Linear Discriminant Analysis (LDA)

(Boston Data)

Intelligent Data Analysis 2020

Martin Russell

PCA vs LDA

- PCA finds the direction of *maximum variance* of the data
- If the data is partitioned into classes PCA does not use this information
- … but it may be that the direction of maximum variance is also the best direction for separating the classes
- If the data is partitioned into classes PCA and you want to separate the classes, need to use **Linear Discriminant Analysis (LDA)**
- Not to be confused with Latent Dirichlet Allocation (LDA)

Linear Discriminant Analysis

- Suppose we have a data set X where each data point belongs to one of K classes $c_1, \dots, c_K$
- Let X_k be the set of data points in X that belong to class c_k.

$$X = X_1 \cup X_2 \cup \cdots \cup X_K$$

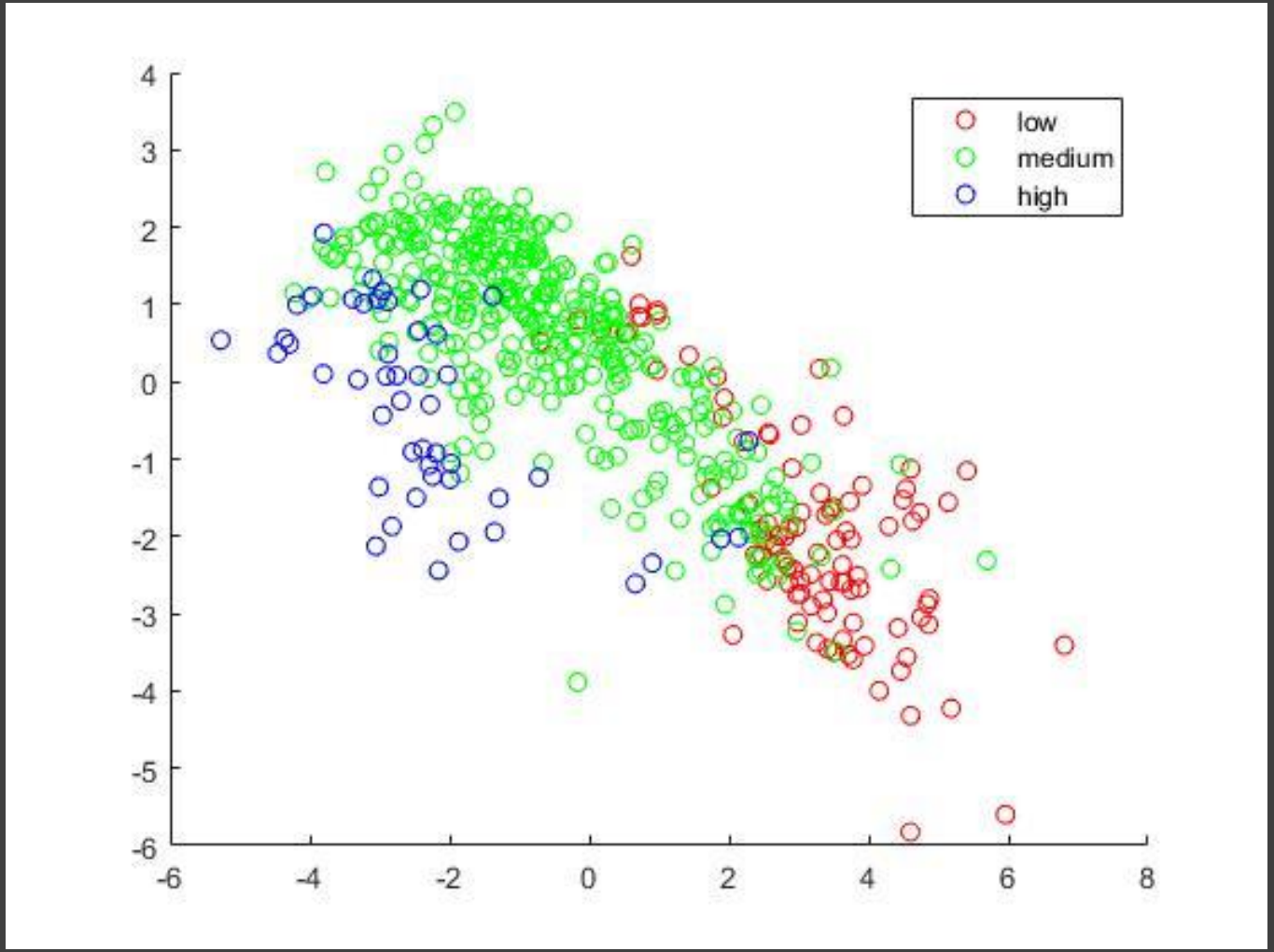
- Let C_k be the covariance matrix of the data set X_k
- Define the average **within-class covariance matrix** C_w by

