# Principal component analysis.

Defn:- The basic idea of PCA is to describe the variation of a set of multivariate data in terms of a set of uncorrelated variables each of which is a particular linear combination of the original variables.

1. compute the principal component to the following covariance Matrix

$$\Sigma = \begin{bmatrix} 1 & 4 \\ 4 & 100 \end{bmatrix}$$

Soln:- Consider

$$\begin{vmatrix} 1-\lambda & 4 \\ 4 & 100-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)(100-\lambda) - 16 = 0$$

$$\lambda^2 - 101\lambda + 84 = 0$$

Solving we get.

$\lambda_1 = 100.16135$

$\lambda_2 = 0.8386.$

Eigen vector :-

(i) $\lambda_1 = 100.16135.$

$$\begin{bmatrix} 1-100.16135 & -4 \\ 4 & 100-100.1635 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -99.16135 & 4 \\ 4 & -0.16135 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$-99.16135x_1 + 4x_2 = 0.$

$4x_1 - 0.16135x_2 = 0.$

$\Rightarrow \quad x_2 = 1$

$x_1 = 0.040345.$

$$e_1 = \begin{bmatrix} \dfrac{x_1}{\sqrt{x_1^2+x_2^2}} \\ \dfrac{x_2}{\sqrt{x_1^2+x_2^2}} \end{bmatrix} = \begin{bmatrix} 0.0404 \\ 0.9992 \end{bmatrix}$$

(ii) $\lambda_2 = 0.8386.$

$$\begin{bmatrix} 0.1614 & 4 \\ 4 & 99.1614 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

$0.1614x_1 + 4x_2 = 0$

$4x_1 + 99.1614x_2 = 0$

Put $x_2 = 1$

$$x_1 = -24.79035$$

$$e_2 = \begin{bmatrix} \dfrac{-24.79035}{\sqrt{1+24.79035^2}} \\ \dfrac{-24.79035}{\sqrt{1+24.79035^2}} \end{bmatrix}$$

$$= \begin{bmatrix} -0.9991 \\ 0.0403 \end{bmatrix}$$

I principal component:

$$P_1 = \frac{\lambda_1}{\lambda_1+\lambda_2} = \frac{100.16135}{100.1613 + 0.8386}$$

$$= 0.9916.$$

II Principal component

$$P_2 = \frac{\lambda_2}{\lambda_1+\lambda_2} = \frac{0.8386}{100.1613 + 0.8386}$$

$$= 0.0833$$

The principal components are

$$Y_1 = e_1' x = 0.04x_1 + 0.999x_2.$$

$$Y_2 = e_2^2 x = -0.9991x_1 + 0.0403x_2.$$