

COL 744 - Machine Learning  
**Assignment 2**

Tanuma Patra

March 17, 2016

1. (a) Dual Objective can be expressed in the form

$$\alpha^T Q \alpha + b^T \alpha + C \quad (0.1)$$

where

$$Q = -.5 * (YY^T) .* (XX^T) \quad (0.2)$$

Number of support vectors(train) = 272

(considered values of alpha between 0.0005 and 0.9999 for support vectors )

- (b) Calculated 'w' is reported in file 'Q1\_A\_w.mat'.

Intercept term , b = -1.5557.

Average Accuracy = 98.33%

- (c) For Gaussian Kernel , K is a m\*m matrix where m is total number of training example where K(i,j) is given as :

$$K(i,j) = \exp(-\gamma * |x^{(i)} - x^{(j)}|) \quad (0.3)$$

and Q is given as

$$Q = -.5 * (YY^T) .* K \quad (0.4)$$

Number of support vectors(train) = 470

(considered values of alpha between 0.0005 and 0.9999 for support vectors - reported in file 'Q1\_B\_svm.mat')

Obtained Accuracy = 98.0745%

Accuracy obtained linear SVM is less than the accuracy obtained in Gaussian Kernel SVM.

- (d) Using LibSVM library :

Number of support vectors linear Kernel = 319

Number of support vectors Gaussian Kernel = 646

Average Accuracy Linear Kernel = 98.202%

Average Accuracy Gaussian Kernel = 98.7136%

The indexes of the support vectors are stored in Q1d\_gaussian\_sv.mat and Q1d\_linear\_sv.mat respectively.

It can be observed that the difference in accuracies obtained from CVX package and libSVM package vary slightly. The libSVM package is slightly more accurate compared to the CVX package. This is because, for libSVM the no. of support vectors being reported are way more compared to CVX as the support vectors right at the border are accounted for in libSVM which is not done in CVX.

- (e) The method followed by me is substituting the unknown values with the average of the remaining values in the 3 respective columns of the train data. The accuracies obtained are as follows:  
Average Accuracy Linear Kernel = 98.202%  
Average Accuracy Gaussian Kernel = 98.7136%  
As can be seen from the values, there is no change in the average accuracy obtained.
2. (a) The script to generate the output has been added to the folder. It takes input from a CVX file (input.csv). The file provides with the index of the digits representation which is used by the program to display the digits at those positions.
- (b) Neural network is created and trained for classes 8 and 3 and trained accordingly. Stochastic gradient descent has been used. The algorithm stops when the difference between the  $j$ th values of two consecutive iterations becomes less than  $10^{-8}$ . Also there is an upper bound to the number of iterations, 35000 for the given data, so that the case of oscillating without converging is taken into account.
- (c) The network has an accuracy of 97.03 percent. The training time is around 3-5 minutes.
- (d) 10 perceptrons have been used in the final layer. The time taken for the training process to complete was 928.217754 seconds. The accuracy fluctuates between 91% to 94% with the accuracy of the last run being 92.3% .