DEPARTMENT OF COMPUTER SCIENCE

Machine Learning

K S Induja Suresh
22cs02indu@pg.cusat.ac.in
January 11, 2023

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis also called LDA is a dimensionality reduction technique for reducing the unnecessary features to make the model more precise.It is mainly used to project high dimensional space to lower dimensional space in order to reduce the resources and the cost.
Usually if there are more features, they tend to increase the accuracy of the data but here it faces the curse of dimensionality (i.e., it cannot tackle all the features) in which it reduces a set of features to another set of features .Not only it reduces the number of features it also find the axis that maximizes the separation between multiple classes

Steps:

- Step 1:Calculate the mean value of class 1 and class 2 (μ_1 and μ_2)
- Step 2:Calculate within the class scatter matrix (S_w) (It captures how well the data is scattered within its class)

$$S_w = S_1 + S_2$$

  S_1 =Covariance matrix of class 1
  S_2 =Covariance matrix of class 2

$$S_1 = \sum (x - \mu_1)(x - \mu_1)^T$$
$$S_2 = \sum (x - \mu_2)(x - \mu_2)^T$$

- Step 3:Calculating class scatter matrix (S_B)

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

- Step 4:Find the best LDA projection vector
  (Finding out using the eigen vectors having largest eigen value)

$$\begin{bmatrix} V_1 \\ V2 \end{bmatrix} = S_w^{-1}(\mu_1 - \mu_2)$$

- Step 5: Dimension Reduction

$$Y = W^T X$$

  where W^T is the projection vector and X is the Input data sample

Example:

Taking a 2D data set:
C_1 = (4,1),(2,4),(2,3),(3,6),(4,4)
C_2 = (9,10),(6,8),(9,5),(8,7),(10,8)
Finding μ_1 and μ_2 :

$$\mu_1 = \frac{4+2+2+3+4}{5}, \frac{1+4+3+6+4}{5}$$
$$\mu_1 = (3, 3.6)$$

Similarly,

$$\mu_2 = (8.4, 7.6)$$

Calculating $S_1 = \sum(x - \mu_1)(x - \mu_1)^T$

$$(x_1 - \mu_1) = \begin{bmatrix} -1 & -1 & -1 & 0 & 1 \\ -2.6 & 0.4 & -0.6 & 2.4 & 0.4 \end{bmatrix}$$

$$(x_1 - \mu_1)^T = \begin{bmatrix} -1 & -2.6 \\ -1 & 0.4 \\ -1 & -0.6 \\ 0 & 2.4 \\ 1 & 0.4 \end{bmatrix}$$

Calculating $(x - \mu_1)(x - \mu_1)^T$ for each matrix. So at last we have a total of 5 matrices

$$\begin{bmatrix} 1 \\ -2.6 \end{bmatrix} * \begin{bmatrix} 1 & -2.6 \end{bmatrix} = \begin{bmatrix} 1 & -2.6 \\ -2.6 & 6.76 \end{bmatrix}$$

Similarly for the other 4 matrices (after performing = $(x - \mu_1)(x - \mu_1)^T$
Results:

$$\begin{bmatrix} 1 & -0.4 \\ -0.4 & 0.16 \end{bmatrix}, \begin{bmatrix} 1 & 0.6 \\ 0.6 & 0.36 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 5.76 \end{bmatrix}, \begin{bmatrix} 1 & 0.4 \\ 0.4 & 0.16 \end{bmatrix}$$

