Linear Discriminant Analysis (LDA)

Prof. G. Panda
Professorial fellow
IIT, Bhubaneswar, India

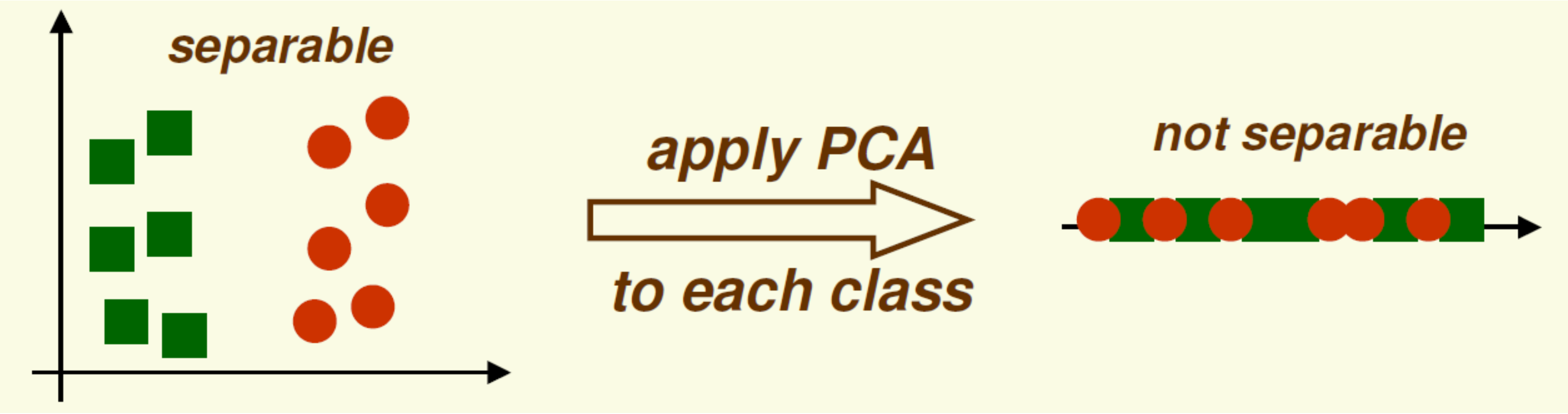
1/18/2019

LDA Objective

- The objective of LDA is to perform dimensionality reduction.
  - So what, PCA does this….
- However, we want to preserve as much of the class discriminatory information as possible.

Recall PCA

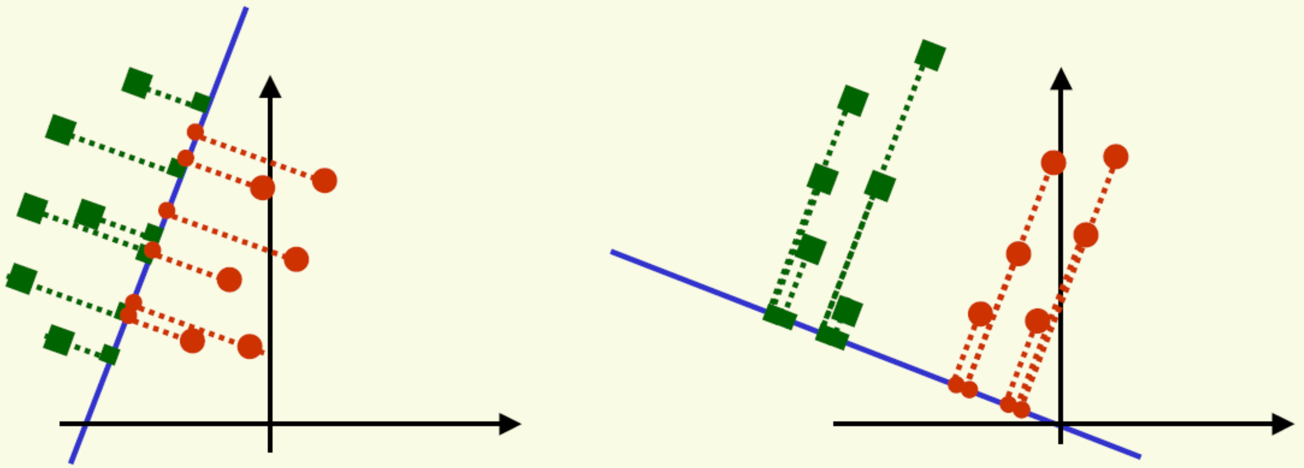
- PCA finds the most accurate *data representation* in a lower dimensional space
- Project data in the directions of maximum variance



LDA Motivations

Main Idea: find projection to a line such that samples from different classes are well separated

Example in 2D



bad line to project to, classes are mixed up

good line to project to, classes are well separated

PCA vs LDA

PCA	LDA
Unsupervised	Supervised
Best represents the data	Best discriminates the data
Project the data in the directions of maximum variance	Project the data that maximizes the class separability
May not be good for classification	Good for classification

Linear Discriminant Analysis (LDA) for two classes

- Suppose we have 2 classes and p-dimensional samples
  - n1 samples come from the first class
  - n2 samples come from the second class
- We seek to obtain a scalar z by projecting the samples onto a line (c-1 space, c=2)
- Of all the possible lines we would like to select the one that maximizes the separability of the scalars.

Linear Discriminant Analysis (LDA) for Two Classes

- The task of LDA is to project on line in the direction v which maximizes

want projected means are far from each other

$$J(v) = \frac{(m_1 - m_2)^2}{S_1^2 + S_2^2}$$

want scatter in class 1 is as small as possible, i.e. samples of class 1 cluster around the projected mean m_1

want scatter in class 2 is as small as possible, i.e. samples of class 2 cluster around the projected mean m_2

LDA Algorithm for Two Classes

- **Step 0:** For a given dataset D={X,Y}, separate samples of class 1 (D1={X_1,Y_1}) and class 2 (D2={X_2,Y_2})
- **Step 1:** Compute the zero mean data and mean vector for each class

$$m_1 = \begin{pmatrix} \bar{X}_1 \\ \bar{Y}_1 \end{pmatrix}, \; m_2 = \begin{pmatrix} \bar{X}_2 \\ \bar{Y}_2 \end{pmatrix}$$

- **Step 2:** Compute scatter matrices S1 and S2 for each class

$$S_1 = \begin{pmatrix} var(X_1) & cov(X_1, Y_1) \\ cov(Y_1, X_1) & var\,(Y_1) \end{pmatrix}$$

$$S_2 = \begin{pmatrix} var(X_2) & cov(X_2, Y_2) \\ cov(Y_2, X_2) & var\,(Y_2) \end{pmatrix}$$

- **Step 3:** Calculate the within-class scatter matrix

$$S_w = S_1 + S_2$$

LDA Algorithm for Two Classes (Contd..)

- **Step 4:** Calculate the inverse of the within-class scatter matrix (S_w^{-1})
- **Step 5:** Calculate the best eigenvector (Direct method)

