INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

EE 657:PRML

ASSIGNMENT

Name	Roll No.
Abhishek Kumar	140108003
Ayush Anshul	140108052

TASK-1

1.A

i. When samples are modelled by separate covariance matrix.

Accuracy obtained for CLASS1('e') is 89%.

Accuracy obtained for CLASS2('c') is 93%.

Accuracy obtained for CLASS3('l') is 100%.

Hence, average accuracy for first classifier is 94%.

ii. When samples are pooled together to get a common diagonal covariance matrix.

Accuracy obtained for CLASS1('e') is 86%.

Accuracy obtained for CLASS2('c') is 85%.

Accuracy obtained for CLASS3('l') is 100%.

Hence, average accuracy for first classifier is 90.33%.

iii. Covariance matrix is forced to be identity matrix.

Accuracy obtained for CLASS1('e') is 87%.

Accuracy obtained for CLASS2('c') is 85%.

Accuracy obtained for CLASS3('l') is 100%.

Hence, average accuracy for first classifier is 90.67%.

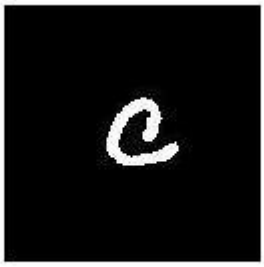
1.B

Four Misclassified image from each classifier together with classifier label and correct label is also in the folder './problem1/misclassified_images'. We observe that the similar images have higher chances of misclassification.

<u>Four samples which are misclassified by classifier 1 are :-</u>

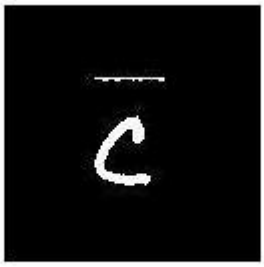
Classifier :1 Class : e, Misclassified to :c

Classifier :1 Class : c, Misclassified to :e



Classifier :1 Class : e, Misclassified to :c

Classifier :1 Class : c, Misclassified to :l

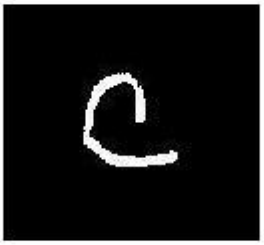


Four samples which are misclassified by classifier 2 are :-

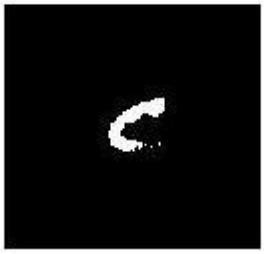
Classifier :2 Class : e, Misclassified to :c

Classifier :2 Class : e, Misclassified to :c

Classifier :2 Class : c, Misclassified to :l



Classifier :2 Class : c, Misclassified to :l

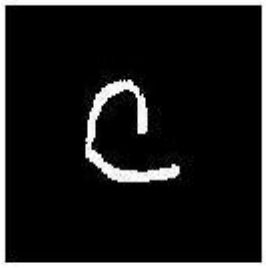


Four samples which are misclassified by classifier 3 are :-

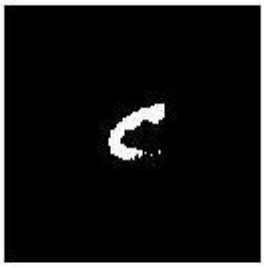
Classifier :3 Class : e, Misclassified to :c

