Chapter 6

PCA (Principal Component Analysis) < LDA, SVD

House Price Prediction

Size	Locatn	YOC	OC Y/N	Builder	Color	Garden	Swimmi	Distance from Metro	Price

Problem of Plenty.
-> Representⁿ problem
-> Overfitting problem

So we need to Reduce Dimension.

* we need to identify Important component.

* PCA is dimension Reduction Technique
* Dimensionality Reduction ->
  * Reduces the dimension of feature space.

    Ex If there are 100 features/colⁿ in dataset and you want to get only 10 features then with dimensionality reduction technique we can achieve this.

  * It transforms dataset which is in n dimension space to n' dimension space where $n' < n$

Why Dimensionality Reduction ->

↳ Normally it is argued that many features gives more accurate result.

↳ However after some point the performance of model decreases (overfitting) with increase in no of features.

↳ This is known as "Curse of Dimensionality"

So Dimension Reduction is crucial.

* PCA enables us to identify the corelation and pattern in a dataset so that it can be transformed into row dataset of significantly lower dimension without loss of important information.

Steps in Performing PCA

step① → get the data

step② → Substract the Mean and Produce New Dataset (Row Data Adjust)

step3 → Calculate the Covariance Matrix.

step 4 → Calculate the Eigen Vectors and Eigen values of the Covariance matrix

step 5 → Choosing Components and forming a feature Vector.

Feature Vector = (eig1, eig2, eig3 ... eign)

step6 → Derive the new data set

Final DataSet = Row feature Vector * Row Data Adjust

* Row Feature Vector → is matrix with the Eigen Vectors in the column transposed so that Eigen Vectors are now in the rows with most significant Eigen Vectors on top.

* Final Datala Set is set with data items in col and dimensions along rows.

step 7 → Getting the old Data back.

step 7 → Getting the old Data back.

Final Data = Row Feature Vector × Row Data Adjust.

Row Data Adjust = Row Feature Vector^{-1} * Final Data.

= Row Feature Vector^{T} * Final Data

Now Original Data = Row Data Adjust + Original Mean

Perform PCA on following Dataset

x	2.5	0.5	2.2	1.9	3.1	2.3	2.0	1.0	1.5	1.1
y	2.4	0.7	2.9	2.2	3.0	2.7	1.6	1.1	1.6	0.9

step 1: given dataset

x	2.5	0.5	2.2	1.9	3.1	2.3	2.0	1.0	1.5	1.1
y	2.4	0.7	2.9	2.2	3.0	2.7	1.6	1.1	1.6	0.9

step 2: calculate mean

$\bar{x} = 1.81$

$\bar{y} = 1.91$

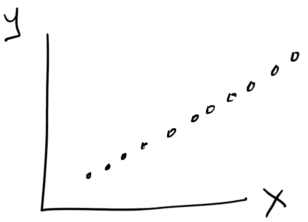
step 3 Adjusted Dataset (So the Dataset Mean is Zero).

$x = x - \bar{x}$	0.69	-1.31	0.39	0.09	1.29	0.49	0.19	-0.81	-0.31	-0.71
$y = y - \bar{y}$	0.49	-1.21	0.99	0.29	1.09	0.79	-0.31	-0.81	-0.31	-1.01

step 4: [– Note →.

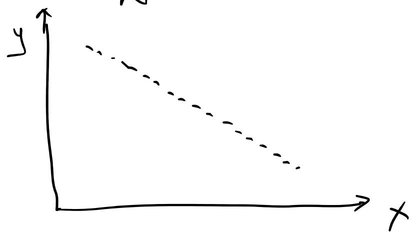
To find Covariance Matrix → It is measure of the extent to which the corresponding elements from two set of ordered data move in same direction.

$$cov(x, y) = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{N}$$



$x\uparrow$ $y\uparrow$

Large +ve covariance betn x & y.



