Principle Component Analysis - Solved Example

→ Given the data in Table, reduce the dimension from 2 to 1 using the Principal 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

Step 1: Calculate Mean

$$\bar{X}_1 = \frac{1}{4}(4+8+13+7) = 8$$

$$\bar{X}_2 = \frac{1}{4}(11+4+5+14) = 8.5$$

Step 2: Calculation of the Covariance matrix

$$S = \begin{bmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Cov(X_2, X_2) \end{bmatrix}$$

$$Cov(X_1, X_1) = \frac{1}{N-1} \sum_{k=1}^{N} (X_{1k} - \bar{X}_1)(X_{1k} - \bar{X}_1)$$

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

$$= 14$$

$$Cov(X_1, X_2) = \frac{1}{N-1} \sum_{k=1}^{N} (X_{1k} - \bar{X}_1)(X_{2k} - \bar{X}_2)$$

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

$$= -11$$

$$Cov(X_2, X_1) = \frac{1}{3}\big((11-8.5)(4-8) + (4-8.5)(8-8) + (5-8.5)(13-8) + (14-8.5)(7-8)\big)$$

$$= -11$$

$$Cov(X_2, X_2) = \frac{1}{3}\big((11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\big)$$

$$= 23$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$ Co-variance matrix.

Step 3:- Eigen values of the Co-variance matrix.

The characteristic equation of the covariance matrix is,

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

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

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

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

$$\lambda = \frac{1}{2}\left(37 \pm \sqrt{565}\right)$$

$$= \underbrace{30.3849}_{\lambda_1}, \underbrace{6.6151}_{\lambda_2}$$

Step 4:- Computation of Eigen vectors

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

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

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

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

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

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

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

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

$$\Rightarrow u_1 = 11t \; ; \; u_2 = (14-\lambda)t$$

$$U_1 = \begin{bmatrix} 11 \\ 14-\lambda \end{bmatrix}$$

Eigen vector for given co-variance matrix S.

→ for principal components, consider the Largest eigen value.

$$\lambda_1 = 30.3849 \checkmark$$

$$\lambda_2 = 6.6151$$

$$U_1 = \begin{bmatrix} 11 \\ 14-\lambda_1 \end{bmatrix}$$

→ To find a unit eigen vector, we compute the length of U_1 which is given by,

$$\|U_1\| = \sqrt{11^2 + (14-\lambda_1)^2}$$

$$= \sqrt{11^2 + (14-30.3849)^2}$$

$$= 19.7348$$

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

$$= \begin{bmatrix} 11/19.7348 \\ (14-30.3849)/19.7348 \end{bmatrix}$$

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

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$ (Compute if necessary)

<u>Step 5</u>:- Computation of first principal components.

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

For Ex1: $$e_1^T \begin{bmatrix} X_{1k} - \bar{X}_1 \\ X_{2k} - \bar{X}_2 \end{bmatrix} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} X_{11} - \bar{X}_1 \\ X_{21} - \bar{X}_2 \end{bmatrix}$$

$$= 0.5574(X_{11} - \bar{X}_1) - 0.8303(X_{21} - \bar{X}_2)$$

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

$$= -4.30535$$

for Ex2:

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

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

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

$$= 0 + 0.8303 \times 4.5 = 3.7363$$

for Ex3:

$$= 0.5574(13-8) - 0.8303(5-8.5)$$

$$= 0.5574(5) - 0.8303(-3.5)$$

$$= 5.6930$$

For Ex4:

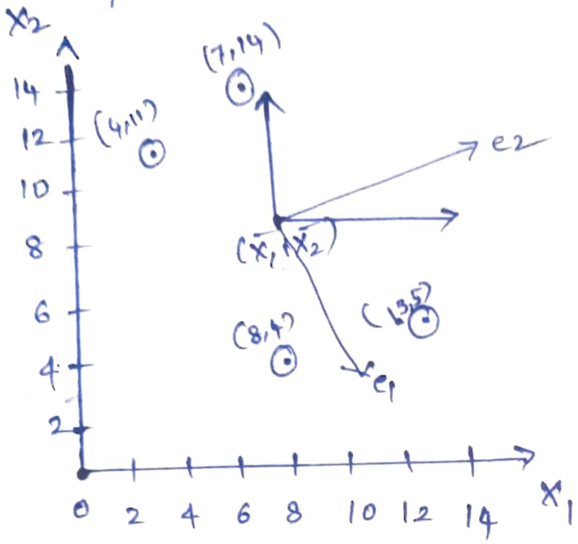
$$= 0.5574(7-8) - 0.8303(14-8.5)$$

$$= -0.5574 - 0.8303(5.5)$$

$$= -5.1240$$

Feature	Ex1	Ex2	Ex3	Ex4
x_1	4	8	13	7
x_2	11	4	5	14
First Principal Components	-4.3052	3.7363	5.6930	-5.1240

Step 6:- Geometrical meaning of first principal components.



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

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