Classifier :3 Class : e, Misclassified to :c

Classifier :3 Class : c, Misclassified to :l



Classifier :3 Class : c, Misclassified to :l



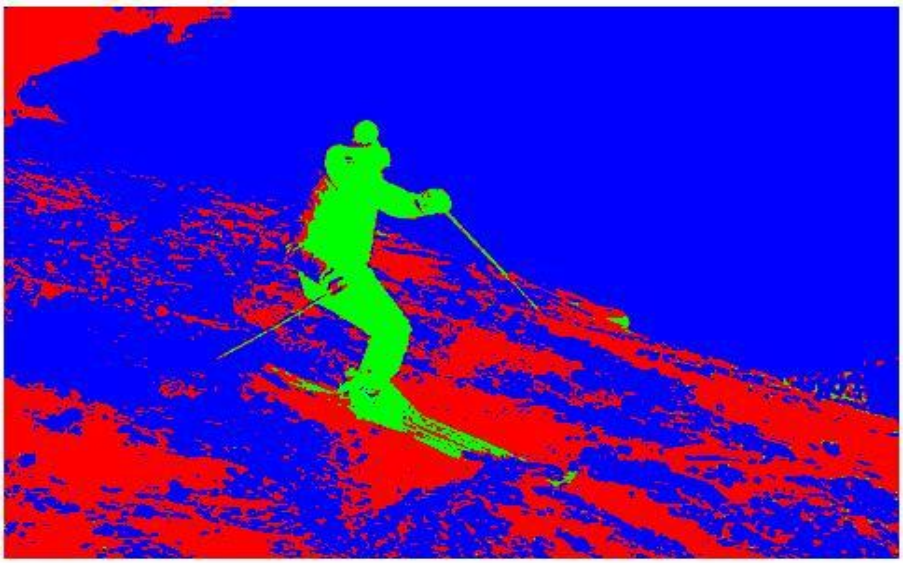
- Code and results of this problem are also saved in the folder './problem1'.To see the results run 'main_file.m' in the folder './problem1'.

TASK-2

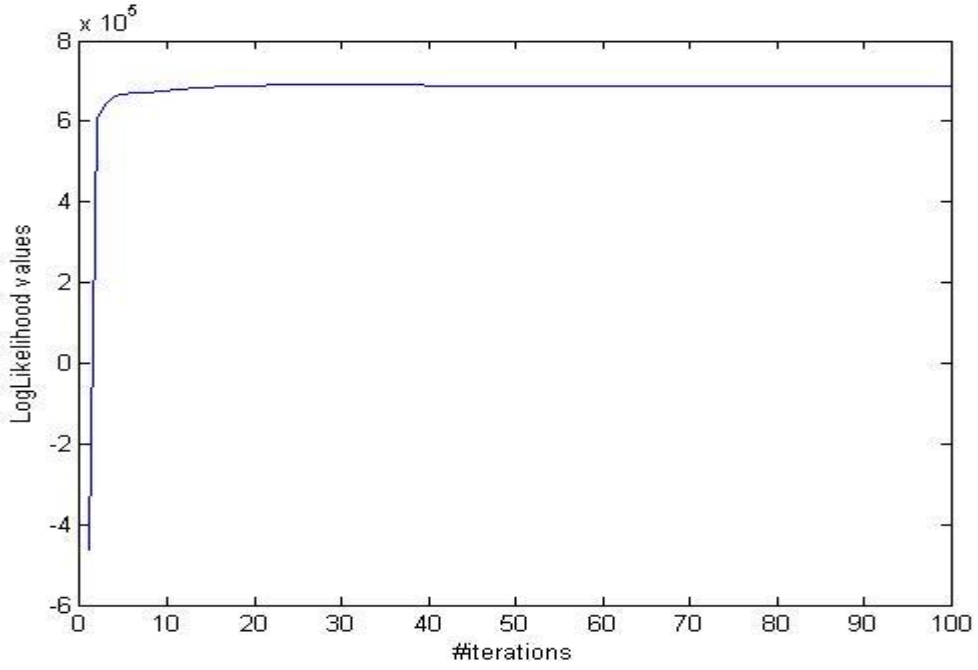
(GMM BASED CLUSTERING)

Segmented output when we ran 100 iterations and after convergence assigned mean of the Gaussian with highest responsibility to each pixel.

Segmented output when we ran 100 iterations and after convergence assigned full R,G,B values of the Gaussian with highest responsibility to each pixel.



Graph of convergence of log-likelihood values (note that we get convergence only after approximately 10 iterations):-



Code and results of this problem are also saved in the folder './problem2'.To see the results run './problem2/main_file.m'.

- './problem2/100ski_image_loglikelihood.jpg' is the graph of log-likelihood vs. number of iterations.
- './problem2/100ski_image_segmented.jpg' is the segmented image after 100 iterations where mean of Gaussian with highest responsibility is assigned to each pixel.
- './problem2/100ski_image_RGB_segmented' is the segmented image after 100 iterations where full R,G,B values of the Gaussian with highest responsibility is assigned to each pixel.

