# (*) Principle Component Analysis :-

(Q) Given the data in table, reduce the dimension from 2 to 1 using principle 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 |

## Solution :

(I) calculate Mean.

$$\overline{x_1} = \frac{4+8+13+7}{4} = 8$$

$$\overline{x_2} = \frac{11+4+5+14}{4} = 8.5$$

(II) calculate covariance matrix,

$$S = \begin{bmatrix} Cov(x_1, x_1) & Cov(x_1, x_2) \\ Cov(x_2, x_1) & Cov(x_2, x_2) \end{bmatrix}$$

$$Cov(x_1, x_1) = \frac{1}{N-1} \sum_{K=1}^{N} (x_{1K} - \bar{x}_1)(x_{1K} - \bar{x}_1)$$

$$= \frac{1}{3}\left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2\right) = 14$$

$$Cov(x_1, x_2) = \cancel{Cov} \frac{1}{N-1} \sum_{K=1}^{N} (x_{1K} - \bar{x}_1)(x_{2K} - \bar{x}_2)$$

$$= \frac{1}{3}\left[(4-8)(11-8.5) + (8-8)(4-8.5) + (13-8)(5-8.5) + (7-8)(14-8.5)\right] = -11$$

$$Cov(x_1, x_2) = Cov(x_2, x_1) = -11$$

$$Cov(x_2, x_2) = \frac{1}{N-1} \sum_{K=1}^{N} (x_{2K} - \bar{x}_2)(x_{2K} - \bar{x}_2)$$

$$= \frac{1}{3}\left[(11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\right]$$

$$= 23$$

$$\therefore S = \begin{bmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Cov(X_2, X_2) \end{bmatrix} = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

(III) Eigen values of the covariance matrix.

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

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

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

$$\therefore \lambda^2 - 37\lambda + 201 = 0$$

$$\therefore \lambda = \frac{1}{2}\left[37 \pm \sqrt{565}\right]$$

$$= 30.3849, 6.6151$$

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

(IV) Computation of the eigen vectors.

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

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

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

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

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

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

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

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

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

$$\therefore \boxed{U = \begin{bmatrix} 11 \\ 14-\lambda \end{bmatrix}}$$

Scanned with CamScanner

$$u_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}$$

* To find a unit eigen vector, we compute the length of $u_1$ which is given by,

$$\|u_1\| = \sqrt{11^2 + (14-\lambda_1)^2} = \sqrt{11^2 + (14-30.3849)^2} = 19.7348$$

$$e_1 = \begin{bmatrix} 11/\|u_1\| \\ (14-\lambda_1)/\|u_1\| \end{bmatrix} = \begin{bmatrix} 11/19.7348 \\ (14-\lambda_1)/19.7348 \end{bmatrix}$$

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

Similarly, calculate $e_2$, $e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$

(V) Computation of first principal components:

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

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

$$= 0.5574(x_{11} - \bar{x}_1) - 0.8303(x_{21} - \bar{x}_2)$$

$$= 0.5574(4-8) - 0.8303(11-8.5)$$

$$= -4.30535$$

| 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 |