$x\uparrow$ $y\downarrow$

Large -ve covariance

Zero covariance]

Covariance Matrix → $C = \begin{bmatrix} 0.61655555 & 0.6154444 \\ 0.6154444 & 0.71655556 \end{bmatrix}$

Step 5: Compute Eigen Values & Vector.

$$C - \lambda I = 0$$

$$\begin{array}{c|cc} & x & y \\ \hline x & 0.61655555-\lambda & \overset{cov(x,y)}{0.6154444} \\ y & \underset{cov(y,x)}{0.6154444} & 0.71655556-\lambda \end{array} = 0$$

$$\lambda^2 - 1.3332\lambda + 0.0630244 = 0$$

$$\boxed{\begin{aligned} \lambda_1 &= 0.0490834 \\ \lambda_2 &= 1.284627712 \end{aligned}}$$

* The largest Eigen Value gives first PCA

Eigen Vector → Let $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $v = \begin{vmatrix} v_1 \\ v_2 \end{vmatrix}$ be Eigen Vectors for λ_1

$$\begin{bmatrix} 0.56747215 & 0.61544444 \\ 0.61544444 & 0.66747216 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$0.56767215 u_1 + 0.6154444 u_2 = 0 \quad — (1)$$

$$0.61544444 u_1 + 0.66747216 u_2 = 0 \quad — (2)$$

Let $u_1 = 1$ in eq (1)

$u_2 = -0.92205266.$

$$\therefore \boxed{u = \begin{bmatrix} 1 \\ -0.922205266 \end{bmatrix}} \begin{array}{l} u_1 \\ u_2 \end{array}$$

for λ_2

$$\begin{bmatrix} -0.66747216 & 0.6154444 \\ 0.6154444 & -0.567472152 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$-0.66747216 v_1 + 0.6154444 v_2 = 0 \quad — (3)$$

$$0.6154444 v_1 + -0.567472152 v_2 = 0 \quad — (4)$$

Consider Eq 3 put $v_2 = 1$

$$v_1 = 0.922205266$$

$$\therefore \boxed{v = \begin{bmatrix} 0.922205266 \\ 1 \end{bmatrix}} \begin{array}{l} v_1 \\ v_2 \end{array}$$

(6) Normalize the Vector → to convert it into Vector of Unit length

How → divide it by length of Vector.

$$u = \frac{1}{\sqrt{u_1^2 + u_2^2}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{1.360213596} \begin{bmatrix} 1 \\ -0.922052616 \end{bmatrix}$$

$$\boxed{u = \begin{bmatrix} 0.7351786533 \\ -0.6778734 \end{bmatrix}}$$

$$v = \frac{1}{\sqrt{v_1^2 + v_2^2}} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{1}{1.360213596} \begin{bmatrix} 0.922052616 \\ 1 \end{bmatrix}$$

$$\boxed{v = \begin{bmatrix} 0.6778734 \\ 0.7357786533 \end{bmatrix}}$$

$$\text{Feature Vector} = \begin{array}{cc} u & v \\ \begin{bmatrix} 0.7357786535 & 0.6778734008 \\ -0.6778734008 & 0.735778635 \end{bmatrix} \end{array}$$

Final Data = Feature Vector^T × Data adjust

$$= \begin{array}{c} u \rightarrow \\ v \rightarrow \end{array} \begin{bmatrix} 0.7357786535 & -0.6678734008 \\ 0.6778734008 & 0.735778635 \end{bmatrix}_{2\times2} *$$

$$\begin{bmatrix} 0.69 & -1.31 & 0.39 & 0.09 & 1.29 & 0.49 & 0.19 & -0.81 & -0.31 & -0.71 \\ 0.49 & -1.21 & 0.99 & 0.29 & 1.09 & 0.79 & -0.31 & -0.81 & -0.31 & -0.71 \end{bmatrix}_{2\times10}$$

1D converted dataset

$$\begin{array}{c} u \rightarrow \\ v \rightarrow \end{array} \begin{bmatrix} 0.1751 & -0.142 & -0.384 & -0.130 & 0.209 & -0.175 & 0.349 & -0.046 & -0.017 & -1.00 \\ 0.827 & -1.777 & -0.992 & 0.274 & 1.875 & 0.912 & -0.099 & 0.046 & 0.017 & 1.00 \end{bmatrix}$$

1D converted data.

final data of (2.5, 2.4) 0.827 → will have contribution of both

Note Here only if λ_2 is used as it is greater & given first PCA

∴ Calculating feature Vector for λ_2 let it be v as calculated above

$$v = \begin{bmatrix} 0.92205266 \\ 1 \end{bmatrix} \qquad \text{feature Vector} = \begin{bmatrix} 0.92205266 \\ 1 \end{bmatrix}$$

only v

Final Dataset = Feature Vector^T × Adjusted Data.

$$= \underbrace{\begin{bmatrix} 0.92205266 & 1 \end{bmatrix}}_{1\times2} * \begin{bmatrix} \_\_\_\_\_\_\_\_ \end{bmatrix}_{2\times10}$$

$$= \begin{bmatrix} \_\_\_\_\_\_\_\_ \end{bmatrix}_{1\times10}$$

Single Row

Original Dataset $= 2 \times 10$

Final Dataset $= 1 \times 10$

Reduced Dimension from 2D to 1D.

