

ECSE 4965: Program Assignment 2

Name: Li Dong

RIN: 661243168

February 22, 2017

It is noted that the image data are converted to a compact .txt file for easy manipulation. The preprocessing functions are also included in the source code.

1.

Plots for \mathbf{W}_k can be seen from Figure 1.

The learning rate is set to 1.

The weights are saved in train_W_GD.txt.

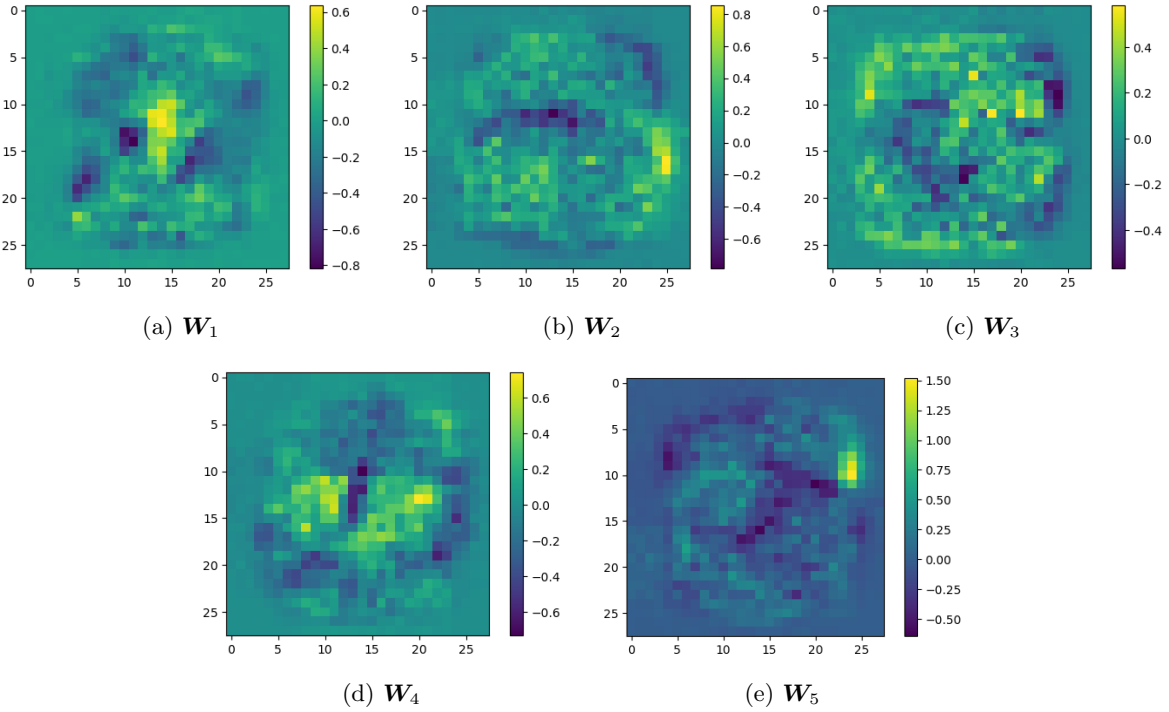


Figure 1: Images for \mathbf{W}_k

2. Plots for \mathbf{W}_k for this problem is omitted since the visual difference is minimal comparing to Figure 1. The weight file is also not provided since the accuracy is no better than problem 1.

The learning rate is set to 0.5, batch size 100 and regularization parameter 1e-3.

3. Gradient descent: By applying the weights trained from 1 to the test data, we can get the following results: average error=4.597%, error of digit 1=1.316%, error of digit 2=5.051%, error of digit 3=6.893%, error of digit 4=2.645%, error of digit 5=7.431%

Stochastic gradient descent: By applying the weights trained from 2 to the test data, we can get the following results: average error=5.259%, error of digit 1=1.598%, error of digit 2=5.859%, error of digit 3=7.184%, error of digit 4=2.950%, error of digit 5=9.180%.