PCA Numerical Example

Let’s understand steps of PCA with an example of two dimensional data.

We start with the following set of 5 points:

$$\mathbf{X} = \begin{bmatrix} 4 & 6 & -1 & -2 & -3 \\ 5 & 9 & -7 & -5 & -9 \end{bmatrix}$$

This is a more compact way of writing (using vectors & matrices); otherwise you will have many summations to deal with.

PCA Numerical Example

Calculate mean:

$$\bar{X} = \begin{bmatrix} 0.8 \\ -1.4 \end{bmatrix}$$

PCA Numerical Example

Center the points by performing $(\mathrm{X} - \bar{X})$

$$\begin{bmatrix} 3.2 & 5.2 & -1.8 & -2.8 & -3.8 \\ 6.4 & 10.4 & -5.6 & -3.6 & -7.6 \end{bmatrix}$$

PCA Numerical Example

The covariance matrix A is: 1/n * (X – $\bar{X}$)(X – $\bar{X}$)^T
where n is the number of points (=5).

$$A = \begin{bmatrix} 12.56 & 24.72 \\ 24.72 & 50.24 \end{bmatrix}$$

PCA Numerical Example

Finding the eigenvalues:

$$\begin{bmatrix} 12.56 & 24.72 \\ 24.72 & 50.24 \end{bmatrix} - \lambda I$$

I is the identity matrix. I = $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

PCA Numerical Example

Finding the eigenvalues:

$$\begin{bmatrix} 12.56 - \lambda & 24.72 \\ 24.72 & 50.24 - \lambda \end{bmatrix}$$

PCA Numerical Example

Finding the eigenvalues:

Determinant of $\begin{bmatrix} 12.56 - \lambda & 24.72 \\ 24.72 & 50.24 - \lambda \end{bmatrix} = 0$

PCA Numerical Example

Finding the eigenvalues:

$$(12.56 - \lambda) * (50.24 - \lambda) \; - \; 24.72^2 = 0$$

PCA Numerical Example

Finding the eigenvalues:

$$(\lambda^2) - 62.8\lambda + 19.936 = 0$$

PCA Numerical Example

Finding the eigenvalues:

Now, we find the roots of this quadratic equation as:

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

PCA Numerical Example

Finding the eigenvalues:

This gives the roots as:

$$\lambda_1 = 62.4809$$
$$\lambda_2 = 0.3190$$

PCA Numerical Example

Calculate the Eigenvector for the 1st Eigenvalue:

$$(A - \lambda_1 I) * X = 0$$

where X = $\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}$

PCA Numerical Example

Calculate the Eigenvector for the 1st Eigenvalue:

$$(A - \lambda_1 I) = \begin{bmatrix} 12.56 & 24.72 \\ 24.72 & 50.24 \end{bmatrix} - \begin{bmatrix} 62.48 & 0 \\ 0 & 62.48 \end{bmatrix}$$

$$(A - \lambda_1 I) = \begin{bmatrix} -49.92 & 24.72 \\ 24.72 & -12.24 \end{bmatrix}$$

PCA Numerical Example

Calculating the Eigenvector for the 1st Eigenvalue:

$$\begin{bmatrix} -49.92 & 24.72 \\ 24.72 & -12.24 \end{bmatrix} * \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$a_1{}^2 + b_1{}^2 = 1$$

From these equations, we can get the 1st eigenvector.

PCA Numerical Example

Calculating the Eigenvector for the 1st Eigenvalue:

$\begin{bmatrix} -49.92 & 24.72 \\ \mathbf{24.72} & \mathbf{-12.24} \end{bmatrix} * \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ➔ 24.72 a_1 = 12.24 b_1 ➔ a_1 = 0.4951 b_1

$a_1{}^2 + b_1{}^2 = 1$ ➔ 0.2452 $b_1{}^2$ + $b_1{}^2$ = 1 ➔ 1.2452 $b_1{}^2$ = 1 ➔ $b_1{}^2$ = 0.8031 ➔ b_1 = + or – 0.8962 ➔ a_1 = 0.4951 x +/- 0.8962 = 0.4437

1st Eigen vector $\begin{pmatrix} 0.4437 \\ 0.8962 \end{pmatrix}$

PCA Numerical Example

Calculate the Eigenvector for the 2nd Eigenvalue:

$$(A - \lambda_2 I) * X = 0$$

where X = $\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$

PCA Numerical Example

Calculate the Eigenvector for the 2nd Eigenvalue:

$$(A - \lambda_2 I) = \begin{bmatrix} 12.56 & 24.72 \\ 24.72 & 50.24 \end{bmatrix} - \begin{bmatrix} 0.03 & 0 \\ 0 & 0.03 \end{bmatrix}$$

$$(A - \lambda_2 I) = \begin{bmatrix} 12.53 & 24.72 \\ 24.72 & 50.21 \end{bmatrix}$$

PCA Numerical Example

Calculate the Eigenvector for the 2nd Eigenvalue:

$$\begin{bmatrix} 12.53 & 24.72 \\ 24.72 & 50.21 \end{bmatrix} * \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$a_2{}^2 + b_2{}^2 = 1$$

From these equations, we can get the 2nd eigenvector.

PCA Numerical Example

Calculate the Eigenvector for the 2nd Eigenvalue:

$\begin{bmatrix} \mathbf{12.53} & \mathbf{24.72} \\ 24.72 & 50.21 \end{bmatrix} * \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ➔ 12.53 a_2 = -24.72 b_2 ➔ a_2 = -1.97 b_2

$a_2{}^2 + b_2{}^2 = 1$ ➔ 3.89 $b_2{}^2$ + $b_2{}^2$ = 1 ➔ 4.89 $b_2{}^2$ = 1 ➔ $b_2{}^2$ = 0.2045
➔ b_2 = + or – 0.4522 ➔ a_2 = -1.97 x +/- 0.4522 = 0.8909

From these equations, we can get the 2nd eigenvector.

2nd Eigen vector $\begin{pmatrix} 0.8909 \\ -0.4522 \end{pmatrix}$

PCA Numerical Notes

The variance of projections in the line of the principal components is equal to the Eigenvalues of the principal components.

Sum of all the Eigenvalues is equal to the total variance of all the variance.

1st eigen value is bigger than 2nd , 2nd is bigger than 3rd and so on.