Adding all the matrices

$$S_1 = \begin{bmatrix} 4 & -2 \\ -2 & 13.18 \end{bmatrix}$$

Finding the mean of the resultant covariance matrix ($S_1/5$): :

$$S_1 = \begin{bmatrix} 0.8 & -0.4 \\ -0.4 & 2.636 \end{bmatrix}$$

Similarly,
Calculating $S_2 = \sum (x - \mu_2)(x - \mu_2)^T$

$$(x_2 - \mu_2) = \begin{bmatrix} 0.6 & -2.4 & 0.6 & -0.4 & 1.6 \\ 2.4 & 0.4 & -2.6 & -0.6 & 0.4 \end{bmatrix}$$

$$(x_2 - \mu_2)^T = \begin{bmatrix} 0.6 & 2.4 \\ -2.4 & 0.4 \\ 0.6 & -2.6 \\ -0.4 & -0.6 \\ 1.6 & 0.4 \end{bmatrix}$$

Calculating $(x - \mu_2)(x - \mu_2)^T$ for each matrix. So at last we have a total of 5 matrices
Results:

$$\begin{bmatrix} 0.36 & 1.44 \\ 1.44 & 5.76 \end{bmatrix}, \begin{bmatrix} 5.76 & -0.96 \\ -0.96 & 0.16 \end{bmatrix}, \begin{bmatrix} 0.36 & -1.56 \\ -1.56 & 6.76 \end{bmatrix}, \begin{bmatrix} 0.16 & 0.24 \\ 0.24 & 0.36 \end{bmatrix}, \begin{bmatrix} 2.56 & 0.64 \\ 0.64 & 0.16 \end{bmatrix}$$

Adding all the matrices

$$S_2 = \begin{bmatrix} 9.2 & -0.2 \\ -0.2 & 13 \end{bmatrix}$$

Finding the mean of the resultant covariance matrix ($S_2/5$): :

$$S_2 = \begin{bmatrix} 1.84 & -0.04 \\ -0.04 & 2.64 \end{bmatrix}$$

Calculating S_w :

$$S_w = S_1 + S_2$$

$$S_w = \begin{bmatrix} 2.64 & -0.44 \\ -0.44 & 5.276 \end{bmatrix}$$

Finding the best LDA projection vector:

$$\begin{bmatrix} V_1 \\ V2 \end{bmatrix} = S_w^{-1}(\mu_1 - \mu_2)$$

Calculating the inverse matrix S_W^{-1} :

$$S_w^{-1} = \begin{bmatrix} 2.64 & -0.44 \\ -0.44 & 5.276 \end{bmatrix}^{-1} = \begin{bmatrix} 0.3841 & 0.0320 \\ 0.0320 & 0.1922 \end{bmatrix}$$

Rearranging the values of S_W^{-1}
Resultant matrix:

$$S_w^{-1} = \begin{bmatrix} 0.1922 & -0.0320 \\ -0.0320 & 0.3841 \end{bmatrix}$$

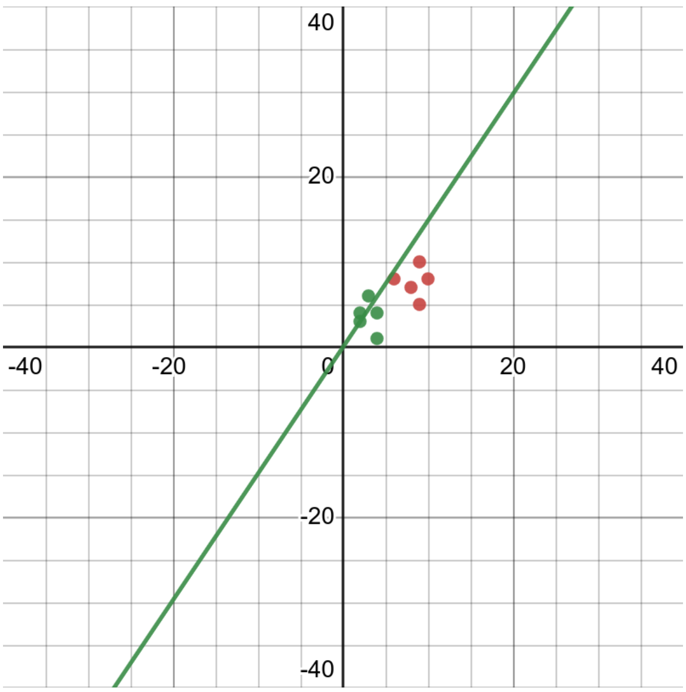
Finding V_1 and V_2

$$S_w^{-1}(\mu_1 - \mu_2) = \begin{bmatrix} 0.1922 & -0.0320 \\ -0.0320 & 0.3841 \end{bmatrix} * \begin{bmatrix} -5.4 \\ -4 \end{bmatrix}$$

