1. dataset

X_1	X_2	D
3	-1	a
4	-2	a
3	0	a
-1	1	b
1	2	b
0	2	b

By using PCA method, reduce the dimension of unsupervised version of the first dataset.

At first, the covariance matrix is found.

$$\mu = [1.66\ 0.33]$$

$$X - \mu = \begin{bmatrix} 1.34 & -1.33 \\ 2.34 & -2.33 \\ 1.34 & -0.33 \\ -2.66 & 0.66 \\ -0.66 & 1.66 \\ -1.66 & 1.66 \end{bmatrix}$$

$$cov = \frac{1}{\mathrm{n}-1} * \sum_{n=1}^{n} (X-\mu)'(X-\mu)$$

$$cov = \frac{1}{5}\begin{pmatrix} 1.34 & 2.34 & 1.34 & -2.66 & -0.66 & -1.66 \\ -1.33 & -2.33 & -0.33 & 0.66 & 1.66 & 1.66 \end{pmatrix}\begin{pmatrix} 1.34 & -1.33 \\ 2.34 & -2.33 \\ 1.34 & -0.33 \\ -2.66 & 0.66 \\ -0.66 & 1.66 \\ -1.66 & 1.66 \end{pmatrix}$$

$$= \begin{pmatrix} 3.87 & -2.67 \\ -2.67 & 2.67 \end{pmatrix}$$

According to PCA method, eigen value (λ) and eigen matrix (E) are found by using the covariance matrix.

$$cov * E = \lambda * E$$

With a small trick,

$$cov - I * \lambda = 0$$

$$\begin{pmatrix} 3.87 - \lambda & -2.67 \\ -2.67 & 2.67 - \lambda \end{pmatrix} = 0$$

$$(3.87 - \lambda) * (2.67 - \lambda) — (-2.67)^2 = 0$$

$$\lambda_1 = 6 \qquad \lambda_2 = 0.53$$

We should solve it for each lambda value. At first, for $\lambda_1 = 6$

$$(cov - I * \lambda) * E = 0$$

$$\begin{bmatrix} 3.87 - 6 & -2.67 \\ -2.67 & 2.67 - 6 \end{bmatrix}\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -2.13 * E_1 - 2.67 * E_2 \\ -2.67 * E_1 - 3.33 * E_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-0.8 * E_1 = E_2$$

And then, for $\lambda_2 = 0.53$

$$\begin{bmatrix} 3.87 - 0.53 & -2.67 \\ -2.67 & 2.67 - 0.53 \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3.34 * E_1 - 2.67 * E_2 \\ -2.67 * E_1 + 2.14 * E_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$E_1 = 0.8 * E_2$$

Eigen vector,

for λ_1 for λ_2

$$E = \begin{bmatrix} E_1 & 0.8 * E_2 \\ -0.8 * E_1 & E_2 \end{bmatrix} * k$$

$$E = \begin{bmatrix} 1 & 0.8 \\ -0.8 & 1 \end{bmatrix} * k$$

In order that the determinant of the eigen vectors matrix is 1,

$$1 * k^2 + 0.64 * k^2 = 1$$
$$k = 0.7809$$

When we write 0.7809 value instead k,

$$E = \begin{bmatrix} 0.7809 & 0.6247 \\ -0.6247 & 0.7809 \end{bmatrix}$$

In order to compute the transformed dataset

$$X * E = \begin{bmatrix} 2.97 & 1.09 \\ 4.37 & 0.94 \\ 2.34 & 1.87 \\ -1.41 & 0.16 \\ -0.47 & 2.19 \\ -1.25 & 1.56 \end{bmatrix}$$

$\mu = [1.09\ 1.3]$

$$((X * E) - \mu)^2 = \begin{matrix} 3.53 & 0.04 \\ 10.75 & 0.13 \\ 1.56 & 0.32 \\ 6.26 & 1.30 \\ 2.44 & 0.79 \\ 5.48 & 0.07 \end{matrix}$$

We pick the first column because it has greater standard deviation, and we eliminate the second column.