$$\vec{v} = S_w^{-1}(m_1 - m_2)$$

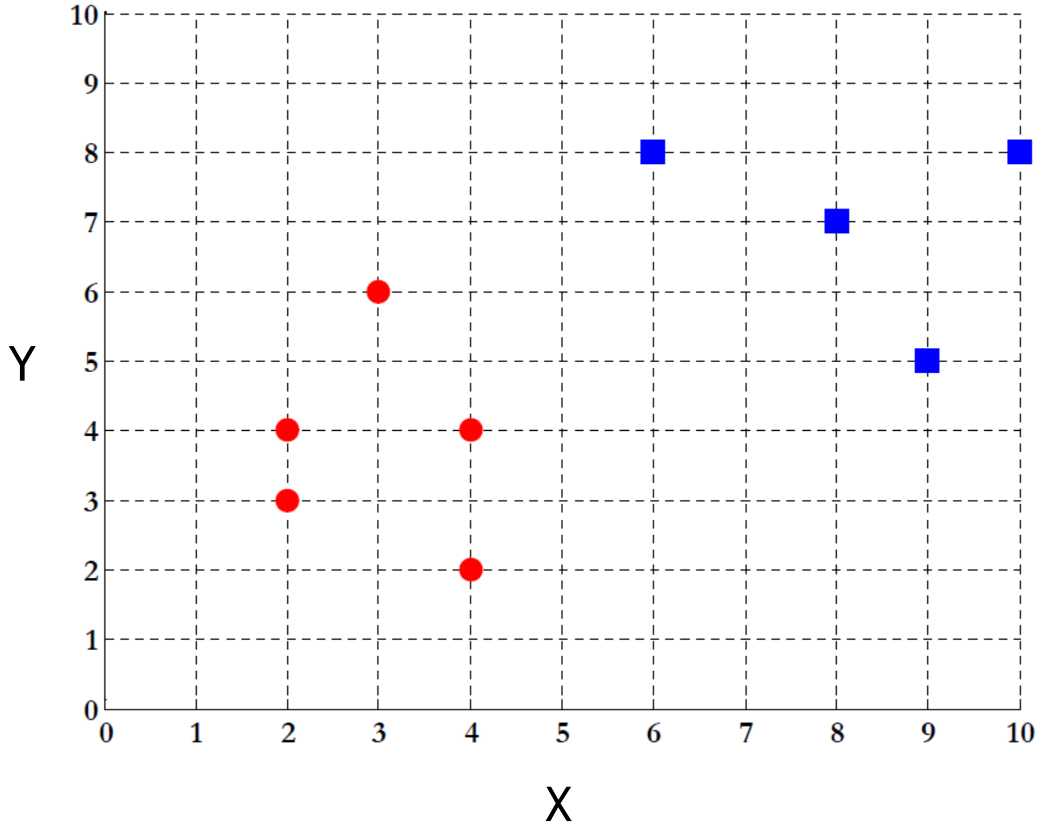
- **Step 6:** Project the samples of each class in the direction of v (Feature reduction)

$$z_1 = D_1 * v$$
$$z_2 = D_2 * v$$

LDA Example –STEP 1

Samples for class-1 (D1)	
X_1	Y_1
4	2
2	4
2	3
3	6
4	4
$\bar{X}_1 = 3$	$\bar{Y}_1 = 3.8$

Samples for class-2 (D2)	
X_2	Y_2
9	10
6	8
9	5
8	7
10	8
$\bar{X}_2 = 8.4$	$\bar{Y}_2 = 7.6$



$$m_1 = \begin{pmatrix} \bar{X}_1 \\ \bar{Y}_1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3.8 \end{pmatrix}$$

$$m_2 = \begin{pmatrix} \bar{X}_2 \\ \bar{Y}_2 \end{pmatrix} = \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix}$$

LDA Example –STEP 1

Zero mean data for class-1	
$\boldsymbol{X_1} - \bar{X}_1$	$\boldsymbol{Y_1}$-$\bar{Y}_1$
1	-1.8
-1	0.2
-1	-0.8
0	2.2
1	0.2
$\bar{X}_1 = 0$	$\bar{Y}_1 = 0$

Zero mean data for class-2	
$\boldsymbol{X_2}$-$\bar{X}_2$	$\boldsymbol{Y_2} - \bar{Y}_2$
0.6	2.4
-2.4	0.4
0.6	-2.6
- 0.4	-0.6
1.6	0.4
$\bar{X}_2 = 0$	$\bar{Y}_2 = 0$

LDA Example –STEP 2

- **Calculate the scatter matrices**

$$S_1 = \begin{pmatrix} var(X_1) & cov(X_1, Y_1) \\ cov(Y_1, X_1) & var\,(Y_1) \end{pmatrix} = \begin{pmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{pmatrix}$$

$$S_2 = \begin{pmatrix} var(X_2) & cov(X_2, Y_2) \\ cov(Y_2, X_2) & var\,(Y_2) \end{pmatrix} = \begin{pmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{pmatrix}$$