TASK-3

(Face Recognition Using PCA)

i. Eigenface images corresponding to top 5 eigen values are :-

Eigen Face: 1

Eigen Face: 2

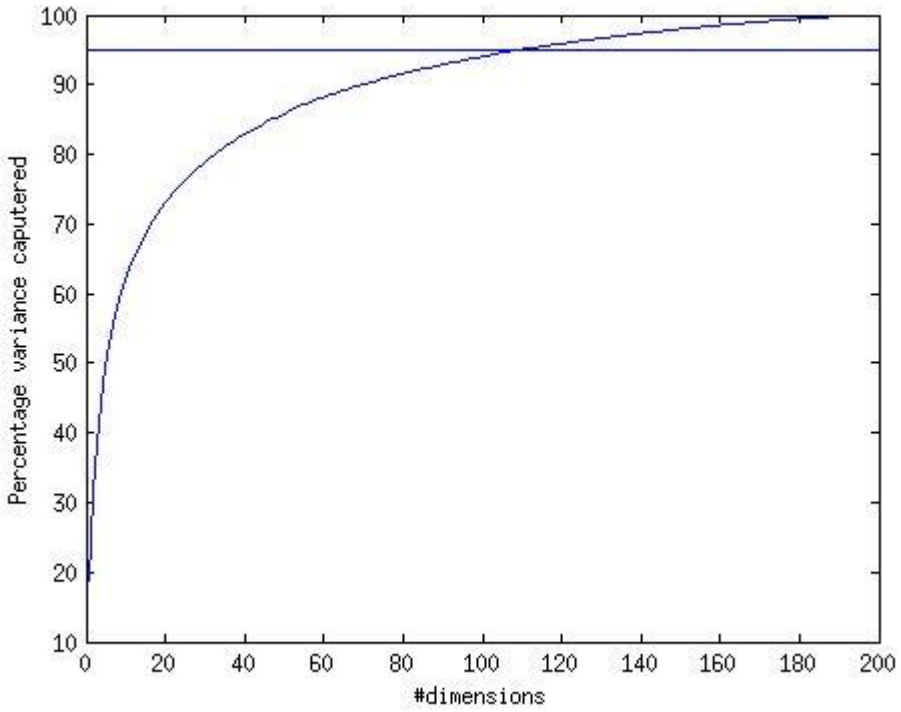
Eigen Face: 3

Eigen Face: 4

Eigen Face: 5

The images are blurred and hence are also called 'ghost faces'. The images are also saved inside the folder './problem3/ TOP5_eigenface_PartA '.

ii. The required graph is :-



Note that 95% variance is captured when eigenvectors corresponding to top 110 eigenvalues are used.

This graph is also saved in './problem3/variance_vs_dimensions'.

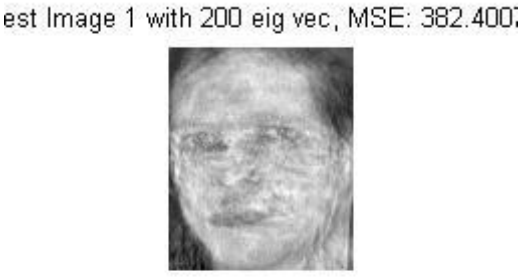
iii.a. For 'face_input_1.pgm', Eigenface corresponding to largest eigen value is:-

In this case the MSE is 1767.2424.

iii.b. For 'face_input_1.pgm', Eigenface corresponding to top 15 eigen value is:-

In this case the MSE is 797.023.

