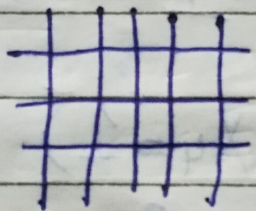
Q What is the idology behind the Principle Component Analysis [PCA]?

Ans [Principle Component Analysis] ⇒ Principle Component Analysis [PCA] Converts the Higher Dimensions data into Lower Dimensions data.

Benefits are :-

1. Faster Execution of Algorithms.
2. Visualization.



10D

Convert → 3D

So Principle Component Analysis basically converts the Higher Columns or Features Data into the Lower Columns or Features data. This helps in providing the best results...

Date..................

Dimensionality Reduction

- Feature Selection
- Feature Extraction
  - Principle Component Analysis [PCA]

Q How Feature Selection Works or How to select the Features [Columns] from the Complete data?

Ans **Feature Selection** ⇒ Feature Selection Approach of Dimensionality Reduction says that selects only those Features from the Complete data which have the higher impact with the Output...

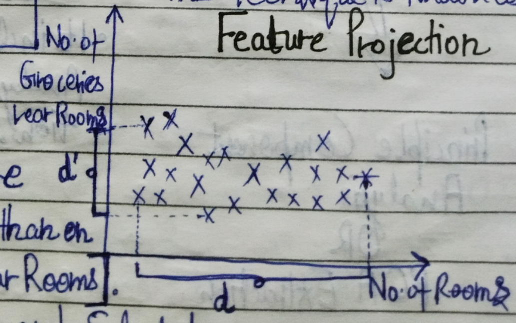
Q What is the Mathematical Trick to identify that which Features have the Higher Impact or are important?

sol: Let's say we have a Room_Purchase Data.

No. of Rooms	No. of Groceries near Rooms	Price [In Lakhs]
3	2	60
4	0	90
5	1	120
2	3	45
1	2	30

To identify which features [Columns] are best, the Mathematical trick is:- let's say I want to select One feature from "No. of Rooms" and "No. of Groceries near Rooms". So the way is -

This technique is Known as **Feature Projection**



The Variance of the portion covered by the datapoints is more in X-axis [No. of Rooms] rather than on the Y-axis [No. of Groceries near Rooms].

So X-axis feature is more effective and Selected....

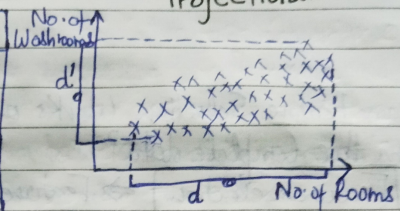
Date..................

Drawback of Feature Selection

Let's say We have changed our data.

No. of Rooms	No. of Washrooms	Price
2	1	30
3	2	50
4	3	90
5	4	120

Applying Feature Projection...



- In this Case, after Applying the Feature Projection, $d \simeq d'$. It means the Variance of No. of Rooms is somewhat equals to d'. In that Case, we cannot Mathematically also decide which feature is important.

- To solve this problem Feature Extraction came into picture...

Q How Feature Extraction is performed or What is the Mathematical [Geomatric Intitution] Intitution of Feature Extraction and how it solves the Problem of Feature Selection ??

Ans

Feature Extraction ⟹ • Feature Extraction is the technique of selecting the features only that are Extracted by the Feature Extraction.
- It includes the Principle Component Analysis which is used to do Extraction.

⇓

Principle Component Analysis
OR
PCA Extraction

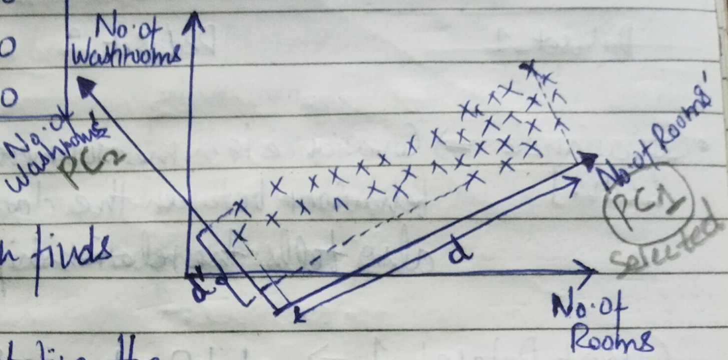
Principle Component Analysis

Date..................

⊙ Mathematical Intitution of PCA [Principle Component Analysis] using the same example of Feature Selection...

No.of Rooms	No.of Washrooms	Price
2	1	30
3	2	50
4	3	90
5	4	120

Applying Feature Projection...



