DD2424 Assignment 1 Report

Chunyang Han

chunyang@kth.se

1 Exercise1

1.Graphs of the total loss and cost function on the training data and the validation data after each epoch of the mini-batch gradient descent algorithm.
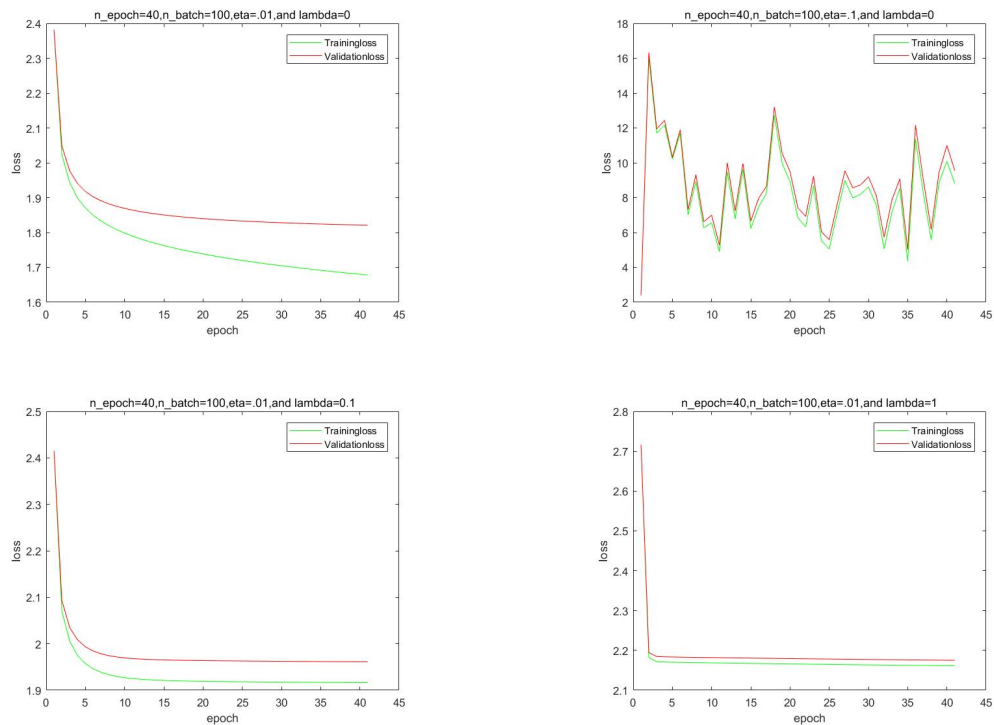

Figure 1: Graphs of loss on the training data and validation data after each epoch

2.Images representing the learnt weight matrix after the completion of training for the following parameter settings, And the final test accuracy.

Figure 2: lambda=0, n epochs=40, n batch=100, eta=.1 ;accuracy=0.2559

3.A short comment on the effect of increasing the amount of regularization and the importance of the correct learning rate.

Figure 3: lambda=0, n epochs=40, n batch=100, eta=.01 ;accuracy=0.3672

Figure 4: lambda=.1, n epochs=40, n batch=100, eta=.01 ;accuracy=0.3337

Figure 5: lambda=1, n epochs=40, n batch=100, eta=.01 ;accuracy=0.2191

-When regularization amount is increasing, the training loss and validation loss drop faster. What's more,the corresponding values of them both increase and become closer to each other when the curves tend to be stable.

-The learning rate determines the speed of learning, if it is too big then we would miss the optimal result, i.e. over shoot; if it is too low, the gradient descent speed would be very slow. Hence, a correct learning rate is very important.