Linear Discriminant Analysis - LDA

Introduction

Linear Discriminant Analysis (LDA) is used to solve dimensionality reduction for data with higher attributes

- Pre-processing step for pattern-classification and machine learning applications.
- Used for feature extraction.
- Linear transformation that maximize the separation between multiple classes.
- “Supervised” - Prediction agent

LDA Applications

- Bankruptcy Prediction.
  - Explain which firms entered bankruptcy vs. survived.
- Face Recognition
  - Used to reduce the number of features to a more manageable level before classification.
- Marketing
  - Distinguish different types of customers and/or products on the basis of collected surveys or other forms of data.

LDA Applications- contd.

- Biomedical Studies.
  - Assessment of severity of state of a patient and prognosis of disease outcome.
- Earth Science

Feature Subspace

Reduces the dimensions of a d-dimensional data set by projecting it onto a (k)-dimensional subspace (where k < d)

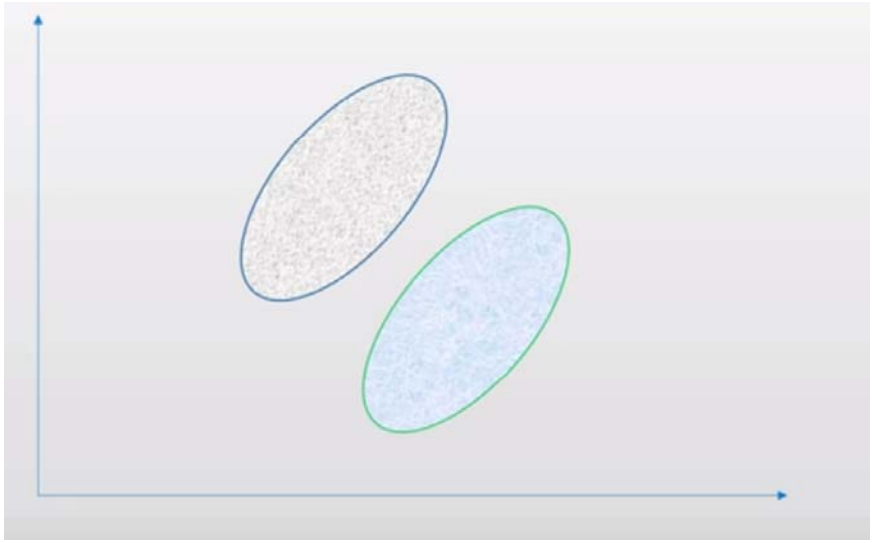
Feature space data is well represented?

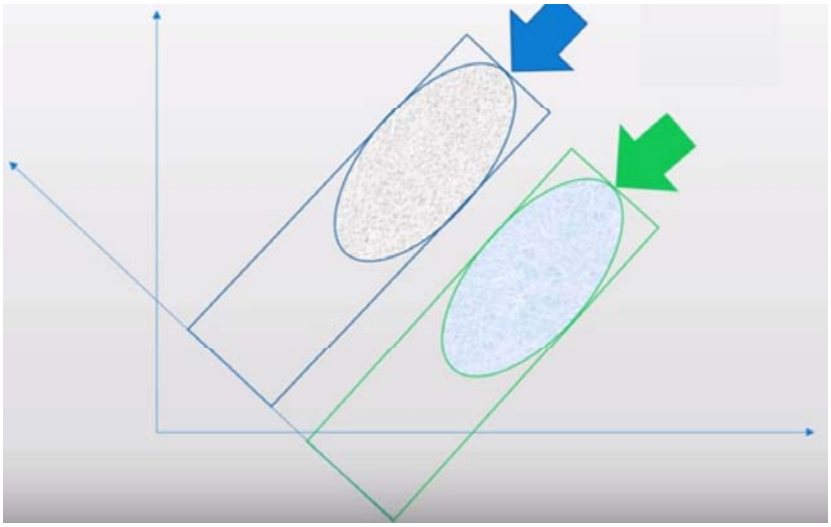
- Compute eigen vectors from dataset
- Collect them in scatter matrix
- Generate *k*-dimensional data from d-dimensional dataset.

Scatter Matrix

- Within class scatter matrix
- In between class scatter matrix

Maximize the between class measure & minimize the within class measure.





LDA steps

1. Compute the d-dimensional mean vectors.
2. Compute the scatter matrices
3. Compute the eigenvectors and corresponding eigenvalues for the scatter matrices.
4. Sort the eigenvalues and choose those with the largest eigenvalues to form a d×k dimensional matrix
5. Transform the samples onto the new subspace.