iii.c. For 'face_input_1.pgm', Eigenface corresponding to all eigen value is:-

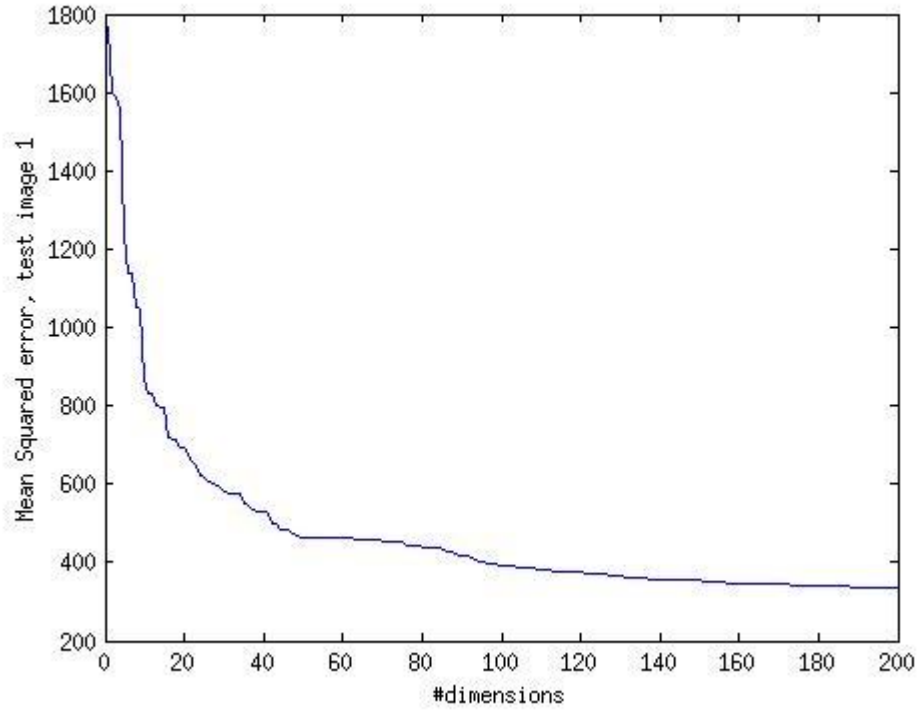


In this case the MSE is 382.400.

MSE decreases as we use more and more eigen vectors but note that even on using all the eigenvectors for 'face_input_1.pgm' we do not get full reconstruction as the given image is not from within the training set.

- './problem3/ face_input_1r_1.pgm' is reconstructed image of 'face _input_1.pgm' using top 1 eigenvalues.
- './problem3/ face_input_1r_15.pgm' is reconstructed image of 'face _input_1.pgm' using top 15 eigenvalues.
- './problem3/ face_input_1r_200.pgm' is reconstructed image of 'face _input_1.pgm' using all eigenvalues.

iv. The graph of MSE vs. the number of Eigen faces used is :-



- This graph is also saved in './problem3/image1_mse'.

MSE decreases sharply initially but as later the eigen values become smaller and smaller the impact becomes lesser. Even on using all eigenvectors we still have a MSE of ~382 as the image is new and not from training set.

iii.a. For 'face_input_2.pgm', Eigenface corresponding to largest eigen value is:-

In this case the MSE is 1766.6327.

iii.b. For 'face_input_2.pgm', Eigenface corresponding to top 15 eigen value is:-

In this case the MSE is 384.1851.

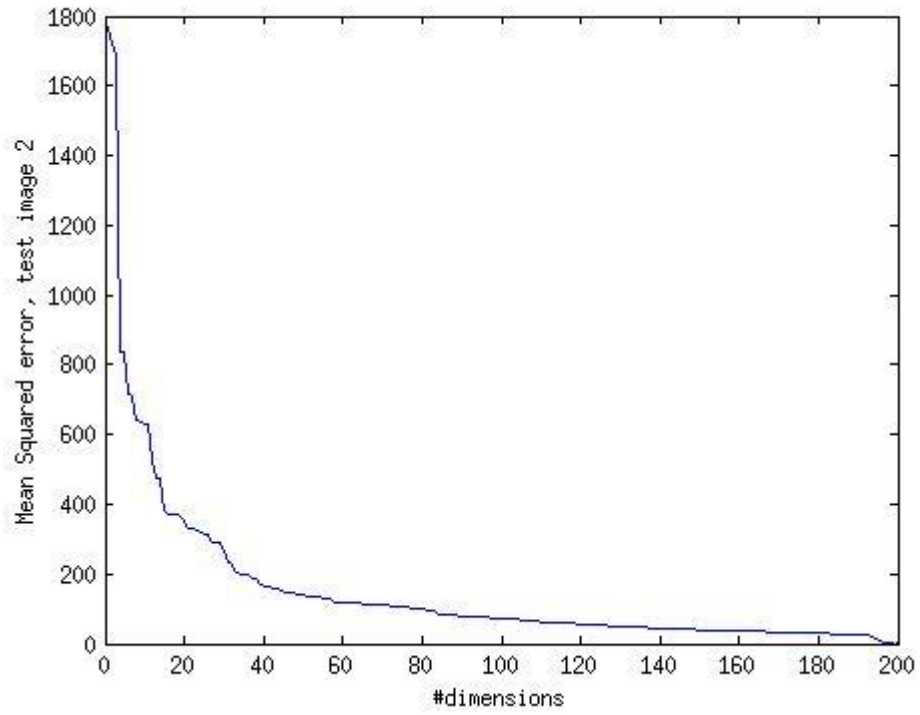
iii.c. For 'face_input_2.pgm', Eigenface corresponding to all eigen value is:

In this case the MSE is 32.197.

