Principal Component Analysis (PCA)

Compiled by,

Dr. Shashank Shetty

Principle Component Analysis – Solved Example

- Given the data in Table, reduce the dimension from 2 to 1 using the Principal Component Analysis (PCA) algorithm.

Feature	Example 1	Example 2	Example 3	Example 4
X_1	4	8	13	7
X_2	11	4	5	14

Principle Component Analysis – Solved Example

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Step 1: Calculate Mean

$$\bar{X}_1 = \tfrac{1}{4}(4 + 8 + 13 + 7) = 8,$$

$$\bar{X}_2 = \tfrac{1}{4}(11 + 4 + 5 + 14) = 8.5.$$

Principle Component Analysis – Solved Example

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$\mathrm{Cov}(X_1, X_1) = \frac{1}{N-1} \sum_{k=1}^{N} (X_{1k} - \bar{X}_1)(X_{1k} - \bar{X}_1)$$

$$= \tfrac{1}{3}\left((4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2\right)$$

$$= 14$$

Principle Component Analysis – Solved Example

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\begin{aligned}
\mathrm{Cov}(X_1, X_2) &= \frac{1}{N-1} \sum_{k=1}^{N} (X_{1k} - \bar{X}_1)(X_{2k} - \bar{X}_2) \\
&= \tfrac{1}{3}((4-8)(11-8.5) + (8-8)(4-8.5) \\
&\quad + (13-8)(5-8.5) + (7-8)(14-8.5) \\
&= -11
\end{aligned}$$

Principle Component Analysis – Solved Example

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\begin{aligned} \mathrm{Cov}(X_2, X_1) &= \mathrm{Cov}(X_1, X_2) \\ &= -11 \end{aligned}$$

$$\begin{aligned} \mathrm{Cov}(X_2, X_2) &= \frac{1}{N-1}\sum_{k=1}^{N}(X_{2k} - \bar{X}_2)(X_{2k} - \bar{X}_2) \\ &= \tfrac{1}{3}\left((11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\right) \\ &= 23 \end{aligned}$$

Principle Component Analysis – Solved Example

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Step 2: Calculation of the covariance matrix.

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$= \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

Principle Component Analysis – Solved Example

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Step 3: Eigenvalues of the covariance matrix

The characteristic equation of the covariance matrix is,

$$0 = \det(S - \lambda I)$$

$$= \begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix}$$

$$= (14-\lambda)(23-\lambda) - (-11)\times(-11)$$

$$= \lambda^2 - 37\lambda + 201$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Principle Component Analysis – Solved Example

Step 3: Eigenvalues of the covariance matrix

The characteristic equation of the covariance matrix is,

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$0 = \det(S - \lambda I)$$

$$= \begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix}$$

$$= (14-\lambda)(23-\lambda) - (-11)\times(-11)$$

$$= \lambda^2 - 37\lambda + 201$$

$$\lambda = \frac{1}{2}(37 \pm \sqrt{565})$$

$$= 30.3849,\ 6.6151$$

$$= \lambda_1,\ \lambda_2 \quad \text{(say)}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Principle Component Analysis – Solved Example

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Step 4: Computation of the eigenvectors

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = (S - \lambda\ I)\ U$$

$$= \begin{bmatrix} 14 - \lambda & -11 \\ -11 & 23 - \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$= \begin{bmatrix} (14 - \lambda\ )u_1 - 11u_2 \\ -11u_1 + (23 - \lambda\ )u_2 \end{bmatrix}$$

$$(14 - \lambda)\, u_1 = 11\, u_2$$

$$(14 - \lambda\ )u_1 - 11u_2 = 0$$

$$-11u_1 + (23 - \lambda\ )u_2 = 0$$

$$\frac{u_1}{11} = \frac{u_2}{14 - \lambda} = t$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Step 4: Computation of the eigenvectors

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\frac{u_1}{11} = \frac{u_2}{14 - \lambda} = t$$

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda \end{bmatrix}.$$

$$u_1 = 11t, \quad u_2 = (14 - \lambda\ )t$$

Principle Component Analysis – Solved Example

Step 4: Computation of the eigenvectors

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}.$$

- To find a unit eigenvector, we compute the length of U_1 which is given by,

$$\|U_1\| = \sqrt{11^2 + (14 - \lambda_1)^2}$$

$$= \sqrt{11^2 + (14 - 30.3849)^2}$$

$$= 19.7348$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Principle Component Analysis – Solved Example

