Coding Project 4: Teaching a Computer to Recognize Written Numbers

Lucas Liu

Abstract

In this project, image classification was conducted and evaluated in MATLAB. The procedure to achieve this involved several computational methods including the Wavelet Transform, Principal Component Analysis, and Linear Discriminant Analysis. The Wavelet Transform was used to detect crucial features of each image, Principal Component Analysis of the Wavelet Transform was done to find a basis to separate the distinct features, and Linear Discriminant Analysis was used to determine the threshold that separates the distinct classes.

1 Introduction

Machine learning is a branch of artificial intelligence that enables computer systems to learn from data and make decisions or predictions based on that learning, without being explicitly programmed. It involves training algorithms to identify patterns and relationships in large datasets, which can be used to automate tasks or provide insights that would be difficult or impossible to obtain using traditional programming techniques. Machine learning is used in a wide range of applications, from image recognition and natural language processing to predictive modeling and fraud detection.

The remainder of the report includes a brief overview of the background behind the computational methods used, the results of the project, and a conclusion of the findings.

2 Theoretical Background

This section consists of a brief overview of the main computational methods used in the project.

2.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is based on the idea of projecting the data onto a lower-dimensional subspace while maximizing the separability of the classes. Given a matrix of data, X, the goal of LDA is to find a suitable projection of the data onto a lower-dimension space that maximizes the distance between inter-class distance while minimizing the distance between intra-class distance. The intra-class scatter matrix measures the spread of the data within each class while the inter-class scatter matrix measures the spread of data between classes.

Given two classes of data X_1 and X_2 with mean column vectors μ_1 and μ_2, the intra-class scatter matrix, S_w, can be described as:

$$S_w = \sum_{j=1}^{2}(x_j - \mu_j)(x_j - \mu_j)^T$$

and the inter-class scatter matrix, S_B, is:

$$S_B = (\mu_2 - \mu_1)(\mu_2 - \mu_1)^T$$

As such, the goal is to find a vector, $\vec{w}$ that maximises S_B and minimizes S_w:

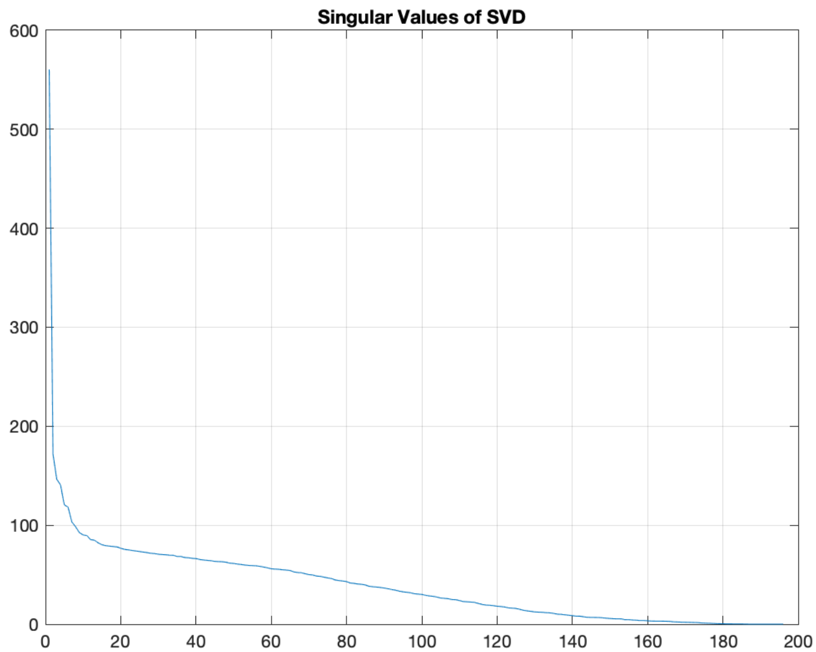
$$\max\left(\frac{\vec{w}^T S_B \vec{w}}{\vec{w}^T S_w \vec{w}}\right)$$

Finally, LDA involves finding a projection matrix that maximizes the ratio of the between-class scatter matrix to the within-class scatter matrix. This can be done using eigenvalue decomposition or singular value decomposition.