MSE decreases as we use more and more eigen vectors and here we get nearly a full reconstruction of image 'face_input_2.pgm' as the test image is from within the training dataset.

- './problem3/ face_input_2r_1.pgm' is reconstructed image of 'face _input_2.pgm' using top 1 eigenvalues.
- './problem3/ face_input_2r_15.pgm' is reconstructed image of 'face _input_2.pgm' using top 15 eigenvalues.
- './problem3/ face_input_2r_200.pgm' is reconstructed image of 'face _input_2.pgm' using all eigenvalues.

iv. The graph of MSE vs. the number of Eigen faces used is :-



- This graph is also saved in './problem3/image2_mse'.

MSE decreases sharply initially but as later the eigen values become smaller and smaller the impact becomes lesser. Here on using all the eigen vectors we observe that MSE nearly approaches zero.

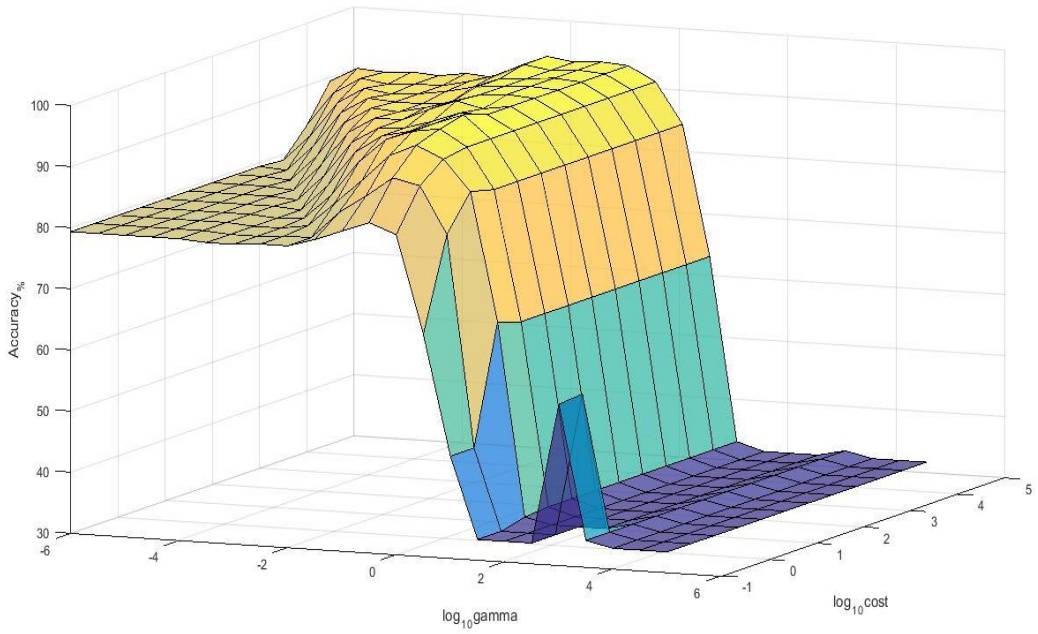
- Code and results of this problem are also saved in the folder '/problem3'. To see the results run './problem3/main_file,m'.

TASK-4

(Support Vector Machines)

Here we used Radial Basis Function Kernel. Varying C(penalty factors/cost) and gamma we plotted a 3D-plot showing the variation of accuracy with them.

The obtained 3D plot is :-



Gamma and Cost factors are varied on log scale. Gamma is varied from 10^{-6} to 10^{5} with a step size of $10^{0.5}$. Cost factor 'C' is varied from 10^{-1} to $10^{4.5}$ with a step size of $10^{0.5}$. On increasing gamma accuracy in prediction decreases due to over-fitting of data.

- Code and results of this problem are also saved in the folder '/problem4'. To see the results run './problem4/main_file.m'. The folder './problem4' also contains the required libsvm binaries to run this file. Accuracy plots are also saved in .fig format to open in matlab.