1 Dimensionality Reduction

The process of converting a set of data having large number of dimensions into data with smaller number of dimensions is called Dimensionality Reduction.

For example, in the below picture instead of working in three dimensional space we can move to two dimensional space and also it is comparatively easy to work.

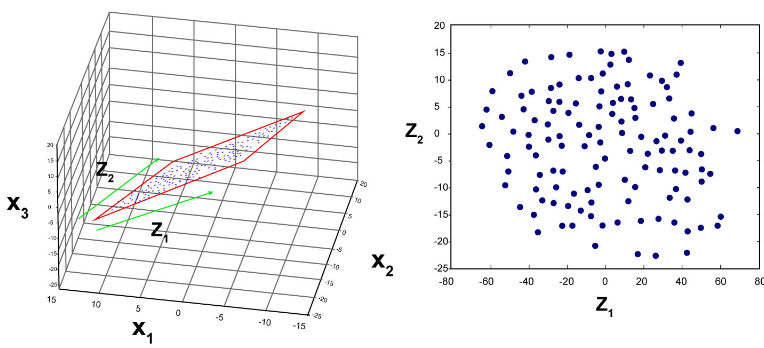

Figure 1

1.1 What are the benefits of Dimension Reduction?

- Compresses the Data and reduces the storage space required.
- Requires lesser computation time.
- Removes redundant features.
- Reduces the dimensions of data to 2D or 3D may allow us to plot and visualize it precisely.
- Potentially reduces the Noise.

1.2 Type of Dimensionality Reduction

The two types of dimensionality reduction are:

Feature Extraction: It finds new features in the data after it has been transformed from a high-dimensional space to a low dimensional space.

Feature Selection: It finds the most relevant features to a problem. This is done by obtaining a subset or key features of the original variables.

1.3 Selecting Features as a Matrix multiplication

For example, for selecting first and third feature we can do the following matrix multiplication:

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

The following matrix multiplication gives the linear combination of features which reduces the dimension:

$$\begin{bmatrix} x_1 \\ 3x_2 \\ 2x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$X^{'} = AX$$

A “new” set of features are selected/extracted from the original one by a matrix multiplication.

2 Principal Component Analysis (PCA):

PCA is a dimensionality reduction method. It is based on feature extraction. PCA refers to a very specific technique for finding a linear transformation (a matrix) that transforms a data set to a lower dimension. So, if the training data set has vectors (data points) of dimension N, the PCA transformation matrix will have dimension M by N and will transform each vector to a lower dimension M by a simple matrix multiply.

PCA is a variance-maximising technique that projects the original data onto a direction that maximizes variance. PCA performs a linear mapping of the original data to a lower-dimensional space such that the variance of the data in the low-dimensional representation is maximized.

2.1 How PCA Works

In mathematical terms, PCA is performed by carrying out the eigen-decomposition of the co-variance matrix[1]. The result would be a set of eigenvectors[2] and a set of eigenvalues[2] which can then be used to describe the original data.

2.2 Steps involved in PCA

PCA is performed by:

- constructing a co-variance matrix
- performing an eigen-decomposition of that matrix to obtain a set of eigenvectors (W)
- columns of W are ordered by the size of their corresponding eigenvalues
- choose the first n columns of W and use it to describe your data

References

[1]Covariance Matrix: https://en.wikipedia.org/wiki/Covariance_matrix
[2]Eigenvalues and eigenvectors:https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors