Principal components from covariance matrix:-

There are different approaches with which one can find principal components. We are going to see PCA from covariance matrix (other approach is PCA from SVD).

Steps:-

1. center the data matrix X_R to X.
2. find covariance matrix C
3. Do eigen decomposition of C ie find the eigen values & eigen vectors of C.

$$C = VSV^T$$

4. find the principal components: these are columns of the matrix XV :- $P = XV$
5. Do dimensionality reduction, if possible.

Note:-

$X_R \longrightarrow$ data matrix

$X \longrightarrow$ centered at origin

$C \longrightarrow$ covariance matrix

$V \longrightarrow$ unit eigenvectors of matrix C (in columns) $\begin{bmatrix} e_1 & e_2 \end{bmatrix}$

$S \longrightarrow$ diagonals are eigen values other elements are zero $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$

1

EX:- Perform PCA on the given matrix & find the reduced matrix. 2

→ $X_R = \begin{bmatrix} 3 & 6 \\ 2 & -1 \\ -2 & 1 \end{bmatrix}$

There are two variables & three observations. It is 2D. $\boxed{n = 3}$

steps

(i) centering the data matrix X_R to X.

Finding mean of columns $\left[\frac{3+2-2}{3}, \frac{6-1+1}{3} \right] = [1, 2]$

Now, centering the data matrix ie considering origin as mean.

$\therefore \quad X = \begin{bmatrix} 3 & 6 \\ 2 & -1 \\ -2 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}$.

$\Rightarrow \quad X = \begin{bmatrix} 2 & 4 \\ 1 & -3 \\ -3 & -1 \end{bmatrix}$ — 0

(ii) finding covariance matrix C.

$$\boxed{C = \frac{X^T X}{n-1}}$$

$$= \frac{1}{3-1} \begin{bmatrix} 2 & 1 & -3 \\ 4 & -3 & -1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & -3 \\ -3 & -1 \end{bmatrix}$$

$$C = \frac{1}{2} \begin{bmatrix} 14 & 8 \\ 8 & 26 \end{bmatrix}$$

$$\boxed{C = \begin{bmatrix} 7 & 4 \\ 4 & 13 \end{bmatrix}} \text{ — 1}$$

For $n \times p$ data matrix, C will be $p \times p$.

2 AAM

(iii) Do eigen decomposition of C
i.e finding eigen values & eigen vectors of C.

$$C = \begin{bmatrix} 7 & 4 \\ 4 & 13 \end{bmatrix}$$

finding eigen values :- $|A - \lambda I| = 0$

$$= \begin{bmatrix} 7-\lambda & 4 \\ 4 & 13-\lambda \end{bmatrix}$$

$$= 91 - 20\lambda + \lambda^2 - 16 = 0$$

$$\lambda^2 - 20\lambda + 75 = 0$$

trace (C) ← 20
i.e sum of diagonal elements

75 → determinant of C $\begin{vmatrix} 7 & 4 \\ 4 & 13 \end{vmatrix}$

$$\Rightarrow \lambda_1, \lambda_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$= \frac{20 \pm \sqrt{(20)^2 - 4(1)(75)}}{2}$$

$$\lambda_1, \lambda_2 = 15, 5$$

Note:-
Always the bigger value should be taken at λ_1 (corresponds to 1st principal component)
i.e $\lambda_1 = 15$ & $\lambda_2 = 5$.

Therefore matrix S becomes

$$\boxed{S = \begin{bmatrix} 15 & 0 \\ 0 & 5 \end{bmatrix}} \quad -(2)$$

(3) AAM

Finding eigen vectors of C.

(i) e_1 corresponding to $\lambda_1 = 15$.

$$\begin{bmatrix} 7-15 & 4 \\ 4 & 13-15 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\begin{bmatrix} -8 & 4 \\ 4 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$-8x + 4y = 0 \Rightarrow 8x = 4y$

$4x - 2y = 0$ $\quad x = \frac{1}{2}y \Rightarrow y = 2x$

let $x = 1$ $y = 2$

$e_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

finding unit vector, dividing by magnitude ie $\sqrt{1^2+4^2} = \sqrt{5}$

$\therefore e_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

(ii) e_2 corresponding to $\lambda_2 = 5$

$$\begin{bmatrix} 7-5 & 4 \\ 4 & 13-5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\begin{bmatrix} +2 & 4 \\ 4 & +8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$+2x + 4y = 0$

$4x + 8y = 0$

$2x = -4y$ $\quad x = 2y.$

let $y = 1$ $x = -2$.

