Machine Learning Assignment

Ashin Babu

January 11, 2023

Linear Discriminant Analysis

LDA is a dimensionality reduction technique used as a pre processing step for the machine learning applications. The dimensions here means the features of the data. When we have a large number of with large number of features in it, we need to reduce the number of features while preserving the essential features in it.
Goal: We need to reduce an N dimensional data onto a smaller subspace k(k$\leq$ n-1)

Example

Let us take a 2 Dimensional dataset:

$$C_1 \Rightarrow X_1 = (x_1, x_2) = (4,1), (2,4), (2,3), (3,6), (4,4)$$
$$C_2 \Rightarrow X_2 = (x_1, x_2) = (9,10), (6,8), (9,5), (8,7), (10,8)$$

Step1:

Compute the within class scatter matrix S_w

$$S_w = S_1 + S_2$$

S_1 is the convariance matrix for the the class C_1
S_2 is the convariance matrix for the the class C_2

Covariance matrix

$$\text{Covariance matrix} = \begin{bmatrix} var(x) & cov(y,x) \\ cov(x,y) & var(y) \end{bmatrix}$$

To calculate the covariance matrix we need to find x variance, y variance and the covariance of x and y
$\mu_x = \frac{1}{n}\sum_1^n x_i$

$\mu_x = \frac{1}{n}\sum_1^n x_i$

$x\ variance = \frac{1}{n}\sum_1^n (x_i - \mu_x)^2$

$y\ variance = \frac{1}{n}\sum_1^n (y_i - \mu_y)^2$

$cov(x,y) = \frac{1}{n}\sum_1^n (x_i - \mu_x)(y_i - \mu_y)$

In our given data:

$\mu_x = 3\ \mu_y = 3.6$

$x\ variance = \frac{1+1+1+0+1}{5} = 0.8$
$y\ variance = \frac{2.6^2+0.4^2+0.6^2+2.4^2+0.4^2}{5} = 2.64$

$$cov(x,y) = \frac{-2.6-0.4-0.4-0.6}{5} = 0.4$$

covariance matrix
$S_1 = \begin{bmatrix} 0.8 & -0.4 \\ -0.4 & 2.64 \end{bmatrix}$

Similiarly we find covariance matrix for second class

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

Hence, $S_w = S_1 + S_2 \Rightarrow$

$S_w = \begin{bmatrix} 2.64 & -0.44 \\ -0.44 & 5.28 \end{bmatrix}$

Step2:

Compute the between class scatter matrix S_B

$S_b = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^\top$

$S_b = \begin{bmatrix} -5.4 \\ -4 \end{bmatrix} \begin{bmatrix} -5.4 & -4 \end{bmatrix} = \begin{bmatrix} 29.16 & 21.6 \\ 21.6 & 16.00\mathbf{0.0.1} \quad Step3: \end{bmatrix}$

Find the best LDA projection vector. Similar to PCA we find the values using eigen vectors having the largest eigen values. we have:

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

$\Rightarrow$

$$\begin{vmatrix} 11.89-\lambda & 8.81 \\ 5.08 & 3.76-\lambda \end{vmatrix} = 0$$

$\lambda = 15.65$

$Substituting\ \lambda$

$$\begin{bmatrix} 11.89 & 8.81 \\ 5.08 & 3.76 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = 15.6 \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 0.91 \\ 0.39 \end{bmatrix}$$

Step 4:

Dimension reduction
In this step after obtaining the projection vector we use the input data to project the reduced data into a plot.

$$y = V^\top X$$

Scalar Value Decompositon

SVD is also a dimensional reduction technique. In this method the given matrix is decomposed into three distinct matrices.

SVD of a m x n matrix is given as:

A = UWV^T

U : mxn matrix of the orthonormal eigenvectors of AA^T
W : a nxn diagonal matrix of the singular values which are the square roots of the eigenvalues of AA^T
V^T :transpose of a nxn matrix containing the orthonormal eigenvectors of A$^T A$

Example:

Find the SVD of

A= $\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

Step 1:

First we have to find the eigen values of AA^T

$$\text{X=A}^T.A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

characteristic equation $\left|A - \lambda I\right|$=0

$\Rightarrow 4 - 4\lambda + \lambda^2 - 1 = 0$

$\lambda = 3, 1$

Step 2: Calculate the diagonal matrix

from the values, calculate the diagonal matrix

$$\text{S=}\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}$$

$$\text{S}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix}$$

Step 3: Calculate Eigen vectors

Using $\lambda = 3$

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 3 \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

$2X_1 + X_2 = 3X_1$

$X_2 = X_2$

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

$Using \lambda = 1$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 1 \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$
$2X_1 + X_2 = X_1$
$X_2 = -X_1$

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

$\Rightarrow$Eigen vector V=
$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Step 4: Calculate the decomposition matrix

U=AV$^T S$

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

$$\Rightarrow \begin{bmatrix} \frac{\sqrt{2}}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Step 5: Check the orthogonality of the decomposition matrix

A=UV$^T S$

$$\Rightarrow \begin{bmatrix} \frac{\sqrt{2}}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$