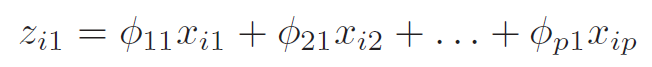
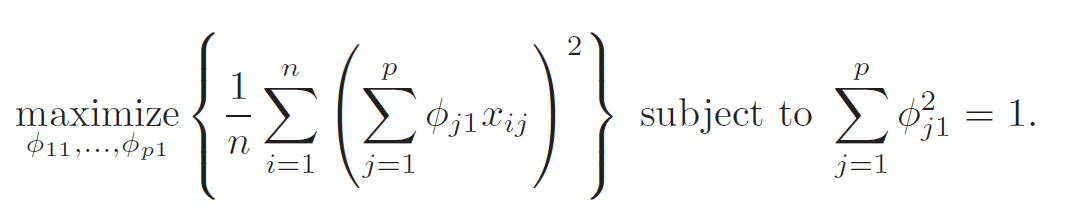


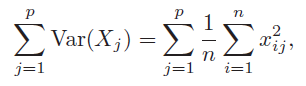
The linear combination of the sample feature values of the form



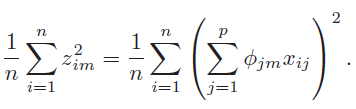
The method to calculate the first principal component loading vector solves the optimization problem



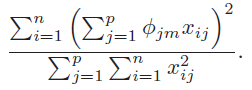
We interested in knowing the proportion of variance explained (PVE) by each principal component. The total variance present in a data set is defined as:

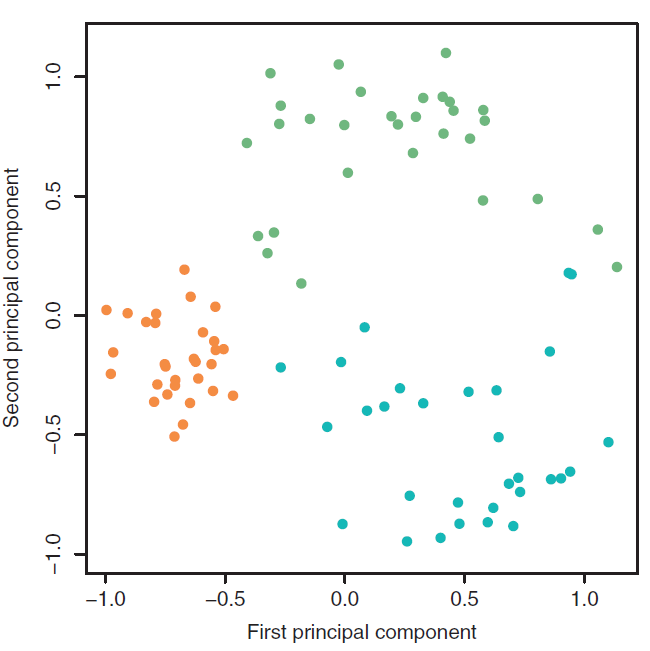
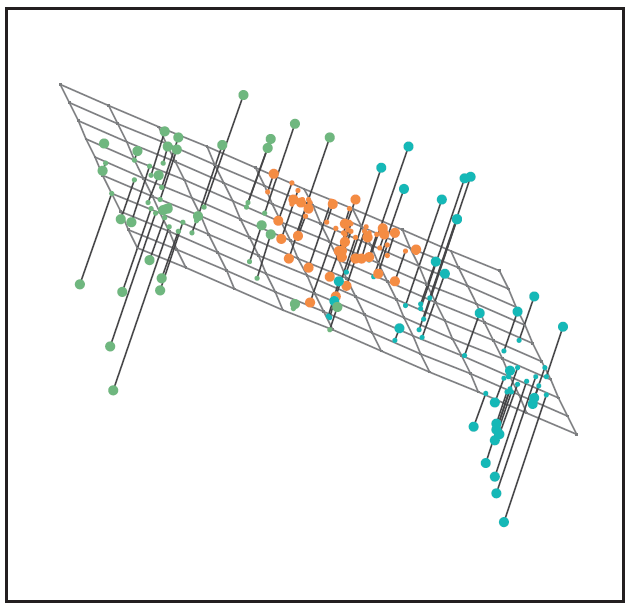


The variance explained by the mth principal component is:



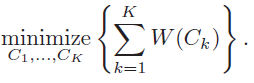
The PVE of the mth principal component is given by:



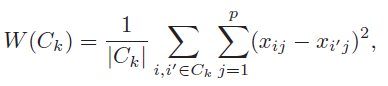


K-mean cluster:

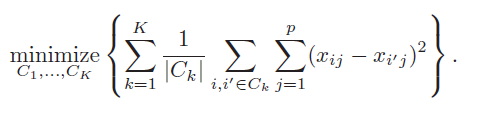
We want to partition the observations into K clusters such that the total within-cluster variation, summed over all K clusters, is as small as possible.

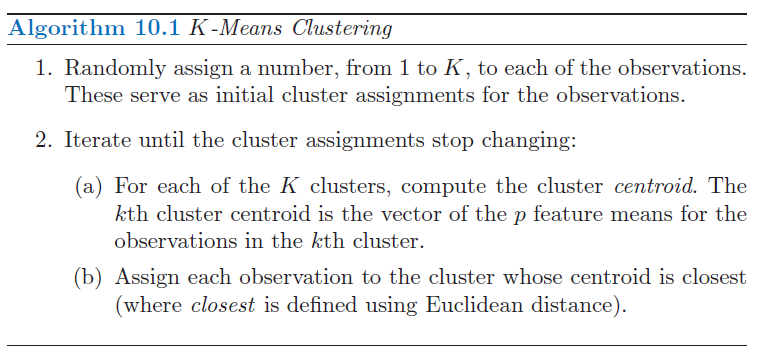


The formula below defines the within-cluster variation. The name of the formula is sqared Euclidean distance.

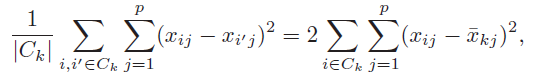


The goal of the methods is to find the minimized value of the within-cluster variation for the kth cluster.





The formula below illuminate why Algorithm 10.1 is guaranteed to decrease the value of the objective at each step



The algorithm will keep running until there is no change for the outcome of the formula above.