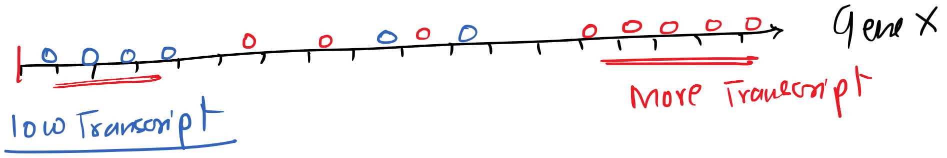
PCA → LDA (Linear Discriminant Analysis)
PCA → SVD (Singular Valued Decomposition)

LDA [Linear Discriminant Analysis]

* Suppose ⇒ We got a Cancer drug.
  - ↳ It works great for some people
  - ↳ But it make it worse for other people.

* How do we decide whom to give the drug?
* May be gene study of patients will be of some help.

① Consider only one gene ⇒ gene X. [1-D]



o ⇒ Drug that work
o ⇒ Drug does not work

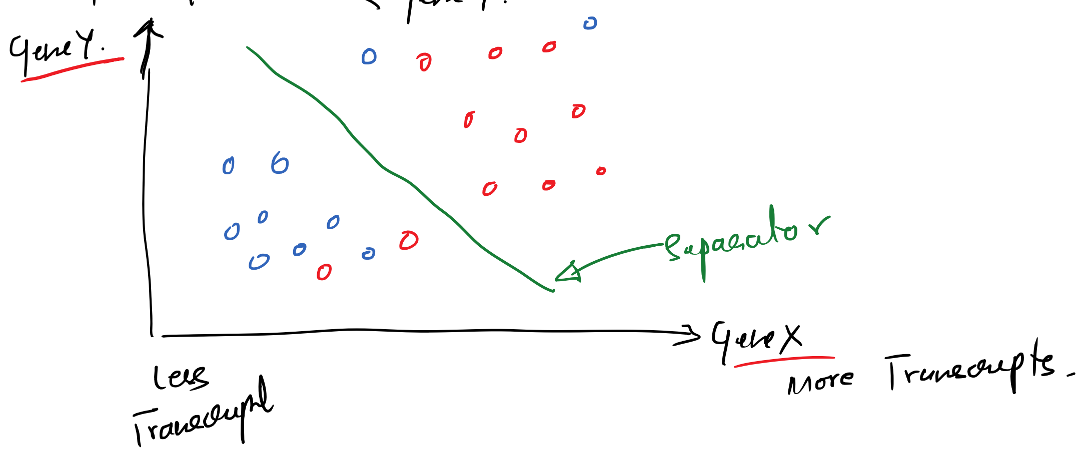
* From above we can see In cases with More Transcript drug is not working.
* Whereas In cases of low Transcript the drug tends to work
* How there are few exceptions (there are overlaps).
* Summary ⇒ gene X does OK job but it has few overlaps.

* Can we do Better?

let us Study 2 genes → Gene X, Gene Y.

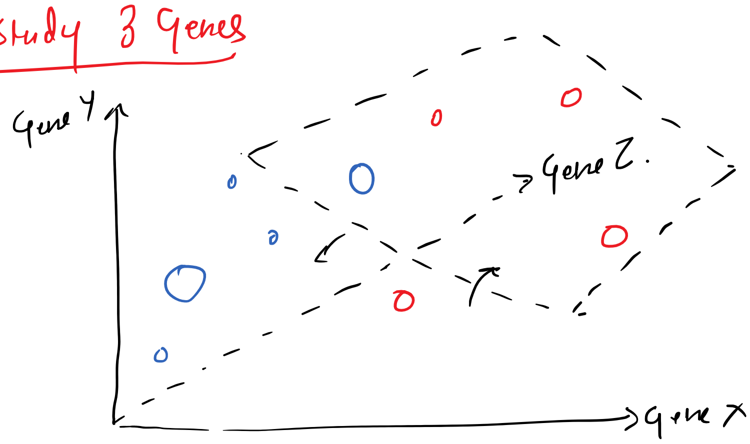
let us Study 2 Genes < Gene X / Gene Y.

