ITM 6285 Data Mining Lab - Principal Component Analysis

Principal Component Analysis, Eigen Values and Eigen Vectors, and Biplot (expected time - 1.5 hour)

**Learning objective:**

1. Understand the task and process of PCA

2. Understand the relationship between PCA and eigen value calculation

3. Calculate the coordinates in the principal component space

4. Generate and interpret the Biplot

**Task 1: Import the data and Background Information**

Import the dataset named “wine.csb”.

We measured 7 attribute of 5 wines (Hedonic, Meat, Dessert, Price, Sugar, Alcohol, and Acidity). One of the tasks is to find out the primary factors that distinguish between expensive wine and affordable wine.

Please use a *MS word document* as the answer sheet.

**Task 2: PCA using eigen value and eigen vector calculation-Data Preparation**

First, get rid of the wine name column, then *scale* the data, name the PCA-ready data set “my.wine”.

Copy the screenshot of my.wine to the answer sheet.

**Task 3: PCA using prcomp**

Do PCA using prcomp, show the R result and generate the screeplot, and then answer what is the primary contributor the PC2?

**Task 4: Calculate coordinates on principal components**

Calculate and copy the coordinates on the principal components to the answer sheet.