The *R* Package ThreeWay For Three-Way Component Analysis

Paolo Giordani[1,⋆], **Henk A.L. Kiers**[2], **Maria Antonietta Del Ferraro**[1]

1. Sapienza University of Rome
2. University of Groningen
⋆Contact author: paolo.giordani@uniroma1.it

Keywords: Multi-way data, Principal Component Analysis, Tucker3, Candecomp/Parafac

Data generally refer to the observations of some variables on a set of units and are stored in a (two-way) matrix, say $\mathbf{X}$ of order $(I \times J)$, where I and J denote the numbers of units and variables, respectively. However, in several situations, the available data consist of some variables collected on a set of units on different occasions and are usually stored in a three-way array, say $\underline{\mathbf{X}}$ of order $(I \times J \times K)$, where K denotes the number of occasions. The array can then be seen as a box in which the ways (or indices) correspond to the vertical, horizontal and depth axis. Multi-way data analysis concerns the cases in which the number of indices is higher than two (three-way data analysis when the number of indices is three). In this work we limit our attention to the three-way case. For more details on multi-way (and three-way) analysis, refer to, *e.g.*, [4, 6].

In order to summarize $\mathbf{X}$ classical Principal Component Analysis (PCA) can be applied. In the three-way framework, PCA is no longer a valuable choice. More specifically, PCA could still be applied for exploring $\underline{\mathbf{X}}$ either by rearranging it into a matrix (aggregating over one of the three ways) or analyzing all the two-way data matrices contained in the three-way array separately. Nonetheless, such strategies fail to discover the existing three-way interaction in the data and, therefore, produce incomplete or even misleading results. In the literature there exist several three-way extensions of PCA. The two most popular techniques are the Tucker3 (T3) method [7] and the Candecomp/Parafac (CP) method [1, 3].

The aim of this work is to illustrate the main features of the *R* [5] package **ThreeWay** [2] . **ThreeWay** offers a suit of functions for performing three-way component analysis. In particular, the most relevant functions are `T3` and `CP`, which implement, respectively, T3 and CP. These and other functions will be described through examples using data sets available in the package.

References

[1] Carroll, J. D. and J. J. Chang (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of eckart-young decomposition. *Psychometrika 35*, 283–319.

[2] Del Ferraro, M. A., H. A. L. Kiers, and P. Giordani (2013). **ThreeWay***: Three-Way Component Analysis*. *R* package version 1.1.1.

[3] Harshman, R. A. (1970). Foundations of the parafac procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics 16*, 1–84.

[4] Kroonenberg, P. M. (2008). *Applied Multiway Data Analysis*. Hoboken, NJ: John Wiley & Sons.

[5] *R* Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: *R* Foundation for Statistical Computing. ISBN 3-900051-07-0.

[6] Smilde, A., R. Bro, and P. Geladi (2004). *Multi-way Analysis: Applications in the Chemical Sciences*. Chichester, UK: John Wiley & Sons.

[7] Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika 31*, 279–311.