



Correcting and clustering preference data in the presence of response-style bias

Mariko Takagishi¹; Michel van de Velden²

¹ Graduate School of Engineering Science, Osaka University

² Department of Econometrics, Erasmus University Rotterdam

Abstract:

Preference data are often obtained in questionnaire-based surveys. Clustering respondents based on content of survey item is useful for discovering latent structures of respondents. However, cluster analysis of preference data may be affected by response styles. Examples of response styles include extreme response style, in which respondents may tend to select ratings at the ends of the scale. A cluster of respondents with an extreme response style can be mistakenly identified as a content-based cluster. To address this problem, we propose a method of clustering respondents based on their indicated preferences for a set of items while correcting for response-style bias. We introduce I-spline basis function to model response style, in which only with non-negative and summation constraint, we can obtain monotone and intuitively reasonable response function estimator. Then we simultaneously correct for response styles and perform a cluster analysis based on the corrected preference data. In the end, our proposed method is applied to empirical survey data from four different Asian countries concerning social values.

Keywords:

Questionnaire-based surveys; Clustering; Response styles; Preference data; Splines