PATTERN RECOGNITION! S.Natchanon (MSc.)

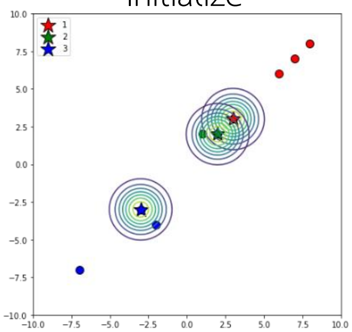
HOMEWORK3 6470177521

PART: GMM

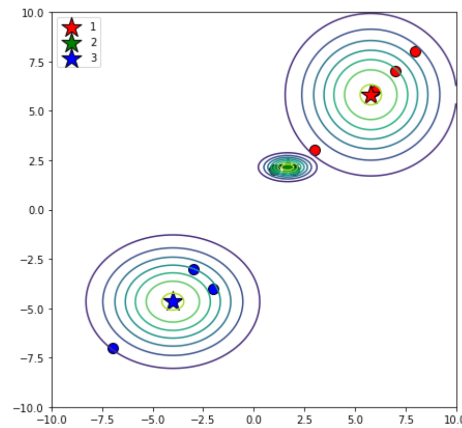
T1: Using 3 mixtures, Repeat three iterations of EM. Write down each w, m, mu, var

- **Ans:** Iteration 1:

Initialize

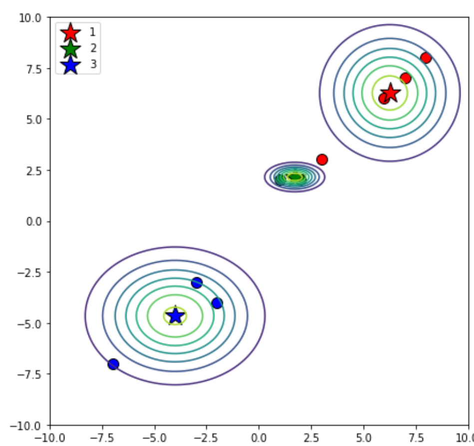



```
Point [1 2] : Assign to G1: 0.12 G2: 0.88 G3: 0.00
Point [3 3] : Assign to G1: 0.73 G2: 0.27 G3: 0.00
Point [2 2] : Assign to G1: 0.27 G2: 0.73 G3: 0.00
Point [8 8] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [6 6] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [7 7] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [-3 -3] : Assign to G1: 0.00 G2: 0.00 G3: 1.00
Point [-2 -4] : Assign to G1: 0.00 G2: 0.00 G3: 1.00
Point [-7 -7] : Assign to G1: 0.00 G2: 0.00 G3: 1.00

Prior(m) = G1: 0.46 G1: 0.21 G1: 0.33

Mean for G1 = 5.79, 5.82
Mean for G2 = 1.68, 2.15
Mean for G3 = -4.00, -4.67
Covariance for G1 = 4.54, 4.29
Covariance for G2 = 0.52, 0.13
Covariance for G3 = 4.67, 2.89
```



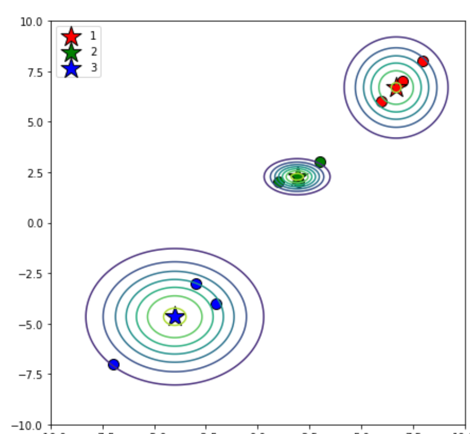
Iteration 2:

```
Point [1 2] : Assign to G1: 0.00 G2: 1.00 G3: 0.00
Point [3 3] : Assign to G1: 0.66 G2: 0.34 G3: 0.00
Point [2 2] : Assign to G1: 0.01 G2: 0.99 G3: 0.00
Point [8 8] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [6 6] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [7 7] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [-3 -3] : Assign to G1: 0.00 G2: 0.00 G3: 1.00
Point [-2 -4] : Assign to G1: 0.00 G2: 0.00 G3: 1.00
Point [-7 -7] : Assign to G1: 0.00 G2: 0.00 G3: 1.00

Prior(m) = G1: 0.41 G1: 0.26 G1: 0.33

Mean for G1 = 6.27, 6.27
Mean for G2 = 1.72, 2.15
Mean for G3 = -4.00, -4.67
Covariance for G1 = 2.95, 2.94
Covariance for G2 = 0.50, 0.13
Covariance for G3 = 4.67, 2.89
```



