Euclidean representations of a set of hierarchies using Multiple Factor Analysis

Marine Cadoret[1*], Sébastien Lê[1], Jérôme Pagès[1]

1. Agrocampus Ouest, Laboratoire de mathématiques appliquées, Rennes (France)
* Contact author: marine.cadoret@agrocampus-ouest.fr

Keywords: Euclidean Representations, Hierarchical Sorting Task, Hierarchy, Multiple Factor Analysis

The aim of this presentation is to propose a new approach for analyzing hierarchies issued from unsupervised classifications performed on the same (statistical) individuals. This issue has already been partially addressed by several authors for the comparison of different classifications methods (Leclerc (1985), Leclerc and Cucumel (1987) or the special issue of Journal of Classification (Vol. 3, 1986) dedicated to the comparison and consensus of classifications).

The starting point of our research framework is the hierarchical sorting task commonly used in psychology and sensory analysis (Egoroff, 2005; Qannari et al., 2010). This method consists in asking subjects to provide each their own hierarchical tree from the same given set of objects. This hierarchical tree is constructed mostly in a binary and descending way: the subjects are asked to divide the objects into two homogeneous groups and then to divide again each of the two groups until they consider the final groups homogeneous. The main feature of this method is that each subject uses his/her own criteria for making these successive divisions. In this kind of experiment were interested into getting a consensus representation of the objects from all the subjects as well as a representation of the subjects, function of the way they classified the objects.

In this talk, we propose a methodology which provides on the one hand a Euclidean representation of the objects and on the other hand a Euclidean representation of the hierarchies (i.e. a subject can be assimilated to his/her hierarchy) linked to the previous one in the manner of Multiple Factor Analysis (MFA; Escofier and Pagès, 1998). This hierarchy representation allows visualizing the different steps taken by each subject and to understand in a certain way his/her cognitive process.

The data associated with a hierarchy j can be gathered in a data table with I rows and Q_j columns (with Q_j the number of levels associated with the hierarchy j). In this case, each level of the hierarchy can be assimilated to a qualitative variable with as many modalities as there are groups for this level.

The data coming from a set of hierarchies can be gathered in a table that juxtaposes the tables associated with each hierarchy. This data table is composed of I rows and $Q = \sum_j Q_j$ columns: each row corresponds to an object and each column to a level associated with a given hierarchy; the columns are grouped by hierarchy.

This kind of table only composed of qualitative variables structured in groups (one group = one hierarchy) can be analyzed by Multiple Factor Analysis: Escofier and Pagès (1998). MFA is applied to a data table in which the same set of individuals (here the objects) are described by several sets of variables (here the hierarchies) structured in groups. MFA balances the influence of each group of variables (i.e. each hierarchy) in the analysis making maximum axial inertia of the clouds associated with the separated analysis of each hierarchy equal to 1.

This methodology will be illustrated with an example in which 24 subjects performed a hierarchical sorting task on 16 advertisements concerning an orange juice.

References

Egoroff, C. (2005). *How to measure tactile perception: a case study on automotive fabrics,* ESN Conference (Madrid, Spain), 25-26 May.

Escofier, B., Pagès, J. (1998). *Analyses factorielles simples et multiples ; objectifs, méthodes et interprétation*, Dunod, Paris.

Leclerc, B.(1985). La comparaison des hiérarchies : indices et métriques. *Mathématiques et Sciences Humaines*, **92**, 5–40.

Leclerc, B., Cucumel, G.(1987). Consensus en classification : une revue bibliographique. *Mathématiques et Sciences Humaines*, **100**, 109–128.

Qannari, E.M., Courcoux, P., Taylor, Y., Buck, D., Greenhoff, K. (2010). *Statistical issues relating to hierarchical free sorting task,* 10th Sensometrics Conference (Rotterdam, the Netherlands), 25-28 July.