

Propuesta de TFG (2017-18)

José Ramón Berrendero

Métodos de análisis multivariante basados en núcleos

Cualquier método estadístico que solo dependa de los datos a través de los productos escalares entre ellos admite una generalización no lineal consistente en sustituir los productos escalares por un núcleo, es decir, una función simétrica y semidefinida positiva. Este tipo de generalización se conoce con el nombre de *truco del núcleo*.

El objetivo de la primera parte de este trabajo es estudiar los resultados matemáticos básicos que se aplican para justificar el truco del núcleo y, en especial, el teorema de Moore-Arosznajn. Una buena referencia para esta parte es el capítulo 2 de Hsing y Eubank (2015).

En la segunda parte el objetivo es estudiar algunas aplicaciones del truco del núcleo. Tal vez la aplicación más conocida son los métodos de clasificación supervisada conocidos como *máquinas de vectores soporte*. Otra aplicación importante es la extensión no lineal del *análisis de componentes principales*, un método muy conocido de reducción de la dimensión. En función del tiempo disponible y de los intereses del estudiante se podrían considerar otros ejemplos.

Una revisión de las aplicaciones más conocidas del truco del núcleo se puede encontrar, por ejemplo, en Hofmann et al. (2008).

Para realizar este trabajo es muy conveniente haber cursado (o estar cursando de manera simultánea a la realización del trabajo) la asignatura *Estadística II*.

Referencias

Hofmann, T., Schölkopf, B. y Smola, A. J. (2008). Kernel methods in machine learning. *Ann. Statist.*, **26**, 1171-1220.

Hsing, T. y Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. John Wiley & Sons.