HW ML 3

Minh Le Quoc

October 2022

1 We will prove that

$$t_n = y(x, w) + \epsilon \Longleftrightarrow W = \left(X^t X\right)^{-1} X^T t$$

2 Problem 1

Now we will minimize:

$$P = \frac{\beta}{2} \sum_{n=1}^{N} \{y(x_n, w) - t_n\}^2$$

Suppose that: $y(x_n, n) = w_1 x_n + w_0$ for linear regression problem

$$x = 1x_1 1x_2 \vdots\vdots 1x_n; \qquad t = t_1 t_2 \vdots t_n; \qquad w = w_0 w_1$$

Then,

$$y = y_1 y_2 \vdots y_n = w_1 x_1 + w_0 w_2 x_2 + w_0 \vdots w_n x_n + w_0 = x.w$$

$$t - y = t_1 - y_1 t_2 - y_2 \vdots t_n - y_n$$

$\longrightarrow t - y_2^2 = (t_1 - y_1)^2 + \cdots + (t_n - y_n)^2$
$= \sum (t_i - y_i)^2 = P$

$$\longrightarrow P = t - y = t - xw = (xw - t)^T (xw - t)$$

Take the derivate of P:

$$\frac{\partial(P)}{\partial(w)} = 2x^T (t - xw) = 0$$

$\leftrightarrow x^t = x^T xw$
$\leftrightarrow w = (x^T x)^{-1} x^T.t$

3 Prove that $X^t X$ is invertable when X full rank

Solution
If X is full rank, X is linear independent.

$$\Rightarrow \vec{v}^T X^T X \vec{v} = \vec{v}^T \overrightarrow{0} = 0 \Rightarrow (X\vec{v})^T X \vec{v} = 0 \Rightarrow (X\vec{v}) \cdot (X\vec{v}) = 0 \Rightarrow X\vec{v} = \overrightarrow{0}$$

So we have: if $\vec{v} \in N\left(X^T X\right) \Rightarrow \vec{v} \in N(X)$

$$\Rightarrow \vec{v} can only be \overrightarrow{0} \Rightarrow N\left(X^T X\right) = N(X) = \{\overrightarrow{0}\}$$

$\Rightarrow$ $X^T X$ is linearly independent; and $X^T X$ is a square matrix $\Rightarrow$ $X^T X$ is invertible