

2018ICSE0621

Sai Ram.K

Part - C.

7 - CSE-10.

Q.1] a) Principle Component Analysis or PCA is the algorithm that minimizes the large data set with use of eigen values.

- It is also a method to reduce dimensions (or dimensionality reduction) to decrease the complexity of ~~and~~ analysis.

b)

X	3	9	12	8
Y	7	5	4	10

Step 1: No. of features, $n = 2$
No. of samples, $N = 4$

Step 2: Computing mean
 $\bar{x} = \frac{3+9+12+8}{4} = 8$; $\bar{y} = \frac{7+5+4+10}{4} = 6.5$

Step 3: Compute Covariance Matrix
Ordered Pairs $\Rightarrow (x, x), (x, y), (y, x), (y, y)$

$$\begin{aligned} \text{(i)} \quad \text{Cov}(x, x) &= \frac{1}{N-1} \sum_{k=1}^N (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \\ &= \frac{1}{4-1} [(3-8)^2 + (9-8)^2 + (12-8)^2 + (8-8)^2] \\ &= 14 \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad \text{Cov}(x, y) &= \frac{1}{4-1} [(3-8)(7-6.5) + (9-8)(5-6.5) + (12-8)(4-6.5) + (8-8)(10-6.5)] \\ &= -4.667. \end{aligned}$$

$$\text{(iii)} \quad \text{Cov}(x, y) = \text{Cov}(y, x) = -4.667$$

$$(iv) \text{Cov}(y, y) = \frac{1}{4-1} \left[(7-6.5)^2 + (5-6.5)^2 + (4-6.5)^2 + (10-6.5)^2 \right]$$

$$= 7$$

→ Covariance Matrix, $S = \begin{bmatrix} \text{Cov}(x, x) & \text{Cov}(x, y) \\ \text{Cov}(y, x) & \text{Cov}(y, y) \end{bmatrix}$

$$S = \begin{bmatrix} 17 & -4.667 \\ -4.667 & 7 \end{bmatrix}$$

Step 1: Eigen Value & Eigen Vector

① Eigen Value

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

$$\therefore \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

$$\Rightarrow \det \left(\begin{bmatrix} 17 & -4.667 \\ -4.667 & 7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right)$$

$$= \det \left(\begin{bmatrix} 17 - \lambda & -4.667 \\ -4.667 & 7 - \lambda \end{bmatrix} \right) = 0$$

$$(17 - \lambda)(7 - \lambda) + (4.667)^2 = 0.$$

$$98 - 14\lambda - 7\lambda + \lambda^2 + 21.781 = 0.$$

$$\lambda^2 - 21\lambda + 119.781 = 0.$$

$$\boxed{\lambda_1 = 10.5 ; \lambda_2 = 10.50}$$

② Eigen vector of λ_1

$$(S - \lambda_1 I) u_1 = 0.$$

$$\therefore \begin{bmatrix} 17 - \lambda_1 & -4.667 \\ -4.667 & 7 - \lambda_1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0.$$

$$\text{We get, } (17 - \lambda_1) u_1 - 4.667 u_2 = 0 \quad \text{--- (1)}$$

$$(-4.667) u_1 + (7 - \lambda_1) u_2 = 0 \quad \text{--- (2)}$$

$$\text{From (1), } \frac{u_1}{4.667} = \frac{u_2}{17 - \lambda_1} = t$$

$$\text{When } t = 1, u_1 = 4.667 \text{ and } u_2 = 17 - \lambda_1$$

$$\therefore \text{Eigen Vector of } u_1 \text{ of } \lambda_1 = \begin{bmatrix} 4.667 \\ 17 - \lambda_1 \end{bmatrix} = \begin{bmatrix} 4.667 \\ 17 - 10.5 \end{bmatrix}$$

$$\therefore \lambda_1 = \begin{bmatrix} 4.667 \\ 3.500 \end{bmatrix}$$