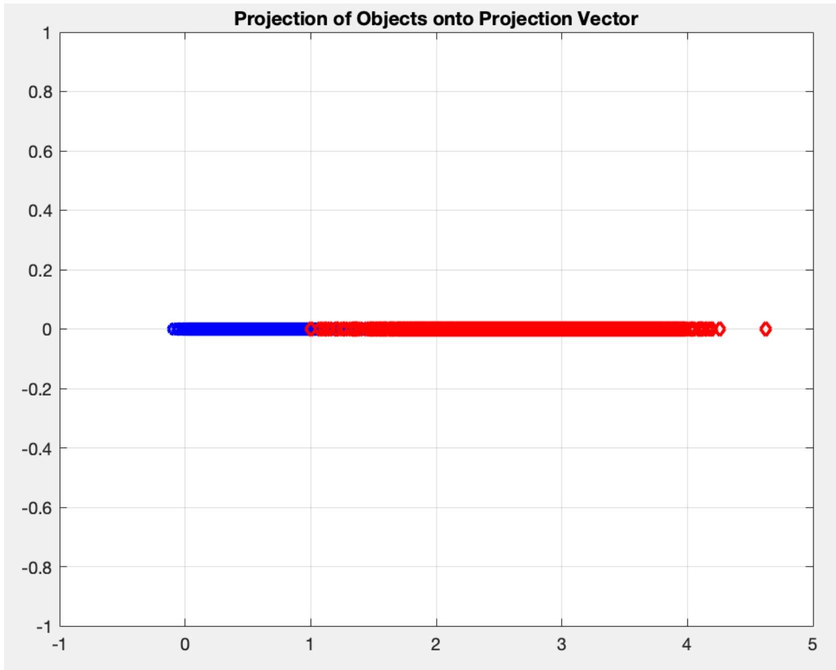
$$S_B w = \lambda_{max} S_w W$$

The resulting projection can be used to transform the data into a lower-dimensional space where the classes are more separable, making it easier to perform classification.

3 Results

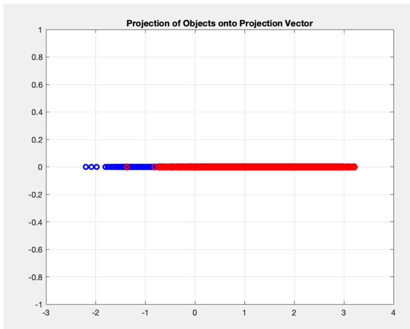


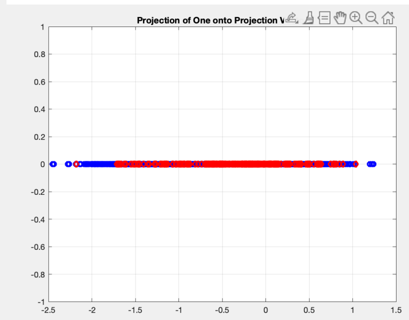
The above plot shows the singular values of the Singular Value Decomposition of the Wavelet Transformed picture matrix. There were a total of 196 singular values of the data matrix's SVD. From the graph, it can be seen that the most significant singular values are present at the very front, followed by a significant drop off in the magnitude of the singular values. This implies that the majority of the crucial information present in the data can be captured in a lower dimensional space. This is beneficial as we can apply the same analysis to the data to yield accurate results while reducing the computational work necessary. Consequently, the first 15 principal components were used in the susbequent analysis.

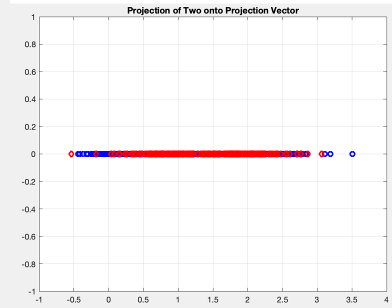


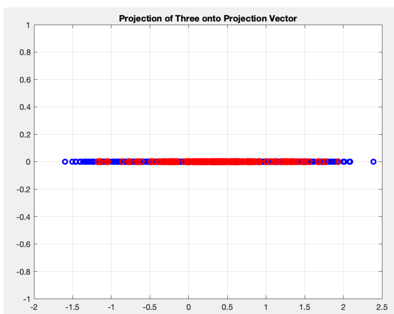
The above plot shows the projections of the objects onto the optimal projection line obtained in the analysis. The images of the written number one are represented in blue while the images of the written number zero are represented in red. As can be seen, there is a distinct separation between the blue and red symbols. As such, the obtained projection vector does achieve our original purpose- maximizing inter-class distances while minimizing intra-class distances.

