In the Name of God

Statistical Pattern Recognition

Homework IV

Principal Component and Fisher Linear Discriminant Analysis

Assignment Date: **19 Azar** Submission Deadline: **1 Dey**

Contents

1 Introduction **2**
- 1.1 Principal Component Analysis (PCA) 2
- 1.2 Fisher Linear Discriminant Analysis (Fisher LDA) 2
- 1.3 Dataset: JAFFE 2

2 Principal Component Analysis (PCA) **2**

3 Fisher Linear Discriminant Analysis (Fisher LDA) **3**

1 Introduction

In this assignment, the goal is to explore and implement dimensionality reduction techniques for face recognition using Principal Component Analysis (PCA) and Fisher Linear Discriminant Analysis (Fisher-LDA).

1.1 Principal Component Analysis (PCA)

PCA is a foundational technique in unsupervised dimensionality reduction, providing a powerful tool for extracting essential features from high-dimensional datasets. The fundamental objective of PCA is to transform the original data into a new set of uncorrelated variables, known as principal components, that capture the maximum variance within the dataset.

1.2 Fisher Linear Discriminant Analysis (Fisher LDA)

Fisher Linear Discriminant Analysis (also called Linear Discriminant Analysis(LDA)) are methods used in statistics, pattern recognition and machine learning to find a linear combination of features which characterizes or separates two or more classes of objects. The resulting combination may be used as a linear classifier, or more commonly, for dimensionality reduction before later classification.

1.3 Dataset: JAFFE

The JAFFE dataset consists of images of different facial expressions from 10 different Japanese female subjects. Each subject was asked to perform 7 facial expressions. The images need to be resized to 64x64 for efficiency.

2 Principal Component Analysis (PCA)

In this project, we aim to implement Principal Component Analysis (PCA) from scratch for facial emotion recognition. You can refer to section 12.1.1 of **Bishop, Pattern Recognition and Machine Learning** for formulations.

Your tasks include data preprocessing, normalization, visualization, computing eigenvalues, eigenfaces, projecting onto a lower-dimensional subspace, reconstructing images, and applying a Bayes classifier. Therefore, follow the instructions:

1. Load facial images and organize the data into a DataFrame.
2. Resize the images to have 64 x 64 dimensions.
3. Flatten your images to have a vector form with dimensions 4096x1. This way you see each image as a vector of features like in your previous homework (In order to visualize the faces at any time, you can reshape these feature vectors back to 64 x 64 dimension).
4. Calculate the mean face then normalize the images by subtracting the mean face from each image.
5. Randomly select one image from each class of emotions and visualize it before and after subtracting the average face.
6. Calculate the covariance matrix of zero-mean faces, and find its eigenvalue and eigen vectors (This is called eigen decomposition of the covarianace matrix).
7. Plot the eigenvalues in descending order.
8. Pick top 10 eigen vectors (also called eigen faces in this case) and visualize them in form of face images.
9. Project the images onto a lower-dimensional subspace (2D and 3D) using the top eigenfaces then visualize the projected data. (while plotting the projected samples, use different colors for each class of emotions)

10. Implement a function to reconstruct the projected faces from the previous step. Don't forget to add the mean face too.
11. Randomly select one image from each class of data and visualize its original and reconstructed forms.
12. You have fully implemented PCA so far, but how many principle components (PCs) are needed to keep 90% of the cumulative variance in data?
13. Project the data using the number of components you found in the previous step.
14. Split the projected data from the previous step into train and test sets, keep 30 percent for test and 70 percent for train.
15. Apply a Bayes classifier (preferably QDA) to the projected data and report the accuracy.
16. (**Bonus**) See how the number of PCs affect the accuracy of the Bayes classifier. Plot the classification accuracy for both training and test sets, in terms of the number of PCs. How do you interpret this plot?

3 Fisher Linear Discriminant Analysis (Fisher LDA)

In this section, we will implement Fisher Linear Discriminant Analysis (LDA) from scratch without considering class labels. You can refer to section 4.1.6 of **Bishop, Pattern Recognition and Machine Learning** for formulations.

Your tasks include whitening the data, computing the between-scatter matrix (SB) and within-scatter matrix (SW), plotting eigenvalues of the separability matrix, plotting separability measure vs. the number of components, choosing an appropriate number of features, projecting data into new subspaces, applying a Bayes classifier with Gaussian parametric estimate, and reporting accuracy.

1. Like for PCA, load facial images, resize them to have 64 x 64 dimensions, and flatten your images to have a vector form with dimensions 4096x1.
2. First whiten your data and then compute between scatter (SB) and within scatter (SW) matrices.
3. Calculate the separability matrix using SB and SW matrices, and find its eigen value and eigen vectors (This is again an eigen decomposition).
4. Plot the eigen values of the separability matrix in descending order.
5. Project the data into a new subspace (the dimension of the new subspace can be anything you wish).
6. Now split the projected data from the previous step into train and test sets, keep 30 percent for test and 70 percent for train.
7. Then apply a Bayes classifier and report the accuracy.
8. (**Bonus**) See how the number of extracted features affect the accuracy of the Bayes classifier. Therefore, plot the classification accuracy in terms of the number of extracted features. How do you interpret this plot?