(DA-Lab)
MS. SHWETA TIWARI
April 26, 2022

DA-Lab
DATA ANALYTICS-LAB

—

By SHWETA TIWARI

Principal Component Analysis with R Programming

Principal component analysis(PCA) in R programming is an analysis of the linear components of all existing attributes. Principal components are linear combinations (orthogonal transformation) of the original predictor in the dataset. It is a useful technique for EDA(Exploratory data analysis) and allows you to better visualize the variations present in a dataset with many variables.

R – Principal Component Analysis

First principal component captures the maximum variance in dataset. It determines the direction of higher variability. **Second principal component** captures the remaining variance in data and is uncorrelated with PC1. The correlation between PC1 and PC2 should be zero. So, all succeeding principal components follow the same concept. They capture the remaining variance without being correlated to the previous principal component.

The Dataset

The dataset **mtcars**(motor trend car road test) comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles. It comes pre-installed with the dplyr package in R.

Principal Component Analysis with R language using dataset

We perform Principal component analysis on **mtcars** which consists of 32 car brands and 10 variables.