

A Robust Tucker3 Model for Compositional Data

V. Todorov^{1,*}, M. A. Di Palma² and M. Gallo²

¹ *United Nations Industrial Development Organization, Vienna, Austria; v.todorov@unido.org*

² *University of Naples-L'Orientale, Naples, Italy; madipalma@unior.it, mgallo@unior.it*

* *Corresponding author*

Abstract. Double counting is inherent to the output concept, therefore it is preferable to use manufacturing value added (MVA) instead to measure the manufacturing production. While the issue of double counting in production statistics is successfully addressed by using MVA, commodity exchange in trade data is still measured as output. The relevance of value added has increased in the recent years due to the unbundling of the production process, where different stages of value chain take place in different countries. We want to represent the export statistics through value added to output ratio using data from international statistical databases. The data sets considered are organized by country, commodity or activity and year (activities are classified according to the International Standard Industrial Classification of all economic activities (ISIC)) and thus they are *three-way compositional data*.

Different methods exist for analysis of multi-way data and we choose Tucker3 because it provides a compromise between parsimonious and flexible models. The Tucker3 method as most of the N-way methods is based on alternating least squares (ALS) which makes it vulnerable to the presence of outliers in the data. Even a single outlying data point can strongly influence the resulting model and the conclusions based on it. A robust version of Tucker3 was presented by Pravdova *et al.* (2001) but it suffers from two main deficiencies. First of all the robust initialization of the algorithm is based on MCD which will not work in high dimensions. And secondly, the method is not suitable for applying on compositional data. We propose to select the initial subset using robust PCA and to transform the compositional data applying *ilr* transformation (Egozcue *et al.*, 2003). Furthermore, since to our knowledge there is no readily available software for computing robust Tucker3 models, we provide implementation of the proposed algorithm in R. The method is compared to its competitors both in terms of its efficiency and the computational effort needed.

Keywords. *Multi-way; Robust; Outliers; Compositional data; Manufacturing value added*

References

- Egozcue, J & Pawlowsky-Glahn, V. & Mateu-Figueras, G. & Barcelo-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, **35**, 279–300.
- Pravdova, V. & Estainne, F. & Walczak, B. & Massart, D.L. (2001). A robust version of the Tucker3 model. *Chemometrics and Intelligent Laboratory Systems*, **59**, 75–88.