2D.



It is better than studying only one gene but still there are overlapping.

let us study 3 Genes



In case of 3 Genes we will need to represent the information in 3-D.

-> Here the separator is a plane.

Suppose we want to study 4 gene $\xrightarrow{\text{Need}}$ 4-D Representation.
-> Can't Draw 4D graph.

* Here more the Dimension, more will be Complication in graph representation.

* PCA reduces dimensions by focusing on the genes with most Variation.

* Normally PCA is useful in plotting data with lot of dimension (or genes) onto simple XY plane.

* Here we are not interested much in identifying genes with most variation.
  Instead we are interested in Maximizing Separability betn the two groups so we can make best classification decision.

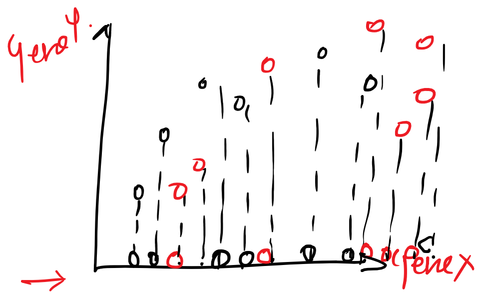
Note

PCA →	Help to identify feature with Max variation by reducing no of Component
LDA →	Helps to build better separator/ classifier by reducing Component

* Best possible Ways

(1) We have study of 2 Genes → Ignore gene Y.

So project all data point on X axis



→ This is not very good idea because it ignores the useful information that genes Y provides.

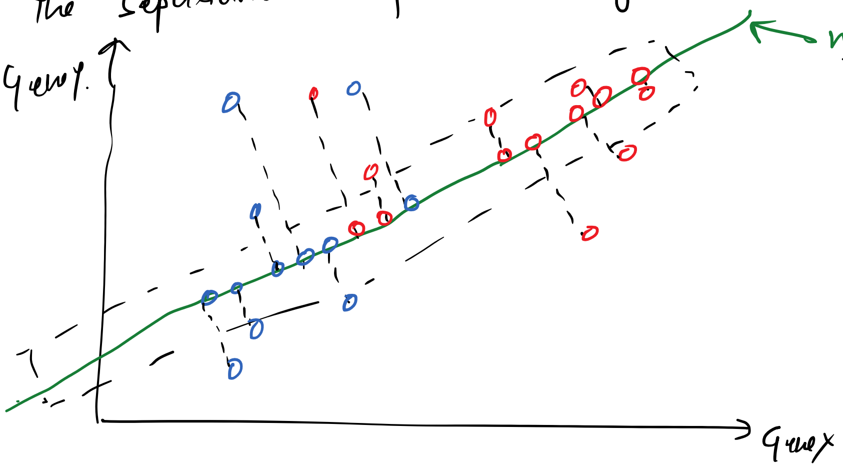
* LDA provides better way.

→ LDA reduces a 2-D graph to 1-D graph.

→ LDA uses both the genes to create a new axis and project the data on new axis in a way to maximize the separation of two categories.

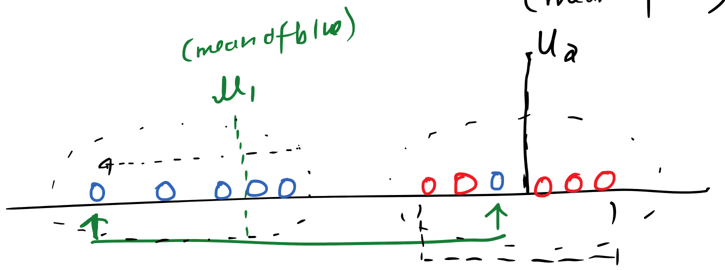
new axis (constructed using gene X info & ...)

the Separation of two categories.

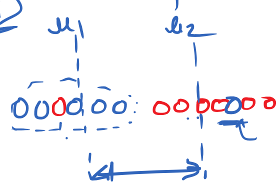


How LDA create a New Axis →

μ_1 and μ_2 are mean of respective class.



Case 1) means of two classes are near.



