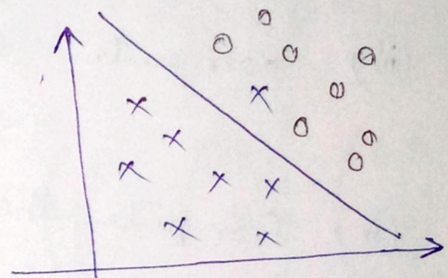
Date - 22 Jan 2025

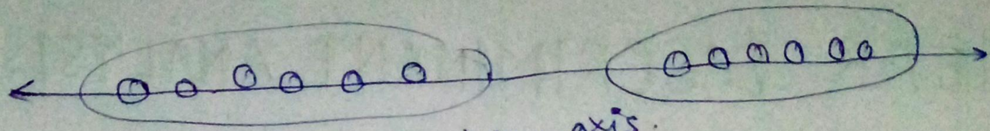
LINEAR DISCRIMINANT ANALYSIS (LDA)

- linear discriminant analysis or normal discriminant analysis or discriminant function analysis is a dimensionality reduction technique that is commonly used for supervised classification problems.
- LDA is used to project the features in higher dimension space into a lower dimension space.



- suppose we have two sets of data points belonging to two different classes that we want to classify
- when the data points are plotted on the 2D-plane, there's no straight line that can separate the two classes of the data points completely.

- Here, LDA uses both the axes (X and Y) to create a new axis and projects data onto a new axis in a way to maximize the separation of two categories and hence, reducing the 2-D graph into a 1-D graph.
- Two criteria are used by LDA to create new axis:
  (i) maximize the distance between means of the two classes
  (ii) minimize the variation within each class



new axis.

* linear discriminant Analysis steps:

(i) Compute the class means of dependent variable.

$$\mu_1 = \frac{1}{N_1} \sum_{x \in \omega_1} x$$

(ii) Derive the covariance matrix of the class variable.

$$S_1 = \sum_{x \in \omega_1} (x - \mu_1)(x - \mu_1)^T$$

(iii) compute the within class-scatter matrix

$$S_w = S_1 + S_2$$

(iv) Compute the between class scatter matrix

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

(v) Compute the eigen values and eigen vectors from the within class and between class. scatter matrix

$$S_w^{-1} S_B \omega = \lambda \omega$$

Here we got λ values for each class

(vi) sort the values of eigen values and select the top k values.

(vi) find the eigen vector corresponds to the. top k eigen vectors.

$$\left(S_w^{-1} S_B - \lambda I\right)\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = 0$$

(vii) Obtain the LDA by taking the dot product of. eigen vectors and original data.

example:

compute the linear discriminant projection for the following two dimensional dataset

samples for class ω_1:

$$X_1 = (x_1, x_2) = \{(4,2), (2,4), (2,3), (3,6), (4,4)\}$$

samples for class ω_2:

$$X_2 = (x_1, x_2) = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$$

(i) The classes mean are:

$$\mu_1 = \frac{1}{N_1} \sum_{x \in \omega_1} x = \frac{1}{5}\left[\begin{pmatrix}4\\2\end{pmatrix} + \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}2\\3\end{pmatrix} + \begin{pmatrix}3\\6\end{pmatrix} + \begin{pmatrix}4\\4\end{pmatrix}\right] = \begin{pmatrix}3\\3.8\end{pmatrix}$$

$$\mu_2 = \frac{1}{N_2} \sum_{x \in \omega_1} x = \frac{1}{5}\left[\begin{pmatrix}9\\10\end{pmatrix} + \begin{pmatrix}6\\8\end{pmatrix} + \begin{pmatrix}9\\5\end{pmatrix} + \begin{pmatrix}8\\7\end{pmatrix} + \begin{pmatrix}10\\8\end{pmatrix}\right] = \begin{pmatrix}8.4\\7.6\end{pmatrix}$$

(ii) Covariance matrix of the first class:

$$S_1 = \sum_{x \in \omega_1} \frac{(x-\mu_1)(x-\mu_1)^T}{N-1}$$

$$= \frac{\left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T + \left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T + \left[\begin{pmatrix}2\\3\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}2\\3\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T + \left[\begin{pmatrix}3\\6\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}3\\6\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T + \left[\begin{pmatrix}4\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}4\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T}{N-1}$$

$$= \begin{pmatrix}1 & -0.25\\ -0.25 & 2.2\end{pmatrix}$$

similarly

$$S_2 = \begin{pmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{pmatrix}$$

(iii) within - class scatter matrix

$$S_W = S_1 + S_2 = \begin{pmatrix} 1 & -0.25 \\ -0.25 & 2.2 \end{pmatrix} + \begin{pmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{pmatrix}$$

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

(iv) between - class scatter matrix.

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

$$= \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix}\right]\left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix}\right]^T$$

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

$$= \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix}$$

(v) find eigen values:

$$S_W^{-1} S_B W = \lambda W$$

$$|S_W^{-1} S_B - \lambda I| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\Rightarrow \left| \begin{pmatrix} 9.2213 - \lambda & 6.489 \\ 4.2339 & 2.9794 - \lambda \end{pmatrix} \right|$$

$$\Rightarrow (9.2213 - \lambda)(2.9794 - \lambda) - 6.489 \times 4.2339 = 0$$

$$\Rightarrow \lambda^2 - 12.2007\lambda = 0$$

$$\Rightarrow \lambda(\lambda - 12.2007) = 0$$

$$\Rightarrow \lambda_1 = 0 \text{ and } \lambda_2 = 12.2007$$

(vi) find eigen vector

$$\left(S_\omega^{-1} S_B - \lambda I \right) \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = 0$$

$$\omega_1 = \begin{pmatrix} -0.5755 \\ 0.8178 \end{pmatrix} \qquad \omega_2 = \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix} = \omega^*$$

$$W^* = S_\omega^{-1}(\mu_1 - \mu_2) = \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix} = \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}$$

(vii) obtain the LDA by taking the dot product of eigen vectors and original data.

x_1	4	2	2	3	4	9	6	9	8	10
x_2	2	4	3	6	4	10	8	5	7	8
1st LD	4.46	3.48	3.06	5.2	5.3	12.35	8.8	10.2	10.19	12.42

example $\begin{bmatrix} 4, & 2 \end{bmatrix}_{1\times 2} \begin{bmatrix} 0.9088 \\ 0.4173 \end{bmatrix}_{2\times 1} = 4.46$