

R Visual Tools for Three-way Data Analysis

Michele Gallo^a, Valentin Todorov^b and Maria Anna Di Palma^a

^aUniversity of Naples-L'Orientale, Naples, Italy, ^bUnited Nations Industrial Development Organization (UNIDO), Vienna, Austria

The standard multivariate analysis addresses data sets represented as two dimensional matrices. In recent years, an increasing number of application areas like chemometrics, computer vision, econometrics and social network analysis involve analysis of data sets that are represented as multidimensional arrays and multiway data analysis becomes popular as an exploratory analysis tool (see [1]). The most popular multiway models are CANDECOP/PARAFAC and TUCKER3 [3]. The results from a three-way analysis can be presented in several different ways (see [3]), the first one being tables of the coefficients or loadings for each mode, either rotated or not. While it is important to inspect the numerical output of the methods for analysis of three-way data (the component matrices and the core array) in order to properly interpret the results, of great help can be different visual representations of these outcomes. The most typical plots are: (i) Pair-wise graphs of the components for each mode separately, (ii) All-components plots which will show all components of a single mode using the levels of the mode as X-axis, (iii) Per-component plot, showing a single component on all modes simultaneously in the same plot, (iv) Joint biplots for Tucker 3 models and (v) Trajectory plots.

We present an R package, **rrcov3way**, implementing a set of functions for the analysis of multiway data sets, including PARAFAC and TUCKER3 as well as their robust alternatives. Apart from basic tools for data handling and preprocessing of multidimensional arrays, tools for display of raw data as well as the models in two and three dimensional plots are provided. Several examples based on the data sets available with the package are used in the presentation to demonstrate the basic usage of the functions and illustrate some of the graphical results obtainable with the software. These graphical procedures, mainly based on [2] and [3], are flexible enough to give the user the possibility to design the graphs according to the needs and the data at hand but at the same time provide suitable default parameters which facilitate their use.

Keywords: Three way, Visualization, R

References

- [1] Acar E, Yener B (2009). Unsupervised Multiway Data Analysis: A Literature Survey. *IEEE Trans. Knowl. Data Eng.*, **21**(1), 6–20.
- [2] Kiers HA (2000). Some procedures for displaying results from three-way methods. *Journal of chemometrics*, **14**(3), 151–170.
- [3] Kroonenberg PM (2008). *Applied multiway data analysis*. Wiley series in probability and statistics, John Wiley and Sons, Hoboken, NJ.