SML Assignment-2

Shobhit Raj (2022482)

1 Question 1

1.1 Assumptions

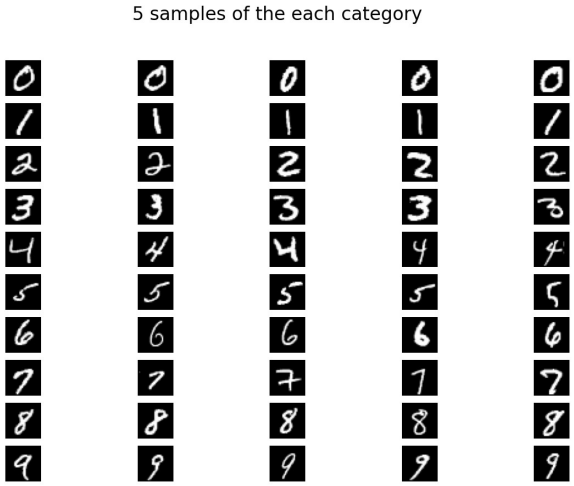
If the determinant of the covariance matrix is zero, then it is added to 'lambda times identity matrix' where lambda is a small constant like .000001, so can take pseudo inverse of this modified matrix while doing QDA.

1.2 Approach

The task involves analyzing the dataset, which consists of images of digits from 0 to 9. The dataset includes 60,000 training samples and 10,000 test samples. The goal is to visualize 5 samples from each class in the training set, vectorize the test set images, and then apply Quadratic Discriminant Analysis (QDA) to classify the digits.

1.2.1 Visualization of Training Samples

To visualize 5 samples from each class in the training set, the dataset is loaded, and a dictionary `class_samples` is initialized to hold the indexes of samples for each class. The code iterates through the dataset and collects 5 samples for each class, storing their indexes in `class_samples`. Then, using Matplotlib, these samples are displayed in a grid format with each row representing a different class.



(a) Visualization of Training Samples

Figure 1: Sample visualization

1.2.2 Vectorization and QDA

Before applying QDA, the images are vectorized to make them 784-dimensional. The mean vector and covariance matrix are computed for each class using only the training set. QDA expressions derived in lectures are used for classification. I've implemented an optimization technique by precomputing and storing certain values like determinant of covariance matrices, the inverse of covariance matrices, and the logarithm of class priors for all 10 classes. This approach effectively reduces the time taken for computations, especially in scenarios where these values are repeatedly needed, such as during the classification of test samples in the execution of QDA.

1.2.3 Classification and Accuracy Evaluation

For each sample in the test set, QDA is applied to determine its class. The accuracy of classification is computed both overall and for each class.

1.3 Results

The overall accuracy achieved on the test set is calculated to be approximately 68.34%. When we normalize the values of pixels in the matrix between 0 and 1, the accuracy comes out to be 80.47%. Additionally, class-wise accuracy is reported, showing the performance of the classifier for each digit class. The accuracy for each class is as follows:

- Class 0 accuracy = 96.0204081632653%
- Class 1 accuracy = 93.74449339207048%
- Class 2 accuracy = 46.80232558139535%
- Class 3 accuracy = 51.18811881188119%
- Class 4 accuracy = 45.21384928716905%
- Class 5 accuracy = 29.596412556053814%
- Class 6 accuracy = 94.36325678496868%
- Class 7 accuracy = 43.28793774319066%
- Class 8 accuracy = 85.42094455852155%
- Class 9 accuracy = 93.1615460852329%

The accuracy ranges from 29% to 96% for individual classes.

2 Question 2

2.1 Assumptions

If the determinant of the covariance matrix is zero, then it is added to 'lambda times identity matrix' where lambda is a small constant like .000001, so can take pseudo inverse of this modified matrix while doing QDA.

2.2 Approach

I performed Principal Component Analysis (PCA) on a dataset consisting of images. The dataset contains images from different classes, and goal is to reduce the dimensionality of the data using PCA and then reconstruct the images to assess the effectiveness of the dimensionality reduction process. Additionally, we will apply Quadratic Discriminant Analysis (QDA) on the reduced data to classify images.

