Steps for PCA Analysis:

Suppose data set is like:

x	y
2	1
3	5
4	5
5	6
6	7
7	8

1. Find the zero mean:

$$Mean\ of\ x = \mu x, Mean\ of\ y = \mu y$$

$$\mu x\ =\ 4.5\ and\ \mu y\ =\ 5.33$$

Then,

x	y	$x - \mu x$	$y - \mu y$
2	1	-2.5	-4.333333
3	5	-1.5	-0.333333
4	5	-0.5	-0.333333
5	6	0.5	0.666667
6	7	1.5	1.666667
7	8	2.5	2.666667

Now A=

-2.5	-4.333333
-1.5	-0.333333
-0.5	-0.333333
0.5	0.666667
1.5	1.666667
2.5	2.666667

2. Find the covariance matrix of A.

$$covA = \frac{1}{m}\ (A * A^T)$$

Here m is number of rows

for this example, it is,

$$covA = \begin{bmatrix} 3.5 & 4.2 \\ 4.2 & 5.9 \end{bmatrix}$$

Md. Hasibur Rahman, Lecturer, Dept. of CSE, BUBT

3. Now we have to find the PC scores (Eigen value) & PC loadings (Eigen vector). Here are the steps with equations

Equation for Eigen vector of the matrix $covA$ is

$covA.\overline{X} = \lambda.\overline{X}$

Here,

$\overline{X} = PC\ loadings\ (eigen\ vector)$

$\lambda = PC\ scores\ (eigen\ value)$

Now go with the steps to find Eigen vectors of a matrix

i. Calculate: $A - \lambda I$

   Here $\begin{bmatrix} 3.5-\lambda & 4.2 \\ 4.2 & 5.9-\lambda \end{bmatrix}$

ii. Now $\det(A - \lambda I) = 0 - - - - - - - - - (eq1)$

   $det \begin{bmatrix} 3.5-\lambda & 4.2 \\ 4.2 & 5.9-\lambda \end{bmatrix} = 0$

iii. From *eq1* we will get the value of $\lambda 1$ & $\lambda 2$

   Here $\lambda 1 = 0.32$ & $\lambda 2 = 9.05$

iv. $Now\ [covA - \lambda I][\overline{X}] = \bar{0} - - - - - - - - - (eq2)$

   As $\lambda 2 > \lambda 1,$ $we\ will\ take\ \lambda 2\ as\ the\ best\ PC\ score$

   So for $\lambda 2$ the *eq2* is

   $\begin{bmatrix} -5.5 & 4.2 \\ 4.2 & -3.15 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

   From this we will get two equations to solve $x1\ and\ x2$ and get the Eigen vector $\overline{X}$

   From above we get

   $-5.5\ x1 + 4.2\ x2 = 0 - - - - - (eq3)$

   $4.2\ x1 - 3.15\ x2 = 0 - - - - - (eq4)$

   Solving *eq3 & eq4* we get $x1 = 0.604$ & $x2 = -0.797$

v. You will find out the (PC1)Eigen vector $\bar{X}\ for\ \lambda 2$

   Here $\bar{X} = \begin{bmatrix} 0.604 \\ -0.797 \end{bmatrix}$

4. Calculate the projected value as *PC1*

$PC1 = A.\overline{(X)}^T$

5. Now plotting the data:

Here in graph, red line is PC1, green line is PC2 and blue dots are our data. We have observed that PC1 is the best one.

The equation to plot for PC1 is

$$0.604x1 - 0.797x2 = 0$$

[All calculations is approximate]

Wow this is awesome!

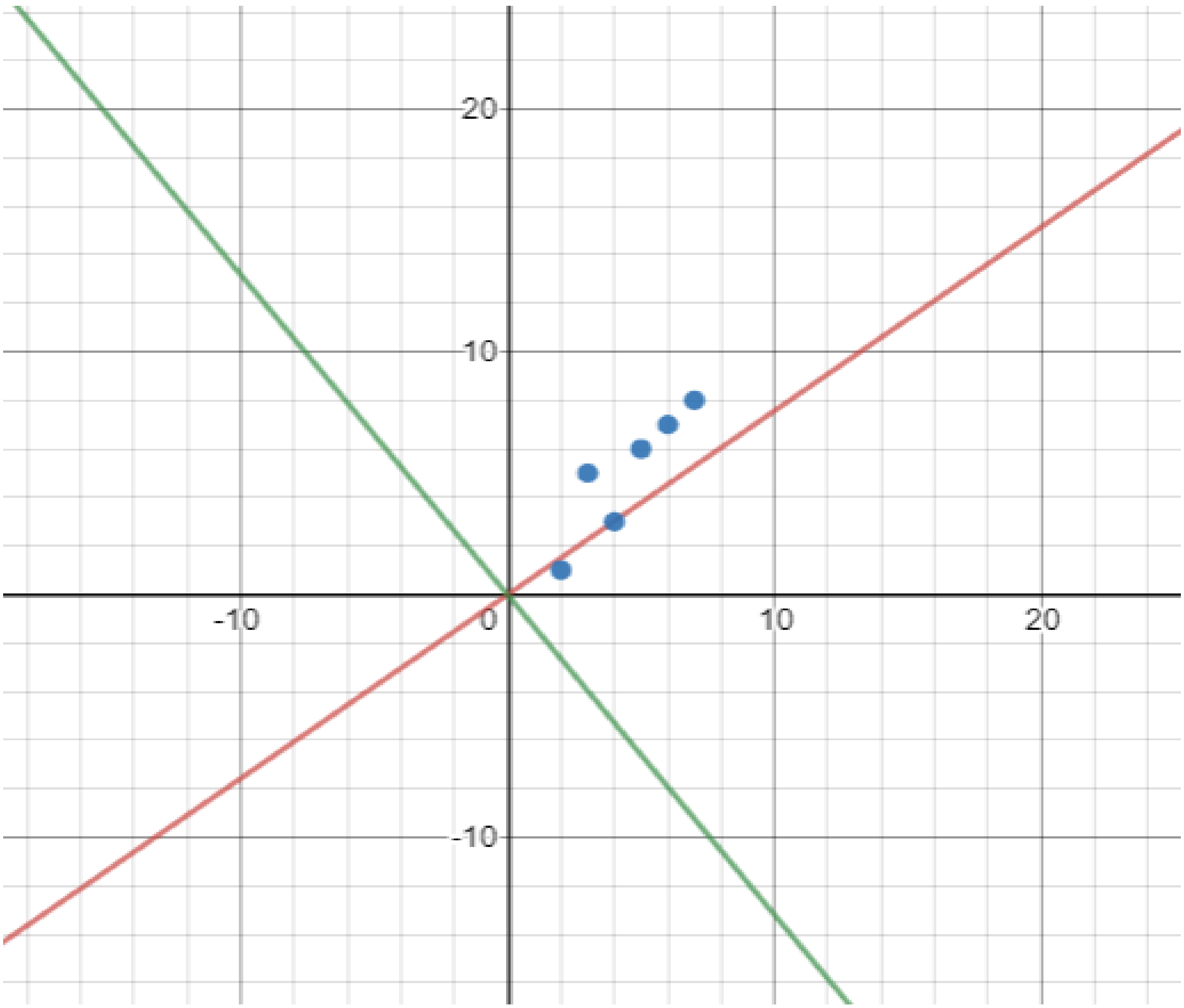

Figure 1: PCA Analysis graph.