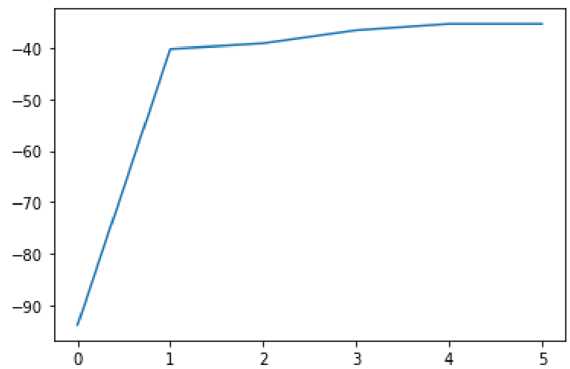
Iteration 3:

```
Point [1 2] : Assign to G1: 0.00 G2: 1.00 G3: 0.00
Point [3 3] : Assign to G1: 0.25 G2: 0.75 G3: 0.00
Point [2 2] : Assign to G1: 0.00 G2: 1.00 G3: 0.00
Point [8 8] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [6 6] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [7 7] : Assign to G1: 1.00 G2: 0.00 G3: 0.00
Point [-3 -3] : Assign to G1: 0.00 G2: 0.00 G3: 1.00
Point [-2 -4] : Assign to G1: 0.00 G2: 0.00 G3: 1.00
Point [-7 -7] : Assign to G1: 0.00 G2: 0.00 G3: 1.00

Prior(m) = G1: 0.36 G1: 0.31 G1: 0.33

Mean for G1 = 6.70, 6.70
Mean for G2 = 1.91, 2.27
Mean for G3 = -4.00, -4.67
Covariance for G1 = 1.74, 1.74
Covariance for G2 = 0.63, 0.20
Covariance for G3 = 4.67, 2.89
```



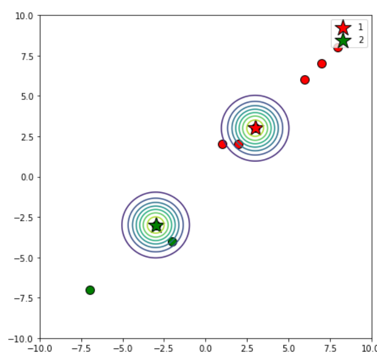
T2: Dose log likelihood of model goes up every iteration?

- **Ans**: Yes, it’s goes up!! (Plot by 6 iterations)



T3: Using 2 mixtures, Repeat three iterations of EM. Write down each w, m, mu, var

- **Ans:** Initialize:

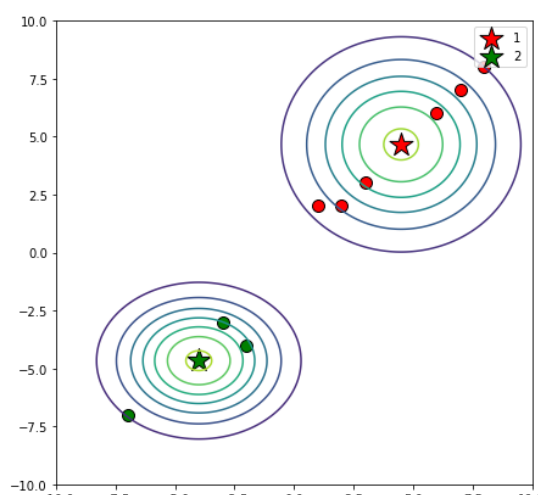


Iteration 1,2,3: (Same for every iteration)

Point [1 2] : Assign to G1: 1.00 G2: 0.00
Point [3 3] : Assign to G1: 1.00 G2: 0.00
Point [2 2] : Assign to G1: 1.00 G2: 0.00
Point [8 8] : Assign to G1: 1.00 G2: 0.00
Point [6 6] : Assign to G1: 1.00 G2: 0.00
Point [7 7] : Assign to G1: 1.00 G2: 0.00
Point [-3 -3] : Assign to G1: 0.00 G2: 1.00
Point [-2 -4] : Assign to G1: 0.00 G2: 1.00
Point [-7 -7] : Assign to G1: 0.00 G2: 1.00

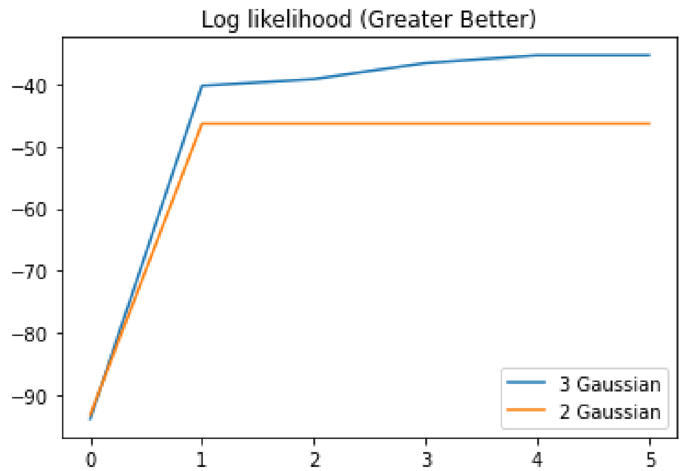
Prior(m) = G1: 0.67 G1: 0.33

Mean for G1 = 4.50, 4.67
Mean for G2 = -4.00, -4.67
Covariance for G1 = 6.92, 5.89
Covariance for G2 = 4.67, 2.89



T4: Which features should be discretized? And What are the criteria?

