Homework 1

Wangqian Miao, UC Davis

April 30, 2018

Instruction

- My python version is 3.6. I run the code on my laptop. Its information is CPU: i5-6200U, RAM size: 8GB.
- I run the code on spyder. So, there are some "#%%" in my code, which means a cell (like Jupyter). I use the cells to debug and test my code.

1 Problem 1: Linear Regression

1.1

The gradient is as following.

$$\nabla f(\boldsymbol{\omega}) = \frac{2}{N}(\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{\omega} - \boldsymbol{y})) + \lambda\boldsymbol{\omega} \tag{1}$$

After doing the normalization of the data, we can get the weight vector under different learning rate when the stop condition is $\epsilon = 0.001$.

Learning Rate	Time(s)	Iter	Result
10^{-7}	too long	–	Not Suggested
10^{-6}	too long	–	Not Suggested
10^{-5}	298	690073	Good
10^{-4}	29	68996	Good
10^{-3}	3	6898	Not Bad
10^{-2}	≈ 0	687	Rough

Table 1: Learning result of differdent learning rate on cpusmall.txt

In conclusion, I will choose $\eta = 10^{-3}, 10^{-4}, 10^{-5}$. The weight vector is in the Appendix.

1.2

The matrix form of MSE is that $\|\boldsymbol{X}\boldsymbol{\omega}^* - \boldsymbol{y}\|^2/N$, in which $\boldsymbol{\omega}^*$ is the weight vector we learn.

In this part, the step size and error control are "$\eta = 0.001, \epsilon = 0.001$". After doing the cross validation, the MSE is 2.593×10^{-4}.

1.3

In this part, the step size and error control are "$\eta = 0.001, \epsilon = 0.001$". On the test dataset the MSE is 6.244×10^{-3}.

2 Problem 2: Logistic Regression

2.1

The gradien is as following.

$$\begin{aligned} \nabla f(\boldsymbol{w}) &= -\frac{1}{N}\sum_{i=1}^{N}\frac{y_i \boldsymbol{x_i}}{1+\exp\left(y_i \boldsymbol{\omega^T x_i}\right)} + \lambda\boldsymbol{\omega} \\ &= -\frac{1}{N}\boldsymbol{X^T k} + \lambda\boldsymbol{\omega} \end{aligned} \tag{2}$$

In this expression, $\boldsymbol{k}$ is a column vector and $k_i = y_i/(1+\exp\left(y_i \boldsymbol{\omega^T x_i}\right))$.

2.2

In this part. the accuracy I get is 93.125% under the parameters as "$\eta = 0.01, \epsilon = 0.001$".

Appendix

$\eta = 10^{-3}$: [[0.00101373] [0.00088652] [0.00039534] [0.00095895] [0.00088306] [0.00046597] [0.00014014] [0.00083379] [0.00069961] [0.00017081] [0.00047319] [0.00054504]]

$\eta = 10^{-4}$: [[0.00017468] [0.00082189] [0.00097859] [0.00109373] [0.00033627] [0.00058696] [0.00052377] [0.00086618] [0.00021178] [0.00029312] [0.00052432] [0.00063001]]

$\eta = 10^{-5}$: [[0.00050466] [0.00102257] [0.00111366] [0.00082273] [0.00041286] [0.00094829] [0.00098085] [0.00076782] [0.0009305] [0.00050751] [0.00020548] [0.0009205]]