LDA Example –STEP 3

- **Calculate the within-class scatter matrix**

$$S_w = S_1 + S_2$$

$$= \begin{pmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{pmatrix} + \begin{pmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{pmatrix}$$

$$= \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}$$

LDA Example –STEP 4

- **Compute the inverse of within-class scatter matrix**

$$S_w^{-1} = \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix}$$

LDA Example –STEP 5

- **Calculate the best eigenvector (Direct method)**

$$\vec{v} = S_w^{-1}(m_1 - m_2)$$

$$= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix}$$

$$= \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}$$

LDA Example –STEP 6

- **Reduce dimensionality and form feature vector**
  - Size of feature vector of class-1: 5×2
  - Size of v: 2×1
  - Resulted feature vector will be of size 5×1

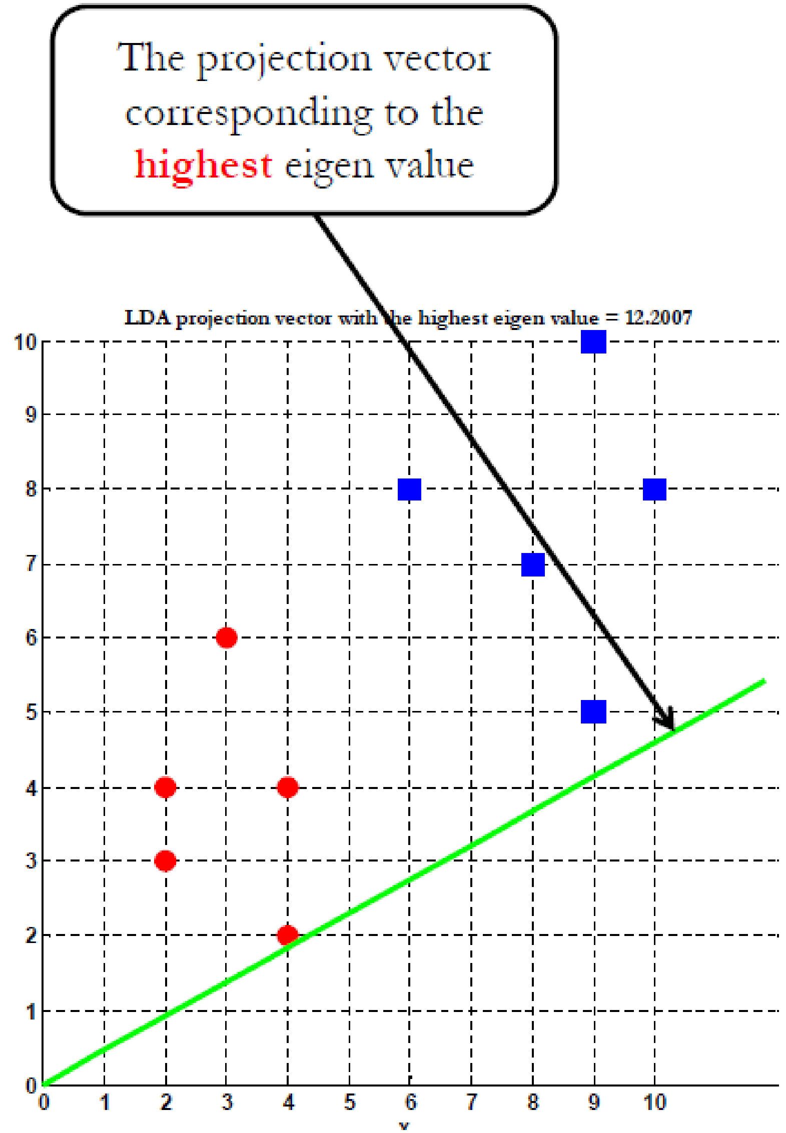
$$z_1 = D_1 * v$$

z_1
4.4698
3.4868
3.0695
5.2302
5.3044

$$z_2 = D_2 * v$$

z_2
12.3522
8.7912
10.2657
10.1915
12.4264

LDA Projection



LDA Example STEP 5 (Another method)

- **Calculate the eigenvalues and eigenvector (Another method)**

$$S_w^{-1} S_b v = \lambda v$$

$$\Rightarrow |S_w^{-1} S_b - \lambda I| = 0$$

Where, $S_b = (m_1 - m_2)(m_1 - m_2)^T$

Thank You