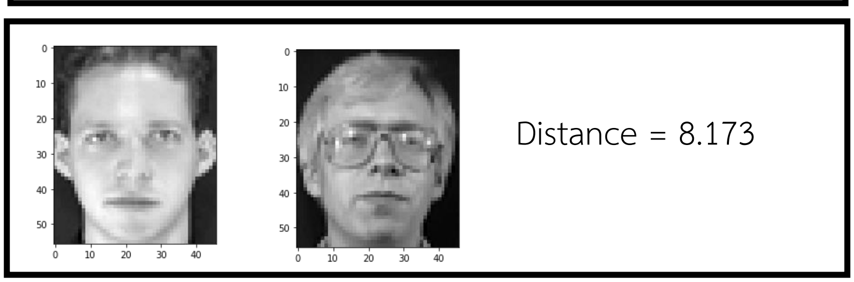
- **Ans**: 3 mixture model has a better result.



PART: PCA & Fisherfaces

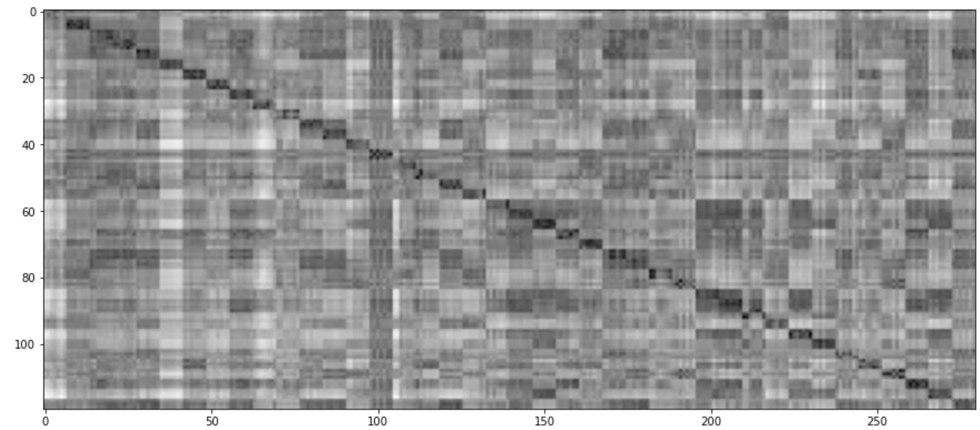
T5: What is the Euclidean distance between first person and second person image (First two images) Does it make sense? Is it useful for face verification?

- **Ans:** If we just look at these two results and know that the same person should has smaller distance than different person. I would say that these number is making no sense and it won't be useful in face verification task.



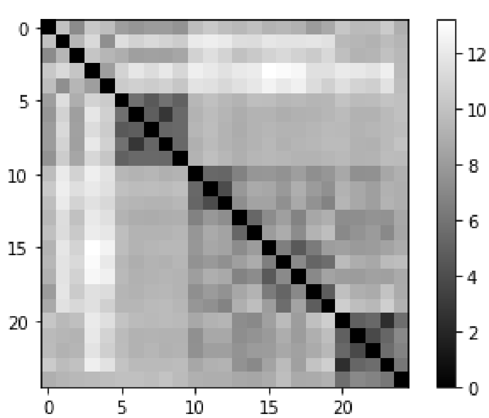
T6: Show the similarity matrix that take T list as train and D list as test

- **Ans:**



T7: From example similarity matrix below, what does the black square suggest about the pictures from person number 2? What do the patterns from person number 1 say about?

- **Ans:** the darker region represents smaller Euclidean distance, that means image from person number 2 (5:10 region) are similar (maybe in the view of color, angle, position), while image from person number 1 (0:5) are not.



A simple face verification system

T8: What is the true positive rate and the false alarm rate for t=10?

