

Clustering de ideologías usando NMF

Víctor Gallego y Roi Naveiro

01/04/2019

En este ejercicio plantearemos un problema de factorización de una matriz no negativa que representará el número de veces que cierta persona ha votado a cierto partido político en los últimos años.

1. Construcción de la matriz de datos. Estos son los votos de cada persona:

- Irene: 1 Podemos, 2 PSOE.
- Pepe: 2 Podemos, 1 PSOE, 1 Ciudadanos.
- Francisco: 6 Vox.
- Luisa: 1 Ciudadanos, 3 PP, 2 Vox.
- Cayetano: 4 Ciudadanos, 1 PP.
- Eva: 2 Podemos.
- Eugenia: 1 Podemos.
- Mario: 1 Vox

2. Factorizar la matrix \mathbf{X} para distintas dimensiones latentes

```
library(NMF)
```

```
d = 3
```

```
pol_nmf <- nmf(X, d)
```

```
fit(pol_nmf)
```

```
FALSE <Object of class:NMFstd>
```

```
FALSE features: 5
```

```
FALSE basis/rank: 3
```

```
FALSE samples: 8
```

```
# Factores
```

```
print("Matrix de factores")
```

```
FALSE [1] "Matrix de factores"
```

```
W <- basis(pol_nmf)
```

```
#dim(W)
```

```
print("Matrix de coeficientes")
```

```
FALSE [1] "Matrix de coeficientes"
```

```
H <- coef(pol_nmf)
```

```
print(H)
```

```
FALSE      irene      pepe      fran      luisa      cayetano
FALSE [1,] 2.220446e-16 2.220446e-16 5.232156e-01 1.744052e-01 2.220446e-16
FALSE [2,] 2.220446e-16 7.486676e-02 2.220446e-16 2.994670e-01 3.743338e-01
FALSE [3,] 1.855397e-01 1.855397e-01 2.220446e-16 2.220446e-16 2.220446e-16
FALSE      eva      eugenia      mario
FALSE [1,] 2.220446e-16 2.220446e-16 8.720260e-02
FALSE [2,] 2.220446e-16 2.220446e-16 2.220446e-16
FALSE [3,] 1.236931e-01 6.184655e-02 2.220446e-16
```

```
#dim(H)
```

3. Interpretar los clusters políticos.

4. Reconstruye la matriz y recomienda partidos a los que votar a las diferentes personas.

```
print("Matrix reconstruida")
```

```
FALSE [1] "Matrix reconstruida"
```

```
V.hat <- fitted(pol_nmf)
```

```
print(V.hat)
```

FALSE		irene	pepe	fran	luisa	cayetano
FALSE Ps		2.000000e+00	2.000000e+00	2.509677e-15	2.498721e-15	2.476619e-15
FALSE PSOE		1.000000e+00	1.000000e+00	1.312927e-15	1.301971e-15	1.279869e-15
FALSE Cs		1.820716e-15	6.000000e-01	1.895695e-15	2.400000e+00	3.000000e+00
FALSE PP		1.227544e-15	4.000000e-01	1.302523e-15	1.600000e+00	2.000000e+00
FALSE Vox		2.587505e-15	2.604129e-15	6.000000e+00	2.000000e+00	2.629426e-15
FALSE		eva	eugenia	mario		
FALSE Ps		1.333333e+00	6.666667e-01	2.412863e-15		
FALSE PSOE		6.666667e-01	3.333333e-01	1.216113e-15		
FALSE Cs		1.806984e-15	1.793251e-15	1.798881e-15		
FALSE PP		1.213811e-15	1.200078e-15	1.205708e-15		
FALSE Vox		2.573773e-15	2.560040e-15	1.000000e+00		