

Variable selection to construct indicators of quality of life for data structured in groups

Amaury Labenne, *IRSTEA*, amaury.labenne@irstea.fr
Marie Chavent, *University of Bordeaux, INRIA Bordeaux Sud-Ouest*,
Vanessa Kuentz-Simonet, *IRSTEA*,
Jérôme Saracco, *Bordeaux INP, INRIA Bordeaux Sud-Ouest*,

Abstract. The analysis and measurement of quality of life may be made via two complementary approaches. The first one, based on survey of individuals, concerns the analysis of levels of life satisfaction. We focus here on the second one, based on national data, which analyses living conditions of people. The aim is to create composite indices of living conditions. According to authors, the components of quality of life are related to different themes (groups of variables): “Family conditions”, “Employment”, “Housing”, . . . For this purpose, dimension reduction methods are particularly suitable. Multiple Factor Analysis (MFA) is a method designed to handle data structured into groups of quantitative variables. In our study, each theme is composed of a group of quantitative and/or categorical variables. Since our data are naturally structured in groups of variables, we develop an extension of MFA for mixed data type, called MFAmix. Thus the principal components from MFAmix are our composite indices for measuring quality of life. However, the creation of these indices raises two questions. How many principal components keep to create indices? How select a limited number of variables to get similar indices for easier interpretation? We propose answers to these questions in this communication.

Keywords. factorial analysis, dimension reduction, mixed data type, quality of life