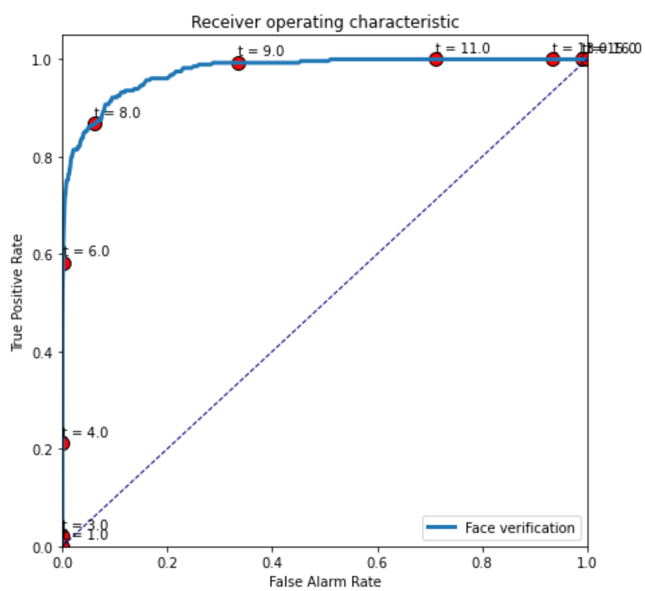
- **Ans:** TP = 279, TN = 5936, FP = 4984, FN = 1

  True positive rate = 0.996

  False Alarm rate = 0.456

T9: Plot RoC curve. What should be the minimum and maximum threshold?

- **Ans:** threshold range should be determined by min and max value of similarity matrix



T10: What is the EER? What is the recall rate at 0.1% false alarm rate?

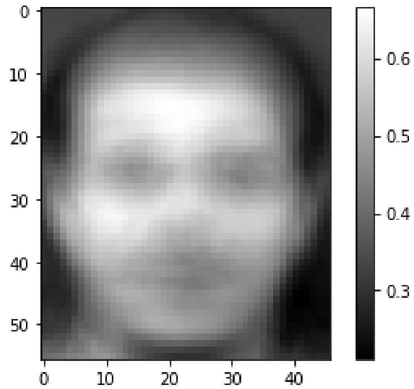
- **Ans:**

```
Equal error rate :  0.08946886446886448
At Threshold :  8.088999999999996
0.1% FAR Precision :  0.9013705993012631
0.1% FAR Recall :  0.9214285714285714
```

Principle Component Analysis

T11: Compute the mean vector from the training images?

- **Ans:**



T12: What is the size of the covariance matrix? What is the rank of the covariance matrix?

- **Ans:** The covariance matrix size is (features, features) 2576, 2576

  The rank of this covariance is (N-1) 119

T13: What is the size of the Gram matrix? What is the rank of Gram matrix? How many non-zero eigenvalues do we expect to get?

- **Ans:** The Gram matrix size is (sample, sample) 120, 120

  The rank of this Gram is (N-1) 119

Then we will get 119 non-zero value out of 120 values from eigenvalues.

T14: Is the Gram matrix also symmetric? Why?

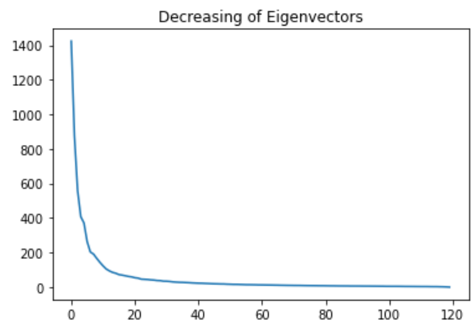
- **Ans:** Yes. Because this is a product of matrix and its transpose.

T15: Compute the eigenvectors and eigenvalues of the Gram. How many non-zero eigenvalues are there?

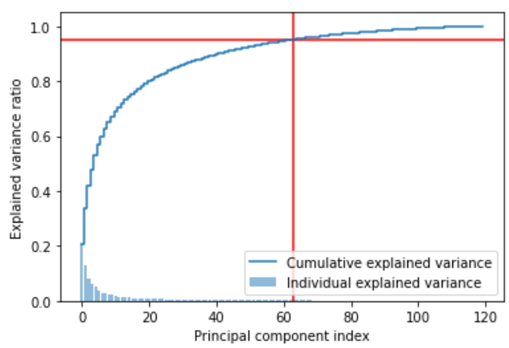
- **Ans:** there are 119 eigenvectors that get non-zero values.

T16: Plot the eigenvalues. If we want to keep at least 95% of variance in the data, how many eigen vectors should we use?

- **Ans:**

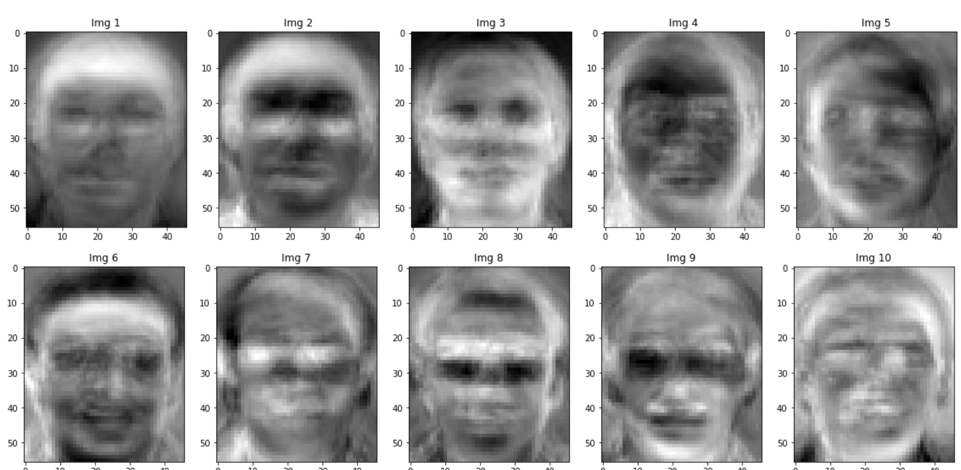


At least 64 Eigenvectors should be used to keep 95% variance of data



T17: Compute eigenfaces. Show the first 10 eigenfaces.