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Step 4: Computation of the eigenvectors

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

- To find a unit eigenvector, we compute the length of U_1 which is given by,

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\|U_1\| = \sqrt{11^2 + (14 - \lambda_1)^2}$$

$$= \sqrt{11^2 + (14 - 30.3849)^2}$$

$$= 19.7348$$

$$e_1 = \begin{bmatrix} 11/\|U_1\| \\ (14 - \lambda_1)/\|U_1\| \end{bmatrix} = \begin{bmatrix} 11/19.7348 \\ (14 - 30.3849)/19.7348 \end{bmatrix} = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

e_1

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Step 5: Computation of first principal components

$$e_1^T \begin{bmatrix} X_{1k} - \bar{X}_1 \\ X_{2k} - \bar{X}_2 \end{bmatrix}$$

$$e_1^T \begin{bmatrix} X_{1k} - \bar{X}_1 \\ X_{2k} - \bar{X}_2 \end{bmatrix} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} X_{11} - \bar{X}_1 \\ X_{21} - \bar{X}_2 \end{bmatrix}$$

$$= 0.5574(X_{11} - \bar{X}_1) - 0.8303(X_{21} - \bar{X}_2)$$

$$= 0.5574(4 - 8) - 0.8303(11 - 8,5)$$

$$= -4.30535$$

Principle Component Analysis – Solve

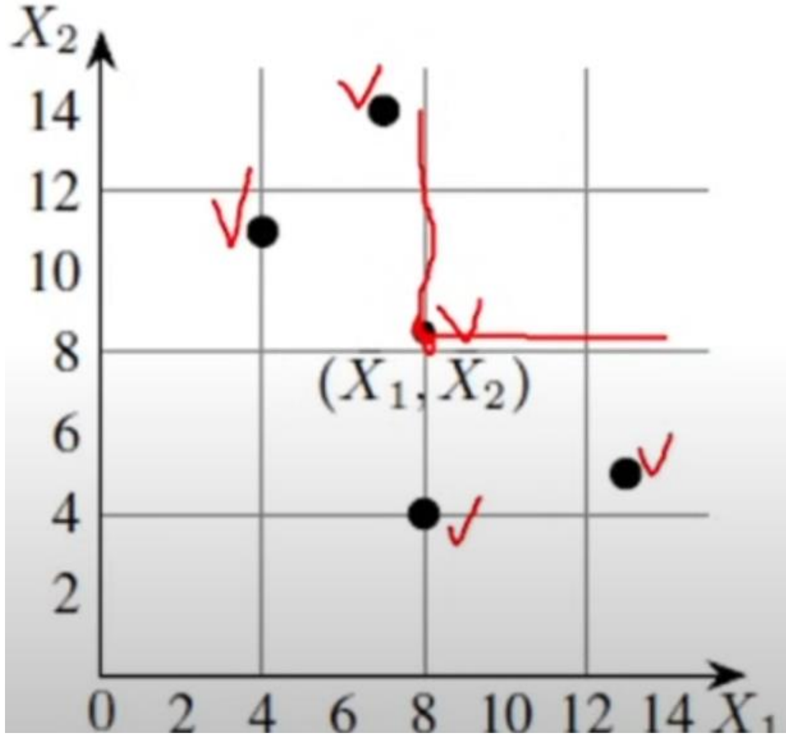
Step 5: Computation of first principal components

F
X_1
X_2

Feature	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14
First Principle Components	-4.3052	3.7361	5.6928	-5.1238

Step 6: Geometrical meaning of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14



