CSCI 567 Homework 1

Shamim Samadi

September 24, 2017

Problem 1. Linear Regression

- **Question 1.1** When this bad scenario happens, the number of ***nonzero*** singular values of the data matrix X is less than D.

- **Question 1.2** Bias solution

$$\frac{\partial RSS(\tilde{w})}{\partial w_0} = 0 \tag{1}$$

$$\sum_{n=1}^{N} [2(y_n - (w_0 + \sum_{d=1}^{D} w_d x_{nd})))(-1)] = 0 \tag{2}$$

$$\sum_{n=1}^{N} y_n = \sum_{n=1}^{N} (w_0 + \sum_{d=1}^{D} w_d x_{nd}) \tag{3}$$

$$\sum_{n=1}^{N} y_n = \sum_{n=1}^{N} w_0 + \sum_{n=1}^{N} \sum_{d=1}^{D} w_d x_{nd} \tag{4}$$

$$\sum_{n=1}^{N} y_n = N w_0 + \sum_{n=1}^{N} \sum_{d=1}^{D} w_d x_{nd} \tag{5}$$

$$\sum_{n=1}^{N} y_n = N w_0 + \sum_{d=1}^{D} \sum_{n=1}^{N} w_d x_{nd} \tag{6}$$

$$\sum_{n=1}^{N} y_n = N w_0 + \sum_{d=1}^{D} w_d \sum_{n=1}^{N} x_{nd} \tag{7}$$

$$b^* = w_0 = \frac{1}{N} \sum_{n=1}^{N} y_n - \sum_{d=1}^{D} w_d (\frac{\sum_{n=1}^{N} x_{nd}}{N}) \tag{8}$$

Now since $\frac{\sum_{n=1}^{N} x_{nd}}{N} = 0; \forall d = 1, 2, ..., D$ (assumption in equation (6)), we have: $w_d(\frac{\sum_{n=1}^{N} x_{nd}}{N}) = 0; \forall d = 1, 2, ..., D$. And thus: $\sum_{d=1}^{D} w_d \frac{\sum_{n=1}^{N} x_{nd}}{N} = 0$, which leads to

$$b^* = \frac{1}{N} \sum_{n=1}^{N} y_n \tag{9}$$

Problem 2. Logistic Regression

- **Question 2.1**

  From the lectures, assuming $\sigma(a) = \frac{1}{1+e^{-a}}$, we know: $\frac{d\sigma(a)}{da} = \sigma(a)[1-\sigma(a)]$
  Also, due to no access to any features ($D = 0$): $\sigma(x_n) = \sigma(\sum_{d=1}^{D} w_d x_{nd} + b) = \sigma(b)$

$$\frac{\partial E}{\partial b} = \frac{\partial}{\partial b} \sum_{n=1}^{N} [y_n \log \sigma(b) + (1-y_n) \log(1-\sigma(b))] = \sum_{n=1}^{N} [y_n(\frac{\frac{\partial}{\partial b}\sigma(b)}{\sigma(b)}) + (1-y_n)(\frac{\frac{\partial}{\partial b}(1-\sigma(b))}{1-\sigma(b)})] \tag{10}$$

$$\frac{\partial E}{\partial b} = \sum_{n=1}^{N} [y_n \frac{\sigma(b)(1-\sigma(b))}{\sigma(b)} + (1-y_n)\frac{-\sigma(b)(1-\sigma(b))}{1-\sigma(b)}] = \sum_{n=1}^{D} (y_n - \sigma(b)y_n - \sigma(b) + y_n\sigma(b)) \tag{11}$$

$$\frac{\partial E}{\partial b} = \sum_{n=1}^{D} (y_n - \sigma(b)) = 0 \to \sum_{n=1}^{N} y_n - \sum_{n=1}^{N} \sigma(b) = 0 \Rightarrow \sigma(b^*) = P(y_n = 1|b) = \frac{1}{N}\sum_{n=1}^{N} y_n \tag{12}$$

  or: $b^* = \log(\frac{1}{N}\sum_{n=1}^{N} y_n) - \log(1 - \frac{1}{N}\sum_{n=1}^{N} y_n)$