

Clustering de Datos Funcionales: revisión y avances recientes

Manuel Oviedo de la Fuente, Universidade da Coruña, CITIC, España

Manuel Febrero Bande, Universidade de Santiago de Compostela, España

Palabras clave: clustering, datos funcionales, clasificación no supervisada

Abstract

El clustering de datos funcionales tiene como objetivo identificar patrones heterogéneos dentro de funciones continuas, como curvas, imágenes y superficies. Su aplicación abarca muchos campos científicos, como la biología, la química, la ingeniería, las ciencias ambientales, la medicina, la psicología y las ciencias sociales. El notable crecimiento en la aplicación del clustering de datos funcionales subraya la necesidad de un enfoque sistemático para desarrollar métodos de clustering eficientes y de fácil uso.

En este estudio, se revisan los métodos de agrupación de datos funcionales como la agrupación jerárquica, DBSCAN, k-means y cambio medio. Estos procedimientos serán probados en escenarios con datos simulados, donde se analizará la selección de los principales parámetros de interés (como el número de clústers), y se evaluará la calidad de las agrupaciones mediante medidas como el índice de Silueta o el índice de Dunn. Por último, se presentarán ejemplos de uso con datos reales: i) curvas horarias de generación en el mercado eléctrico español, y ii) análisis de patrones de movimiento de los ciudadanos en redes de transporte público (como el metro), con el objetivo de responder a consultas sobre la carga media de viajeros en una parada, en función del tipo de día y franja horaria. Para estos análisis, se utilizará el paquete de R fda.usc.

Referencias

Jacques, J., Preda, C. Functional data clustering: a survey. *Adv Data Anal Classif* 8, 231–255 (2014). [10.1007/s11634-013-0158-y](https://doi.org/10.1007/s11634-013-0158-y)

Febrero-Bande M, Oviedo de la Fuente M (2012). Statistical Computing in Functional Data Analysis: The R Package fda.usc. *Journal of Statistical Software*, 51(4), 1-28. <https://www.jstatsoft.org/v51/i04/>

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2023). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.6, <https://CRAN.R-project.org/package=cluster>

Zhang, M., & Parnell, A. (2023). Review of clustering methods for functional data. *ACM Transactions on Knowledge Discovery from Data*, 17(7), 1-34. doi.org/10.1145/3581789