Linear Discriminant Analysis

Dr. Avinash Kumar Singh

Founder, Robotics and Artificial Intelligence Training Academy

Senior Researcher Montpellier University France

Discussion Points

- Linear Discriminant Analysis
  - Objective
  - Steps
  - Use Case – Face Recognition

Objective

- Used to reduce the dimension of multivariate data, while preserving as much of the relevant information as possible.
  - If we have samples represented in the m dimensional space, $\{x_1, x_2, x_3, x_4, x_5, x_6, \ldots x_M\}$, our intension is to find out k features, such that $k \leq m$ that preserve most of the variance present in the data.
- However, we want to preserve as much of the class discriminatory information as possible.

Derivation

- In many physical, biological, and statistical convention it is desirable to represent a system of points with the help of a line or plane.
- Therefore, we can represent the equation of line with the linear combination of these points(variables).

$$y_k = \sum_{i=1}^{m} ak_i x_i \quad \text{Where K=1,2,3, ……. p}$$

$$\begin{array}{c} C_1 \\ \\ C_2 \\ \\ C_n \end{array} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ y_p \end{bmatrix} = \begin{bmatrix} a1_1 & a1_2 & a1_3 & . & a1_m \\ a2_1 & a2_2 & a2_3 & . & a2_m \\ a3_1 & a3_2 & a3_3 & . & a3_m \\ . & . & . & . & . \\ ap_1 & ap_2 & ap_3 & . & ap_m \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ x_m \end{bmatrix}$$

Steps used in LDA: [Step-1: Mean Calculation]

- Calculate mean of each class$(\mu_1, \mu_2, \dots, \mu_n)$ and global mean (μ):

Let us assume, we have samples having m features. We have stored all these in a matrix called training samples T having N rows and M columns (where N represents the population and M represents the Features).So in that regard we have data like.

μ_1	$a1_1$	$a1_2$	$a1_3$	.	$a1_m$
	$a2_1$	$a2_2$	$a2_3$	.	$a2_m$
μ_2	$a3_1$	$a3_2$	$a3_3$	.	$a3_m$
	.	.	.	.	.
μ_n	ap_1	ap_2	ap_3	.	ap_m

Step-1: Calculate Between Class and Within Class Scatter Matrix

$a1_1$	$a1_2$	$a1_3$	.	$a1_m$
$a2_1$	$a2_2$	$a2_3$	.	$a2_m$
$a3_1$	$a3_2$	$a3_3$	.	$a3_m$
.	.	.	.	.
ap_1	ap_2	ap_3	.	ap_m

- Between Class Scatter Matrix
  - Represents the distance between the mean of two/n classes.

$$SB = \sum_{i=1}^{n} N_i(\mu_i - \mu) * (\mu_i - \mu)^T$$

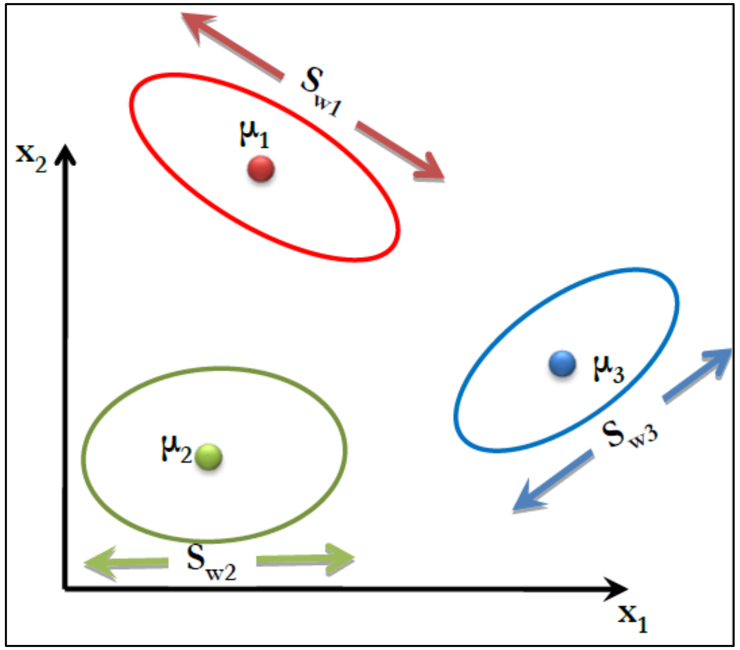
- Within Class Scatter Matrix
  - Represents the variance of each class.

