Principal Component Analysis

Mannem Charan AI21BTECH11019

Abstract

This report consists of my basic understanding of one of the popular feature extraction techniques ”Principal Component Analysis”.

1 Principal Component Analysis

PCA, Principal Component Analysis is an unsupervised learning algorithm used for dimenstionality reduction.It helps in removing the unwanted features in the dataset by using some mathematical methods.It is one of the popular methods in exploratory data analysis and predictive modelling.Here we try to convert correlated features into a set of linearly uncorrelated features.The task of PCA is to find the features/Principal Components which better represents the dataset.

2 Understanding PCA

In PCA, we try to construct new features(Principal Components) from the original features and we will retain features which have high variance.Reducing the dimensionality of the model often comes at the accuracy of model, so we will try to remove features which has less effect on accuracy.Basically it includes following steps,

1) Standardization of the data.
2) Constructing the covariance matrix.
3) Computing Eigen vectors and Eigen values of covariance matrix to produce Principal components.
4) Creating a new feature vector from the principal components.
5) Update the data with new feature vector

2.1 Standardization of the data

In this step we will try to scale down all the features into the same region so that each attribute contributes equally in the analysis with retaining it’s characteristics.So mathematically it is done by subtracting each value in an attribute with the mean of attribute and dividing it with standard deviation of the attribute.

$$x := \frac{x-\mu}{\sigma} \tag{2.1}$$

After this step all the features will be transformed to same scale.(Note : After doing standardization, the mea

2.2 Constructing the covariance matrix

To understand this step, first we will try to understand what is covariance of two variables.Covariance is statistics concept which gives an idea of how two variables are related to each other in the data given. Let say x and y are two random variables (Attributes in this context),then the covariance will be

$$Cov\,(x, y) = E\left\{(x - \mu_x)\left(y - \mu_y\right)\right\} \tag{2.2}$$

where,

- E is expectation operator
- μ_x is mean of random variable x.
- μ_y is mean of random variable y.

1) When the value of covariance is 0, it implies that both the variables x and y are independent.
2) When the value of covariance is > 0 , it implies that the variables are positively related means that when x increases y tend to increase.
3) And similarly When the value of covariance is < 0 , it implies that the variables are negatively related means that when x increases y tend to decrease and vice-versa.

So, in this step we will try to construct a covariance matrix C, which consists of covariance of every feature pair.It will be an kXk matrix with k as the no. of dimensions and it can be calculated as,

$$C = \frac{X^T X}{N} \tag{2.3}$$

where, X is the input matrix consisting of k columns/features and N rows/data points.

2.3 Computing Eigen Vectors and Eigen values of covariance matrix to produce Principal components

Now to understand why we need to find eigen vectors of covariance matrix, first we will try to see what actually are Principal Components.

Principal Components : Principal Components are explicitly nothing but linear combination of orginal feature set. These are sorted in the sense of how much information they are retaining.These principal components are independent/uncorrelated with each other and meant to have as much as information possible in the early members so that we can neglect the later members.And these new variables doesn't have any real meaning since they are constructed as linear combination of features.

Mathematically speaking, the principal components are the lines in k-dimensional space which captures most information of the data and these lines are orthogonal to each other.First we will understand what we mean by **capturing the information**,for that let us take a data point x_1 and unit vector u_1 .Now if we take inner product of these two

$x_1^T u_1$, this will be the projection of x_1 on u_1.And this projection is the measure of how much the vector u_1 is capable of capturing information about x_1. This will be 0 if both are orthogonal and maximum when both are collinear.
So as we are saying that we want capture as much as information as possible, we will maximize the average sum of **squares** of projections of each data point on this unit vector u_1 i.e.,

$$\underset{u_1^T u_1 = 1}{\arg\max} \frac{1}{N} \sum_{i=1}^{N} \left(x_i^T u_1 \right)^2 \tag{2.4}$$

Now if we solve the objective function further we will get,

$$\frac{1}{N} \sum_{i=1}^{N} \left(x_i^T u_1 \right)^2 = \frac{1}{N} \sum_{i=1}^{N} x_i^T u_1 x_i^T u_1 \tag{2.5}$$

$$= \frac{1}{N} \sum_{i=1}^{N} u_1^T x_i x_i^T u_1 \tag{2.6}$$

$$= u_1^T \frac{1}{N} \sum_{i=1}^{N} x_i x_i^T u_1 \tag{2.7}$$

$$= u_1^T C u_1 \tag{2.8}$$

Now using lagrange multipliers we will introduce new variable λ_1, so the objective function now becomes,

$$J = u_1^T C u_1 - \lambda_1 \left(u_1^T u_1 - 1 \right) \tag{2.9}$$

Now the gradient of above objective function will be

$$\delta_{u_1} J = \left(C + C^T \right) u_1 - 2\lambda_1 u_1 \tag{2.10}$$

$$= 2Cu_1 - 2\lambda_1 u_1 = 0 \tag{2.11}$$

$$\implies Cu_1 = \lambda_1 u_1 \tag{2.12}$$

The above equation is satisfied when u_1 is Eigen vector of covariance matrix and λ_1 is Eigen value of the Eigen vector u_1.So in other words, we need to find eigen vectors of the covariance matrix.

2.4 Creating a new feature vector from the principal components

As we discussed in previous section, the eigen vectors will be the principal components and we will sort the eigen vectors w.r.t eigen values i.e., more the eigen value more will be the information it captures.So we will choose the principal components with more eigen values and this will be the new feature vector.

2.5 Update the data with new feature vector

In this step we just try to change the data according to the new feature set which we computed using eigen vectors of covariance matrix.So we will try to change the axes of orginal data to principal components.Mathematically the new data set can calculated as,

$$X' = NewFeatureVector^T X^T \tag{2.13}$$

3 Questions

1) What are principal components?
2) What covariance signifies?
3) What is the objective function in PCA?
4) What eigen values of principal components represents?
5) What are the applications of PCA?

4 Solutions

1) Principal components are the linear combination of orginal features and these new features are uncorrelated to each other.
2) The sign of Covariance of two variables gives the information about how the variables are related in the data.
   - If it is positive , both are positively related
   - If it is negative, both are negatively related
   - If it is 0, then both are independent
3) The objective function of PCA is

   $$J = u^T Cu - \lambda_1 \left(u^T u - 1\right) \tag{4.1}$$

   where,
   a) C is covariance matrix of input matrix
   b) u is the principal component
4) The eigen value represents how much the information is captured by that eigen vector/principal component.Using these eigen values we will sort the eigen vectors.
5) The Applications of PCA are,
   - It is used as a dimensionality reduction technique for various AI applications like computer vision,image compression etc.
   - It can be used to find hidden patterns in the data.Some fields where PCA is used are data mining,Finance,etc.