PATTERN RECOGNITION

PROGRAMMING ASSIGNMENT – 2

GROUP 16 – REPORT

ED14B057-MUHAMMEDH AJWAHIR

SHANMUKH CHAITHANYA

SHUBHAM

Question 5

In the image classification datasets PCA is done on the whole dataset.

* We computed the 48 dimensional mean vector
* Computed covariance matrix
* Computed the eigen vectors and the corresponding eigenvalues
* Sorting the eigen values in decreasing order
* Choose k eigenvectors with largest eigen values
* And the samples were transformed to those k eigen vectors

Variance is calculated as –

* We know J is the sum of k eigen values, where J is the mean square error cost.
* If J’ is the sum of all the eigen values of the covariance matrix then the variance retained can be calculated as-
  + (1-J/J’)\*100

For the dataset given to us- 99.1% of variance is retained for **31 dimensions.**

98.19% of variance is retained for **25 dimensions**

96.9% of variance is retained for **20 dimensions**

So, we choose 31 dimensions and check the behaviour by further reducing the dimensions

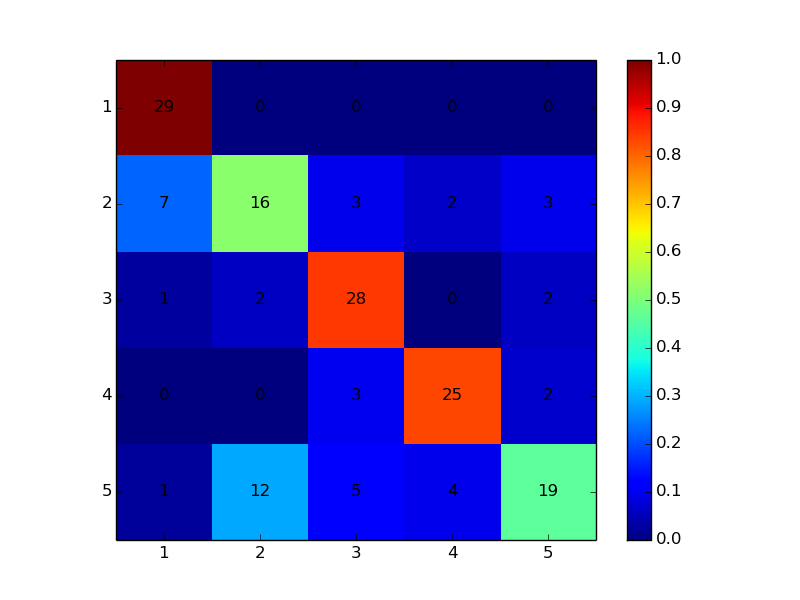
Classificaition with C-SVM with Gaussian kernel

For 48 dim

Accuracy: 71.3414634146

parameters used : gamma value= 1e-11, cost 90

it is seen that if we change the gamma value the accuracy goes down

****