# Pattern Recognition and Machine Learning

## Task - 2: Face Recognition - Principal Component Analysis

Use the input images of different individuals to calculate Eigenfaces and reduce dimensionality of the face recognition problem.

**Given**

- 5 images each of 40 individuals
- Size of each image is 112*92 pixels

- 1 image of 2 individuals to build using Eigenfaces
- Size of each such image is 112*92 pixels

**Things to consider:**

- Out of the available features, which ones are correlated?
- Which features/variables are the most significant in describing the full data set?

**Concepts:**

- Dimensionality Reduction
- Variance, Covariance
- Linear Algebra
  - Matrix Multiplication
  - Eigenvalues and Eigenvectors
  - Orthogonal Diagonalization (Gram Schmidt Process)
  - Singular Value Decomposition (SVD)
  - Projection of a vector on another

**Linear Algebra Background:**

- The matrices $\boldsymbol{A^T A}$ and $\boldsymbol{A A^T}$ share the same non-zero eigenvalues.
- To get from an eigenvector $\vec{\boldsymbol{v}}$ of $\boldsymbol{A^T A}$ to an eigenvector of $\boldsymbol{A A^T}$ , we need to multiply $\vec{\boldsymbol{v}}$ on the left by A and to get from an eigenvector $\vec{\boldsymbol{w}}$ of $\boldsymbol{A A^T}$ to an eigenvector of $\boldsymbol{A^T A}$, we need to multiply $\vec{\boldsymbol{w}}$ on the left by $\boldsymbol{A^T}$.

**Principal Component Analysis:**

The covariance matrix **S** is equal to

$$\mathbf{S} = \frac{1}{n}\boldsymbol{B}\boldsymbol{B^T} \tag{1}$$

where B is the m x n matrix whose $i^{th}$ column is $\vec{x_i} - \vec{\mu_i}$
m is the number of variables/features
n is the number of observations

The matrix **S** is orthogonally diagonalizable. The eigenvectors of this matrix are called *principal components* of the data set. In the case of image processing, the eigenvectors are called eigenfaces (as each eigenvector of the image data set is itself an image).

**Interpretation from PCA**:

The direction given by the first principal direction (eigenvector corresponding to the highest eigenvalue) explains/accounts for the maximum amount of variance (eigenvalue/sum of all eigenvalues) of the data set. And similarly, the second principal direction accounts for the next maximum amount of variance.

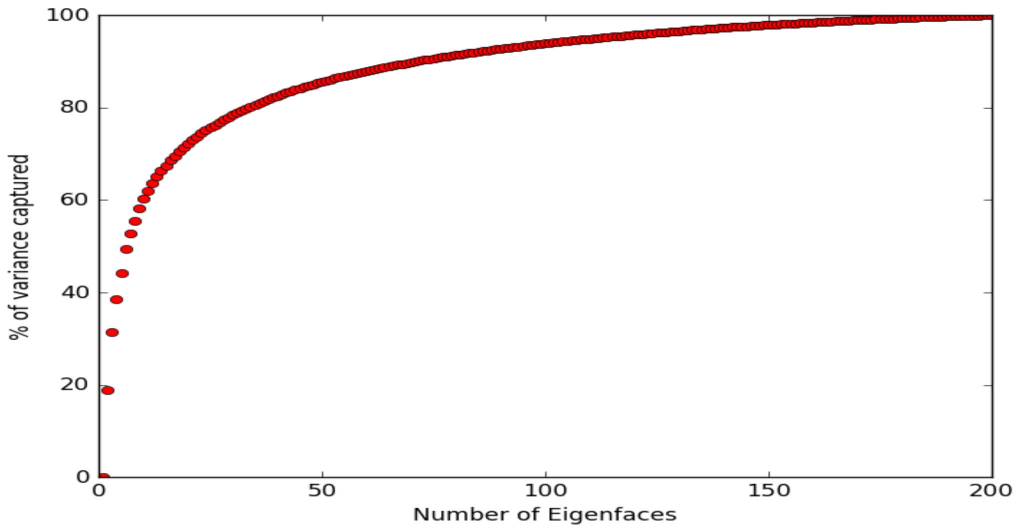
All the eigenvectors are orthogonal to each other, forming a new basis in the dimension they span in.