$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

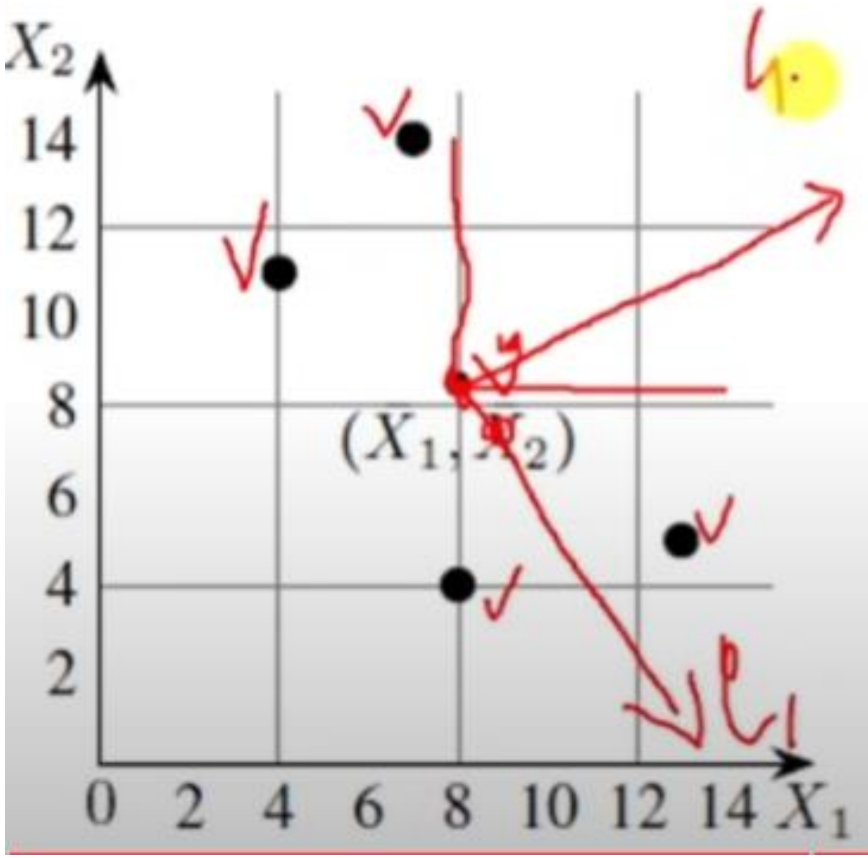
$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$ $S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Step 6: Geometrical meaning of first principal components

F	Ex :
X_1	4
X_2	11



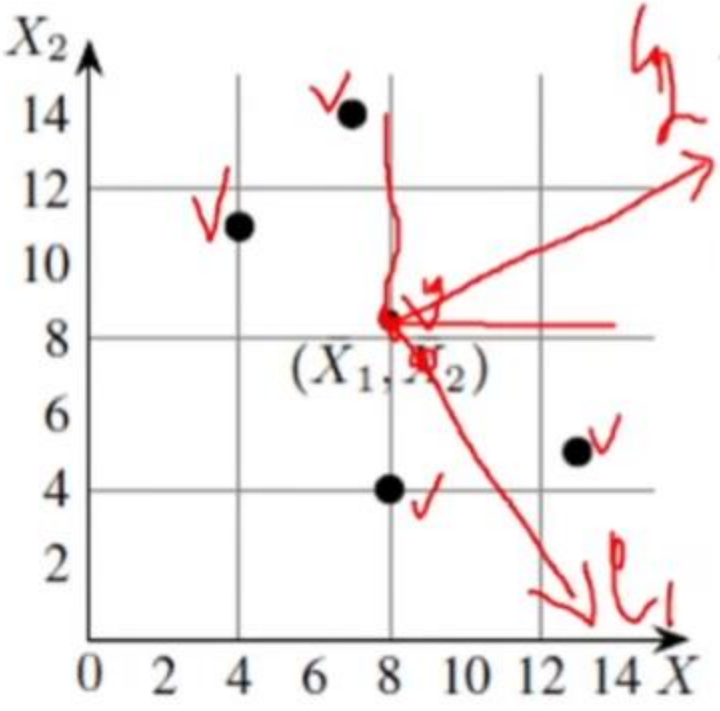
$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

