# Assignment 2 COL-774

Abhishek Agarwal - 2014MCS2114

March 16, 2016

## 1 Question 1

### 1.1 Part a

C = 1 has been used.
Support vectors obtained from training using linear kernel have been attached in a separate file.
Number of support vectors when threshold is 1e-4 : 281.

### 1.2 Part b

Vector w has been saved in a separate file.
Value of b is: 1.5557.
Accuracy obtained : 98.33/100

### 1.3 Part c

C = 1 has been used.
Support vectors obtained from training using gaussian kernel have been attached in a separate file.
Number of support vectors when threshold is 1e-4 : 435.
Value of b is : 0.8377
Accuracy obtained : 98.72/100
Accuracy obtained here is more than the linear kernel.

### 1.4 Part d

Support vectors (indices) obtained from training using linear and gaussian kernel have been attached in separate files for libSvm.
Number of support vectors when linear kernel : 329.
Number of support vectors when gaussian kernel : 663.
Accuracy obtained when linear kernel : 97.8177/100
Accuracy obtained when gaussian kernel : 98.7163/100
LibSvm performed faster than the CVX package and give almost similar accuracy in both the cases. Here also accuracy obtained using gaussian is more than the linear kernel.
C=1 was used here too for both kernels. All calculations for gaussian were done with gamma = 2.5e-4.

# 2 Question 2

## 2.1 Part a

The code for such a situation can be found in render.m.
Will take file-name(test4,train5,etc.) and index as input.

## 2.2 Part b

The data for classes 3 and 8 from the original MNIST data have been extracted using extract38.m.
Stopping criteria used is: error-difference to be accounted as 0.0001 between two successive error costs.

## 2.3 Part c

Training time for binary classification: 1458 seconds
Accuracy obtained against the test data: 97.43/100

## 2.4 Part d

Used script extract.m to segregate features and labels on original MNIST data.
Number of units in output layer used for multi-class: 10
Training time: 8168.02 seconds
Accuracy against the test data: 93.4/100
Training time is more as the size of data is more and dimensions increase many-fold here.