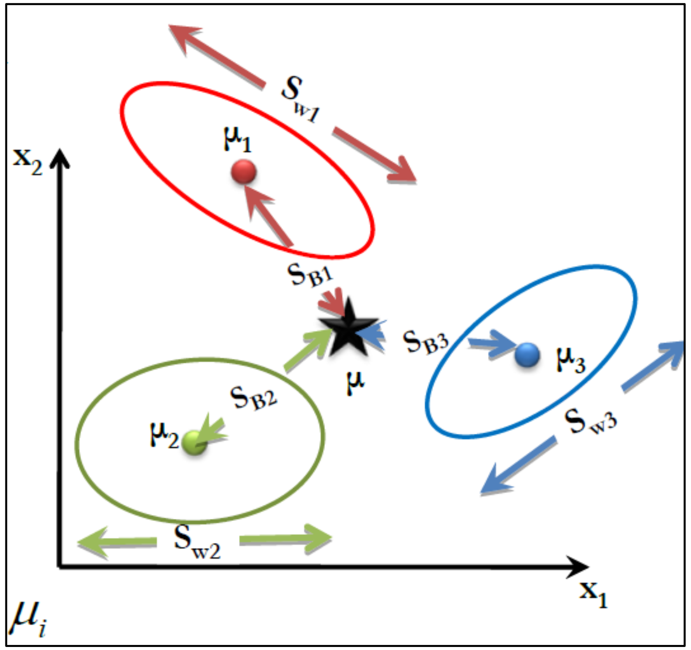
$$SW = \sum_{i=1}^{n} (s_i - \mu_i) * (s_i - \mu_i)^T$$

Step-3: Calculate Cost Function

Sir Ronald Aylmer Fisher

- Cost Function:
  - Minimize the within class scatter matrix while maximize the between class scatter matrix. $Cost(\Psi) = \frac{SB}{SW}$





https://en.wikipedia.org/wiki/Ronald_Fisher

Step-4: Eigen Vector and Eigen Value Decomposition

- Linear discriminant components are the Eigen values and Eigen vectors, those are computed on the basis of cost function (Ψ), calculated in the previous step.

$$\Psi_{m*m}$$

Eigen Value (λ_{m*m})

Eigne Vector (Ω_{m*m})

Step:4 Eigen Vector Selection

- Sort the eigen values in the descending order and the eigen vector as well.
- Select the k components out of m based on how much variance you want to preserve.

W=

$w1_1$	$w1_2$	$w1_3$		$w1_m$
$w2_1$	$w2_2$	$w2_3$		$w2_m$
$w3_1$	$w3_2$	$w3_3$		$w3_m$
.	.	.		.
wm_1	wm_2	wm_3		wm_m

m*m

$W_k =$

$w1_1$	$w1_2$	$w1_3$
$w2_1$	$w2_2$	$w2_3$
$w3_1$	$w3_2$	$w3_3$
.	.	.
wm_1	wm_2	wm_3

