Machine Learning Homework week 3

Thiều Ngọc Mai

September 2023

Problem 1:

The maximum likelihood function of the Linear Regression is:

$$
\begin{aligned}
p(t|x, w, \sigma^2) &= \prod_{n=1}^{N} \mathcal{N}(t_n|y(x_n, w), \sigma^2) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(t_n - y(x_n, w))^2}{2\sigma^2}\right) \\
\rightarrow \log p(t|x, w, \sigma^2) &= \sum_{n=1}^{N} \left(-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(t_n - y(x_n, w))^2}{2\sigma^2}\right)
\end{aligned}
$$

Since $-\frac{1}{2}\log(2\pi\sigma^2)$ and σ^2 are constant, to maximize the maximum likelihood function, we would maximize $-\sum_{n=1}^{N} t_n - y(x_n, w)^2$ which equal to minimize $\sum_{n=1}^{N}\{t_n - y(x_n, w)\}^2$

$$
\begin{aligned}
L &= \sum_{n=1}^{N}\{t_n - y(x_n, w)\}^2 \\
&= \|t - y\|^2 \\
&= \|t - xw\|^2 \\
\frac{\partial L}{\partial w} &= \frac{\partial(t - xw)}{\partial w} \frac{\partial L}{(t - xw)} = 0 \\
&\leftrightarrow 2X^T(t - Xw) = 0 \\
&\leftrightarrow w = (X^T X)^{-1} X^T t
\end{aligned}
$$

Problem 2:

Suppose $\vec{v} \in N(X^TX)$ and X is full rank (i.e $N(X) = \{0\}$):

$$\begin{aligned} X^TX\vec{v} &= \vec{0} \\ \leftrightarrow \vec{v}^T X^TX\vec{v} &= \vec{v}\vec{0} \\ \leftrightarrow \vec{v}^T X^TX\vec{v} &= 0 \\ \leftrightarrow (X\vec{v})^T X\vec{v} &= 0 \\ \leftrightarrow \|X\vec{v}\|^2 &= 0 \\ \leftrightarrow X\vec{v} &= \vec{0} \end{aligned}$$

Then, $\vec{v} \in N(X^TX)$ and $\vec{v} \in N(X) \rightarrow$ N(X$^T X) = N(X) = \{0\}$
Thus,X^TX is invertible