

Segmentación de mercados: proceso

Jordi López Sintas

9 de enero de 2015

Modelos de segmentación: proceso

1. El procedimiento estándar

Fase exploratoria: la clasificación jerárquica

La clasificación jerárquica con el procedimiento de Ward minimiza la suma de las distancias euclidianas al cuadrado entre el individuo i y el centro del segmento al que se asigna, $\sum_k \sum_{i \in k} \sum_j = (x(i, j) - x(j', j))^2$. El primer sumatorio calcula el error de la observación i en el segmento k para todas las variables, el segundo realiza el cálculo para todos los individuos en el segmento k y, finalmente, el último sumatorio calcula la suma del error para todos los segmentos (Wishart 1998).

Funciona de la siguiente manera

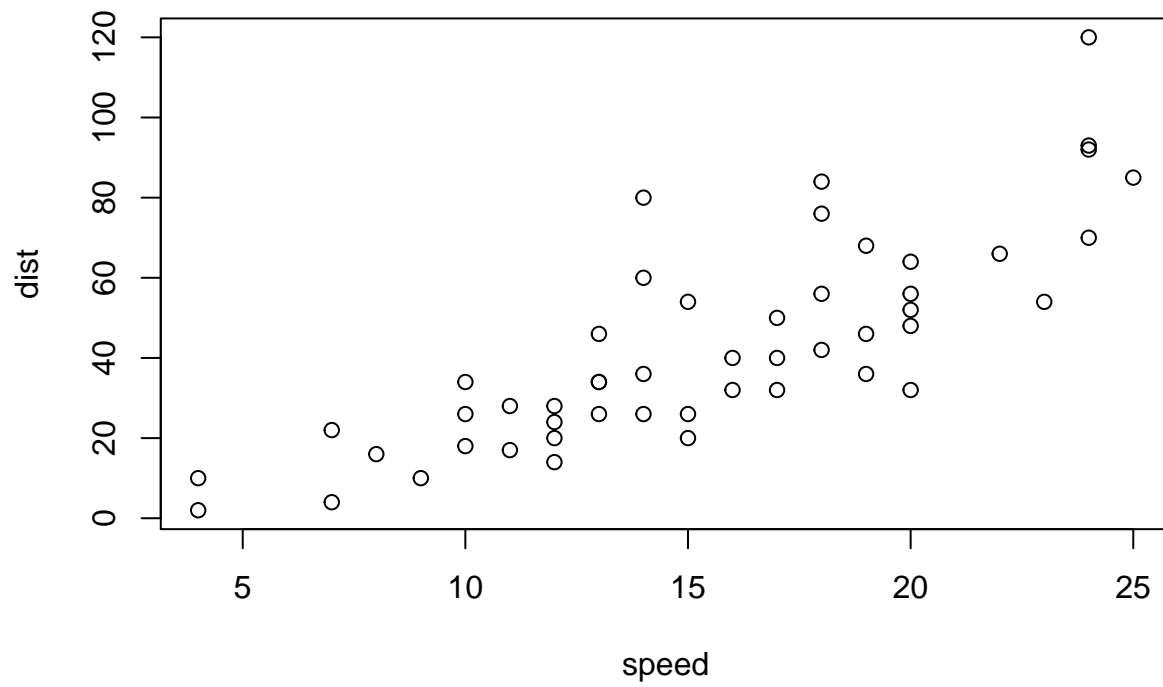
1. Computa la matriz de las distancias euclidianas al cuadrado para todo par de observaciones, i, i' , $d^2(i, i') = \sum_k (x(i, j) - x(j', j))^2$.
2. Agrupa las dos observaciones o grupos i y i' más cercanos, es decir, cuya agrupación minimiza el incremento en el error, E . Inicialmente serán aquellos casos cuya distancia, $d^2(i, i')$, sea mínima.
3. Transforma la matriz inicial de distancias euclidianas al cuadrado D^2 en E^2 , la cual contendrá el error al cuadrado de la unión del nuevo segmento $i \cup i'$.
4. Repite los pasos 2 y 3 y cada vez forma un nuevo grupo con las observaciones o grupos cuya unión resulte en un incremento mínimo en el error E^2 .
5. Finaliza cuando todos los casos se hayan agrupado en un solo segmento

El procedimiento minimiza el cuadrado de la suma de las desviaciones entre los individuos y el centro del grupo al que se le ha asignado. Realiza un proceso de aglomeración con tantas fases o etapas como individuos ha de clasificar, n , de manera que una vez clasificados en su totalidad se minimiza una medida de la heterogeneidad, $\min \sum_{SQD-I}$. En la etapa inicial existen tantos grupos como individuos debe agrupar. A partir de ese momento en cada etapa se formará un nuevo segmento agrupando a dos de los segmentos ya formados en etapas anteriores o un grupo y un individuo aun no clasificado, aquéllos más parecidos entre sí, de forma que se minimice en cada agrupación el incremento en la suma de las diferencias entre el individuo agrupado y la media del grupo al que se asigna, $\min \left\{ \sum_i x^2(i, j) \right\}$. En la etapa inicial la suma del cuadrado de las distancias es cero y se va incrementando a medida que se realizan las agrupaciones.

```
summary(cars)
```

```
##      speed      dist
## Min.   : 4.0    Min.   : 2.00
## 1st Qu.:12.0    1st Qu.: 26.00
## Median :15.0    Median : 36.00
## Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
## Max.   :25.0    Max.   :120.00
```

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.