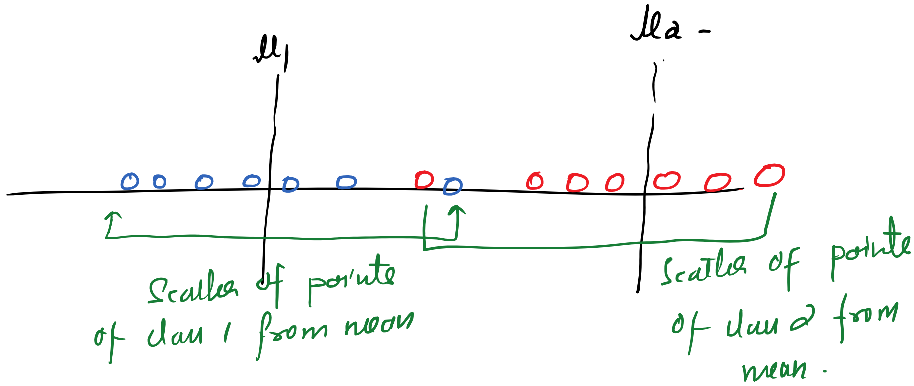
→ the points in each class are widely scattered

The new axis is created according to two criteria (considered simultaneously)

① Maximize the distance between mean of two classes.)

② Minimize the variation (LDA calls the scatter) and is represented by s^2) for each category.

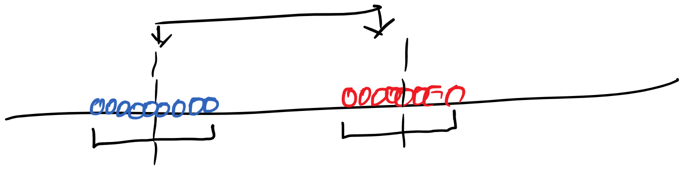
Case 2) Scatter of points within class is large



S_B → Between class Scatter.

case 3) Scatter of points within class is small.

$S_B \Rightarrow$ Between class Scatter.



S_w — Scatter of points is small within class.

within class scatter

Goal -> $S_B \Rightarrow$ Maximize
$S_w \Rightarrow$ Minimize.
} must be focused on both simultaneously.

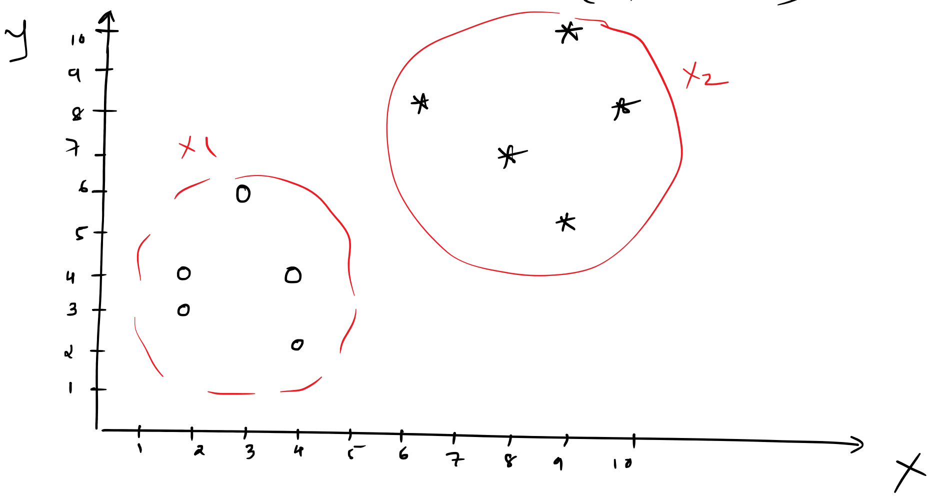
Mathematical Representation $\Rightarrow$ $\dfrac{(\mu_1-\mu_2)^2}{S_{w1}^2+S_{w2}^2}$ = $\dfrac{\text{Square of diff between two mean.}}{\text{Sum of scatter within each category.}}$

Ideally $\dfrac{(\mu_1-\mu_2)^2}{S_{w1}^2+S_{w2}^2}$ $\begin{matrix} \Rightarrow \text{Large} \\ \Rightarrow \text{Small.} \end{matrix}$

Consider ⇒ Compute the (Linear Discriminant projection) for the following two Dimensional Dataset.