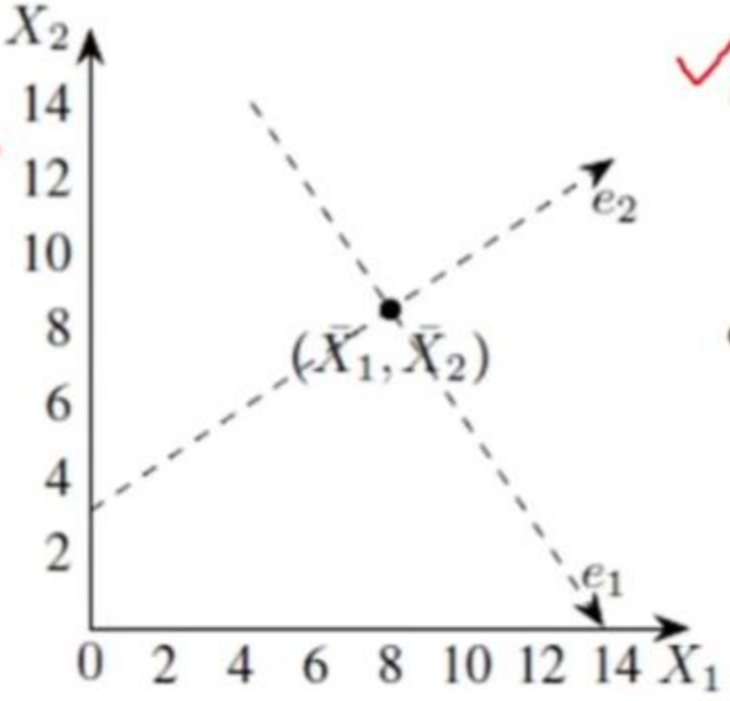
$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

Principle Component Analysis – Solved Example

Step 6: Geometrical meaning of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14