$= \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

finding unit vector, dividing by magnitude $\sqrt{5}$.

$\therefore e_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$\therefore$ matrix V becomes

$$V = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

4 $\Rightarrow \boxed{V = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}$ — 3 AAM

Thus it can verified that:

$$C = V S V^T$$

$$\Rightarrow C = \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} 15 & 0 \\ 0 & 5 \end{bmatrix}}_{S} \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}}_{V^T.}$$

(iv) finding Principal components

$$P = XV$$

$$= \begin{bmatrix} 2 & 4 \\ 1 & -3 \\ -3 & -1 \end{bmatrix} \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}}\begin{bmatrix} 10 & 0 \\ -5 & -5 \\ -5 & 5 \end{bmatrix}$$

$$P = \begin{bmatrix} 4.47 & 0 \\ -2.24 & -2.24 \\ -2.24 & 2.24 \end{bmatrix}$$

PC_1 $\quad$ PC_2

(v) finding variance explained.

For this consider the matrix S

$$S = \begin{bmatrix} 15 & 0 \\ 0 & 5 \end{bmatrix}$$

Total variance $= \lambda_1 + \lambda_2 = 20$

5 AAM

6

% variance explained by $PC_1 = \frac{15}{20} \times 100 = 75\%$.

% variance explained by $PC_2 = \frac{5}{20} \times 100 = 25\%$.

This shows that PC_1 is more important.
If we are ready to work with PC_1 only ie reducing our dimension from 2 to 1, then

(a) $$P = \begin{bmatrix} 4.47 & 0 \\ -2.24 & 0 \\ -2.24 & 0 \end{bmatrix}$$

Now getting the data projected on 1D only ie reversing the process

$\because P = X V$

$\Rightarrow X = P.V^T \longrightarrow$ reduced matrix with P as one column only

let the reduced matrix is denoted as $X_1 = P.V^T$

$$= \begin{bmatrix} 4.47 & 0 \\ -2.24 & 0 \\ -2.24 & 0 \end{bmatrix} \cdot \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} 4.47 & 8.94 \\ -2.24 & -4.48 \\ -2.24 & -4.48 \end{bmatrix}$$

$$X_1 = \begin{bmatrix} 2 & 4 \\ -1 & -2 \\ -1 & -2 \end{bmatrix}$$

Now bringing the matrix back ie moving from center at origin. 5 AAM

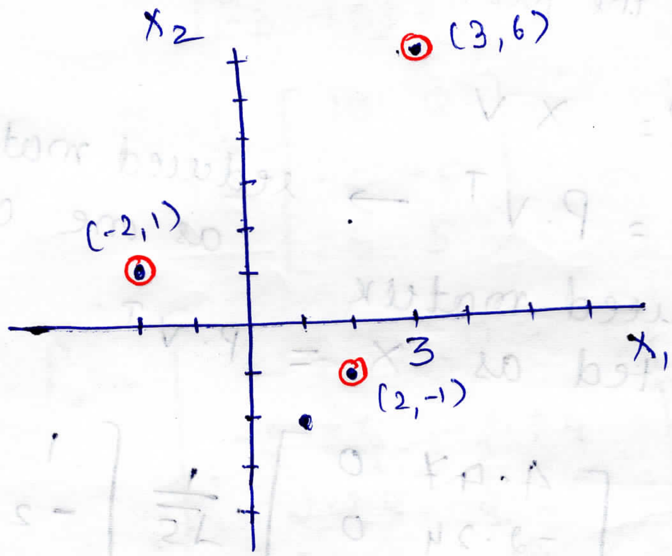
$$X_{R_1} = \begin{bmatrix} 2 & 4 \\ -1 & -2 \\ -1 & -2 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}$$

$$X_{R_1} = \begin{bmatrix} 3 & 6 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$ reduced data matrix

$$X_R = \begin{bmatrix} 3 & 6 \\ 2 & -1 \\ -2 & 1 \end{bmatrix}$$ original data matrix.

