Homework 1

Mehul Yesminkumar

September 25, 2017

Question 1.1

$D + 1 > N$. When feature is greater than total data.

Question 1.2

$$RSS = \sum_{n=1}^{N}[y_n - (b + \sum_{d=1}^{D} w_d x_{n,d})]^2$$

$$\frac{\partial RSS}{\partial b} = 2\sum_{n=1}^{N}[-y_n + b + \sum_{d=1}^{D} w_d x_{n,d}] = 0$$

$$\implies -\sum_{n=1}^{N} y_n + \sum_{n=1}^{N} b + \sum_{n=1}^{N}\sum_{d=1}^{D} w_d x_{n,d} = 0$$

$$\implies -\sum_{n=1}^{N} y_n + \sum_{n=1}^{N} b + \sum_{d=1}^{D} w_d \sum_{n=1}^{N} x_{n,d} = 0 \qquad but.. \sum_{n=1}^{N} x_{n,d} = 0$$

$$\implies \sum_{n=1}^{N} b = \sum_{n=1}^{N} y_n$$

$$Hence. \quad b = \frac{\sum_{n=1}^{N} y_n}{N}$$

Question 2.1

$$
\begin{aligned}
\min_{w,b} \epsilon(w,b) &= \min_{w,b} -\sum_{n=1}^{N}\{y_n \log(\sigma(w^T x_n + b)) + (1-y_n)\log(1-\sigma(w^T x_n + b))\} \\
&= \min_{w,b} -\sum_{n=1}^{N}\{y_n \log(\sigma(b)) + (1-y_n)\log(1-\sigma(b))\} \\
\frac{\partial}{\partial b}\min_{w,b} \epsilon(w,b) &= -\sum_{n=1}^{N}\{\frac{y_n}{\sigma(b)} + \frac{(-1)(1-y_n)}{1-\sigma(b)}\} \\
&= -\sum_{n=1}^{N}\{y_n(1-\sigma(b)) - (1-y_n)\sigma(b)\} \\
&= -\sum_{n=1}^{N}\{y_n - \sigma(b)\} = 0 \qquad \implies \sum_{n=1}^{N} y_n = \sum_{n=1}^{N}\sigma(b) \\
&\implies \sigma(b) = \frac{1}{N}\sum_{n=1}^{N} y_n \qquad but...\sigma(b) = \frac{1}{1+e^{-b}} \\
&Hence. \quad b = \log(\sum_{n=1}^{N} y_n) - \log(N - \sum_{n=1}^{N} y_n)
\end{aligned}
$$