S-k-means

Saptarshi Chakraborty and Swagatam Das

17 June 2021

Load The Libraries

```
library(igraph)
```

```
##
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```

```
library(rgl)
library(MASS)
library(mlbench)
```

```
## Warning: package 'mlbench' was built under R version 4.0.3
```

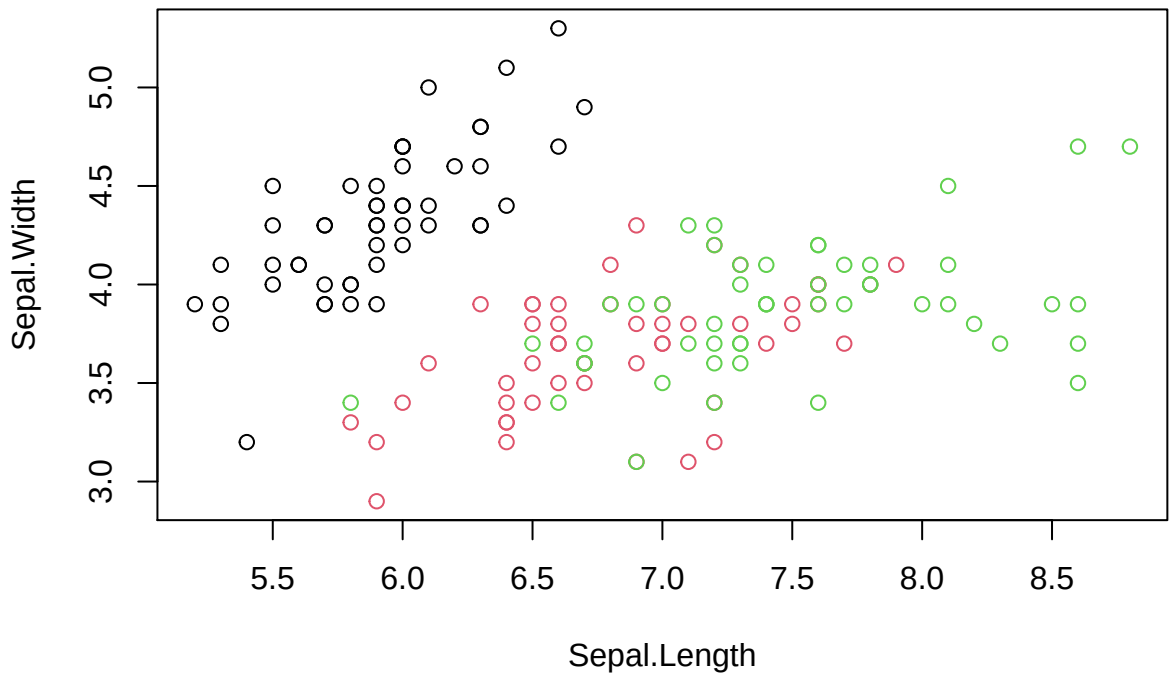
Load the Source Codes

```
source('functions.R')
```

Implementation on Iris Dataset

```
# Load the data
data(iris)
X=iris
X=data.matrix(X)

# Seperate out the class labels
toss=X[,5]
X=X[,-5]
n=dim(X)[1]
X=X-min(X)+1 # Makes every cooridinate positive
# Plot the data
plot(X,col=toss)
```




```
numclus=3 # Number of Clusters
sa=sample(n,numclus) # Sample the cluster centroids uniformly from the data
M=X[sa,] # Construct the initial centroid matrix
```

Run the S-k-means algorithm

```
l=s.kmeans(X,M,30) # Run the S-k-means algorithm with 30 iteartions
```

```
l
```

```
## $label
##   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3
##  [75] 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 2 2 2 2 2 3 2 2 2 2
## [112] 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 3 3 2 2 2 2 2 2 2 2 2 2 2 2 2
## [149] 2 2
##
## $centroids
##      Sepal.Length Sepal.Width Petal.Length Petal.Width
## [1,]     5.895693    4.311645     2.355714    1.141470
## [2,]     7.516836    3.906011     6.444279    2.946591
## [3,]     6.794072    3.625149     5.171628    2.217253
```

Compare between the ground truth and obtained cluster centroids

```
compare(l$label,toss,'adjusted.rand')
```

```
## [1] 0.885697
```