$\rightarrow X_1 = (x_1, y_1) = \{(4,2), (2,4), (2,3), (3,6), (4,4)\} \Rightarrow o$

$\rightarrow X_2 = (x_2, y_2) = \{(9,10), (6,8), (9,5), (8,7), (10,8)\} \Rightarrow *$



Step 1) Find Class Means

$$\mu_1 = \frac{1}{N_1} \sum x \in X_1$$

$$= \frac{1}{5}\left[\begin{pmatrix}4\\2\end{pmatrix} + \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}2\\3\end{pmatrix} + \begin{pmatrix}3\\6\end{pmatrix} + \begin{pmatrix}4\\4\end{pmatrix}\right] = \frac{1}{5}\begin{pmatrix}15\\19\end{pmatrix}$$

$$= \begin{pmatrix}3\\3.8\end{pmatrix}$$

$$u_2 = \frac{1}{N_2} \sum x \in X_2$$

$$= \frac{1}{5}\left[\begin{pmatrix}9\\10\end{pmatrix} + \begin{pmatrix}6\\8\end{pmatrix} + \begin{pmatrix}9\\5\end{pmatrix} + \begin{pmatrix}8\\7\end{pmatrix} + \begin{pmatrix}10\\8\end{pmatrix}\right] = \frac{1}{5}\begin{pmatrix}42\\38\end{pmatrix}$$

$$= \begin{pmatrix}8.4\\7.6\end{pmatrix}$$

Step 2) Covariance Matrix of Both Class.

Covariance Matrix of Class 1. (X_1).

$$\boxed{S_1} = \frac{1}{N-1} \sum_{x \in X_1} (x - \mu_1)(x - \mu_1)^T$$

$$\boxed{S_1} = \frac{1}{N-1}\sum_{x \in X_1}(x-\mu_1)(x-\mu_1)^T$$

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

$$\left. + \left[\begin{pmatrix}3\\6\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^2 + \left[\begin{pmatrix}4\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^2\right]$$

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

Covariance matrix of Second class.

$$\boxed{S_2} = \frac{1}{N-1}\sum_{x \in X_2}(x-\mu_2)(x-\mu_2)^T.$$

$$= \frac{1}{4}\left[\left[\begin{pmatrix}9\\10\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}6\\8\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}9\\5\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2\right.$$

$$\left. + \left[\begin{pmatrix}8\\7\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}10\\8\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2\right]$$

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

Step 3 Within Class Scatter Matrix (Minimize)

$$S_W = S_1 + S_2$$

$$= \begin{bmatrix}1 & -0.25\\ -0.25 & 2.2\end{bmatrix} + \begin{bmatrix}2.3 & -0.05\\ -0.05 & 3.3\end{bmatrix}$$

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

Step 4 -> Between class Scatter Matrix (maximize).

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

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

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

$$S_B = \begin{bmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{bmatrix}$$

Step 5 The LDA projection is then obtained as solution of the generalized Eigen Value Problem

$$S_w^{-1} S_B w = \lambda w.$$

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

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

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

$$= \left| \begin{pmatrix} 9.2213 - \lambda & 6.489 \\ 4.2339 & 2.9794 - \lambda \end{pmatrix} \right| \quad \text{—— I}$$

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

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

$$\lambda \Rightarrow \boxed{\begin{matrix} \lambda_1 = 0 \\ \lambda_2 = 12.2007 \end{matrix}}$$ 2 Eigen Values.

Now By Substituting $\lambda = \lambda_1$ in I we get

Eigen Vector 1 = w_1 for Eigen = $\begin{pmatrix} -0.5755 \\ 0.8178 \end{pmatrix}$

Eigen vector 1 = $\underline{\underline{w_1}}$ = $\begin{pmatrix} \\ 0.8178 \end{pmatrix}$
for Eigen value λ_1

By Substituting $\lambda = \lambda_2$ in I we get

Eigen Vector 2 for Eigen value λ_2 = w_2 = $\begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}$

Step 6 Final Data = $y = w^T X$

w ← Projection Vector; X ← Input Data.

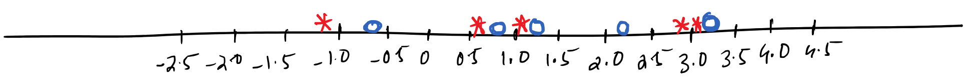
For Projection Vector w_1

Final Data of $X_1 = \underline{\underline{w_1^T}}\ \underline{\underline{X_1}} = [-0.5755 \quad 0.8178] \left[\begin{pmatrix} 4 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix} \right]$

