

SEMINARIO

Lucía Trapote Reglero

Universidad de Valladolid

Extending TCLUS T to higher dimensions

Abstract: Outliers distort traditional clustering, leading to unreliable partitions. Robust methods like TCLUS T address this by extending trimming to multi-cluster settings. While effective in low dimensions, TCLUS T struggles in moderate- to high-dimensional spaces due to parameter estimation complexity. Robust Linear Grouping (RLG) offers an alternative by assuming clusters lie near lower-dimensional subspaces, yet it fails when subspaces intersect or errors are non-isotropic. We propose a robust method extending TCLUS T by integrating the High Dimensional Data Clustering (HDDC) framework, incorporating trimming and eigenvalue constraints. This approach bridges TCLUS T and RLG through a careful adaptation of implementation steps. We present its theoretical properties, a feasible algorithm, and a strategy for selecting input parameters. In addition, we briefly discuss an extension to mixtures of t -distributions to handle data from heavy-tailed distributions. The methodology's performance is demonstrated via a simulation study, proving its effectiveness in complex, moderate-dimensional scenarios.

Joint work with Luis Ángel García Escudero and Agustín Mayo Íscar.

- [1] L. A. García-Escudero, A. Gordaliza, R. San Martín, S. Van Aelst y R. Zamar (2009), Robust linear clustering, *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 71(1), 301-318. DOI:10.1111/j.1467-9868.2008.00682.
- [2] C. Bouveyron, S. Girard y C. Schmid (2007), High-dimensional data clustering, *Computational Statistics & Data Analysis*, 52(1), 502-519. DOI:10.1016/j.csda.2007.02.009.
- [3] L. A. García-Escudero, A. Gordaliza, C. Matrán y A. Mayo-Íscar (2008), A general trimming approach to robust cluster analysis, *Annals of Statistics*, 36(3), 1324-1345. DOI:10.1214/07-AOS515.

**Seminario IMUVa, Edificio LUCIA
Lunes 6 de Julio de 2026 (11:00)**