$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

$$\overline{X_1} = 8$$

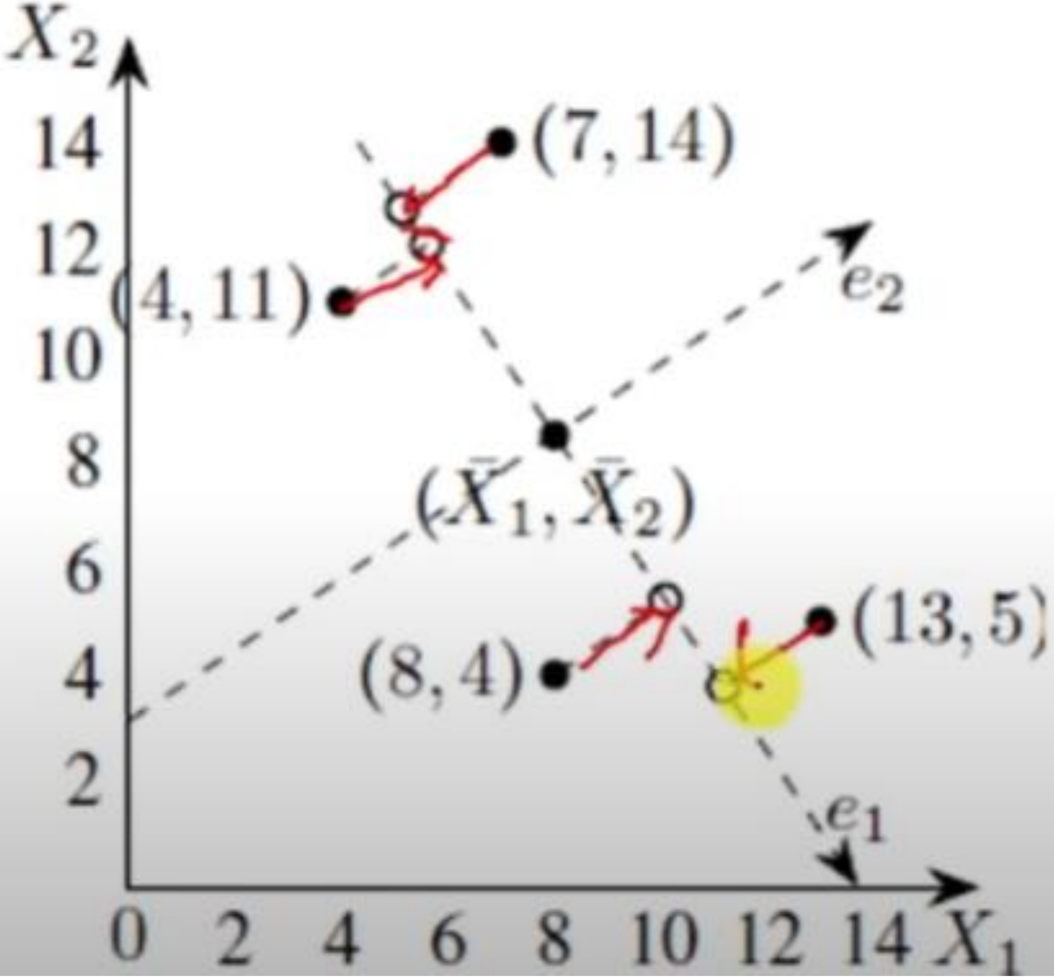
$$\overline{X_2} = 8.5$$

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix} \quad S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Step 6: Geometrical meaning of first principal components



PCA

- Principal component analysis,or PCA, is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set.
- Reducing the number of variables of a data set naturally comes at the expense of accuracy, but the trick in dimensionality reduction is to trade a little accuracy for simplicity. Because smaller data sets are easier to explore and visualize, and thus make analyzing data points much easier and faster for machine learning algorithms without extraneous variables to process.

- So, to sum up, the idea of PCA is simple: **reduce the number of variables of a data set, while preserving as much information as possible.**

What Are Principal Components?

- Principal components are new variables that are constructed as linear combinations or mixtures of the initial variables.
- These combinations are done in such a way that the new variables (i.e., principal components) are uncorrelated and most of the information within the initial variables is squeezed or compressed into the first components.
- So, the idea is 10-dimensional data gives you 10 principal components, but PCA tries to put maximum possible information in the first component, then maximum remaining information in the second and so on, until having something like shown in the scree plot below.

PCA

- What Is Principal Component Analysis?

- Principal component analysis, or PCA, is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set.

HOW DO YOU DO A PRINCIPAL COMPONENT ANALYSIS?

1. Standardize the range of continuous initial variables
2. Compute the covariance matrix to identify correlations
3. Compute the eigenvectors and eigenvalues of the covariance matrix to identify the principal components
4. Create a feature vector to decide which principal components to keep
5. Recast the data along the principal components axes

Working of PCA:

- PCA works on a process called **Eigenvalue Decomposition** of a covariance matrix of a data set. The steps are as follows:
- First, calculate the covariance matrix of a data set.
- Then, calculate the eigenvectors of the covariance matrix.
- The eigenvector having the highest eigenvalue represents the direction in which there is the highest variance. So this will help in identifying the first principal component.
- The eigenvector having the next highest eigenvalue represents the direction in which data has the highest remaining variance and also orthogonal to the first direction. So, this helps in identifying the second principal component.
- Like this, identify the top ‘k’ eigenvectors having top ‘k’ eigenvalues to get the ‘k’ principal components.

Agglomerative Hierarchical Clustering Solved Example

- Consider the following set of 6 one dimensional data points:
- 18, 22, 25, 42, 27, 43
- Apply the agglomerative hierarchical clustering algorithm to build the hierarchical clustering **dendogram**.
- Merge the clusters using **Min distance** and update the proximity matrix accordingly.
- Clearly show the **proximity matrix** corresponding to each iteration of the algorithm.

Agglomerative Hierarchical Clustering Solved Example

- Step – 1

	18	22	25	27	42	43
18	0	4	7	9	24	25
22	4	0	3	5	20	21
25	7	3	0	2	17	18
27	9	5	2	0	15	16
42	24	20	17	15	0	1
43	25	21	18	16	1	0

- Step – 1

	18	22	25	27	42	43
18	0	4	7	9	24	25
22	4	0	3	5	20	21
25	7	3	0	2	17	18
27	9	5	2	0	15	16
42	24	20	17	15	0	1
43	25	21	18	16	1	0

(42, 43)

Agglomerative Hierarchical Clustering Solved Example

- Step – 2

	18	22	25	27	42, 43
18	0	4	7	9	24
22	4	0	3	5	20
25	7	3	0	2	17
27	9	5	2	0	15
42, 43	24	20	17	15	0

Agglomerative Hierarchical Clustering Solved Example

- Step – 2

	18	22	25	27	42, 43
18	0	4	7	9	24
22	4	0	3	5	20
25	7	3	0	2	17
27	9	5	2	0	15
42, 43	24	20	17	15	0

