

# Visualization and Classification via Contiguity Analysis

Ludovic Lebart

CNRS-ENST, 46 rue Barrault, 75013, Paris, France  
lebart@enst.fr

**Keywords:** Contiguity Analysis, Unfolding, Local and Partial Correlation, Classification, Principal Component Analysis

## Abstract

In a first part, we consider the case of a set of multivariate observations, ( $n$  objects described by  $p$  variables, leading to a  $(n, p)$  matrix  $\mathbf{X}$ ), having an *a priori* graph structure. The  $n$  observations are the vertices of a symmetric graph  $G$ , whose associated  $(n, n)$  matrix is  $\mathbf{M}$  ( $m_{ii'} = 1$  if vertices  $i$  and  $i'$  are joined by an edge,  $m_{ii'} = 0$  otherwise). Such situation occurs when vertices represent time-points, geographic areas, or the vertices of *k-nearest neighbours graph* derived from a distance between observations. Contiguity Analysis, confronting local and global variances, provides a straightforward generalization of Linear Discriminant Analysis. It enables to point out the levels responsible of the observed patterns (*local* versus *global* level).

The second part is devoted to the situation in which  $\mathbf{M}$  and the graph structure are not external, but derived from the data matrix  $\mathbf{X}$  itself. Some interesting possibilities of exploration of data are sketched. The idea of deriving a metric likely to highlight the existence of clusters dates back to the works of Art et al. (1982) and Gnanadesikan et al. (1982). We present here the contribution of contiguity analysis to such approaches.

The third part deals with external graphs built from a set of instrumental variables. In that context, contiguity analysis provides a powerful tool for a non-parametric analysis of partial correlations.

## References

- Art D., Gnanadesikan R., Kettenring J.R. (1982) Data Based Metrics for Cluster Analysis, *Utilitas Mathematica*, **21** A, 75-99.
- Burtschy B., Lebart L. (1991) Contiguity analysis and projection pursuit. In : *Applied Stochastic Models and Data Analysis*, R. Gutierrez and M.J.M. Valderrama, Eds, World Scientific, Singapore, 117-128.
- Gnanadesikan R., Kettenring J.R., Landwehr J.M. (1982) Projection Plots for Displaying Clusters, in *Statistics and Probability, Essays in Honor of C.R. Rao*, G. Kallianpur, P.R. Krishnaiah, J.K.Ghosh, eds, North-Holland.
- Lebart, L. (2000): Contiguity Analysis and Classification, In: W. Gaul, O. Opitz and M. Schader (Eds):*Data Analysis*. Springer,Berlin, 233--244.