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

Factor analysis with variable selection via group L_0 penalty

N. Shimada[1] and M. Yamamoto[1,2]

[1]Osaka University, Graduate School of of Human Sciences, 1-2 Yamadaoka, Suita, Osaka 565-0871, Japan
[2]RIKEN AIP, Japan

Factor analysis with variable selection can be a valuable tool when the estimation is unstable due to irrelevant variables or a large number of observed variables. One such approach is the penalized maximum likelihood method with a weighted group lasso penalty, which applies a weighted penalty function to each variable in the factor loadings for variable selection[1]. However, this method has limitations in that it may underestimate the factor loadings due to the penalty. Also, the method assumes a normal distribution for the observed data, which can lead to bias in non-normal cases. To address these issues, we propose a novel method for variable selection in factor analysis that directly constrains the number of variables to be selected using the group L_0 norm. By directly constraining the loadings, our method can overcome the problem of underestimation of factor loadings. Additionally, our method is based on Matrix Decomposition Factor Analysis[2], which avoids making any assumptions about the underlying probability distributions in the parameter estimation. As a result, our method can produce less biased estimates than existing methods when analyzing data that follows a non-normal distribution, such as the t-distribution. We evaluated the performance of our proposed method through simulation by comparing it to existing methods.

Key words : Matrix Decomposition Factor Analysis (MDFA), variable selection , group $L0$ norm, factor analysis

[1] Hirose, K. and Konishi, S., "Variable selection via the weighted group lasso for factor analysis models," *The Canadian Journal of Statistics*, vol. 40, no. 2, pp. 345–361, 2012.
[2] Adachi, K. and Trendafilov, N. T., "Some mathematical properties of the matrix decomposition solution in factor analysis," *Psychometrika*, vol. 83, pp. 407–424, 2018.