(42, 43), (25, 27)

Agglomerative Hierarchical Clustering Solved Example

- Step – 3

	18	22	25, 27	42, 43
18	0	4	7	24
22	4	0	3	20
25, 27	7	3	0	15
42, 43	24	20	15	0

(42, 43), ((25, 27), 22)

Agglomerative Hierarchical Clustering Solved Example

- Step – 4

	18	22, 25, 27	42, 43
18	0	4	24
22, 25, 27	4	0	15
42, 43	24	15	0

Agglomerative Hierarchical Clustering Solved Example

- Step – 5

	18, 22, 25, 27	42, 43
18, 22, 25, 27	0	15
42, 43	15	0

Agglomerative Hierarchical Clustering Solved Example

- Step – 6

	18, 22, 25, 27, 42, 43
18, 22, 25, 27, 42, 43	0

Agglomerative Hierarchical Clustering Solved Example

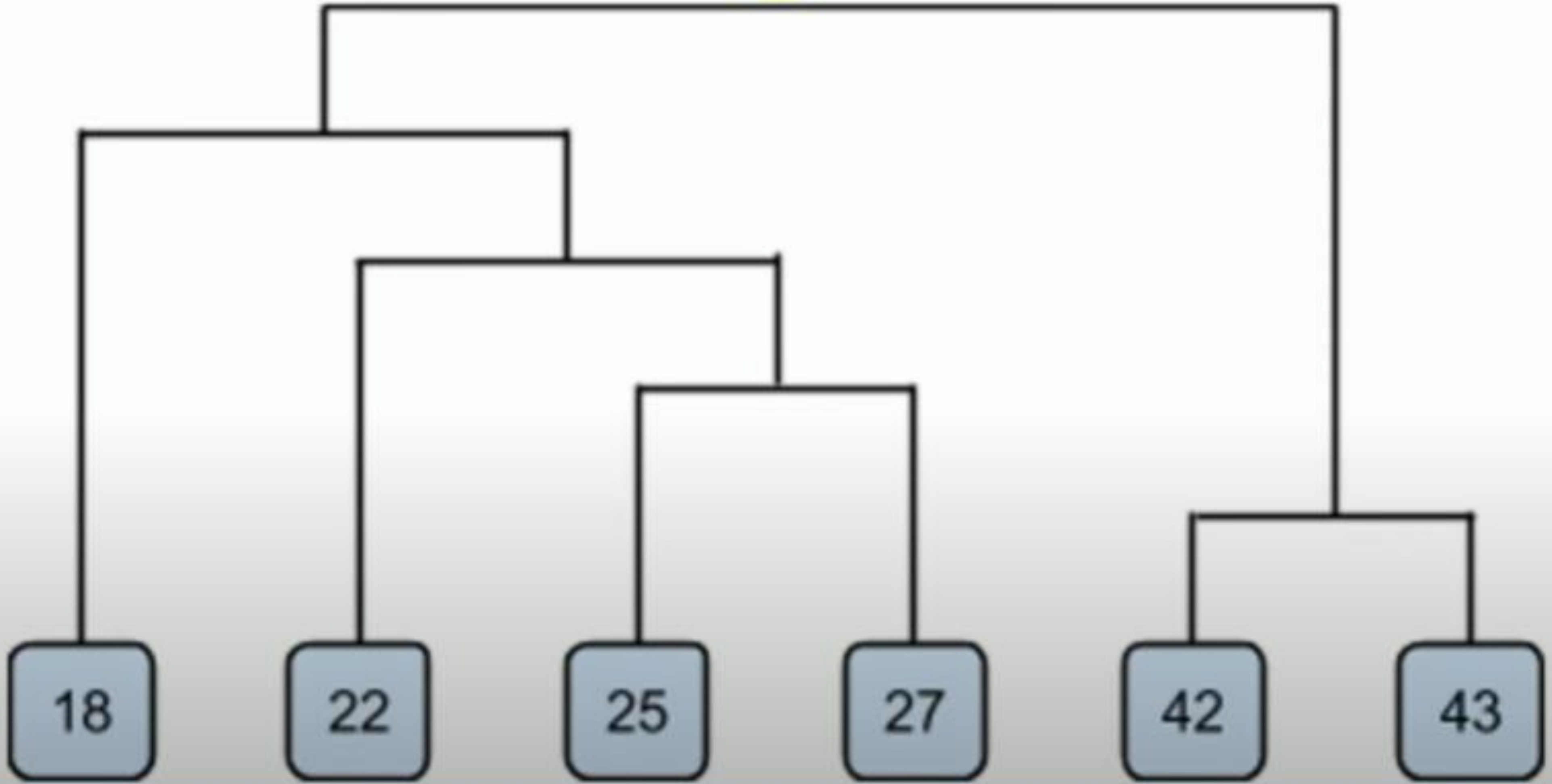
- Step – 5

	18, 22, 25, 27	42, 43
18, 22, 25, 27	0	15
42, 43	15	0

((42, 43), (((25, 27), 22), 18))

Agglomerative Hierarchical Clustering Solved Exar

- **Dendrogram** ((42, 43), (((25, 27), 22), 18



Clusters using a Single Link Technique

Sample No.	X	Y
P1	0.40	0.53
P2	0.22	0.38
P3	0.35	0.32
P4	0.26	0.19
P5	0.08	0.41
P6	0.45	0.30

Agglomerative Methods Solved Example 1

Clusters using a Single Link Technique Example - 1

Problem Definition:

For the given dataset find the clusters using a single link technique. Use Euclidean distance and draw the Dendrogram.

Clusters using a Single Link Technique Example - 1

Step 1: Compute the distance matrix

- So we have to find the Euclidean distance between each and every points.
- Let $A(x_1, y_1)$ and $B(x_2, y_2)$ are two points.
- Then Euclidean distance between

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Sample No.	X	Y
P1	0.40	0.53
P2	0.22	0.38
P3	0.35	0.32
P4	0.26	0.19
