DSSV 2023 Data Science, Statistics & Visualisation
EC DA 2023 European Conference on Data Analysis

Exploring the measurement model in (high-dimensional) multigroup data: Regularized joint latent variable analysis

K. Van Deun[1] and T. Lê[1]

[1]Tilburg University, Department of Methodology & Statistics, Warandelaan 2, 5000LE Tilburg, Netherlands

Exploring multi-group data for similarities and differences in the measurement model is a substantial part of the research conducted in the behavioral and social sciences: It is relevant to studying measurement invariance of scales and group-differences in multivariate relations (e.g., biological pathways in psychopathology). Yet, currently available methods are restrictive in their use and do not scale up with the increasing complexity of current research paradigms. On the one hand, these methods cannot handle data with small sample sizes relative to the number of variables while such high-dimensional data (e.g., thousands of biomarkers for several dozen of patients or word counts obtained by text mining of tweets) are more and more used as a result of digitalization. On the other hand, users of software for exploratory multigroup methods might encounter issues such as the need to fix some parameters beforehand, a lack of convergence, and inconclusive results. Here, we propose a regularized latent variable method that addresses these issues by building on a strong computational framework: The resulting method yields solutions that are constrained to show simple structure and similarity of the loadings over groups when supported by the data. The minimal required input by the user is restricted to the data and number of latent variables. Interpretation is eased by exact zero loadings and, when measurement invariance holds, exactly equal loadings over groups. The (comparative) performance of the method is evaluated in a simulation study and illustrated on empirical data.