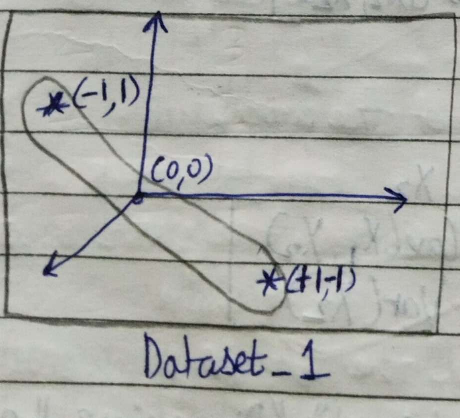
- In Principle Component Analysis the PCA moves the axis and then finds the next axis or new axis.
- In these new axis it got after rotating the original axis, these new axis are termed as Principle Components.

As, we can see that there are 2 PCs, PC_1 and PC_2 [new axis].

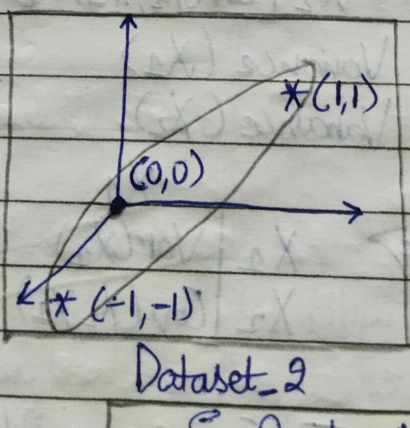
- Finding the Variance at new PC_1 and PC_2. the $d > d'$.
- It means PC_1 will be selected and all the data are manuplated as per the PC_1.

Q How this Principle Component Analysis [PCA] finds the PCs or Principle Components [Vectors]? What is the Mathematics behind it?

Ans ① Firstly the Covariance is better than the Variance because the problem with the Variance is that.



Dataset_1



Dataset_2

As, per the Variances:- the Variance of Dataset_1 is

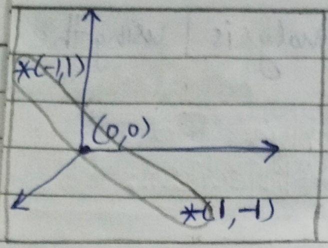
$$= \frac{1+0+-1}{3} \Rightarrow 0$$

The Variance of Dataset_2 is

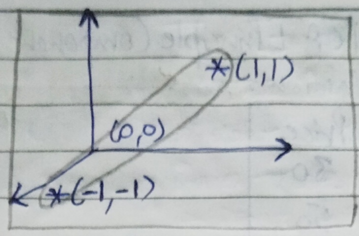
$$\Rightarrow \frac{1+0-1}{3} \Rightarrow 0$$

So, As per the Variance is consemed both Datasets are same. But it is not true at all...

Date.........



Dataset_1



Dataset_2

- This happened because the Variance is not able to identify the relationship between the axis.

to Solve this Covariance come into Picture...

- Covariance [Cov.] ⇒ Covariance is a measurement of Statistics which tells the behaviour between the data points available in the axis and also tells the relationship between the axis...

Cov. of Dataset_1 ⇒ $\frac{-1+0-1}{3} \Rightarrow \frac{-2}{3}$ [Shows the Negative Relationship between the axis...]

Cov. of Dataset_2 ⇒ $\frac{1+0+1}{3} \Rightarrow \frac{2}{3}$ [Shows the Positive Relationship within the axis..]

(2) Covariance Matrix ⇒ Let say We have the data as :-

Feature_1	Feature_2	Target_Column
—	—	—
—	—	—

Assuming that :-
Feature_1 = X_1
Feature_2 = X_2

$$\begin{array}{c|cc} & X_1 & X_2 \\ \hline X_1 & Cov(X_1,X_1) & Cov(X_1,X_2) \\ X_2 & Cov(X_2,X_1) & Cov(X_2,X_2) \end{array} \Longrightarrow \begin{array}{c|cc} & X_1 & X_2 \\ \hline X_1 & Cov(X_1,X_1) & Cov(X_1,X_2) \\ X_2 & Cov(X_2,X_1) & Cov(X_2,X_2) \end{array}$$

And $Cov(X_1,X_1) = Variance(X_1)$
$Cov(X_2,X_2) = Variance(X_2)$

Cov Matrix Will be ⟹

$$\begin{array}{c|cc} & X_1 & X_2 \\ \hline X_1 & Var(X_1) & Cov(X_1,X_2) \\ X_2 & Cov(X_2,X_1) & Var(X_2) \end{array}$$

Things to be Remember:- In Covariance Matrix of "n" Dimensions the Diagonal Values are Variance and non diagonal are Covariance Values...

Spiral

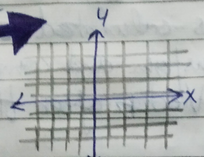
Date...............

Covariance Matrix ⇒ Covariance Matrix is essential because firstly tells the relationship positive or Negative between the axis and it also tells about the Variance [Spread] is there of the data on the axis...

③ Eigen Decomposition from the Covariance Matrix...

