MACHINE LEARNING

ATHIRA K VASAVAN

October 2022

Linear Discriminant Analysis(LDA)

Linear Discriminant Analysis(LDA) is very common technique that is commonly used for supervised classification problems in machine learning.it is used for modelling difference in group that is separating two or more classes.Although the logistic regression algorithm is limited to only two class,LDA is applicable for more than two classes of classification problem. LDA technique is to project the original data set on to a lower dimensional space.

The criteria of LDA is:

- it maximizes the distance between means of two classes
- it minimizes the variance within the individual class

Why LDA

LDA can also be used in data prepocessing to reduce the no.of features, just as PCA, which reduce the computing cost significantly.

Steps:

- Step 1:
  To calculate the separation between different class,the mean value of class 1 and class 2 (μ_1 and μ_2)
- Step 2:
  Compute within class scatter matrix (S_w)

  $$S_w = S_1 + S_2$$

  S_1 is the Covariance matrix for the class 1
  S_2 is the Covariance matrix for the class 2

so,let's now find the covariance matrices of each class

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

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

- Step 3:
  compute between class scatter matrix (S_B)

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

- Step 4:
  Find the best LDA projection vector
  (similar to principal component analysis(PCA) we find this using eigen vectors having largest eigen value)

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

- Step 5:
  Dimension Reduction

$$Y = W^T X$$

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

Example:

C_1 = (4,1),(2,4),(2,3),(3,6),(4,4)
C_2 = (9,10),(6,8),(9,5),(8,7),(10,8)
compute μ_1 and μ_2 : μ_1 is the mean of class c_1 which is computed by

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

Similarly,

μ_2 is the mean of class c_2 which is computed by

$$\mu_2 = \frac{9+6+9+8+10}{5}, \frac{10+8+5+7+8}{5}$$

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

After calculating S_1 and S_2

$$(x_1 - \mu_1) = \begin{bmatrix} 1 & -1 & -1 & 0 & 1 \\ -2.6 & .4 & -.06 & 2.4 & .4 \end{bmatrix}$$

$$(x_2 - \mu_2) = \begin{bmatrix} .6 & -2.4 & .6 & -.4 & 1.6 \\ 2.4 & .4 & -2.6 & -.6 & .4 \end{bmatrix}$$

Calculating $(x - \mu_1)(x - \mu_1)^T$ for each matrix we get total of 5 matrices

$$\begin{bmatrix} -5.4 \\ -4 \end{bmatrix} * \begin{bmatrix} -5.4 & -4 \end{bmatrix} = \begin{bmatrix} 29.16 & 21.6 \\ 21.6 & 16 \end{bmatrix}$$

Adding all the matrices and finding the average to find the covariance matrix

$$S_1 = \begin{bmatrix} .80 & -.40 \\ -.40 & 2.6 \end{bmatrix}$$

Similarly,

$$S_2 = \begin{bmatrix} 1.84 & -.04 \\ -.04 & 2.64 \end{bmatrix}$$

Calculating S_w :

$$S_w = S_1 + S_2$$

$$= \begin{bmatrix} .80 & -.40 \\ -.40 & 2.6 \end{bmatrix} + \begin{bmatrix} 1.84 & -.04 \\ -.04 & 2.64 \end{bmatrix}$$

$$= \begin{bmatrix} 2.64 & -.44 \\ -.44 & 5.28 \end{bmatrix}$$

Calculating class scatter matrix,
$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$

$$S_B = \begin{bmatrix} 29.16 & 21.6 \\ 21.6 & 16 \end{bmatrix}$$

Finding the best LDA projection vector: Substituting the values in the equation we get,

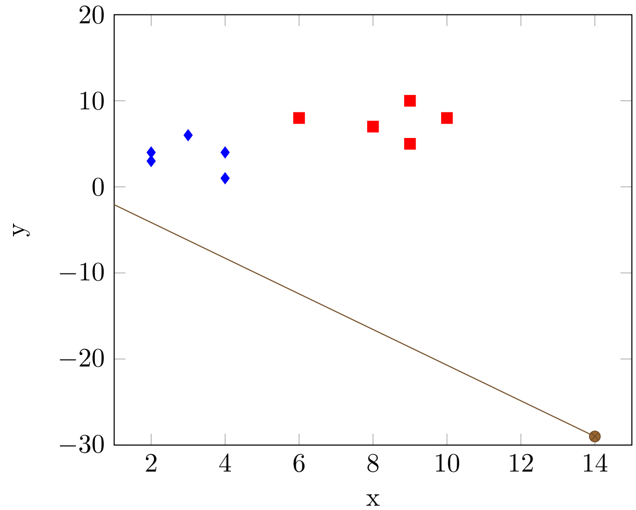
$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} .91 \\ .39 \end{bmatrix}$$

Reducing the value to fit the graph:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 14 \\ -29 \end{bmatrix}$$

After Performing dimension reduction:

- blue denotes C_1
- red denotes C_2



Singular Value Decomposition (SVD)

The Singular Value Decomposition of a matrix is a factorization of the matrix into three matrices.The dimensionality reduction tool deals with the factorization of a matrix to three other matrices U,D and V respectively.

