LINEAR DISCRIMINANT ANALYSIS (LDA)

Thasneem M H

January 2023

1 Introduction

Linear Discriminant Analysis (LDA) is one of the commonly used dimensionality reduction techniques that are used in machine learning to solve more than two-class classification problems. It is also considered as a pre-processing step for modeling differences in ML and also in applications of pattern classification. LDA is similar to PCA but LDA in addition finds the axes that maximises the separation between multiple classes. The goal of LDA is to project feature space (N-dimensional data) onto a smaller subspace k(k¡=n-1) while maintaining the class discriminatory information.

Lets take an example and workout step by step. Lets take a 2D dataset.

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

Step 1: compute within class scatter matrix(S_w)
$S_w = S_1 + S_2$

S_1 is covariance matrix for class C_1 and S_2 is covariance matrix for class C_2 So lets now find the covariance matrices of each class.

$\mu 1 = (4 + 2 + 2 + 3 + 4/5, 4 + 4 + 3 + 6 + 4/5)$

$\mu 1 = [3 \qquad 3.6]$
$similarly \mu 2 = [8.4 \qquad 7060]$

$Covariance matrix of the first class :$

S$_1 = \sum_{x \epsilon \omega_1} (x - \mu_1) (x - \mu_1)^T =$

$$\begin{pmatrix} 0.8 & -0.4 \\ -0.4 & 2.6 \end{pmatrix}$$

$Covariance matrix of the second class :$

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

Within-class scatter matrix:

$$\mathrm{S}_w = S_1 + S_2 = \begin{pmatrix} 0.8 & -.4 \\ -.4 & 2.6 \end{pmatrix} + \begin{pmatrix} 1.84 & -.04 \\ -.04 & 2.64 \end{pmatrix}$$

$$=\begin{pmatrix} 2.64 & -.44 \\ -.4 & 5.28 \end{pmatrix}$$

Step 2: compute between class scatter matrix

$$\mathrm{S}_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

$$= \begin{pmatrix} -5.4 \\ -4 \end{pmatrix} \begin{pmatrix} -5.4 & -4 \end{pmatrix}$$

$$= \begin{pmatrix} 29.16 & 21.6 \\ 21.6 & 16 \end{pmatrix}$$

$Step 3 : find the best LDA projection vector.$

$Similar to PCA we find this using Eigenvectors having largest eigenvalue.$

$$\mathrm{S}_W^{-1} S_B w = \lambda w$$

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

$We can also find the optimal projection using another method ie;$

$$w^* = S_w^{-1}(\mu_1 - \mu_2) = \begin{pmatrix} 0.1921 & -.032 \\ -.03 & 0.38 \end{pmatrix}^{-1} \left[\begin{pmatrix} -5.4 \\ -4 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]$$

$$= \begin{pmatrix} .30 & .01 \\ .01 & .18 \end{pmatrix} \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix}$$

$$= \begin{pmatrix} .91 \\ .39 \end{pmatrix}$$

2 Singular Value Decomposition(SVD)

Suppose A is an m $\times$ n matrix with rank r. The matrix AA^T will be m $\times$ m and have rank r. The matrix A^TA will be n $\times$ n and also have rank r. Both matrices A^TA and AA^T will be positive semi definite, and will there fore have r (possible repeated) positive eigenvalues, and r linearly independent corresponding eigenvectors. As the matrices are symmetric, these eigenvectors will be orthogonal, and we can choose them to be ortho-normal.

We call the eigenvectors of A^TA corresponding to its non-zero eigen-values v_1, . . . , v_r. These vectors will be in the row space of A. We call the eigenvectors of AA^T corresponding to its non-zero eigenvalues u_1, . . . , u_r.These vectors will be in the column space of A. Now, these vectors have a remarkable relation. Namely,

Av$_1 = \sigma u_1, Av_2 = \sigma_2 u_2, ..., Av_r = \sigma_r u_r$

$where \sigma_1, ...\sigma_r$ are positive numbers called the singualr values of the matrix A.

This relation lets us write

$$\text{A}\big(v_1....v_r\big)=\big(u_1....u_r\big)\begin{pmatrix}\sigma_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \sigma_r\end{pmatrix}$$

This gives us a decomposition AV=U$\sum$

$Notingthatthecolumnsof V areorthonormalwecanrightmultiplybothsidesofthisequalityby$V^T to get A $= U \sum V^T$. This is the singular value decomposition of A.

If we want to we can make V and U square. We just append orthonormal vectors v_{r+1}, . . . , v_n in the nullspace of A to V,and orthonormal vectors u_{r+1}, . . . , u_m in the left-nullspace of A to M. We'll still get AV = U$\sum$and A = U$\sum V^T$.

This singular value decomposition has a particularly nice representation if we carry through the multiplication of the matrices:

A=$\bigcup \sum V^T = u_1 \sigma_1 v_1 + ... + u_r \sigma_r v_r^T$

$Eachofthese \backslash pieces" hasrank1.Ifweorderthesingularvalues$

$\sigma_1 \geq \sigma_2.....\sigma_r$

then the singular value decomposition gives A in r rank 1 pieces in order of importance.

2.1 SVD example

Let's find the singular value decompositions by taking a rank 2 unsymmetric matrix
A=$\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$

A is not symmetric, and there will be no orthogonal matrix Q that will make $Q^{-1}AQ$ diagonal. We need two different orthogonal matrices U and V .

We find these matrices with the singular value decomposition. So, we want to compute A^T A and its eigenvectors.

$$\mathrm{A}^T A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

andso,

$$\begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 9 = \lambda^2 - 10\lambda + 16 = (\lambda - 8)(\lambda - 2)$$

*So,*A^T A has eigenvalues 8 and 2. The corresponding eigenvectors will be

$$\mathrm{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \qquad v_2 = \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$Now, to find the vectors u1 and u2 we multiply v1 and v2 by A:$

$$Av_1 = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$$
$$Av_2 = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 2\sqrt{2} \end{pmatrix}$$

$So, the unit vectors u1 and u2 will be:$

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$The singular values will be 2\sqrt{2} = \sqrt{8} \;\; and \;\; \sqrt{2}$. This gives us the singular value decomposition:

$$\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{pmatrix} \left(\begin{smallmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{smallmatrix} \right)$$