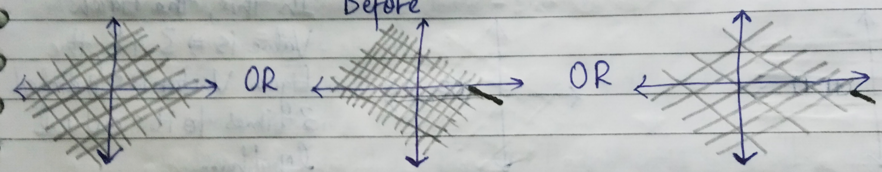
Matrix

$$\begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}_{2\times 2} \xrightarrow{\text{Transform}}$$



Before

When we apply the Matrix on to the Linear Algebra it transforms the grids in any of the possible ways...

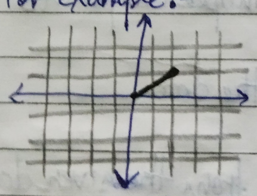


After [Rotate] After [Rotate and Sqeez) After (Rotate and Expand]

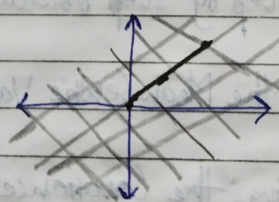
Eigen Vectors ⇒ Eigen Vectors are those Vectors of Linear Algebra on which after applying the Covariance Matrix the direction of that Vectors will remains the same but the magnitude will change.

"$\vec{V}$"

For example:-



Original



After Applied Matrix [Covariance Matrix]

$$A\vec{V} = \lambda \vec{V}$$

(It means applying Matrix on the Vector is like multiplying the Vector with a standard Value.)

As, we can see that the Vector in the original Algebra when represented in the Updated Linear Algebra, the Direction is excately same but the Magnitude is different!

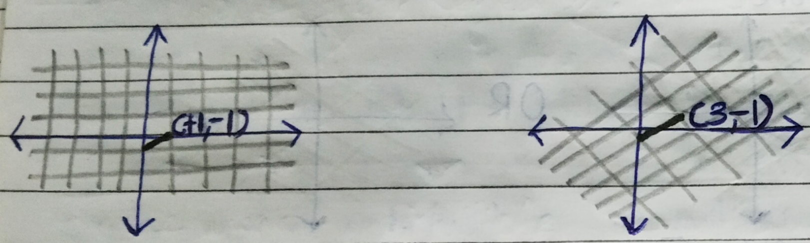
Spiral

Date.................

Thing to be Remember:- The Eigon Vectors are always same number in as the Number of features present in data. For example ⇒ For 2D Data there are 2 Eigon Vectors always there...

$\boxed{\text{Eigon Values}}$ ⇒ Eigon Values are the Distance [Magnitude] of the Eigon Vectors. It is calulated as the Distance of Vector before transform and compares with the Distance of the Vector after transform...

For example:-



In this, the Eigon Value is ⇒ 3 i.e; the Eigon Vector is stretched 3 times to its actual length...

Q Why we are Finding the Eigon Vectors in the Covariance?

Ans After applying the Covariance Matrix on Linear Algebra, then finding the Eigon Vectors is important because The Highest or Largest Eigon Vector will have the highest Variance on the data.

Q Please tell me the Step by Step Procedure of Principle Component Analysis?

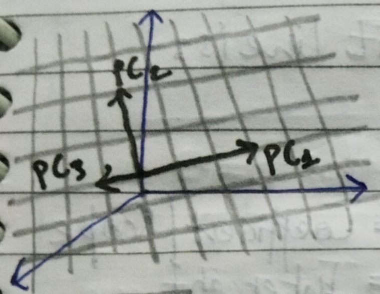
Ans $\boxed{\text{Step 1}}$ ⇒ Finding the Mean Centric Value of the data.

$\boxed{\text{Step 2}}$ ⇒ Calculating the Covariance Matrix from the given data.

$\boxed{\text{Step 3}}$ ⇒ Finding the Eigon Vectors and the Values from the Linear Algebra.

Date..........

Step 4 ⇒ Finding the Highest Eigen ~~Value~~ Vectors. For example ⇒ We have 3 Dimensional Data, it means 3 Eigen Vectors will be there which are classified as Higher Length is PC_1, then PC_2 and last will be PC_3.



- We remove the PC_3 and now we have the choice to convert the 3 Dimensional data into 2 Dimensional or 1 Dimensional data.

Step 5 ⇒ Let's say we want to Convert into 2 Dimensional Data, then:-
we selects the PC_1 and PC_2.

- Shape of Unit Vectors is (2,3) {2 Vectors and 3 Dimensions}.
- Size of Actual Data is ⇒ (400,3) {400 Rows and 3 Columns}.
- Transpose the shape of Unit Vectors. It will become ⇒ (3,2).
- Do Dot Product of Transpose Unit Vectors with the Shape of Actual data.

⇒ (400,3) . (3,2) ⟹ (400,2) {So the 3 dimensional data is Converted into 2 dimensional data.}