- **Ans:**



T18: From the image, what do you think the first eigenvector captures? What about the second eigenvector? Do you think biggest variance are capture in these two eigenfaces?

- **Ans:** The first eigenfaces captures the variance around hair position.

  The second eigenfaces captures the variance of hair, eyes, and shirt.

  Just these two eigenfaces, they capture many important parts of person faces but they still explain below 40% of variance of all data. That make sense because there still has other important parts such as nose and mount of person to be captured.

PCA subspace and face verification system

T19: Find the projection values of all images. Keep the first k = 10 projection values. Report EER and the recall rate at 0.1% FAR?

- **Ans:**

```
Equal error rate :  0.0784798534798535
At Threshold :  4.784
0.1% FAR Precision :  0.9026036644165862
0.1% FAR Recall :  0.9428571428571428
```

T20: What is the k that gives the best EER? Try k = 5-14

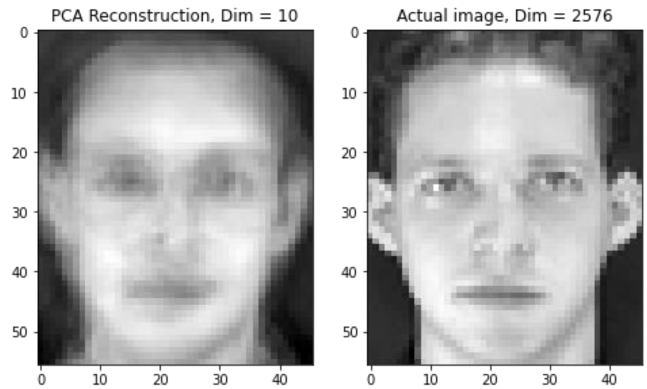
- **Ans:** K = 11 gives the best results.

```
K =  11
Equal error rate :  0.0779761904761905
At Threshold :  4.96
0.1% FAR Precision :  0.9029539607912711
0.1% FAR Recall :  0.9321428571428572
```

(OPTIONAL) PCA reconstruction

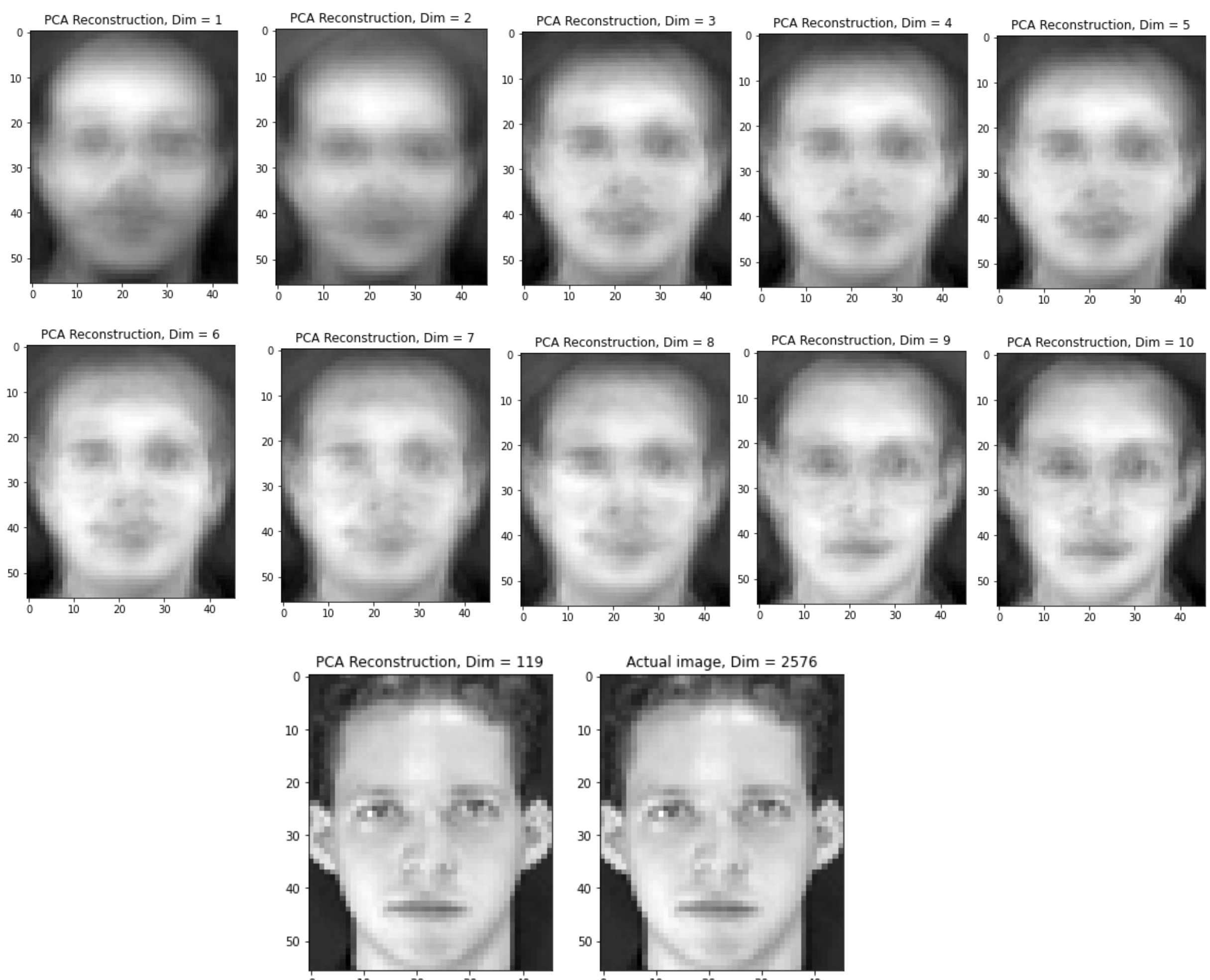
OT1: Reconstruct the first image using this procedure. Use k = 10 what is the MSE?

