = \Sigma_{n=1}^{n} log(N(t_n \mid y(x_n,w), \beta^{-1}))$$

$$= \frac{-\beta}{2} \Sigma_{n=1}^{n} (y(x_n,w) - t_n)^2 + \frac{N}{2} log\beta - \frac{N}{2} log(2\pi)$$

$$max \quad logp(t \mid x,w,\beta) = -max \frac{\beta}{2} \Sigma_{n=1}^{n} (y(x_n,w) - t_n)^2$$

$$= min \frac{1}{2} \Sigma_{n=1}^{n} (y(x_n,w) - t_n)^2$$

We minimize $P = \frac{1}{2} \Sigma_{n=1}^{n} (y(x_n,w) - t_n)^2$ to find w. Suppose

$$X = \begin{bmatrix} 1 & x1 \\ 1 & x2 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix}$$

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

$$\longrightarrow P = \| X\mathbf{w} - \mathbf{t} \|_2^2$$

We have:

$$\bigtriangledown P_{\mathbf{w} = X^T(X\mathbf{w} - t) = X^T X\mathbf{w} - X^T \mathbf{t}}$$

Let:

$$\bigtriangledown P_{\mathbf{w}=0}$$

We have:

$$X^T X\mathbf{w} - X^T\mathbf{t} = 0 \Longleftrightarrow w = (X^T X)^{-1} X^T\mathbf{t}$$

IV. Proof $X^T X$ invertable when X full rank

We have:

$$X^T X = \begin{bmatrix} 1 & n\Sigma_{n=1}^{n} x_i \\ \Sigma_{n=1}^{n} x_i & \Sigma_{n=1}^{n} (x_i)^2 \end{bmatrix}$$

We can see that $X^T X$ is a 2x2 matrix so we just prove that rank($X^T X$) = 2, Then we can conclude that $X^T X$ invertable

$$X = \begin{bmatrix} 1 & x1 \\ 1 & x2 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix}$$

X full rank so rank(X) = 2

$$X^T = \begin{bmatrix} 1 & 1 & ....1 \\ x_1 & x_2 & ....x_n \end{bmatrix}$$

Because X full rank so $rank(X)$ = $rank(X^T)$ = $rank(X^T X)$ = 2

So $X^T X$ invertable when X full rank