$$C_w = \frac{C_1 + C_2 + \cdots + C_K}{K}$$

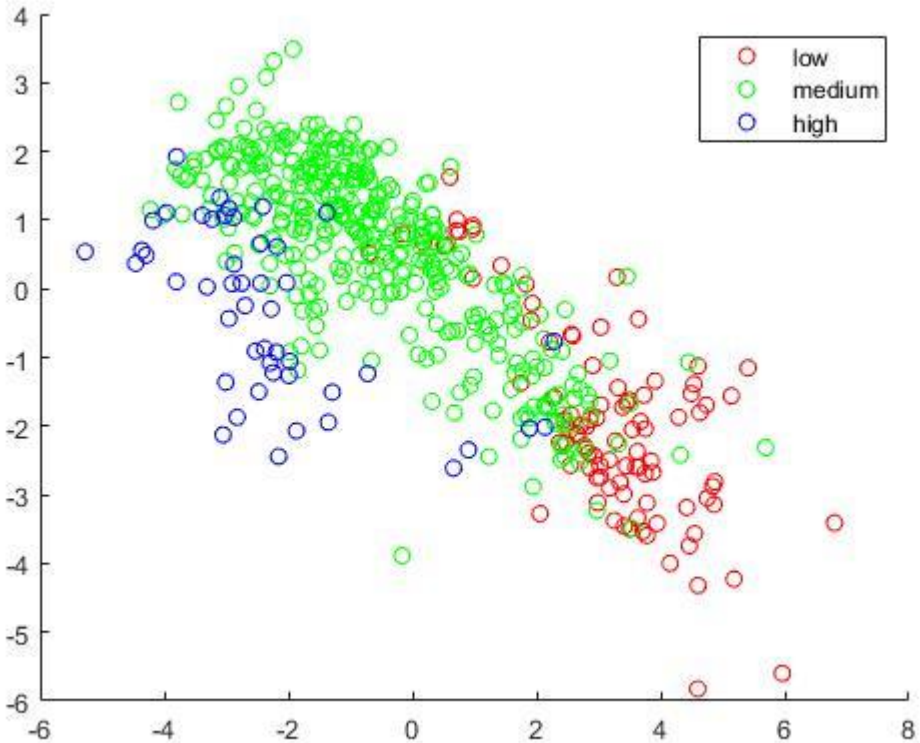
Linear Discriminant Analysis (continued)

- Now let X_b be the matrix with K rows whose k^{th} row is the average of the vectors in X_k
- The **between-class covariance matrix** is the covariance matrix C_b of X_b
- Intuitively, to separate the classes we want to find a direction in the vector space that, simultaneously,
  - Maximises the between-class variance (C_b)
  - Minimises the within-class variance (C_w)
- The vector that has this property is the eigenvector of $C_b C_w^{-1}$ corresponding to the biggest eigenvalue

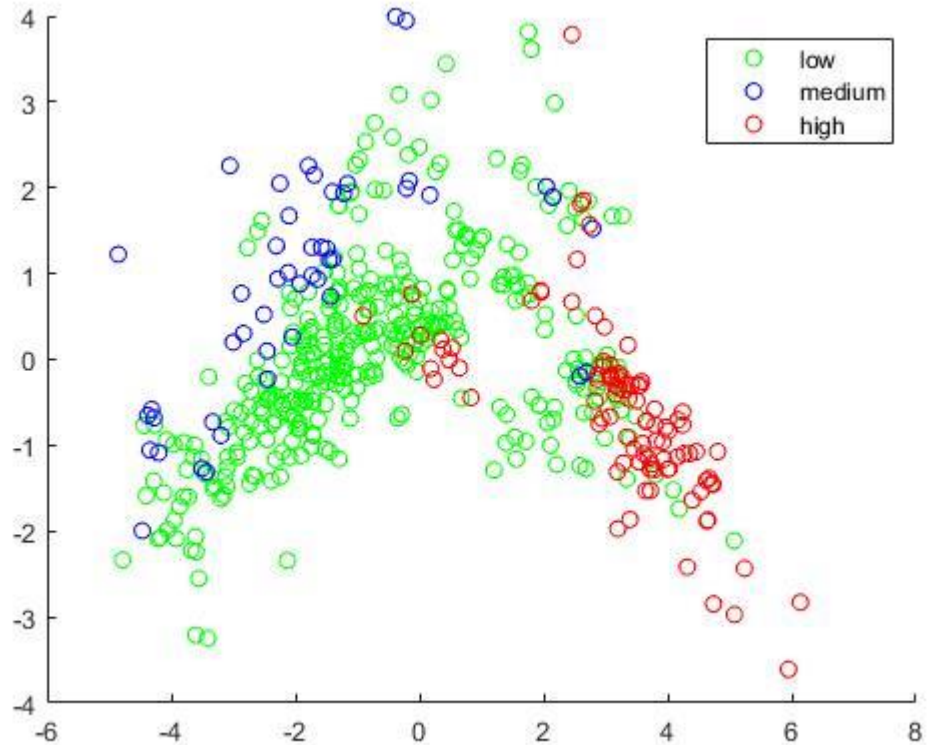
Projection onto 1st two LDA eigenvectors



Comparison of PCA and LDA



LDA



PCA

Things to remember

- PCA finds the direction of maximum variance in a data set
- LDA finds the direction of maximum class separation in a dataset that is partitioned into classes
- Learn the methods for both techniques