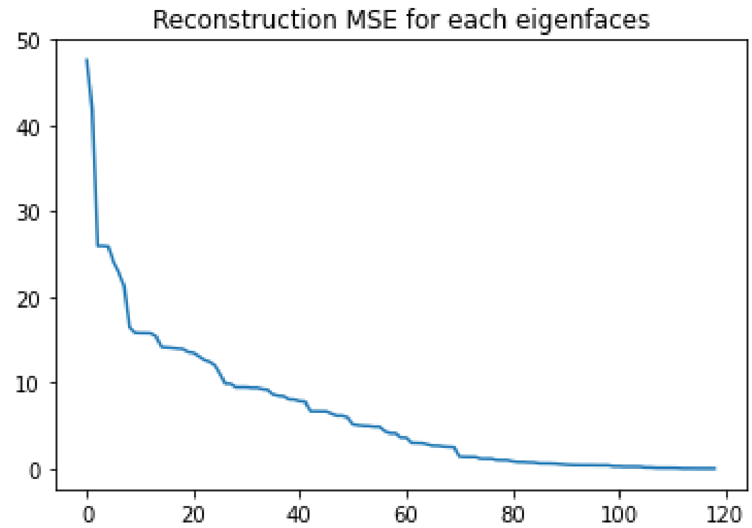
- **Ans:** MSE = 15.838



OT2: For k values of 1-10 and 119, show the reconstructed images. Plot the MSE.

- **Ans:**





OT3: Consider if we want to store 1m images. How much space do we need? If we compress by using the first 10 eigenvalues, how much space do we need?

- **Ans:** [0, 255] 8bits int for non-compress image. 32bit float for compress image

  - 1m images with 2576 values each. Need <u>2456Mb</u> space.
  - For compress image. Need <u>100.6Kb</u> space for eigenfaces. Additional <u>10Kb</u> for meanface, And <u>38.2Mb</u> for compress images itself.

Linear Discriminant Analysis (LDA)

T21: How many PCA dimensions do we need to keep in order to make Sw full rank?

- **Ans:** Sw rank is capped by (number of data – number of class), so to make Sw full rank PCA must reduce dimensions to <u>80</u> (120 samples – 40 classes)

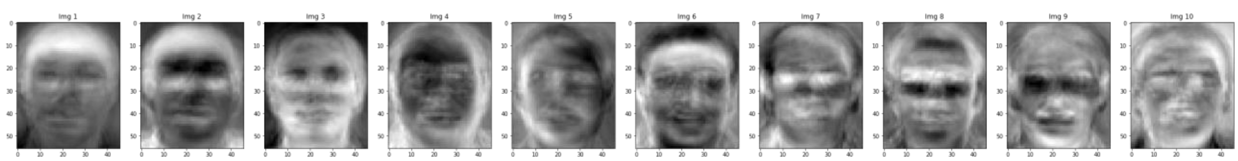
T22: Project the original input to the PCA subspace. Find the LDA projections. Is inv(Sw)Sb asymmetric? How many non-zero eigenvalues are there?

