ECE 6337 — Project Report

Fall 2024

Fredy Orellana

Introduction

This report documents the solutions and findings for three problems focusing on multivariate Gaussian data generation, principal component analysis (PCA), and histogram equalization in image processing. The goal of this project was to apply theoretical concepts learned in class to practical problems, analyzing the results and drawing meaningful conclusions. Each problem builds upon fundamental concepts in data analysis and image processing, showcasing their real-world applications.

Problem 1: Multivariate Gaussian Data

The objective of this task was to generate random samples from multivariate Gaussian distributions with specified parameters, visualize the data in 3D, and analyze the relationship between the covariance matrix and the resulting data point-cloud shape. A function was implemented to generate 1,000 samples from two distributions: (a) with $\mu = [0, 0, 0]$ and $\Sigma = \text{diag}([2, 1.5, 2.5])$, and (b) with $\mu = [0, 0, 0]$ and a non-diagonal covariance matrix containing correlations.

The 3D scatter plot for dataset (a) is shown in Figure 1. The plot depicts an ellipsoidal shape aligned with the coordinate axes, indicating that the variables are uncorrelated due to the diagonal covariance matrix.

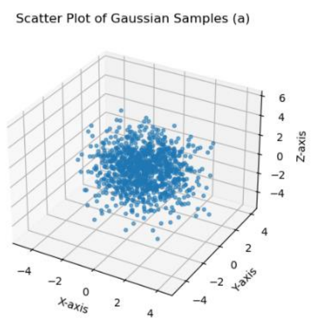

Figure 1: 3D scatter plot for dataset a.

Conversely, the scatter plot for dataset (b), shown in Figure 2, exhibits a rotated ellipsoid. This rotation reflects the presence of correlations between variables as defined by the off-diagonal terms in the covariance matrix.

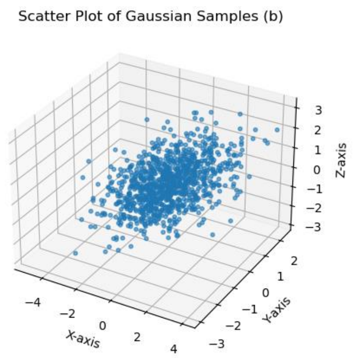

Figure 2: 3D scatter plot for dataset b.

These figures demonstrate how the covariance matrix influences the orientation and spread of multivariate Gaussian data.

Problem 2: Principal Component Analysis (PCA)

The second task required implementing PCA to project the 3D data from Problem 1 onto a 2D subspace, thereby reducing the dimensionality while preserving the variance in the most significant directions. A custom function, project_data, was developed to compute the covariance matrix, perform eigen-decomposition, and construct a transformation matrix from the top d2 eigenvectors.

The PCA projection was first applied to the dataset from Problem 1(a). The 2D scatter plot of the projected data is shown in Figure 3, while the covariance matrix of the projected data was found to retain variance primarily along the selected principal components.

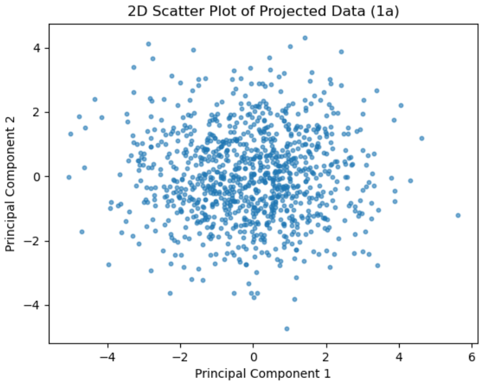

Figure 3: 2D projection scatter plot for dataset a.

Similarly, the PCA projection for the dataset from Problem 1(b) resulted in the scatter plot shown in Figure 4, which retained the rotated shape reflecting the original data's correlation structure.

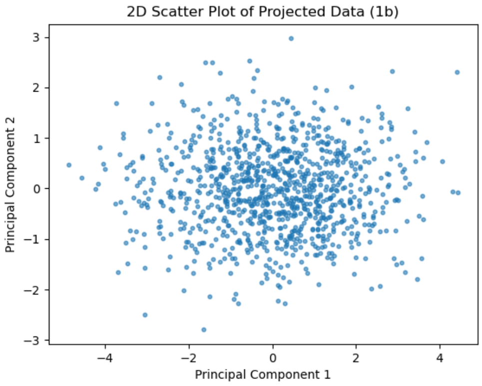

Figure 4: 2D projection scatter plot for dataset b.

