Artificial Intelligence

Dr. Mubashir Ahmad (Ph.D.)

Principal Component Analysis

- Principal Component Analysis is an unsupervised learning algorithm that is used for the dimensionality reduction in machine learning. It is a statistical process that converts the observations of correlated features into a set of linearly uncorrelated features with the help of orthogonal transformation. These new transformed features are called the Principal Components. It is one of the popular tools that is used for exploratory data analysis and predictive modeling. It is a technique to draw strong patterns from the given dataset by reducing the variances.
- PCA generally tries to find the lower-dimensional surface to project the high-dimensional data.
- PCA works by considering the variance of each attribute because the high attribute shows a good split between the classes, and hence it reduces the dimensionality. Some real-world applications of PCA are ***image processing, movie recommendation systems, and optimizing the power allocation in various communication channels.*** It is a feature extraction technique, so it contains the important variables and drops the least important variable.

Principal Component Analysis

- The PCA algorithm is based on some mathematical concepts such as:
- Variance and Covariance
- Eigenvalues and Eigen factors

- Some common terms used in the PCA algorithm:

- **Dimensionality:** It is the number of features or variables present in the given dataset. More easily, it is the number of columns present in the dataset.

- **Correlation:** It signifies how strongly two variables are related to each other. Such as if one changes, the other variable also gets changed. The correlation value ranges from -1 to +1. Here, -1 occurs if variables are inversely proportional to each other, and +1 indicates that variables are directly proportional to each other.

Principal Component Analysis

- **Orthogonal:** It defines that variables are not correlated to each other, and hence the correlation between the pair of variables is zero.

- **Eigenvectors:** If there is a square matrix M and a non-zero vector v is given. Then v will be the eigenvector if Av is the scalar multiple of v.

- **Covariance Matrix:** A matrix containing the covariance between the pair of variables is called the Covariance Matrix.

Steps Principal Component Analysis

1. Getting the dataset

Firstly, we need to take the input dataset and divide it into two subparts X and Y, where X is the training set, and Y is the validation set.

2. Representing data into a structure

Now we will represent our dataset into a structure. Such as we will represent the two-dimensional matrix of independent variable X. Here each row corresponds to the data items, and the column corresponds to the Features. The number of columns is the dimensions of the dataset.

3. Standardizing the data

In this step, we will standardize our dataset. Such as in a particular column, the features with high variance are more important compared to the features with lower variance.
If the importance of features is independent of the variance of the feature, then we will divide each data item into a column with the standard deviation of the column. Here we will name the matrix as Z.

Steps Principal Component Analysis

4. Calculating the Covariance of Z

To calculate the covariance of Z, we will take the matrix Z, and will transpose it. After transpose, we will multiply it by Z. The output matrix will be the Covariance matrix of Z.

5. Calculating the Eigen Values and Eigen Vectors

Now we need to calculate the eigenvalues and eigenvectors for the resultant covariance matrix Z. Eigenvectors or the covariance matrix are the directions of the axes with high information. The coefficients of these eigenvectors are defined as the eigenvalues.

6. Sorting the Eigen Vectors

In this step, we will take all the eigenvalues and sort them in decreasing order, which means from largest to smallest. And simultaneously sort the eigenvectors accordingly in matrix P of eigenvalues. The resultant matrix will be named as P*.

Steps Principal Component Analysis

7. Calculating the new features Or Principal Components

Here we will calculate the new features. To do this, we will multiply the P* matrix to the Z. In the resultant matrix Z*, each observation is the linear combination of original features. Each column of the Z* matrix is independent of each other.

8. Remove less or unimportant features from the new dataset.

The new feature set has occurred, so we will decide here what to keep and what to remove. It means, we will only keep the relevant or important features in the new dataset, and unimportant features will be removed.

Principal Component Analysis Solved Example 1

- Given the data in Table, reduce the dimension from 2 to 1 using the Principal Component Analysis (PCA) algorithm.

Feature	**Example 1**	**Example 2**	**Example 3**	**Example 4**
$\mathbf{X_1}$	4	8	13	7
$\mathbf{X_2}$	11	4	5	14

Principal Component Analysis Solved Example 1

Step 1: Calculate Mean

$$\bar{X}_1 = \tfrac{1}{4}(4 + 8 + 13 + 7) = 8,$$
$$\bar{X}_2 = \tfrac{1}{4}(11 + 4 + 5 + 14) = 8.5.$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = 8$$
$$\overline{X_2} = 8.5$$

Principal Component Analysis Solved Example 1

Step 2: Calculation of the covariance matrix.

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\mathrm{Cov}(X_1, X_1) = \frac{1}{N-1}\sum_{k=1}^{N}(X_{1k} - \bar{X}_1)(X_{1k} - \bar{X}_1)$$

$$= \tfrac{1}{3}\left((4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2\right)$$

$$= 14$$

$$\mathrm{Cov}(X_1, X_2) = \frac{1}{N-1}\sum_{k=1}^{N}(X_{1k} - \bar{X}_1)(X_{2k} - \bar{X}_2)$$

$$= \tfrac{1}{3}((4-8)(11-8.5) + (8-8)(4-8.5)$$

$$+ (13-8)(5-8.5) + (7-8)(14-8.5)$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$N = 4$

$$\boxed{\overline{X_1} = 8 \quad \overline{X_2} = 8.5}$$

Principal Component Analysis Solved Example 1

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\begin{aligned} \mathrm{Cov}(X_2, X_1) &= \mathrm{Cov}(X_1, X_2) \\ &= -11 \end{aligned}$$

$$\mathrm{Cov}(X_2, X_2) = \frac{1}{N-1} \sum_{k=1}^{N} (X_{2k} - \bar{X}_2)(X_{2k} - \bar{X}_2)$$

$$= \frac{1}{3}\left((11 - 8.5)^2 + (4 - 8.5)^2 + (5 - 8.5)^2 + (14 - 8.5)^2\right)$$

$$= 23$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\boxed{\begin{aligned} \overline{X_1} &= 8 \\ \overline{X_2} &= 8.5 \end{aligned}}$$

