



# CONGRESO DE JÓVENES INVESTIGADORES

Real Sociedad Matemática Española

Universidad de Murcia, del 7 al 11 de Septiembre de 2015

---

## Detección de Outliers Multivariantes Usando Proyecciones Aleatorias

Paula Navarro Esteban<sup>1</sup>, Juan Antonio Cuesta Albertos<sup>1</sup>, Alicia Nieto Reyes<sup>1</sup>

En el análisis de conjuntos de datos, es importante la detección de outliers o datos atípicos. Existen múltiples métodos para detectar outliers en datos multivariantes, no obstante la mayoría de ellos requieren estimar la matriz de covarianzas. Cuanto mayor es la dimensión, más difícil es la estimación de la matriz debido a la dispersión en dimensiones grandes. Para evitar estimar dicha matriz, se propone un método para detectar outliers en datos normales multivariantes basado en las proyecciones aleatorias. Éste consiste en proyectar los datos en subespacios unidimensionales donde se aplica un método de detección univariante adecuado. Además, se presenta un método para calcular el número de proyecciones necesarias para conseguir determinados niveles de efectividad. Para ilustrar el procedimiento, se estudian conjuntos de datos simulados.

<sup>1</sup>Departamento de Matemáticas, Estadística y Computación  
Universidad de Cantabria, España  
Avda. Los Castros s/n, 39005 Santander (Cantabria)  
paula.navarro@unican.es  
juan.cuesta@unican.es  
alicia.nieto@unican.es