P5	0.08	0.41
P6	0.45	0.30

$$
\begin{pmatrix}
 & P1 & P2 & P3 & P4 & P5 & P6 \\
P1 & 0 & & & & & \\
P2 & 0.23 & 0 & & & & \\
P3 & 0.22 & 0.14 & 0 & & & \\
P4 & 0.37 & 0.19 & 0.13 & 0 & & \\
P5 & 0.34 & 0.14 & 0.28 & 0.23 & 0 & \\
P6 & 0.24 & 0.24 & 0.10 & 0.22 & 0.39 & 0
\end{pmatrix}
$$

Clusters using a Single Link Technique Example - 1

Step 2: Merging the two closest members.

- Here the **minimum value is 0.10** and hence we combine P3 and P6 (as 0.10 came in the P6 row and P3 column).
- Now, form clusters of elements corresponding to the minimum value and update the distance matrix.

Clusters using a Single Link Technique Example - 1

Sample No.	X	Y
P1	0.40	0.53
P2	0.22	0.38
P3	0.35	0.32
P4	0.26	0.19
P5	0.08	0.41
P6	0.45	0.30

$$\begin{pmatrix} & P1 & P2 & P3 & P4 & P5 & P6 \\ P1 & 0 & & & & & \\ P2 & 0.23 & 0 & & & & \\ P3 & 0.22 & 0.14 & 0 & & & \\ P4 & 0.37 & 0.19 & 0.13 & 0 & & \\ P5 & 0.34 & 0.14 & 0.28 & 0.23 & 0 & \\ P6 & 0.24 & 0.24 & 0.10 & 0.22 & 0.39 & 0 \end{pmatrix}$$

Clusters using a Single Link Technique Example - 1

Now we will update the Distance Matrix:

$$
\begin{array}{c|cccccc}
 & P1 & P2 & P3 & P4 & P5 & P6 \\
\hline
P1 & 0 & & & & & \\
P2 & 0.23 & 0 & & & & \\
P3 & 0.22 & 0.14 & 0 & & & \\
P4 & 0.37 & 0.19 & 0.13 & 0 & & \\
P5 & 0.34 & 0.14 & 0.28 & 0.23 & 0 & \\
P6 & 0.24 & 0.24 & 0.10 & 0.22 & 0.39 & 0
\end{array}
\qquad
\begin{array}{c|ccccc}
 & P1 & P2 & P3, P6 & P4 & P5 \\
\hline
P1 & 0 & & & & \\
P2 & 0.23 & 0 & & & \\
P3, P6 & 0.22 & 0.14 & 0 & & \\
P4 & 0.37 & 0.19 & 0.13 & 0 & \\
P5 & 0.34 & 0.14 & 0.28 & 0.23 & 0
\end{array}
$$

