

IS 622 Week 9 HW

Ben Arancibia

October 22, 2015

7.4.1

Consider two clusters that are a circle and a surrounding ring, as in the running example of this section. Suppose:

- i. The radius of the circle is c .
- ii. The inner and outer circles forming the ring have radii i and o , respectively.
- iii. All representative points for the two clusters are on the boundaries of the clusters.
- iv. Representative points are moved 20% of the distance from their initial position toward the centroid of their cluster.
- v. Clusters are merged if, after repositioning, there are representative points from the two clusters at distance d or less.

In terms of d , c , i , and o , under what circumstances will the ring and circle be merged into a single cluster?

The two clusters will be merged into a single cluster when $i+o = c$.

In order for $i+o = c$ then i must be similar size to o .

In terms of d, c, i , and o the ring and circle will be merged into a single cluster when

$$c = i + (d)o.$$

7.5.1

Using the cluster representation of Section 7.5.1, represent the twelve points of Fig. 7.8 as a single cluster. Use parameter $k = 2$ as the number of close and distant points to be included in the representation. Hint: Since the distance is Euclidean, we can get the square of the distance between two points by taking the sum of the squares of the differences along the x - and y -axes.

```
points <- c(2,2,3,4,5,2,9,3,12,3,11,4,10,5,12,6,6,8,4,8,4,10,7,10)
points <- matrix(points, nrow=12, ncol=2, byrow=TRUE)
points <- as.data.frame(points)
```

points

```
##      V1 V2
## 1     2  2
## 2     3  4
## 3     5  2
## 4     9  3
## 5    12  3
## 6    11  4
## 7    10  5
## 8    12  6
## 9     6  8
## 10    4  8
## 11    4 10
## 12    7 10
```

```
colnames(points) <- c("x", "y")
```

```
(cl <- kmeans(points, 1))
```

```
## K-means clustering with 1 clusters of sizes 12
```

```
##
```

```
## Cluster means:
```

```
##      x      y
```

```
## 1 7.083333 5.416667
```

```
##
```

```
## Clustering vector:
```

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

```
##
```

```
## Within cluster sum of squares by cluster:
```

```
## [1] 237.8333
```

```
## (between_SS / total_SS =  -0.0 %)
```

```
##
```

```
## Available components:
```

```
##
```

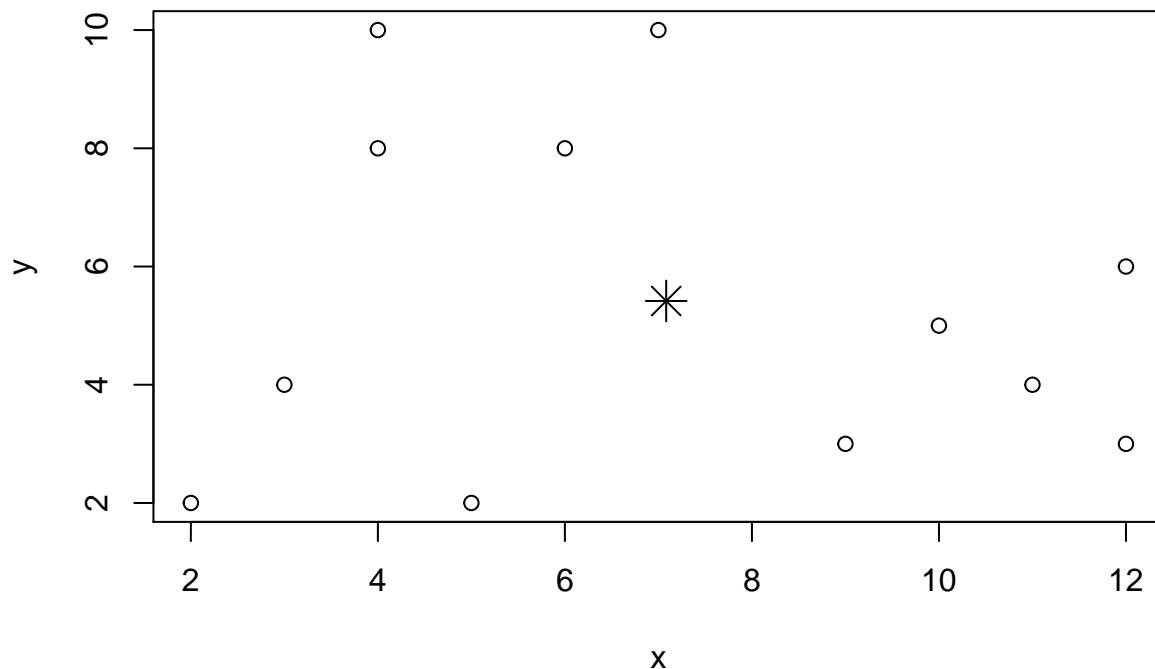
```
## [1] "cluster"      "centers"      "totss"      "withinss"
```

```
## [5] "tot.withinss" "betweenss"    "size"      "iter"
```

```
## [9] "ifault"
```

```
plot(points, col = cl$cluster)
```

```
points(cl$centers, col = 1:2, pch = 8, cex = 2)
```



```
cl$centers
```

```
##      x      y
```

```
## 1 7.083333 5.416667
```