ECSE 4965: Program Assignment 1

Name: Li Dong
RIN: 661243168

February 8, 2017

a. The closed form solution for $\boldsymbol{\Theta}$ is $\boldsymbol{\Theta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$, where

$$\boldsymbol{X} = \begin{Bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \vdots \\ \boldsymbol{x_m} \\ \boldsymbol{1} \end{Bmatrix}_{m+1\times 10} , \boldsymbol{y} = \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \\ 1 \end{Bmatrix}_{m+1\times 1} .$$

$$\boldsymbol{\Theta} = \begin{Bmatrix} 0.43672635 \\ 0.73828806 \\ 0.24621396 \\ 0.78105699 \\ 0.01109724 \\ 0.96073401 \\ 0.17140767 \\ 0.44954264 \\ 0.81574254 \\ 0.92060012 \\ 0.08530772 \end{Bmatrix}$$

b. With learning rate= 0.005 and tolerance= 0.01, the iteration ends at 6939 and the following $\boldsymbol{\Theta}$ is obtained via hand-coded gradient descent method.

$$\boldsymbol{\Theta} = \begin{Bmatrix} 0.43653088 \\ 0.73809711 \\ 0.24602746 \\ 0.78086202 \\ 0.01091228 \\ 0.96054382 \\ 0.17120545 \\ 0.44934754 \\ 0.81555151 \\ 0.92039673 \\ 0.09528786 \end{Bmatrix}$$

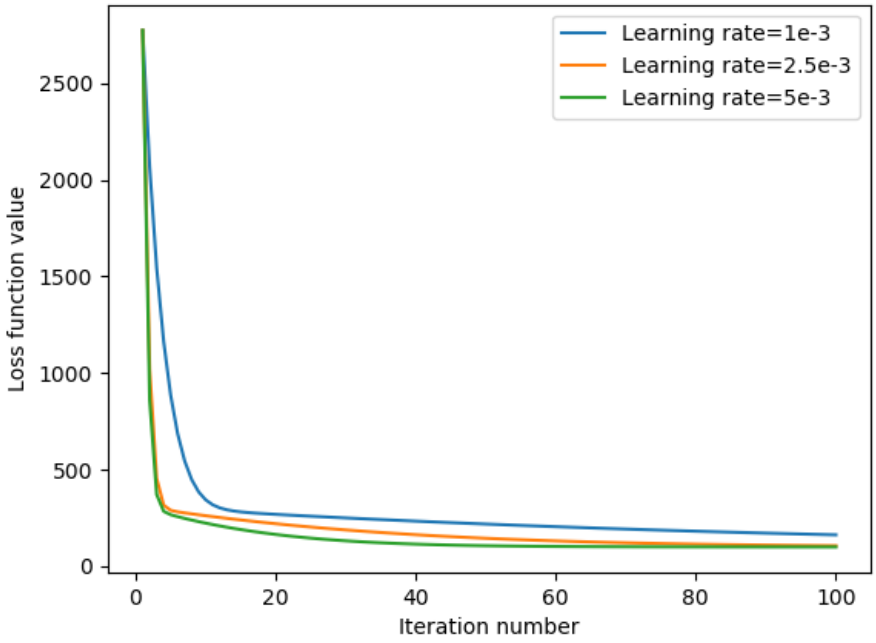

Figure 1

The loss function value change as a function of iteration number with three different learning rates are shown in Figure 1.

c. With learning rate= 0.05, tolerance= 0.01 and batch size= 15, the iteration ends at 461141 and the following $\boldsymbol{\Theta}$ is obtained via hand-coded stochastic gradient descent.

$$\boldsymbol{\Theta} = \begin{Bmatrix} 0.43532825 \\ 0.73783443 \\ 0.24686164 \\ 0.7818615 \\ 0.01174988 \\ 0.96017151 \\ 0.17176041 \\ 0.44843518 \\ 0.81546001 \\ 0.92034 \\ 0.09502804 \end{Bmatrix}$$

The loss function value change as a function of iteration number with three different batch sizes are shown in Figure 2. The learning rate is fixed to 0.05.

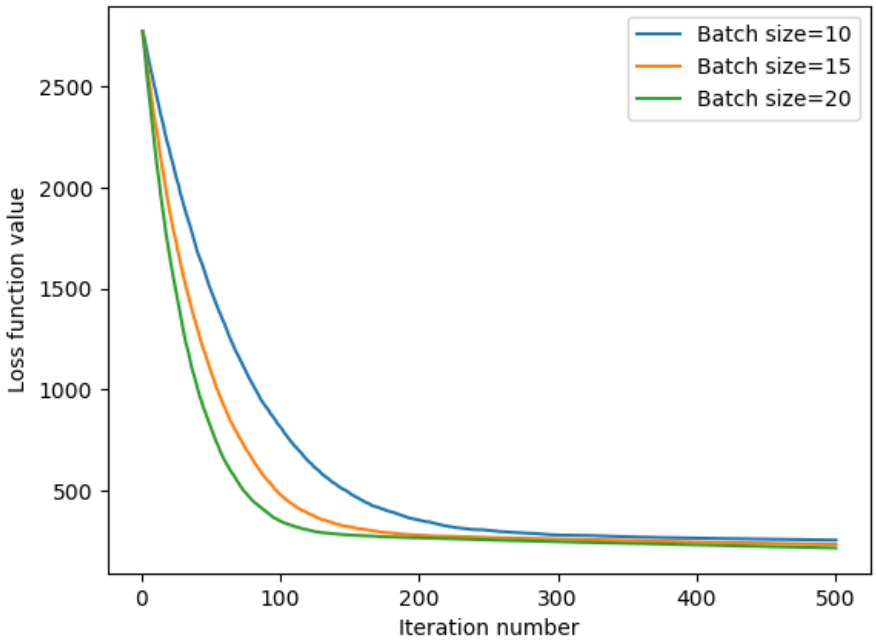
Batch size=10
Batch size=15
Batch size=20
2500
2000
1500
1000
500
Loss function value
0
100
200
300
400
500
Iteration number

Figure 2