m*k

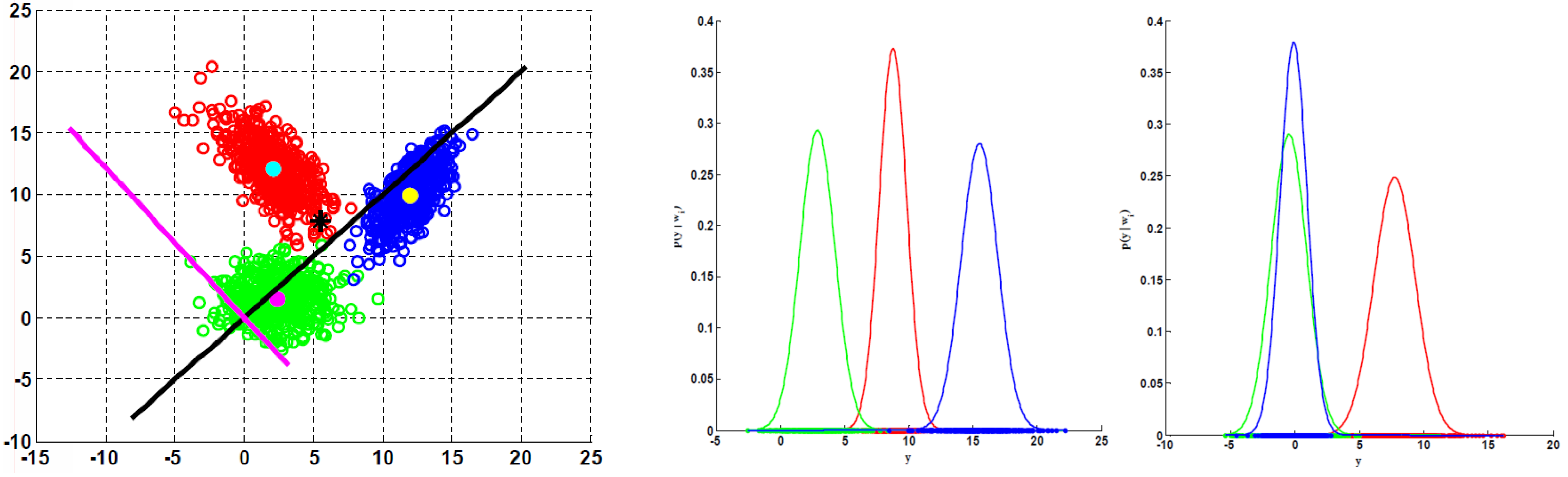
Step 5: Projecting the data to principal directions

$$ProjectedSamples_{p*k} = D_{p*m} * W_{m*k}$$

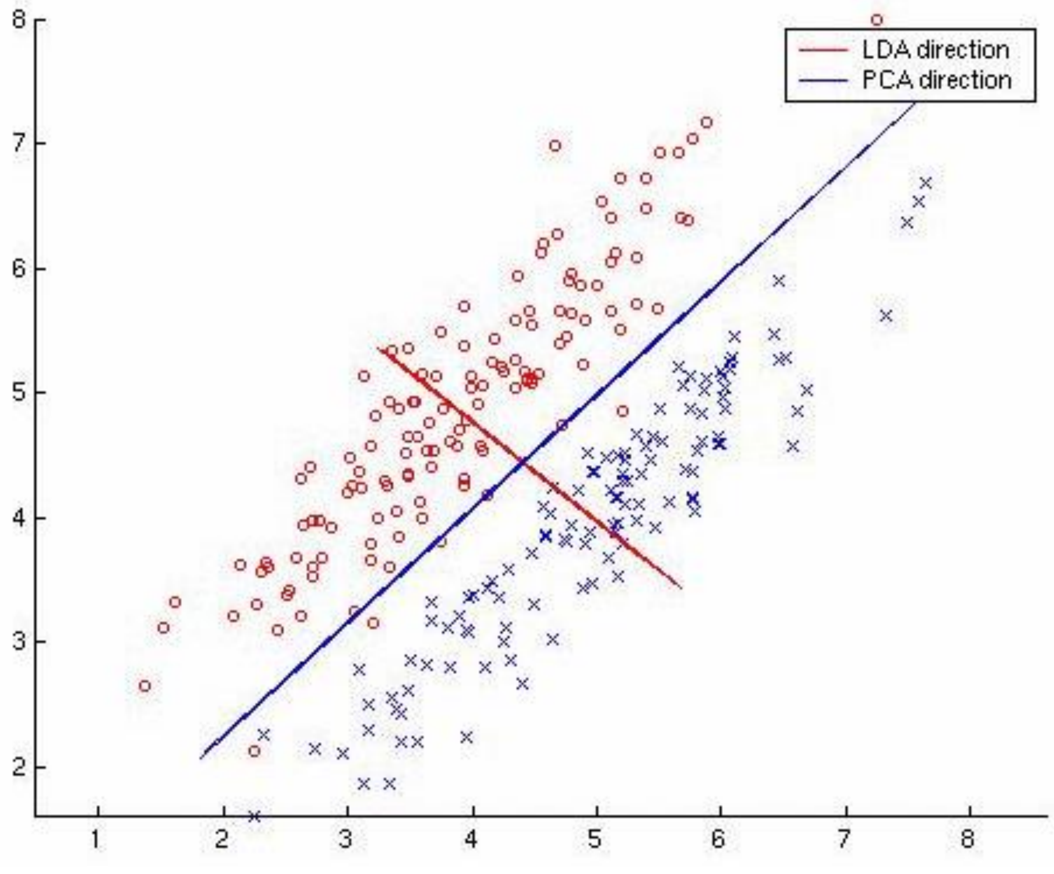
$$\begin{bmatrix} a1_1 & a1_2 & a1_3 & . & a1_m \\ a2_1 & a2_2 & a2_3 & . & a2_m \\ a3_1 & a3_2 & a3_3 & . & a3_m \\ . & . & . & . & . \\ ap_1 & ap_2 & ap_3 & . & ap_m \end{bmatrix} \times \begin{bmatrix} w1_1 & w1_2 & w1_3 \\ w2_1 & w2_2 & w2_3 \\ w3_1 & w3_2 & w3_3 \\ . & . & . \\ wm_1 & wm_2 & wm_3 \end{bmatrix}$$

