

# Analysis of two multiblocks tables: DO-ACT

Myrtille Vivien

Lab. de Probabilité et Statistiques

Dpt Sciences Mathématiques

Université Montpellier 2

France

vivien@math.univ-montp2.fr

Robert Sabatier

Lab. Physique Moléculaire et Structurale

Faculté de Pharmacie

Université Montpellier 1

France

sabatier@pharma.univ-montp1.fr

**Keywords:** Multiblocks analysis, multiway analysis, ACT (STATIS) method, RV coefficient, Tucker Inter-Battery method, Hilbert-Schmidt scalar product.

## Abstract

Many generalizations of standard linear multivariate analysis like Principal Component Analysis (PCA) or Canonical Correlation Analysis (CCA) have been proposed for study three or more sets of variables, that is to say a multiblock table. In this paper, a multiblock table (or a multiblock) is a set (or group) of matrices measured with the same observations. Most of the extensions of CCA determine the dimension (or rank) of the model step by step, using linear combinations of the variables of each matrix and optimizing a criterion. Only few methods optimize a global criterion and generate in a right way a solution with fixed rank. In the case of two multiblocks, some methods exist in the chemometric literature for building regression models. But if the problem is the simultaneous analysis of a pair of multiblock tables, only a few methods exist, but working only with two three-way arrays.

The purpose of this paper is to introduce a new approach for finding common dimensions inside two multiblocks tables with different length (different number of variables in each tables) and for describing each one of them. This global approach, referred as DO-ACT (DOuble ACT), is closely related to the STATIS (or ACT) strategy (Lavit *et al.*, 1994) and Tucker inter-battery method (Tucker, 1958) and is made of three successive steps: *interstructure*, *compromise*, *intrastructure*. The first is based on the optimization of a criterion using the Hilbert-Schmidt scalar product. During the second step, two compromises (one per multiblock) are defined and in the third step the compromise is analyzed and each table is compared to the compromise by mean of graphical representations of the observations and variables. The DO-ACT methodology is built to take into account the block structure (two multiblocks) of the data and can be useful to analyze large data sets where the measurements are organized with an important number of tables.

This new method will be described and an application on a real dataset will be made, using a S-Plus 6.0 program.

## References

Lavit, Ch., Escoufier, Y., Sabatier, R. & Traissac, P. (1994). The ACT (STATIS method). *Computational Statistics & Data Analysis*, 18, 997-119.

Tucker, L.R. (1958). An inter-battery method of factor analysis. *Psychometrika*, 23, 111-136.