$= [-0.668 \quad 2.12 \quad 1.3022 \quad 3.18 \quad 0.9688]$

Final Data of $X_2 = w_1^T X_2 = [-0.5755 \quad 0.8178] \left[\begin{pmatrix} 9 \\ 10 \end{pmatrix} \begin{pmatrix} 6 \\ 8 \end{pmatrix} \begin{pmatrix} 9 \\ 5 \end{pmatrix} \begin{pmatrix} 8 \\ 7 \end{pmatrix} \begin{pmatrix} 10 \\ 8 \end{pmatrix} \right]$

$= [2.9985 \quad 3.0894 \quad -1.0905 \quad 1.1206 \quad 0.7874]$



Here the projection vector w_1 constructed using Smaller Eigen value leads to bad Separability.
$(\underline{\underline{\lambda_1 = 0}})$

For Projection Vector $\underline{\underline{w_2}}$

final Data $X_1 = W_2^T X_1$

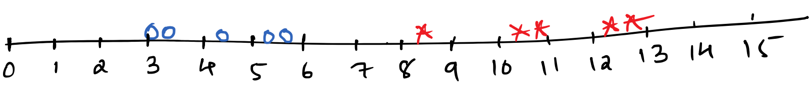
o $= \begin{bmatrix} 0.9088 & 0.4173 \end{bmatrix} \left[\begin{pmatrix} 4 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix} \right]$

$= \begin{bmatrix} 4.4698 & 3.4868 & 3.0695 & 5.2302 & 5.3044 \end{bmatrix}$.

Final Data $X_2 = W_2^T X_2$

$*$ $= \begin{bmatrix} 0.9088 & 0.4173 \end{bmatrix} \left[\begin{pmatrix} 9 \\ 10 \end{pmatrix} \begin{pmatrix} 6 \\ 8 \end{pmatrix} \begin{pmatrix} 9 \\ 5 \end{pmatrix} \begin{pmatrix} 8 \\ 7 \end{pmatrix} \begin{pmatrix} 10 \\ 8 \end{pmatrix} \right]$

$= \begin{bmatrix} 12.3522 & 8.7912 & 10.2657 & 10.1957 & 12.4264 \end{bmatrix}$.



Here the projection Vector (W_2) corresponding to larger Eigen value λ_2 leads to good separability.

This is LDA → when the 2D data points reduced to 1-D data points

→ and also using projection Vector were able to predict a good separator of classes.

SVD [Singular Valued Decomposition]

* We normally use 2D matrix to represent Data Values where column represents features and row represents samples data points

* Matrix computation with all the values in matrix sometime become redundant or computationally expensive.

* we need to represent matrix in a form such that the most important part of matrix which is needed for further computation could be Extracted easily.

* This can be done by SVD

SVD Theoram -> A rectangular matrix A_{mn} can be decomposed it into product of 3 matrices.

$$\boxed{A_{mn} = U_{m \times m} \, \Sigma_{m \times n} \, V^T_{n \times n}}$$

where $U_{m \times m}$ ⇒ Orthogonal Matrix.

$\Sigma_{m \times n}$ ⇒ Diagonal matrix

$V^T_{n \times n}$ ⇒ Transpose of orthogonal matrix $V_{n \times n}$

Columns of U are the orthonormal Eigen Vectors of AA^T

& column of U are the orthonormal Eigen Vectors of AA^T & Σ is diagonal matrix.

The elements of Σ are the square roots of Eigen values of

U & V in decreasing order.

Ex Find the SVD of $\underline{A} = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}_{2\times 3}$

Note we have to Decompose the matrix A into $\underline{U \Sigma V^T}$.

To find U
Σ
V^T.

Solution → To find U :- $AA^T = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$

for Eigen Values ⇒ $\begin{vmatrix} 11-\lambda & 1 \\ 1 & 11-\lambda \end{vmatrix} = 0$ —— (I)

$\Rightarrow \lambda^2 - 22\lambda + 120 = 0$

$\boxed{\lambda_1 = 10}$ $\boxed{\lambda_2 = 12}$

For Eigen Vector → $\begin{bmatrix} 11-\lambda & 1 \\ 1 & 11-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Case 1) $\boxed{\text{Substitute } \lambda = \lambda_1 = 10}$

$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\frac{x_1}{1} = -\frac{x_2}{1} = 1$

$\therefore x_1 = 1 \quad x_2 = -1$

Eigen Vector $\underline{X_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}}$

(cned) Substitute $\lambda = \lambda_2 = 12$

$$\begin{bmatrix} 11-\lambda & 1 \\ 1 & 11-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Eigen Vector $X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

* In U the Eigen Vector generated by larger Eigen value will be the first Column.

$$\therefore U = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Eigen vector of $\lambda = 12$

Eigen Vector of $\lambda = 12$.

