Robust sparse PCA for spatial data

Patricia Puchhammer[1], Ines Wilms[2], Peter Filzmoser[1]

[1] TU Wien - Institute of Statistics and Mathematical Methods in Economics, Wiedner Hauptstr. 8, Vienna, Austria;
e-mail: patricia.puchhammer@tuwien.ac.at

[2] Maastricht University - Institute of Statistics and Mathematical Methods in Economics, Wiedner Hauptstr. 8, Vienna, Austria;
e-mail: peter.filzmoser@tuwien.ac.at

Our goal is to introduce a robust PCA (Principal Component Analysis) method which takes into account spatial dependence among the observations. Specifically, we want to use robust spatial covariance estimators like the ssMRCD estimator [Puchhammer & Filzmoser, 2023] as basis for PCA. The ssMRCD estimator provides N many robust covariance matrix estimates for each part of a spatial partition using additional smoothing across space to properly address the spatial context. For a set of covariance matrices Σ_i, for $i = 1, \ldots, N$, we will also get a set of loadings representing the same variables leading to N many p-dimensional loadings. To simplify the interpretation and visualization of the loadings, we revert to sparsity in the most common way, i.e., the L1 norm. Moreover, also group-wise sparsity is included, where the groups consist of the loadings of one variable across spatial units, respectively. We start by defining the objective function for the loadings with explained variance penalized with entry-wise and group-wise sparsity. For the k-th PC we also need additional orthogonality constraints.

To solve this non-convex non-separable optimization problem, we develop an algorithm based on the Alternating Direction Method of Multipliers [Boyd et al., 1991]. We further illustrate the usefulness of the method in data examples and simulations and provide advanced visualization techniques for spatial PCA.

References:

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends® in Machine learning, 3(1), 1–122.

Puchhammer, P., & Filzmoser, P. (2023). Spatially smoothed robust covariance estimation for local outlier detection. Journal of Computational and Graphical Statistics,1-13.

Keywords: *Spatially Smoothed Mrcd Estimator, Robustness*