2018ICSE0621
Sai Ram

Date
Page

2018ICSE0621
Sai Ram. K
7-CSE-10.

Part-C

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 => $(x,x), (x,y), (y,x), (y,y)$

(i) $Cov(x,x) = \frac{1}{N-1} \sum_{k=1}^{N} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)$

$= \frac{1}{4-1} \left[(3-8)^2 + (9-8)^2 + (12-8)^2 + (8-8)^2 \right]$

$= 14$

(ii) $Cov(x,y) = \frac{1}{4-1} \left[(3-8)(7-6.5) + (9-8)(5-6.5) + (12-8)(4-6.5) + (8-8)(10-6.5) \right]$

$= -4.667$.

(iii) $Cov(x,y) = Cov(y,x) = -4.667$

2018ICSE0621
Sai Ram.

classmate
Date
Page

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

$= 7$

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

$$S = \begin{bmatrix} 14 & -4.667 \\ -4.667 & 7 \end{bmatrix}$$

Step 4: Eigen Value & Eigen Vector

(1) Eigen Value

We know $det(S-\lambda I)=0$ $\quad \because \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

$\Rightarrow det\left(\begin{bmatrix} 14 & -4.667 \\ -4.667 & 7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$

$= det\left(\begin{bmatrix} 14-\lambda & -4.667 \\ -4.667 & 7-\lambda \end{bmatrix}\right) = 0$

$(14-\lambda)(7-\lambda)-(4.667)^2 = 0.$

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

$\cancel{\lambda^2 - 21\lambda + 119.781 = 0.}$ $\quad \lambda^2 - 21\lambda + 76.219 = 0$

$\cancel{\boxed{\lambda_1 = 10.5 \; ; \; \lambda_2 = 10.50}}$ $\quad \boxed{\begin{matrix}\lambda_1 = 16.334 \\ \lambda_2 = 4.666\end{matrix}}$ *choosing λ_1

(2) Eigen vector of λ_1

$(S-\lambda_1 I)U_1 = 0.$

$\therefore \begin{bmatrix} 14-\lambda_1 & -4.667 \\ -4.667 & 7-\lambda_1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0.$

We get, $(14-\lambda_1)u_1 - 4.667u_2 = 0$ ---- (1)

$(-4.667)u_1 + (7-\lambda_1)u_2 = 0$ ---- (2)

From (1), $\frac{u_1}{4.667} = \frac{u_2}{14-\lambda_1} = t$

when $t=1$, $u_1 = 4.667$ and $u_2 = 14-\lambda_1$

$\therefore$ Eigen Vector of u_1 of $\lambda_1 = \begin{bmatrix} 4.667 \\ 14-\lambda_1 \end{bmatrix} = \begin{bmatrix} 4.667 \\ 14-16.33 \end{bmatrix}$

$\cancel{\therefore \lambda_1 = \begin{bmatrix} 4.667 \\ -2.33500 \end{bmatrix}}$ $\quad \boxed{\lambda_1 = \begin{bmatrix} 4.667 \\ -2.334 \end{bmatrix}}$