PCA

Algorithm:

Step-1: Get data

Step-2: Substract the mean

original data

Data =

x	y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9

Data Adjust =

x	y
.69	.49
-1.31	-1.21
.39	.99
.09	.29
1.29	1.09
.49	.79
.19	-.31
-.81	-.81
-.31	-.31
-.71	-1.01

Step-3: Calculate covariance Matrix

$$cov = \begin{bmatrix} .616555556 & .615444444 \\ .615444444 & .716555556 \end{bmatrix}$$

Step-4: Calculate the eigenvectors & eigenvalues of the covariance matrix

$$eigenValues = \begin{bmatrix} .0490833989 \\ 1.28402771 \end{bmatrix}$$

$$eigenvectors = \begin{bmatrix} -.735178656 & -.677873399 \\ .677873399 & -.735178656 \end{bmatrix}$$

Step-5: Choosing components and forming a feature vector

$$\text{FeatureVector} = (eig1, eig2, \dots eign)$$

In our example,

$$\begin{pmatrix} -.677873399 & -.735178656 \\ -.735178656 & .677873399 \end{pmatrix}$$

we are choosing $\begin{pmatrix} -.677873399 \\ -.735178656 \end{pmatrix}$ after sorting

Step-6: Deriving new Dataset

1. Once we have chosen the components (eigen vectors) that we wish to keep in our data and formed a feature vector.

   ⇒ take the transpose of the vector and multiply it on the left of the original dataset, transposed.

⇒ FinalData = RowFeatureVector × RowDataAdjust

where RowFeatureVector is the matrix with the eigen vectors in the column transposed.

RowDataAdjust is the mean-adjusted data transposed

$$\text{eigenVector} = \begin{bmatrix} -.677873399 \\ -.735178656 \end{bmatrix}$$

⇒ Transpose of eigen Vector = $\begin{bmatrix} -.677873399 & -.735178656 \end{bmatrix}$

Let your →

⇒ Transpose of mean-adjusted data =

Let your →

$$\begin{bmatrix} .69 & -1.31 & .39 & .09 & 1.29 & .49 & .19 & -.81 & -.31 & -.71 \\ .49 & -1.21 & .99 & .29 & 1.09 & .79 & -.31 & -.81 & -.31 & -1.01 \end{bmatrix}$$

New data is x · y.

$$\begin{pmatrix} -.677873399 \times .69 \end{pmatrix} + \begin{pmatrix} -.735178656 \times .49 \end{pmatrix} = -.827970186$$

$$\begin{pmatrix} -.677873399 \times (-1.31) \end{pmatrix} + \begin{pmatrix} -.735178656 \times (-1.21) \end{pmatrix} = 1.77758033$$

x

Transformed data ⟸

$$\begin{bmatrix} -.827970186 \\ 1.77758033 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix}$$

finall we choose X for representation of X & Y.