(P3, P6)

Now we will update the Distance Matrix:

$$\begin{pmatrix} & P1 & P2 & P3,P6 & P4 & P5 \\ P1 & 0 & & & & \\ P2 & 0.23 & 0 & & & \\ P3,P6 & 0.22 & 0.14 & 0 & & \\ P4 & 0.37 & 0.19 & 0.13 & 0 & \\ P5 & 0.34 & 0.14 & 0.28 & 0.23 & 0 \end{pmatrix} \quad \begin{pmatrix} & P1 & P2 & P3,P6,P4 & P5 \\ P1 & 0 & & & \\ P2 & 0.23 & 0 & & \\ P3,P6,P4 & 0.22 & 0.14 & 0 & \\ P5 & 0.34 & 0.14 & 0.28 & 0 \end{pmatrix}$$

{(P3, P6), P4}

Now we will update the Distance Matrix:

$$\begin{array}{c|cccc} & P1 & P2 & P3, P6, P4 & P5 \\ \hline P1 & 0 & & & \\ P2 & 0.23 & 0 & & \\ P3, P6, P4 & 0.22 & 0.14 & 0 & \\ P5 & 0.34 & 0.14 & 0.28 & 0 \end{array} \qquad \begin{array}{c|ccc} & P1 & P2, P5 & P3, P6, P4 \\ \hline P1 & 0 & & \\ P2, P5 & 0.23 & 0 & \\ P3, P6, P4 & 0.22 & 0.14 & 0 \end{array}$$

{(P3, P6), P4} and (P2, P5)

Now we will update the Distance Matrix:

$$\begin{array}{cccc} & P1 & P2,P5 & P3,P6,P4 \\ P1 & 0 & & \\ P2,P5 & 0.23 & 0 & \\ P3,P6,P4 & 0.22 & 0.14 & 0 \end{array}$$

$$\begin{array}{ccc} & P1 & P2,P5,P3,P6,P4 \\ P1 & 0 & \\ P2,P5,P3,P6,P4 & 0.22 & 0 \end{array}$$

[{(P3, P6), P4}, (P2, P5)]

Now we will update the Distance Matrix:

$$\begin{array}{cccc} & P1 & P2, P5 & P3, P6, P4 \\ P1 & 0 & & \\ P2, P5 & 0.23 & 0 & \\ P3, P6, P4 & 0.22 & 0.14 & 0 \end{array}$$

$$\begin{array}{ccc} & P1 & P2, P5, P3, P6, P4 \\ P1 & 0 & \\ P2, P5, P3, P6, P4 & 0.22 & 0 \end{array}$$

[{(P3, P6), P4}, (P2, P5)]

[{(P3, P6), P4}, (P2, P5)] , P1

So now we have reached to the solution, the dendrogram for those question will be as follows:

[{(P3, P6), P4}, (P2, P5)] , P1

Dendogram of the cluster formed

Agglomerative clustering

- Agglomerative clustering is a hierarchical clustering technique where clusters are built incrementally by merging smaller clusters into larger ones. It starts with each data point as a single cluster and iteratively merges the closest pairs of clusters until only one cluster remains, forming a hierarchical tree-like structure known as a dendrogram.
- The process of agglomerative clustering typically involves the following steps:
- **Initialization**: Assign each data point to its own cluster.
- **Calculate distances**: Compute the distance or dissimilarity between each pair of clusters. This can be done using various distance metrics, such as Euclidean distance or Manhattan distance, depending on the nature of the data.
- **Merge clusters**: Merge the two closest clusters into a single cluster. The definition of "closest" can vary depending on the linkage criterion being used (e.g., complete linkage, single linkage, average linkage).
- **Update distance matrix**: Recalculate the distances between the new cluster and all other clusters.
- **Repeat**: Repeat steps 2-4 until only a single cluster remains, or until a stopping criterion is met (e.g., a predetermined number of clusters, a threshold distance).