

Computational Methods in Statistics I

Dimension Reduction: An Application In Facial Recognition

Joe Sinotte

October 20, 2014

Outline

- Motivation and Statement of the Problem
- Methodology
 - PCA
 - SP
- Results
- Conclusion

Motivation

- Image recognition is widely used in the world today.
 - Law Enforcement
 - Mapping
 - Social Media
- Thus, the ability to correctly perform image recognition is applicable in many fields of study.

Statement of the Problem

Data:

- two sets of 200 images taken of 40 individuals.
- each individual has five different images in each set.

Problem: Given the data, can a program be written to correctly match an image from one set to a person in the other set?

Dimension Reduction

- Data sets for images are often quite large.
- It is often convenient to use transformations to reduce the dimensionality of the images.
- Two methods of dimension reduction will be used:
 - PCA
 - SP

Principal Component Analysis

- The data set is multiplied by the transpose of the first K columns of an orthogonal matrix found by using Singular Value Decomposition.
- This method allows the user to choose the number of dimensions that give the most descriptive information of the data. The superfluous data are eliminated from the rest of the application.

Simple Projection

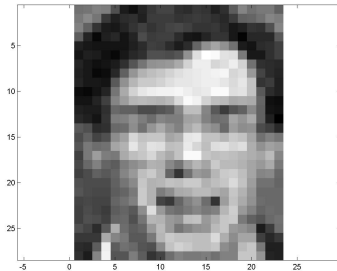
- The matrix used to transform the data is derived by selecting the first K columns of an identity matrix which, in this example, is of dimension $(s_1 \times s_2) = 644$.
- The data are then transformed by premultiplying Y_{test} and Y_{train} by the transpose of this matrix.

Nearest Neighbor

- A column (image) was selected at random from Y_{test} , with the goal being to find a match in Y_{train} .
- The Nearest Neighbor Classifier was used to find the match, where the NN is the image in Y_{train} that minimizes the normed distance between itself and the column of Y_{test} .
- The matching is successful if the two images correspond to the same individual.

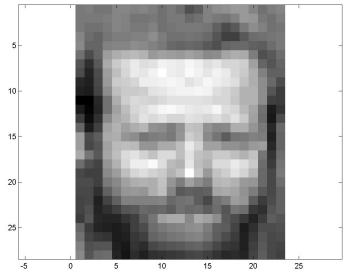
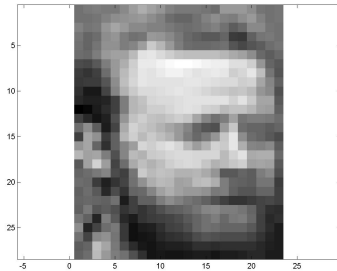
Match with SP

Here is an example of a match using the Simple Projection Method:

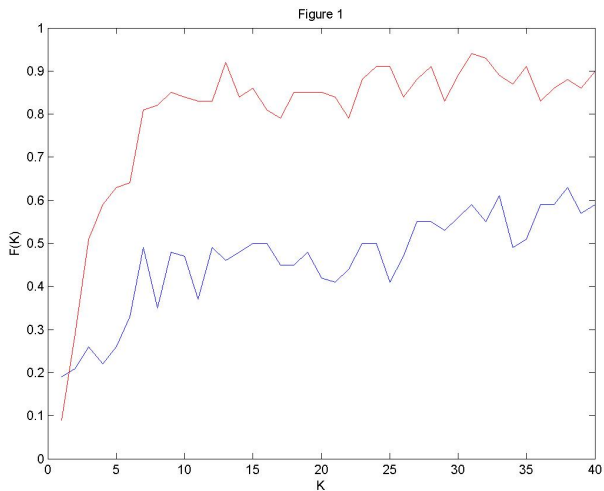


Match with PCA

Here is an example of a match using the Principal Component Analysis:



Results



Conclusions

- Overall, the exercise is a success.
- SP is better for small dimensional data.
- With larger dimensional data sets, PCA becomes much more accurate than SP.