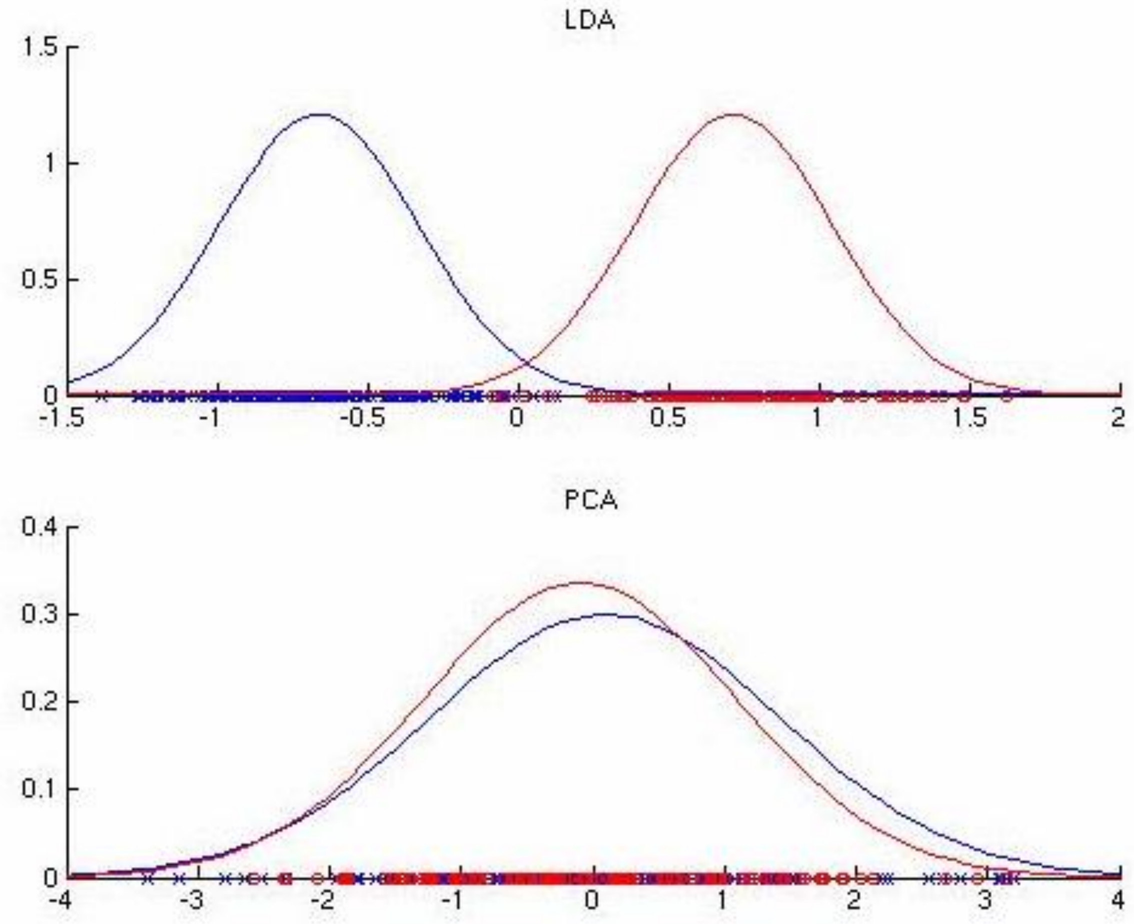
LDA : Another use case

$$\lambda = \begin{bmatrix} 4508.2089 \\ 1878.8511 \end{bmatrix}$$



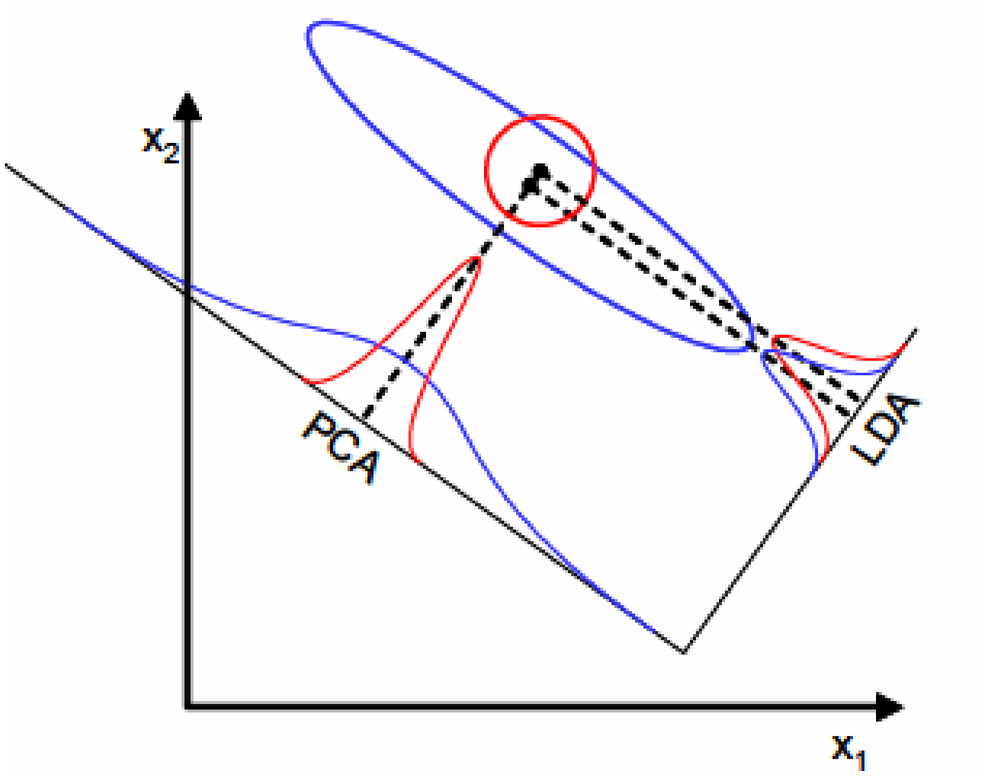
PCA VS LDA





When to use LDA and When PCA

- LDA will fail when the discriminatory information is not in the mean but rather in the variance of the data



Limitation