

Ryan Gosiaco
912819444
5/1/2018

STA 141C HW 1 Report

1.1) The weight vector I got from each step in order from largest step size to smallest is
[0.06769835, 0.04512519, 0.00702038, #0.00073691, 7.40524083e-05, 7.40886683e-06].

1.2) The average MSE I obtained was 5.53147692e-05.

1.3) I tried using a step size of 0.01 but was unable to obtain an MSE due to memory issues with my implementation of gradient descent.

2.1)

Handwritten mathematical derivations for a logistic regression cost function:

$$f(w) = \frac{1}{n} \sum_{i=1}^n \log(1 + \exp\{-y_i w^T x_i\}) + \frac{1}{2} \|w\|^2$$
$$\frac{\partial f(w)}{\partial w_j} = \frac{1}{n} \sum_{i=1}^n \frac{1}{1 + \exp\{-y_i w^T x_i\}} \exp\{-y_i w^T x_i\} (-y_i x_{ij}) + w_j$$
$$\|w\|^2 = w_1^2 + w_2^2 + \dots + w_p^2 \quad \left(\frac{e^x}{e^x} \right) \frac{e^{-x}}{1 + e^x} = \frac{1}{1 + e^x}$$
$$\frac{\partial f}{\partial w_j} = \frac{1}{n} \sum_{i=1}^n \frac{-y_i x_{ij}}{1 + \exp\{y_i w^T x_i\}} + w_j$$
$$\left[\frac{1}{n} \sum_{i=1}^n \frac{-y_i x_{i1}}{1 + \exp\{y_i w^T x_i\}} + w_1, \dots, \frac{1}{n} \sum_{i=1}^n \frac{-y_i x_{ip}}{1 + \exp\{y_i w^T x_i\}} + w_p \right]^T$$

2.2) Unfortunately I was unable to implement this part due to time constraints.