- **Ans:** inv(Sw)Sb is non-symmetric matrix, so Eigen component will be complex number. And LDA rank is capped by min(class-1, N-class), so there are <u>39</u> non-zero eigenvalues out of 80 values.

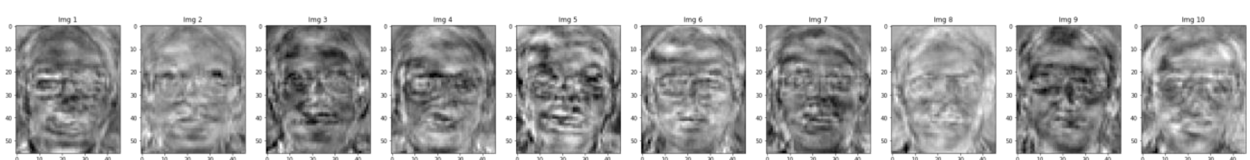
T23: Plot the first 10 LDA eigenfaces compare with PCA.

- **Ans:**

PCA projection



LDA projection



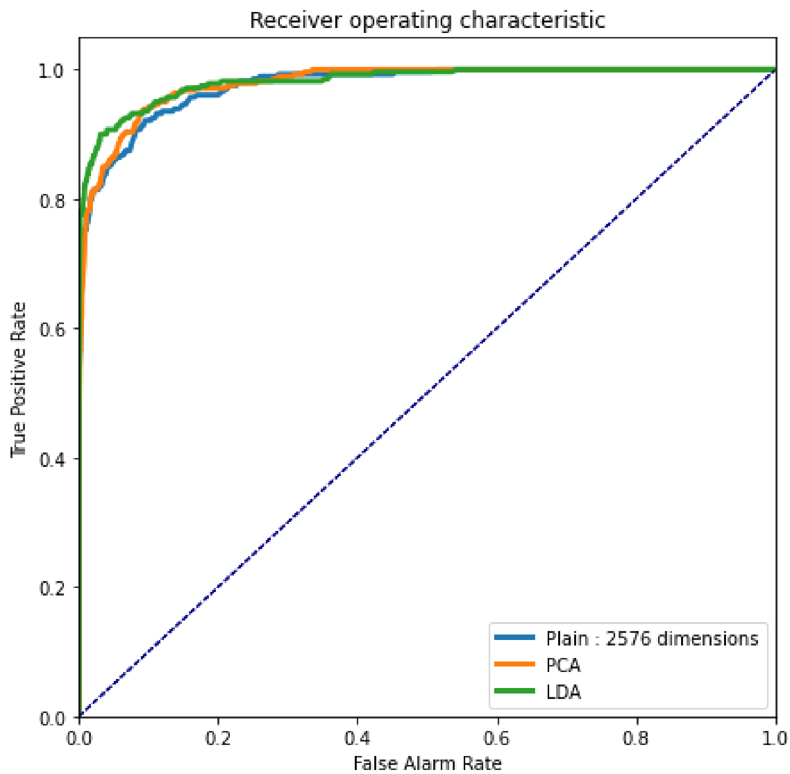
T24: Calculate the fisherfaces projection of all images. What is the EER and recall rate at 0.1% FAR?

- **Ans:**

```
Equal error rate :  0.07211538461538461
At Threshold :  3.661
0.1% FAR Precision :  0.9026665493267623
0.1% FAR Recall :  0.9392857142857143
```

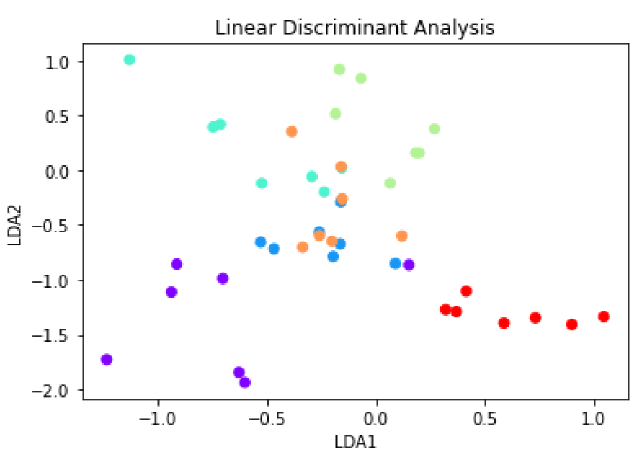
T25: Plot the Roc of all three experiments on the same axes.

