Theory of PCA (Principal Component Analysis)

Kavish Nag
24070126085
AIML B1

Principal Component Analysis (PCA) is a statistical technique used for dimensionality reduction. It transforms a dataset with many correlated variables into a smaller set of new variables called **principal components**.

The main objective of PCA is to reduce the number of dimensions while preserving as much information as possible.

Key Concepts

1. **Variance**
   PCA assumes that directions with higher variance contain more useful information about the structure of the data.

2. **Orthogonal Components**
   Each principal component is perpendicular to the others, ensuring no redundant information.

3. **Eigenvalues and Eigenvectors**
   PCA computes eigenvectors of the covariance matrix.

   - Eigenvectors represent directions of maximum variance.
   - Eigenvalues indicate how much variance is captured.

Example of PCA

Consider a dataset with features:

- Height
- Weight
- BMI

These features are strongly correlated. PCA may transform them into:

- Principal Component 1: Overall body size
- Principal Component 2: Body proportion

Instead of using three correlated variables, the model can use two components that capture most of the variation, reducing complexity.

Inferences from Datasets

Iris Dataset

Inference

The Iris dataset has only four features, which already represent the data effectively. Applying PCA reduced the dimensionality to fewer components, but some class separating information was lost. Since the dataset is already low dimensional, PCA did not improve performance and slightly reduced accuracy. This shows that PCA is not always beneficial for small feature spaces.

Heart Attack Dataset

Observed accuracies:

- Without PCA: about 0.799
- With PCA: about 0.716

Inference

The heart attack dataset contains medically relevant features such as troponin and blood pressure that directly influence classification. PCA combines these features into new components, which may reduce interpretability and remove subtle distinctions useful for prediction. Because the dataset has only moderate dimensionality, dimensionality reduction caused a drop in accuracy instead of improvement.

Overall Conclusion

From the experimental results, we observed that PCA reduced the number of features by compressing the information into fewer principal components. However, PCA does not consider class labels while transforming data. In the Iris dataset, dimensionality reduction caused a small drop in accuracy because some discriminative information was lost when reducing from four features to two components. Similarly, in the Heart Attack dataset, PCA slightly decreased performance because the original feature space already had a manageable number of dimensions and important medical features contributed directly to classification. PCA is more beneficial when datasets contain many redundant or noisy features, but in low or moderate dimensional datasets it may remove useful information, leading to reduced accuracy.