2.2.1 Processing of dataset and PCA

I start by selecting 100 samples from each class in the dataset. These samples are then arranged into a 784×1000 matrix, denoted as X. To prepare the data for PCA, I center it by subtracting the mean from each feature. PCA is applied to the centered data matrix X. We calculate the covariance matrix $S = XX^T$ and then find its eigenvectors and eigenvalues. These are sorted in descending order to form the matrix U.

2.2.2 Reconstruction

I then perform dimensionality reduction by projecting the data onto the subspace spanned by the top p eigenvectors (U_p). The reconstructed data (X_{recon}) is obtained by multiplying U_p with the projected data ($Y = U_p^T X$). The mean squared error (MSE) between the original and reconstructed data is calculated to evaluate the quality of reconstruction.

2.2.3 Visualization

For various values of p (5, 10, 20), I have visualized the reconstructed images. The process involves reshaping each column of the reconstructed data into a 28×28 image. By plotting these images, we observe how the reconstruction improves as the number of retained principal components increases.

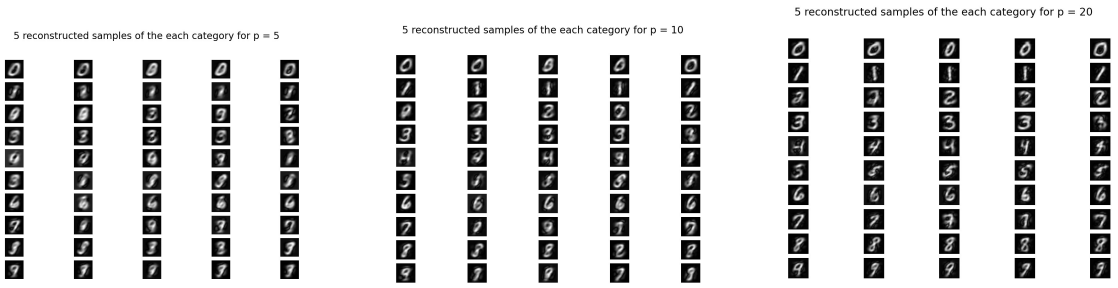


Figure 2: Reconstructed images for $p = 5$, $p = 10$, and $p = 20$.

2.2.4 Classification using QDA

I then apply QDA on the reduced data (Y) for different values of p. The reduced test set (X_{test}) is transformed into the reduced space using the same transformation matrix U_p. QDA is then performed on the reduced test set to classify images. I have evaluated the accuracy on the test set as well as the per-class accuracy for each value of p.

2.3 Results

The MSE between X and X_{recon} is $8.076723385001319 \times 10^{-22}$.

The visual inspection of reconstructed images reveals that as the number of retained principal components (p) increases, the reconstructed images more closely resemble their original counterparts. The accuracy on the test set as well as the per-class accuracy for each value of p are as follows:

For $p = 5$:

- Class 0 accuracy = 87.86%
- Class 1 accuracy = 88.72%
- Class 2 accuracy = 8.33%
- Class 3 accuracy = 45.84%
- Class 4 accuracy = 37.37%
- Class 5 accuracy = 33.30%
- Class 6 accuracy = 67.75%
- Class 7 accuracy = 57.20%
- Class 8 accuracy = 53.39%
- Class 9 accuracy = 26.66%

Accuracy for $p = 5$ is 51.07%.

For $p = 10$:

- Class 0 accuracy = 91.84%
- Class 1 accuracy = 91.28%
- Class 2 accuracy = 60.56%
- Class 3 accuracy = 55.35%
- Class 4 accuracy = 52.55%
- Class 5 accuracy = 42.49%
- Class 6 accuracy = 73.80%
- Class 7 accuracy = 69.36%
- Class 8 accuracy = 56.88%
- Class 9 accuracy = 51.83%

Accuracy for $p = 10$ is 65.12%.

For $p = 20$:

- Class 0 accuracy = 94.49%
- Class 1 accuracy = 94.01%
- Class 2 accuracy = 76.36%
- Class 3 accuracy = 70.69%
- Class 4 accuracy = 66.90%
- Class 5 accuracy = 69.96%
- Class 6 accuracy = 86.53%
- Class 7 accuracy = 79.96%
- Class 8 accuracy = 69.30%
- Class 9 accuracy = 75.62%

Accuracy for $p = 20$ is 78.65%.