Steps for Principal Component Analysis –

Given the following data, use PCA to reduce the dimensions from 2 to 1.

Feature	Ex 1	Ex 2	Ex 3	Ex 4
x	4	8	13	7
y	11	4	5	14

Step 1 :–

No. of Features = 2 = n

No. of Samples = 4 = N

Step 2 :– Computation of mean of variables

mean of x $\quad \bar{x} = \frac{4+8+13+7}{4} = 8$

mean of y $\quad \bar{y} = \frac{11+4+5+14}{4} = 8.5$

Step 3 :– Computation of Covariance matrix

ordered pairs – (x, x) (x, y) (y, x) (y, y)

if n variables – n^2 ordered pairs

$$cov(x,x) = \frac{1}{N}\sum_{k=1}^{N}(x_i - \bar{x})^2$$

M T W T F S S

Page No.:

Date:

YOUVA

(i) Covariance of all ordered pairs.

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

$$= \frac{1}{4-1}\left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2\right]$$

$$cov(x,x) = 14$$

$$cov(x,y) = \frac{1}{4-1}\left[(24-8)(11-8.5) + (8-8)(4-8.5) + (13-8)(5-8.5) + (7-8)(14-8.5)\right]$$

$$cov(x,y) = -11$$

$$cov(y,x) = \frac{1}{4-1}\left[\right. \quad cov(x,y) = -11$$

$$cov(y,y) = \frac{1}{4-1}\left[(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\right]$$

$$= 23$$

(ii) Covariance matrix — $n \times n = 2 \times 2$

$$S = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix}$$

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

Step 4 :- Eigen Value, Eigen vector, Normalised eigen vector.

(i) Eigen value —

$$\text{determinant}(S - \lambda I) = 0$$

cov. matrix, Eigen value, Identity matrix

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

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

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

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

$$\lambda = 30.3849, 6.6151$$

$\lambda_1 > \lambda_2$

$\lambda_1 = 30.3849$ = First Principal component

$\lambda_2 = 6.6151$

(ii) Eigen vector of λ_1

$$(S - \lambda_1 I) u_1 = 0 \qquad u_1 = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

↳ eigen vector

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

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

$$(14-\lambda_1) u_1 - 11 u_2 = 0 \qquad \text{—— (1)}$$

$$(-11 u_1) + (23 - \lambda_1) u_2 = 0 \qquad \text{—— (2)}$$

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

When $t = 1$

$$u_1 = 11$$
$$u_2 = 14 - \lambda_1$$

Eigen Vector(λ_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) Normalise the eigen vector U_1

$$e_1 = \begin{bmatrix} 11/\sqrt{11^2+(-16.384)^2} \\ -16.3849/\sqrt{11^2+(16.38)^2} \end{bmatrix} \rightarrow \text{Normalised eigen vector}$$

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

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

Step5:- Derive new dataset

	Ex1	Ex2	Ex3	Ex4
First PC1	P_{11}	P_{12}	P_{13}	P_{14}

$$P_{11} = e_1^T \begin{bmatrix} 4-8 \\ 11-8.5 \end{bmatrix} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix}$$

(mean of x: 8; mean of y: 8.5)

$$= -4.3052$$

$$P_{12} = e_1^T \begin{bmatrix} 8-8 \\ 4-8.5 \end{bmatrix} = \begin{bmatrix} 0.5574 & -0.83903 \end{bmatrix} \begin{bmatrix} 0 \\ -4.5 \end{bmatrix}$$

$$= 3.7361$$

M T W T F S S

Page No.: YOUVA

Date:

$$P_{13} = e_1^T \begin{bmatrix} 13-8 \\ 5-8.5 \end{bmatrix}$$

$$= 5.6928$$

$$P_{14} = e_1^T \begin{bmatrix} 7-8 \\ 14-8.5 \end{bmatrix}$$

$$= -5.1238$$

	Ex1	Ex 2	Ex3	Ex4
PC1	-4.3052	3.7361	5.6928	-5.1238