

Q. Given the following data, use PCA to reduce the dimension⁽¹⁾ from 2 to 1.

Feature	Example 1	Example 2	Example 3	Example 4
X	4	8	13	7
Y	11	4	5	14

Ans:- No. of features, $n = 2$
No. of samples, $N = 4$

Step 2: Computation of mean of variable

$$\bar{X} = \frac{4 + 8 + 13 + 7}{4} = 8$$

$$\bar{Y} = \frac{11 + 4 + 5 + 14}{4} = 8.5$$

Step 3: Computation of Covariance matrix of ordered pairs $(X, X), (X, Y), (Y, X), (Y, Y)$

$$\text{Cov}(X, Y) = \frac{1}{N-1} \sum_{k=1}^N (x_k - \bar{X})(y_k - \bar{Y})$$

$$\begin{aligned} \text{Cov}(X, X) &= \frac{1}{4-1} \left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2 \right] \\ &= 14 \end{aligned}$$

$$\begin{aligned} \text{Cov}(X, Y) &= \frac{1}{4-1} \left[(4-8)(11-8.5) + (8-8)(4-8.5) \right. \\ &\quad \left. + (13-8)(5-8.5) + (7-8)(14-8.5) \right] \\ &= -11 \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y, Y) &= \frac{1}{4-1} \left[(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 \right. \\ &\quad \left. + (14-8.5)^2 \right] \\ &= 23 \end{aligned}$$

Covariance matrix

$$C = \begin{bmatrix} \text{Cov}(X, X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Cov}(Y, Y) \end{bmatrix}$$

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

Step ④: Eigen value, Eigen vector, Normalized Eigen vector

(i) Eigen Value

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

$$\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = 0$$

$$\det \left(\begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix} \right) = 0$$

$$(14-\lambda)(23-\lambda) - (-11 \times -11) = 0$$

$$\lambda^2 - 37\lambda + 201 = 0 \Rightarrow \frac{1}{2a} \sqrt{b^2 - 4ac}$$

$$\lambda = 30.3849, 6.6151$$

$$\lambda_1 \neq \lambda_2$$

$$\lambda_1 = 30.3849 \Rightarrow \text{First Principal Component}$$

$$\lambda_2 = 6.6151$$

(ii) Eigen vector of λ_1

$$(C - \lambda_1 I) u_1 = 0$$

$$\begin{bmatrix} 14 - \lambda_1 & -11 \\ -11 & 23 - \lambda_1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} (14 - \lambda_1) u_1 - 11 u_2 \\ -11 u_1 + (23 - \lambda_1) u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(14 - \lambda_1) u_1 - 11 u_2 = 0$$

$$-11 u_1 + (23 - \lambda_1) u_2 = 0$$

$$\frac{u_1}{11} = \frac{u_2}{14 - \lambda_1} = t$$

when $t = 1$

$$u_1 = 11$$

$$u_2 = 14 - \lambda_1$$

$$\text{Eigen vector } u_1 \text{ of } \lambda_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ 14 - 30.3849 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -16.3849 \end{bmatrix}$$

(iii) Normalized Eigen vector

$$e_1 = \frac{1}{\sqrt{11^2 + (-16.3849)^2}} \begin{bmatrix} 11 \\ -16.3849 \end{bmatrix}$$

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

λ_2

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

389.465

Step ⑤ : Derive new dataset

	EX1	EX2	EX3	EX4
First Principal Component PC1	p_{11}	p_{12}	p_{13}	p_{14}

$$p_{11} = e_1^T \begin{bmatrix} 4-8 \\ 11-8.5 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix}$$

$$= -4.3052$$

$$p_{12} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} 8-8 \\ 4-8.5 \end{bmatrix}$$

$$= 3.7361$$

$$p_{13} = 5.6928$$

$$p_{14} = -5.1238$$

	EX1	EX2	EX3	EX4
PC1	-4.3052	3.7361	5.6928	-5.1238