By comparing Figures 1 and 3, and Figures 2 and 4, it is evident that the PCA successfully preserved the key variance directions of the original data despite reducing dimensionality.

Problem 3: Histogram Equalization

The third task focused on applying histogram equalization to grayscale images to enhance their contrast. The provided images, Geospatial and MRI, were loaded and their intensity histograms analyzed. The original geospatial image, shown in Figure 5, had most pixel intensities clustered at a single value, resulting in poor contrast.

Figure 5: Original geospatial image.

Similarly, the MRI image, shown in Figure 6, exhibited limited contrast, though it had a broader intensity range.

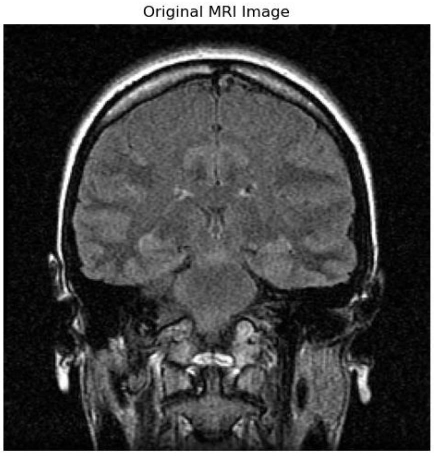

Figure 6: Original MRI image.

Histogram equalization was applied to both images, yielding the enhanced images shown in Figures 7 and 8.

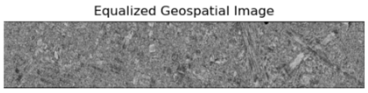

Figure 7: Equalized geospatial image.

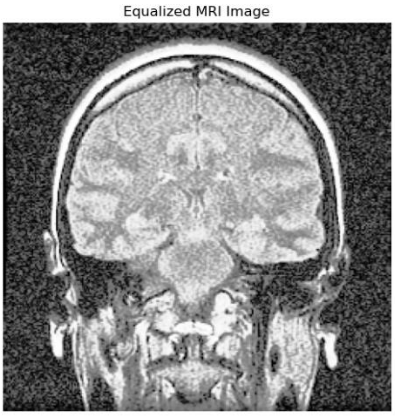

Figure 8: Equalized MRI image.

The geospatial image's contrast improved, revealing previously indistinct features. The histograms of the geospatial image before and after equalization, shown in Figures 9, illustrate how pixel intensities were redistributed from a narrow cluster to a more uniform distribution.

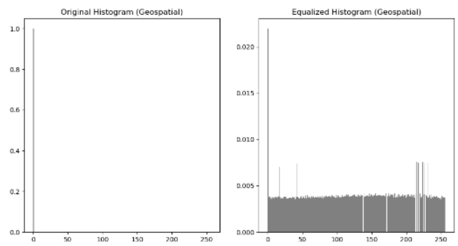

Figure 9: Geospatial histogram before and after equalization.

Similarly, the histograms for the MRI image, shown in Figures 10, demonstrate a similar transformation, with the equalized histogram showing a flatter distribution that enhances contrast.

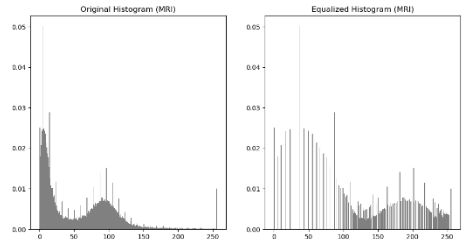

Figure 10: MRI histogram before and after equalization.

These results highlight the utility of histogram equalization in improving image quality, especially for images with poor initial contrast. The geospatial image benefited greatly due to its extreme lack of contrast, while the MRI image's enhancement was more subtle.

Discussion and Conclusion

This project provided practical insights into three key concepts: multivariate Gaussian data generation, PCA, and histogram equalization. In Problem 1, the covariance matrix was shown to directly determine the orientation and spread of Gaussian data, as visualized in Figures 1 and 2. Problem 2 illustrated how PCA can effectively reduce dimensionality while retaining significant variance, as seen in the transformed datasets in Figures 3 and 4. Finally, Problem 3 demonstrated the powerful impact of histogram equalization on enhancing image contrast, as evidenced by the transformations in Figures 5 through 10.

The findings emphasize the practical applications of these techniques in fields ranging from statistical analysis to image processing.