Principal Component Analysis Solved Example 1

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$= \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = \mathbf{8}$$

$$\overline{X_2} = \mathbf{8.5}$$

Principal Component Analysis Solved Example 1

Step 3: Eigenvalues of the covariance matrix

The characteristic equation of the covariance matrix is,

Determinant

$$0 = \det(S - \lambda I)$$

$$= \begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix}$$

$$= (14-\lambda)(23-\lambda) - (-11)\times(-11)$$

$$= \lambda^2 - 37\lambda + 201$$

$\Rightarrow$

$$0 = \lambda^2 - 37\lambda + 201$$

I is the identity matrix

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

mean

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

Cov $\rightarrow$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Principal Component Analysis Solved Example 1

$\lambda^2 - 37\lambda + 201 = 0$ Quadratic equation

- ⇨ a = 1, b = -37, c = 201
- ⇨ $b^2 - 4ac = (-37)^2 - 4(1)(201)$
- ⇨ 1369 – 804
- ⇨ 565

Assume, $X = -\frac{b}{2a} \pm \frac{\sqrt{b^2-4ac}}{2a}$

So $\lambda = -\frac{-37}{2} \pm \frac{\sqrt{565}}{2}$

- ⇨ $\lambda = \frac{37}{2} \pm \frac{\sqrt{565}}{2}$
- ⇨ $\lambda = \frac{1}{2}(37 \pm \sqrt{565})$
- ⇨ $\lambda 1 = 30.3849, \ \lambda 2 = 6.6151$

Roots of quadratic equation

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

mean
$$\overline{X_1} = 8$$
$$\overline{X_2} = 8.5$$

Cov →
$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

eigen values
$$\lambda_1 = 30.3849$$
$$\lambda_2 = 6.6151$$

Principal Component Analysis Solved Example 1

Step 4: Computation of the eigenvectors

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = (S - \lambda\ I)\ U$$

$$= \begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

We consider the first equation

$$= \begin{bmatrix} (14-\lambda\ )u_1 - 11u_2 \\ -11u_1 + (23-\lambda\ )u_2 \end{bmatrix} \begin{matrix} 1 \\ 2 \end{matrix}$$

$$(14-\lambda\ )u_1 - 11u_2 = 0$$

$$-11u_1 + (23-\lambda\ )u_2 = 0$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

mean

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

Cov

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

U is the eigen vector
I is the identity matrix
S is the covariance

Principal Component Analysis Solved Example 1

$$(14-\lambda)u_1 - 11u_2 = 0$$

$$-11u_1 + (23-\lambda)u_2 = 0$$

$$\frac{u_1}{11} = \frac{u_2}{14-\lambda} = t$$

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\frac{u_1}{11} = \frac{u_2}{14-\lambda} = t$$

$$u_1 = 11t, \quad u_2 = (14-\lambda)t$$

$$U_1 = \begin{bmatrix} 11 \\ 14-\lambda \end{bmatrix}$$

Asume

$t=1$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Principal Component Analysis Solved Example 1

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Step 4: Computation of the eigenvectors

- To find the first principal components, we need only compute the eigenvector corresponding to the largest eigenvalue.
- In the present example, the largest eigenvalue is λ_1.
- So, we compute the eigenvector corresponding to λ_1.

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

We need only one principal componant then we use λ1, if you need another principal componant than you should use λ2

Principal Component Analysis Solved Example 1

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}.$$

$\Rightarrow U = \begin{bmatrix} 11 \\ 14 - \lambda \end{bmatrix}$

- To find a unit eigenvector, we compute the length of U_1 which is given by,

$$||U_1|| = \sqrt{11^2 + (14 - \lambda_1)^2}$$
$$= \sqrt{11^2 + (14 - 30.3849)^2}$$
$$= 19.7348$$

Length of the Unit eigen vector U1

$$e_1 = \begin{bmatrix} 11/||U_1|| \\ (14 - \lambda_1)/||U_1|| \end{bmatrix}$$
$$= \begin{bmatrix} 11/19.7348 \\ (14 - 30.3849)/19.7348 \end{bmatrix}$$
$$= \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

$$\boxed{e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}}$$

Principal Component Analysis Solved Example 1

Step 5: Computation of first principal components

$$e_1^T \begin{bmatrix} X_{1k} - \bar{X}_1 \\ X_{2k} - \bar{X}_2 \end{bmatrix} = [0.5574 \quad -0.8303] \begin{bmatrix} X_{11} - \bar{X}_1 \\ X_{21} - \bar{X}_2 \end{bmatrix}$$

$$= 0.5574(X_{11} - \bar{X}_1) - 0.8303(X_{21} - \bar{X}_2)$$

$$= 0.5574(4 - 8) - 0.8303(11 - 8.5)$$

$$= -4.30535$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Principal Component Analysis Solved Example 1

Step 5: Computation of first principal components

Feature	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14
First Principle Components	-4.3052	3.7361	5.6928	-5.1238

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Principal Component Analysis Solved Example 1