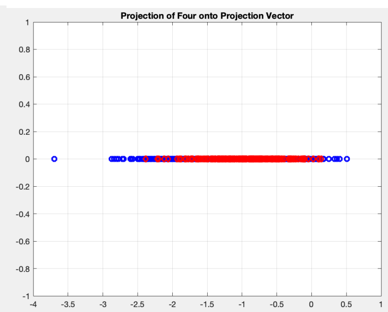
Using the same algorithm used to generate the projection vector above, new images of hand written numbers were given to see if the algorithm can classify them. The process revealed the following:

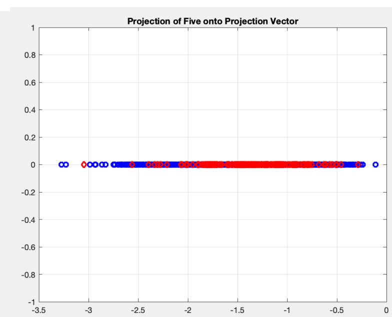

Projection of Objects onto Projection Vector

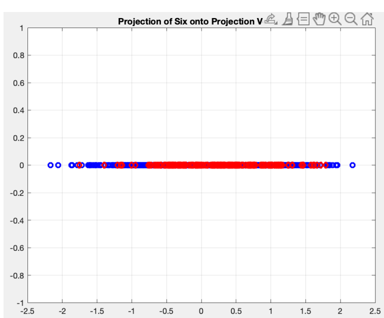

Projection of One onto Projection V

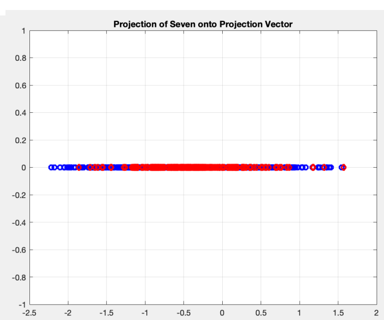

Projection of Two onto Projection Vector

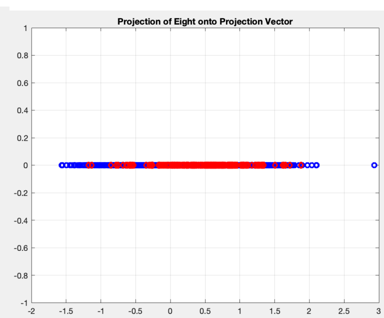

Projection of Three onto Projection Vector

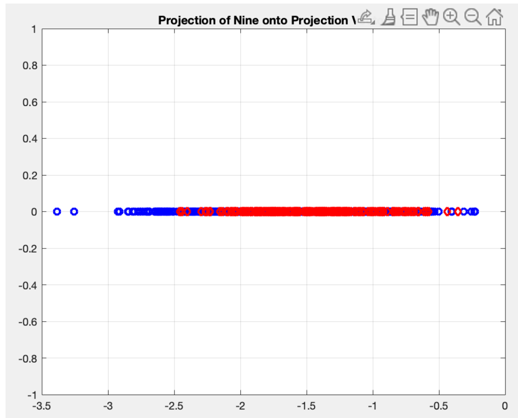

Projection of Four onto Projection Vector


Projection of Five onto Projection Vector


Projection of Six onto Projection V


Projection of Seven onto Projection Vector


Projection of Eight onto Projection Vector



In all of the plots as shown previously, one number, represented in red, is projected onto the projection vector and all of the other images, represented in blue, are also projected on to the projection vector. As can be seen, when comparing one specific written number class to all other written numbers, there is an evident overlap in the projection of the classes and there is no distinct separation or grouping. The best inter-class separation occurs in the first plot, zero against all other numbers, but in all other cases, there is overlapping between the projected points of the different classes.

4 Conclusion

In summary, this project used several computational methods in an algorithm to classify hand written numbers. The Wavelet Transform was used to detect key features of images, the Singular Value Decomposition was then used to find the principal components of the transformed data, and Linear Discriminant Analysis was used to find the best projection line to distinguish between objects of different classes. Upon further inspection, the algorithm is more adept at separating between two specific number classes rather than comparing one class to all of the other classes. For more accurate classification, this problem can be approached in a variety of different ways, e.g. neural networks, decision trees, random forests, etc.