All the new/test vectors can be written in terms of these eigenvectors. We can choose the top $k$ eigenvectors where $k < m$ to construct the test vector. The value of $k$ is chosen based on how much percentage of variance of original data set we need to capture.

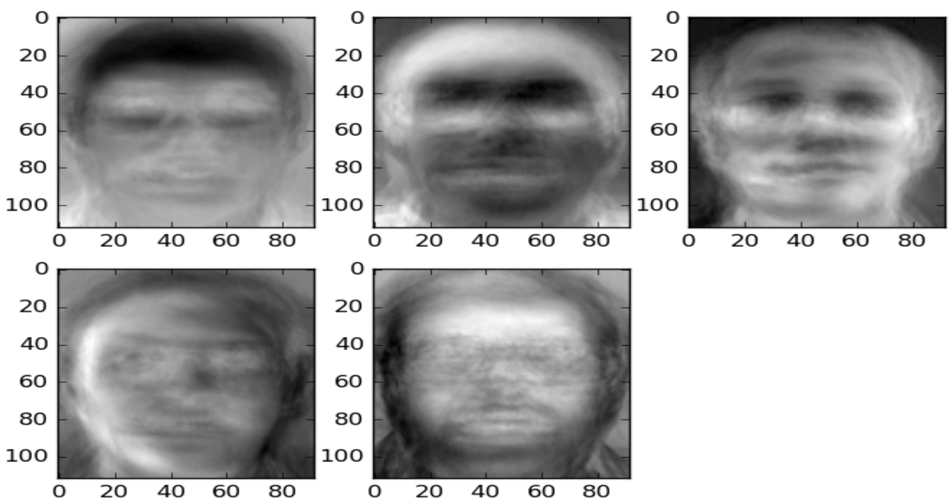
**Summary:**

- Out of the 200 eigenvalues, **110** are required to retain 95 percent variance of the original data set in the reduced space.

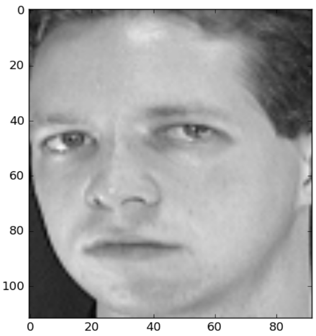
- Graph of number of eigenfaces vs percentage of variance captured:



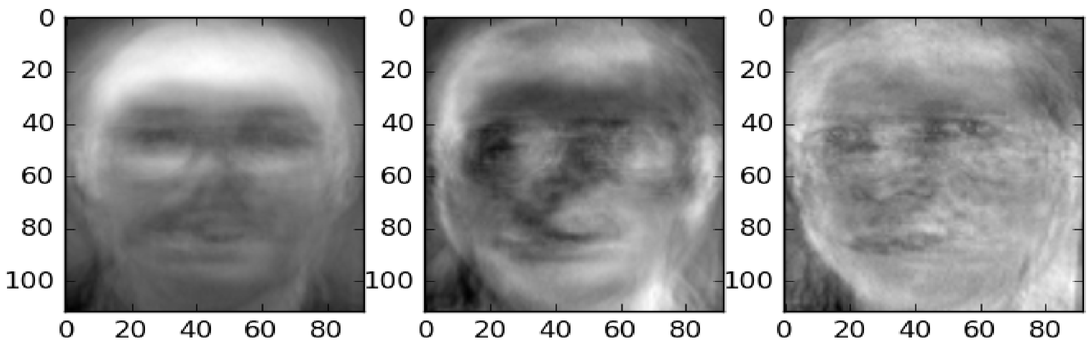

- Images correspoding to the top 5 eigenfaces:



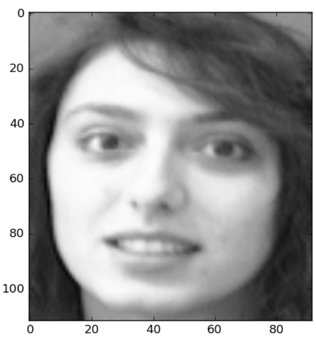
- Input Image 1:



- Input Image 1 reconstructed with top 1, 15 and 200 eigenvectors respectively:



- Input Image 2:



- Input Image 2 reconstructed with top 1, 15 and 200 eigenvectors respectively: