**CS542 Machine Learning**

**Homework 4**

**Chih Wei Tung (U17550453)**

**6.2**

(201)

So, the learning algorithm depends on Gram matrix

**7.3**

Given a data set of two data point and , maximum margin hyperplane is determined by solving

,

Solve

Taking the derivative

For b, use (216) and (217)

Use (220)

**7.4**

From figure 4.1 and (7.4)

From (7.16) and (7.7)

Use this with (7.8) and (7.10)

**Programming**

MINST\_data.mat file contain 4 matrixes, train\_samples, train\_samples\_labels, test\_samples and test\_samples\_labels. Each row in both train and test samples are digits of handwriting, and both train and test sample labels are labels of 0 to 9.

1. **One versus one**

The accuracy is 93.4%.

1. All the data in partitions should be cluster to an array[0:10]
2. Separate data to subset
3. Do the training of SVM
4. Get the prediction and calculate the confusion matrix and accuracy

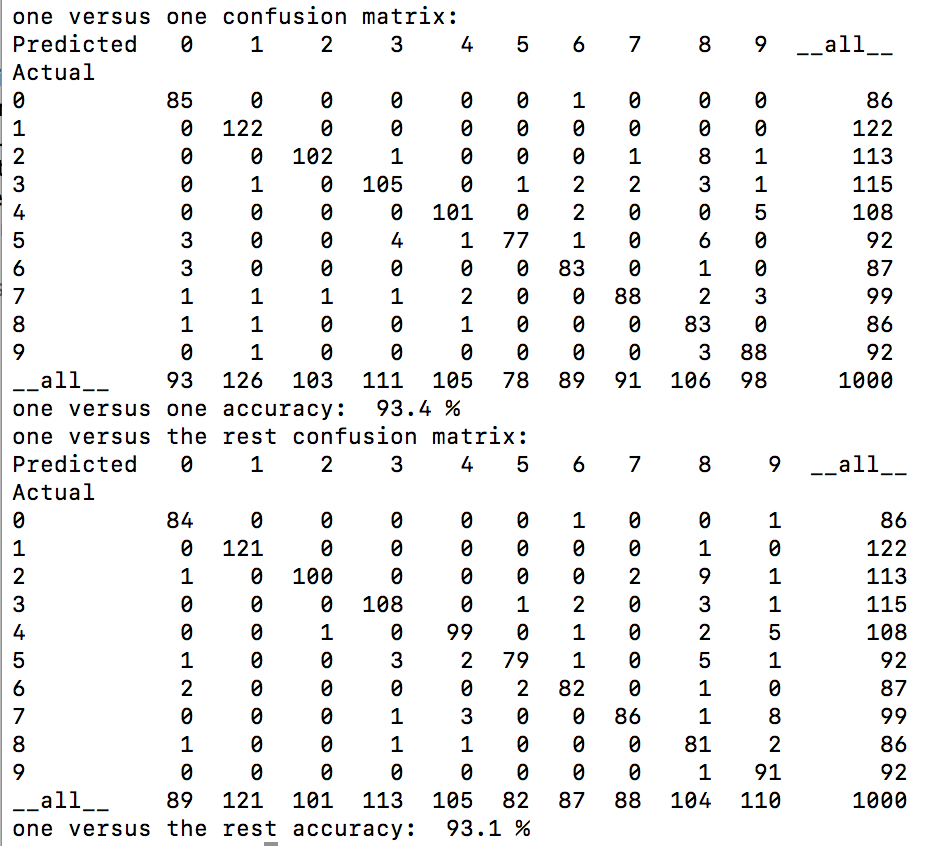
**One versus the rest**

The accuracy is 93.1%

1. All the data in partitions should be cluster to an array[0:10]
2. Separate data to subset
3. Build data that are [0] to the rest, [1] to the rest, …, [9] to the rest
4. Do the training of SVM
5. Get the prediction and calculate the confusion matrix and accuracy

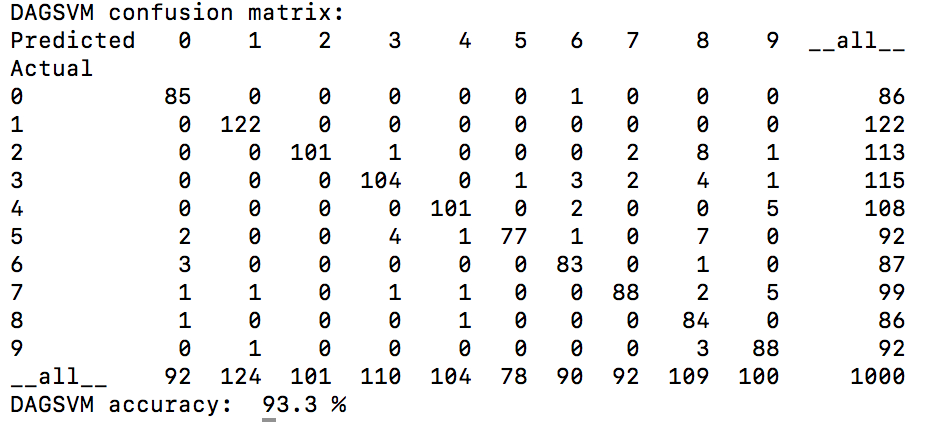
The confusion matrix and accuracy of on versus one and one versus the rest are

shown as below.



1. **DAGSVM**
2. All the data in partitions should be cluster to an array[0:10]
3. Separate data to subset
4. Do the training of SVM
5. Prediction is based on using decision tree

The confusion matrix and accuracy of DAGSVM is shown as below.

****