MACHINE LEARNING

Nguyen Quoc Huy
MSV 11205484

HW week3

1 Problem 1

Prove that

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

Now we will minimize:

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

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

$$x = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}; \qquad t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}; \qquad w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

Then,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 x_1 + w_0 \\ w_2 x_2 + w_0 \\ \vdots \\ w_n x_n + w_0 \end{bmatrix} = x.w$$

$$t - y = \begin{bmatrix} t_1 - y_1 \\ t_2 - y_2 \\ \vdots \\ t_n - y_n \end{bmatrix}$$

$$\begin{aligned} \longrightarrow \|t-y\|_2^2 &= (t_1 - y_1)^2 + \cdots + (t_n - y_n)^2 \\ &= \sum (t_i - y_i)^2 = P \end{aligned}$$

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

Take the derivative of P:

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

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

2 Problem 4

Suppose

$$X^T v = 0$$

Then of course

$$XX^T v = 0$$

Conversely, suppose

$$XX^T v = 0$$

Then

$$v^T XX^T v = 0$$

so that

$$(X^T v)^T (X^T v) = 0$$

This implies

$$X^T v = 0$$

Hence, we have proved that

$$X^T v = 0$$

if and only if v is in the nullspace of

$$XX^T$$

But

$$X^T v = 0$$

and

$$v0$$

if and only if X has linearly dependent rows. Thus,

$$XX^T$$

has nullspace 0 (i.e.

$$XX^T$$

is invertible) if and only if X has linearly independent rows.