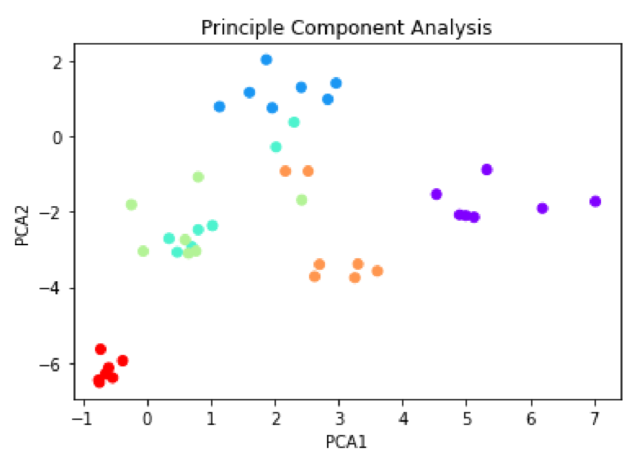
- **Ans:**



OT4: Plot the first two LDA dimensions of the test images (6 people 7 images each). Observe the clustering of between each person. Compare with PCA

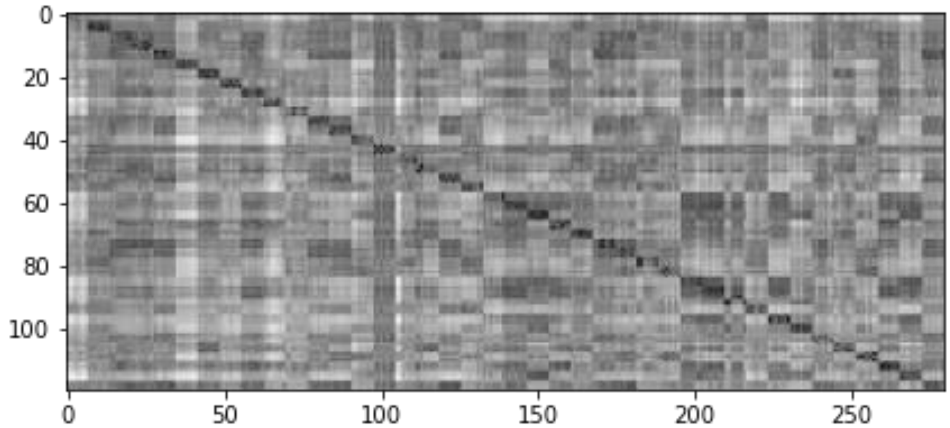
- **Ans:**



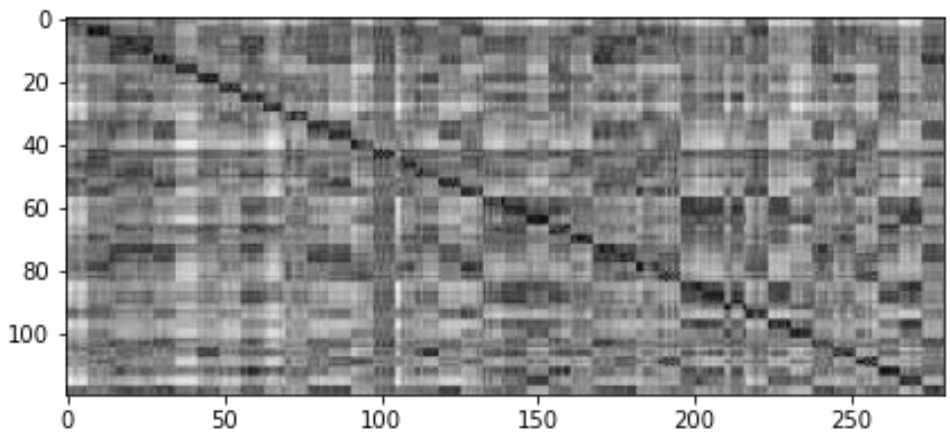


Note: Compare similarity matrix

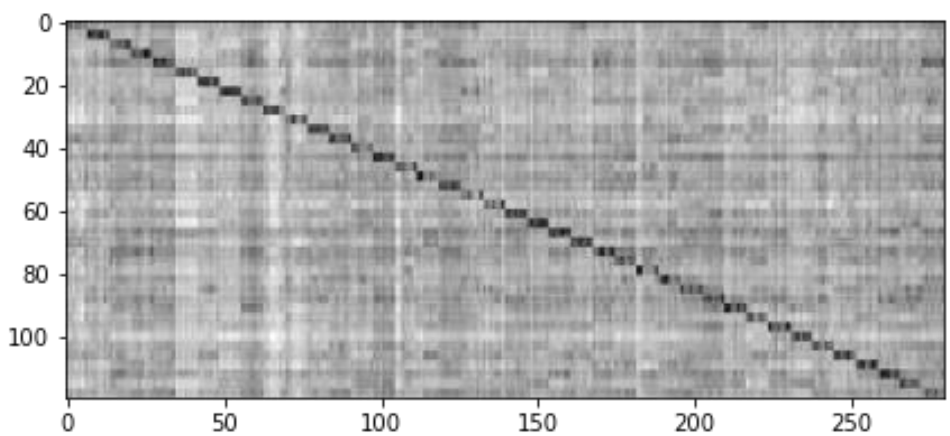
None



PCA



LDA



LDA get the cleanest similarity matrix and very easy to classify by just looking at this visualization.