



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

## TAREA 5

### EJERCICIOS

1. Genera artificialmente datos de acuerdo con el modelo factorial

$$\mathbf{X} = \mathbf{AF} + \mathbf{L} + \mathbf{m}.$$

Emplea, por ejemplo, una matriz  $A$  de dimensiones  $15 \times 4$  —lo que supone que quince variables resultan generadas por cuatro factores ocultos—. Incluye variables con gran y pequeña comunalidad, y toma un tamaño de muestra suficiente como para estimar debidamente la matriz de covarianzas (por ejemplo,  $N \geq 300$ ). Emplea la función `factanal` para estimar el modelo y juzga los resultados. Prueba con diferente número de factores. ¿Es de esperar que la matriz  $\hat{A}$  que obtienes no coincida con la especificada por tí?

2. Haz un análisis factorial sobre `coches.dat`. Rota los factores. Compara con la solución obtenida mediante análisis en componentes principales. ¿Puedes aventurar interpretaciones para algunos de los factores?
3. Los datos sobre elecciones en `elec2001.dge` utilizados en la tarea precedente pueden verse como una tabla de contingencia: el tipo de datos para los que es útil el Análisis de Correspondencias. Lleva a cabo tal análisis, y representa en un plano de modo simultáneo filas (municipios) y columnas (partidos). Comenta lo que observes.
4. Sobre el resultado del análisis en componentes principales de las elecciones de 2.001 que llevaste a cabo en la tarea anterior, realiza un biplot.

### AYUDAS, SUGERENCIAS Y COMPLEMENTOS

- Te serán de utilidad las funciones `princomp` o `prcomp` y `factanal`. La función `prcomp` requiere como input una matriz. Puedes convertir una *data-frame* en matriz utilizando la función `as.matrix`.
- En la librería MASS de R hay las rutinas `corresp` y `mca` para hacer Análisis de Correspondencias. Puedes emplear también las funciones en el guión (que vimos en clase) `ac3.R`, escrita como ilustración. No te limites a emplearla como una caja negra: asegúrate de entender lo que haces.

- Observa que la función `factanal` incluye un contraste formal de diagonalidad de la matriz de correlación residual. Si el número de factores escogido es “adecuado”, la matriz “residual”  $\hat{\Sigma} - \hat{A}\hat{A}'$  debería ser aproximadamente diagonal. Contrastando esta hipótesis, tenemos un contraste indirecto de que el número de factores es apropiado (si se rechazara la hipótesis, tendríamos motivo para incrementar el número de factores comunes).
- Tienes una función `biplot` en R que admite como input el resultado de un análisis en componentes principales.  
En cualquier caso, podrías escribir tú una función que represente biplots con facilidad, tomando como modelo lo que se te facilitó en `ac3.R`.
- Tienes una pequeña bibliografía al final. Algunos libros, aunque muy antiguos, son excelentes: [8] es uno de ellos. Los libros [4],[11], [12] y [2] son manuales de Análisis Multivariante en general. El libro [3] es otro manual general con bastante atención al Análisis Factorial. Monografías útiles son [1], [10] y [9].  
Sobre Análisis de Correspondencias simple, en clase seguimos sobre todo a [3] y a [11]. Libros que te interesarán si quieres profundizar son [7], [6] y [5].  
Todo está en Biblioteca.

## Referencias

- [1] R. Cattell. *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. Plenum Press, 1978.
- [2] C. Chatfield and A.J. Collins. *Introduction to Multivariate Analysis*. Chapman & Hall, London, 1980.
- [3] C.M. Cuadras. *Métodos de Análisis Multivariante*. Eunibar, Barcelona, 1981.
- [4] W.R. Dillon and M. Goldstein. *Multivariate Analysis: Methods and Applications*. Wiley, New York, 1984.
- [5] B. Escofier and J. Pages. *Análisis Factoriales Simples y Múltiples. Objetivos, Métodos e Interpretación*. Servicio Editorial de la UPV/EHU, Bilbao, 1984.
- [6] M. Greenacre. *Theory and Applications of Correspondence Analysis*. Academic Press, New York, 1984.
- [7] M. Greenacre. *Correspondence Analysis in Practice*. Academic Press, New York, 1993.
- [8] H.H. Harman. *Modern Factor Analysis*. The Univ. of Chicago Press, 1960. Hay traducción española de la tercera edición.
- [9] R.A. Johnson and D.W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, 1992.
- [10] S.A. Mulaik. *The Foundations of Factor Analysis*. McGraw-Hill, 1972.
- [11] D. Peña. *Análisis de Datos Multivariantes*. McGraw-Hill, 2002.
- [12] G.A.F. Seber. *Multivariate Observations*. Wiley, New York, 1984.