Laboratory Work 9 Report: Nonlinear Dimension Reduction

Your Name

June 23, 2024

1 Introduction

In this laboratory work, the focus was on applying various nonlinear dimension reduction techniques and evaluating their effectiveness in clustering analysis. The methods used include PCA, TruncatedSVD, Factor Analysis, t-SNE, MDS, and Kernel PCA. The primary objective was to reduce the dimensionality of the data while preserving meaningful relationships among data points for clustering purposes.

2 Results

2.1 Inertia for 2D

Method	Inertia
PCA	406.16
TruncatedSVD	406.16
Factor Analysis	176.48
t-SNE	12287.63
MDS	751.34
Kernel PCA	406.16

Table 1: Inertia values for 2D dimensionality reduction methods

2.1.1 Comments

- **PCA, TruncatedSVD, Kernel PCA**: These methods exhibit similar inertia values, suggesting comparable clustering results in 2D space.

- **Factor Analysis**: Lower inertia indicates more compact clusters compared to linear methods.

- **t-SNE**: Highest inertia indicates more dispersed clusters, typical of its nonlinear nature.

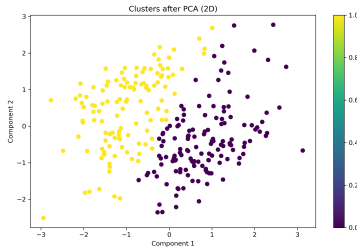- **MDS**: Moderate inertia implies moderate clustering quality.

Figure 1: Clusters after PCA (2D)
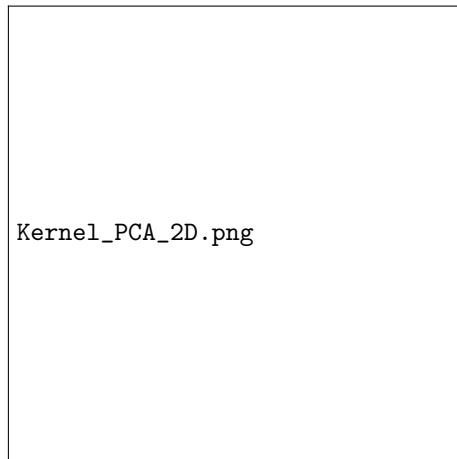

Figure 2: Clusters after t-SNE (2D)

Figure 3: Clusters after Kernel PCA (2D)
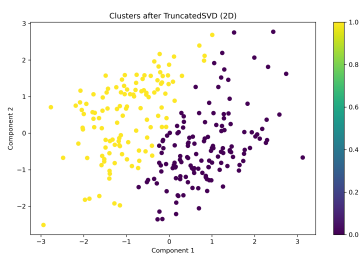
Figure 4: Clusters after Factor Analysis (2D)

Figure 5: Clusters after TruncatedSVD (2D)

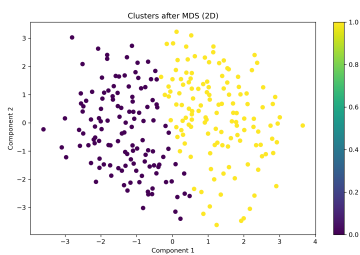Figure 6: Clusters after MDS (2D)

2.2   Inertia for 3D

Method	Inertia
PCA	655.58
TruncatedSVD	655.58
Factor Analysis	220.71
t-SNE	3761.99
MDS	946.10
Kernel PCA	655.58

Table 2: Inertia values for 3D dimensionality reduction methods

2.2.1   Comments

- **PCA, TruncatedSVD, Kernel PCA**: Consistent inertia values across methods indicate stable clustering in 3D.

- **Factor Analysis**: Lower inertia suggests better clustering than PCA and TruncatedSVD.

- **t-SNE**: Highest inertia indicates more scattered clusters, typical of its nonlinear mapping.

- **MDS**: Moderate inertia indicates moderate clustering performance.
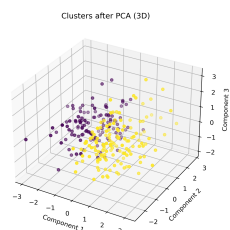

Figure 7: Clusters after PCA (3D)
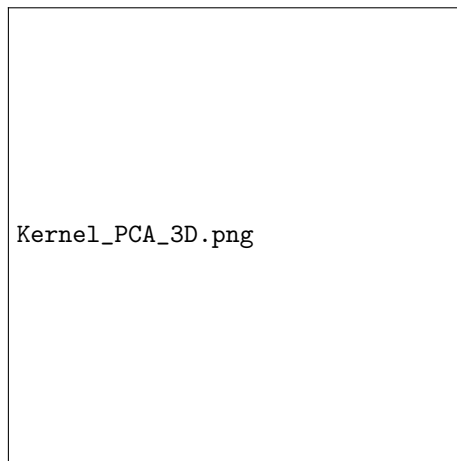
Figure 8: Clusters after t-SNE (3D)

Kernel_PCA_3D.png

Figure 9: Clusters after Kernel PCA (3D)

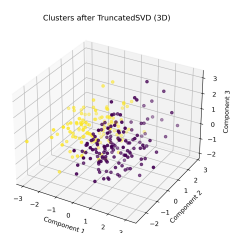Figure 10: Clusters after Factor Analysis (3D)

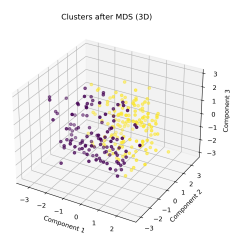Figure 11: Clusters after TruncatedSVD (3D)

Figure 12: Clusters after MDS (3D)

3 Conclusion

The analysis of various nonlinear dimension reduction techniques reveals that:

6

- Methods like PCA, TruncatedSVD, and Kernel PCA provide consistent results in both 2D and 3D dimensions, suggesting reliable clustering performance.

- Factor Analysis shows promise with lower inertia values, indicating potentially more concentrated clusters.

- t-SNE, while effective for visualizing high-dimensional data, tends to produce more dispersed clusters, as evidenced by its higher inertia values.

- MDS demonstrates moderate performance, providing a balance between linear and nonlinear mapping approaches.

Overall, the choice of dimensionality reduction method should consider the specific characteristics of the data and the desired clustering outcomes.