Mathematics behind SVD:

the SVD of m*n matrix A is given by the formula

$$A = UDV^T$$

U:m*n matrix of the orthonormal eigenvectors of AA^T
D: n*n diagonal matrix of the singular values which are the square roots of the eigenvalues of AA^T
V^T :transpose of n*n matrix containing the orthonormal eigenvectors of $A^T A$
where U and V are orthogonal matrices with the orthonormal eigenvectors,
S denotes the diagonal matrix
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 eigenvalues of $A^T A$ by finding out the value of λ from the equation

- Step 3: After converting the eigenvalues , calculate the eigenvectors 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

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

find A^T

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

compute

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

$$\begin{bmatrix} 4 & 6 \\ 6 & 13 \end{bmatrix}$$

Calculating the eigenvalues of a matrix by finding out lambda by solving the equation

$$A^T A - \lambda I = 0$$

$$\begin{bmatrix} 4-\lambda & 6 \\ 6 & 13-\lambda \end{bmatrix} = 0$$

By solving we get the equation :

$$(4-\lambda)(13-\lambda) - 6 * 6 = 0$$

$$52 - 17\lambda + \lambda^2 - 36 = 0$$

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

Solving we get

$$\lambda = 16, 1$$

Finding the eigenvectors for each eigenvalues be the matrix equation .
let

$$(A^T A - \lambda I)X = 0$$

let

$$\lambda = 16$$

$$A^T A - 16I = 0$$

$$\begin{bmatrix} 4-16 & 6 \\ 6 & 13-16 \end{bmatrix} * \begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0$$

For finding the eigenvector, we need to find the null space such that AB=0

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

$$-12x1 + 6x2 = 0$$

$$-12x1 = -6x2$$

$$2x1 = x2$$

$$x1/x2 = 1/2$$

similarly

$$6x1 - 3x2 = 0$$

$$6x1 = 3x2$$

$$2x1 = x2$$

$$x1/x2 = 1/2$$

By applying the rules of Row Echelon form, the matrix can be simplified into

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

similarly let

$$\lambda = 1$$

$$A^T A - 1I = 0$$

$$\begin{bmatrix} 4-1 & 6 \\ 6 & 13-1 \end{bmatrix} * \begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0$$

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

$$3x1 + 6x2 = 0$$

$$x1 = -2x2$$

$$x1/x2 = -2/1$$

similarly

$$6x1 + 12x2 = 0$$

$$x1 = -2x2$$

$$x1/x2 = -2/1$$

By applying the rules of Row Echelon form, the matrix can be simplified into

$$\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

the singular value of

$$\sigma 1 = \sqrt{\lambda 1} = \sqrt{16} = 4$$

$$\sigma 2 = \sqrt{\lambda 2} = \sqrt{1} = 1$$

if

$$\sigma$$

is a singular value of A.it's square is an eigenvalues of

$$A^T * A$$

Normalised eigenvector divided the vector by its length ie.

$$V1 = |\frac{x1}{x1}|$$

$$= \sqrt{x1^2 + x2^2}$$

$$= \sqrt{1 + 2^2}$$

$$= \sqrt{5}$$

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

similarly

$$V2 = |\frac{x1}{x1}|$$

$$= \sqrt{x1^2 + x2^2}$$

$$= \sqrt{-2^2 + 1^2}$$

$$= \sqrt{5}$$

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

since calculate U

$$u1 = 1/\sigma 1 AV1$$

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

$$= 1/4 \begin{bmatrix} \frac{2}{\sqrt{5}} + 6\sqrt{5} \\ 0 + 4\sqrt{5} \end{bmatrix}$$

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

$$u2 = 1/\sigma 1 AV2$$

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

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

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

since U

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

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

$$A = UDV^T$$

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

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

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

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

$$= \begin{bmatrix} \frac{10}{\sqrt{5}} & \frac{15}{\sqrt{5}} \\ 0 & \frac{10}{\sqrt{5}} \end{bmatrix}$$

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