Conclusion :-

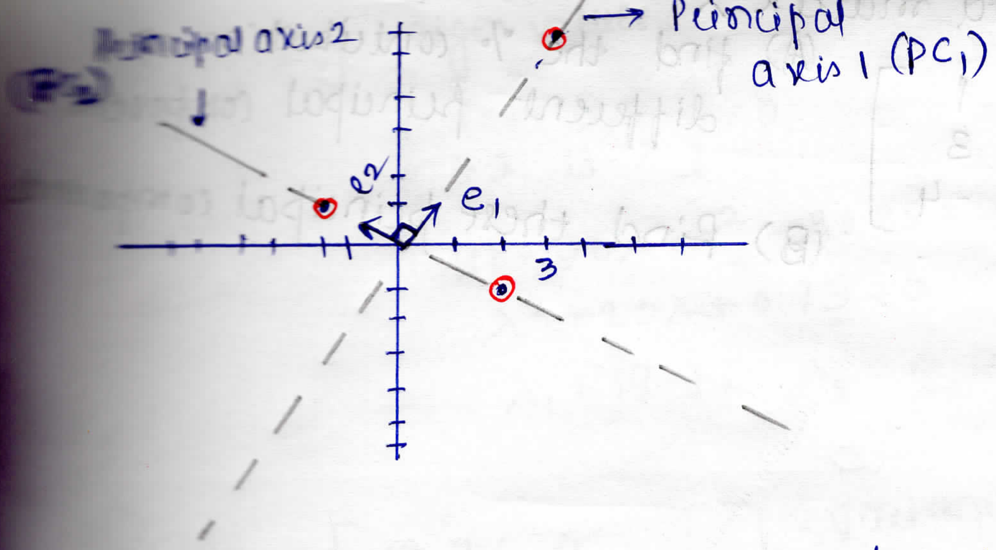
if we draw the data points of original data matrix



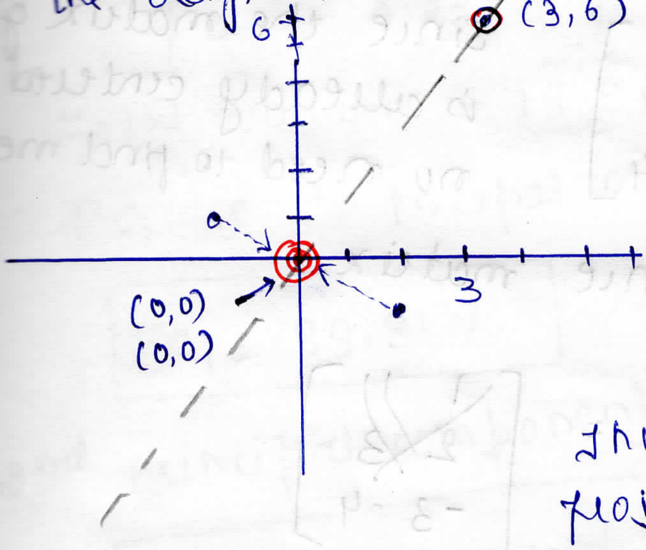
- fig ①

Then by finding variance (eigen values) we found eigen vectors. ~~which are the principal com~~

The larger eigen value contributes more & its corresponding PC is taken & the other is neglected & we found the new data matrix in reduced dimension



But the ~~original~~ new data points are



→ fig ②

$$XR_1 = \begin{bmatrix} 3 & 6 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Thus pts (-2,1) & (2,-1) are projected on PC_1 & collapsed at (0,0) compare fig ① & ②

Thus now you need only one line ie 1D.

~~Load~~

Home work :-

Ex : $XR = \begin{bmatrix} 2 & 3 \\ 1 & -2 \\ 3 & 8 \end{bmatrix}$

Ex : $XR = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ -3 & -4 \end{bmatrix}$

8 AAM