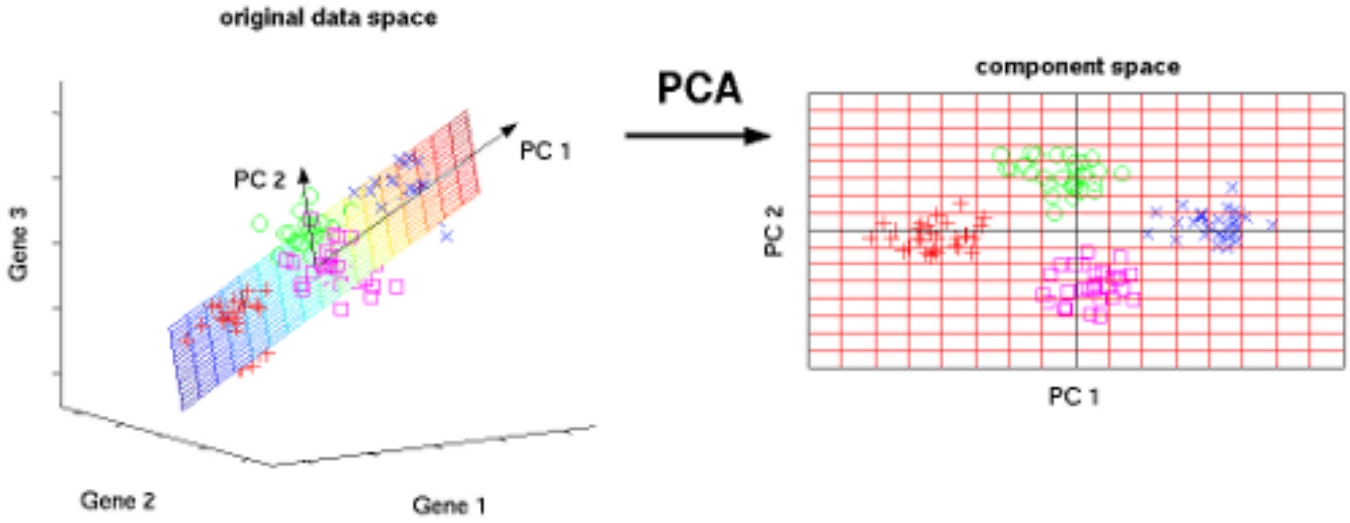
Principal Component Analysis

Epoch IIT Hyderabad

Arin Aggarwal
MA22BTECH11006

1 Introduction

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms a set of features in a dataset into a smaller number of features called principal components while at the same time trying to retain as much information in the original dataset as possible.



2 Working

1) The first step standardizing the range of the continuous initial variables so that each one of them contributes equally to the analysis. If there are large differences between the ranges of initial variables, those variables with larger ranges will dominate over those with small ranges.

$$Standardized\ Value = \frac{value - mean}{standard\ deviation}$$

2) Compute the Covariance Matrix between all features.

$$cov_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n}$$

where n = number of samples.

3) Eigenvectors and eigenvalues need to be computed from the covariance matrix in order to determine the principal components of the data. Principal components are new variables that are constructed as linear combinations or mixtures of the initial variables. These combinations are done in such a way

that the new variables are uncorrelated and most of the information within the initial variables is squeezed or compressed into the first components. There are as many PCs as there is dimensionality of data. Principal components are less interpretable and don't have any real meaning since they are constructed as linear combinations of the initial variables. Geometrically, principal components represent the directions of the data that explain a maximal amount of variance.

The first PC is the line with best fit for the data. The subsequent ones are the ones which are uncorrelated to PC1, i.e. they are orthogonal to each other.

Then the eigenvectors and eigenvalues are computed. The highest eigenvalue accounts for the highest variance. The ratio of variance explained by each value is determined the dividing the value by sum of eigenvalues.

4) Features with lower contribution may be discarded, hence reducing dimensionality of the data. Finally we are left with our **feature vectors**.

5) In the last step, we use the feature vector to get our final dataset. This is done as follows

$$Final\ Dataset\ =\ Feature\_Vector^T * Standardized\_Original\_Dataset^T$$