Reducing the value to fit the graph:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} -0.9098 \\ -1.3526 \end{bmatrix}$$

After Performing dimension reduction:



Singular Value Decomposition (SVD)

They are dimensionality reduction tool that mainly deals with the factorization of a matrix to three other matrices U,S and V respectively.

$$A = U\Sigma V^T$$

where U and V are orthogonal matrices with the orthonormal eigen vectors and Σ denotes the diagonal matrix.
It is related to polar decomposition. There are many areas where SVD can be applied mainly.,Image Compression,Image Recovery etc. SVD can be used to show the original value of matrix as a linear combination of low rank matrices.

Steps:

- Step 1: Convert the matrix to a square matrix by multiplying with the transpose.
- Step 2 : Find the eigen values of $A^T A$ by finding out the value of λ from the equation
- Step 3: After converting the eigen values , calculate the eigen vectors for each value using Row-Echelon form
- Step 4: For finding U in the equation, we need to multiply the inverse of S and V on both sides of the equation

Example:

Taking a matrix of size 2*2

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$$

Converting the matrix A to a square matrix by multiplying the matrix with the transpose of it.

$$A^T A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} * \begin{bmatrix} 2 & 0 \\ 3 & 2 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 13 & 6 \\ 6 & 4 \end{bmatrix}$$

Calculating the eigen values of a matrix by finding out lambda by solving the equation

$$A^T A - \lambda I = \begin{bmatrix} 13-\lambda & 6 \\ 6 & 4-\lambda \end{bmatrix}$$

By solving we get the equation :

$$\lambda^2 - 17\lambda + 16 = 0$$

Solving we get

$$\lambda = 16, 1$$

Finding the eigen vectors from eigen values

$$\lambda = 16$$

$$A^T A - 16I = \begin{bmatrix} -12 & 6 \\ 6 & -3 \end{bmatrix}$$

For finding the eigen vector, we neeed to find the null space such that AB=0

$$\begin{bmatrix} -12 & 6 \\ 6 & -3 \end{bmatrix} * \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Finding the eigen vectors from eigen values

$$\lambda = 1$$

$$A^T A - 1I = \begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix}$$

For finding the eigen vector, we neeed to find the null space such that AB=0

$$\begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix} * \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Performing normalization with the vector to get :

$$V_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

Similarly,

$$V_2 = \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

hence,

$$V = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{16} & 0 \\ 0 & \sqrt{1} \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

Finding U:

$$\mu_1 = \frac{1}{\sigma_1} AV_1$$

$$\sigma_1 = \sqrt{16} = 4$$

$$\mu_1 = \frac{1}{4} \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

we get,

$$\mu_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

Similarly for μ_2 ,

$$\mu_2 = \frac{1}{\sigma_2} AV_2$$

$$\sigma_2 = \sqrt{1} = 1$$

$$\mu_2 = 1 \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

we get,

$$\mu_2 = \begin{bmatrix} \frac{-1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

The matrix thus obtained is converted to get final U matrix

$$U = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$V^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

Hence the value of matrix A can be decomposed into 3 matrices:

$$A = U\Sigma V^T$$

$$\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

Proving the above Left Hand Side and Right Hand Side are correct:

$$U\Sigma = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} * \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{8}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{4}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$U\Sigma V^T = \begin{bmatrix} \frac{8}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{4}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$$

Therefore,

$$A = U\Sigma V^T$$