Now we need to Normalize the matrix $\Rightarrow$ divide by length of respective Vector.

$$* \boxed{U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}$$

length of Vector $\lambda = 12$
$\sqrt{1^2+1^2}$
$= \sqrt{2}$

length of Vector $\lambda = 10$
$\sqrt{(1)^2+(-1)^2} = \frac{1}{\sqrt{2}}$

Step 2 To find V

$$A^T A = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}$$

Here for Eigen Values

$$\lambda^3 - 22\lambda^2 + 132\lambda = 0$$

$$\therefore \boxed{\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &= 10 \\ \lambda_3 &= 12 \end{aligned}}$$

We need to find Eigen Vectors for the 3 Eigen values.

Case 1) Eigen Vector for $\lambda_1 = 0$

$$\Rightarrow \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0.$$

By Cramers Rule

$$\frac{x_1}{\begin{vmatrix} 0 & 2 \\ 10 & 4 \end{vmatrix}} = \frac{x_2}{\begin{vmatrix} 10 & 2 \\ 0 & 4 \end{vmatrix}} = \frac{x_3}{\begin{vmatrix} 10 & 0 \\ 0 & 10 \end{vmatrix}}$$

$$= \frac{x_1}{-20} = -\frac{x_2}{40} = \frac{x_3}{100} = \frac{-1}{20}$$

$$x_1 = 1 \quad x_2 = 2 \quad x_3 = -5$$

$$\text{Eigen Vector } X_1 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$$

Case 2) Eigen Vector for $\lambda = 10$

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

By Cramers Rule

$$\frac{x_1}{\begin{vmatrix} 0 & 4 \\ 4 & -8 \end{vmatrix}} = -\frac{x_2}{\begin{vmatrix} 0 & 4 \\ 2 & -8 \end{vmatrix}} = \frac{x_3}{\begin{vmatrix} 0 & 0 \\ 2 & 4 \end{vmatrix}} \Rightarrow \frac{x_1}{-16} = \frac{-x_2}{-8} = \frac{x_3}{0} = \frac{-1}{8}$$

$\therefore x_1 = 2 \quad x_2 = -1 \quad x_3 = 0$

$$\text{Eigen Vector} = X_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

Case 3) Eigen Vector for $\lambda_3 = 12$

$$\begin{bmatrix} -2 & 0 & 2 \\ 0 & -2 & 4 \\ 2 & 4 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

By Cramers Rule

$$\frac{x_1}{\begin{vmatrix} 0 & 2 \\ -2 & 4 \end{vmatrix}} = \frac{-x_2}{\begin{vmatrix} -2 & 2 \\ 0 & 4 \end{vmatrix}} = \frac{x_3}{\begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix}} \Rightarrow \frac{x_1}{4} = \frac{-x_2}{-8} = \frac{x_3}{4} = \frac{1}{4}$$

$\therefore x_1 = 1 \quad x_2 = 2 \quad x_3 = 1$

$$\text{Eigen Vector } X_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\text{Now } V = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -1 & 2 \\ 1 & 0 & 5 \end{bmatrix}$$

Eigen Vector for $\lambda = 12$ — Eigen Vector for $\lambda = 10$ — Eigen Vector for $\lambda = 0$.

Now Normalizing the matrix $\rightarrow$ Divide by length of Vector.

$$V = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30} \\ 2/\sqrt{6} & -1/\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & -5/\sqrt{30} \end{bmatrix}$$

length of Vector $\sqrt{6}$ — length of Vector $\sqrt{6}$ — length of Vector $\sqrt{30}$

$\begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}$

$$V^T = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 2/\sqrt{5} & -1/\sqrt{5} & 0 \\ 1/\sqrt{30} & 2/\sqrt{30} & -5/\sqrt{30} \end{bmatrix}$$

Step 3 To find Σ (or D)

$$\Sigma = \begin{vmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \\ 0 & 0 & \sqrt{0} \end{vmatrix}_{3\times 3}$$

diag matrix $\Rightarrow$ The diag elements are square root of Eigen values in Decreasing order.

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

$$= \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 2/\sqrt{5} & -1/\sqrt{5} & 0 \\ 1/\sqrt{30} & 2/\sqrt{30} & -5/\sqrt{30} \end{bmatrix}}_{V^T}$$

$=$