Mixed Principal Components Analysis

Rafik Abdesselam

CREM UMR CNRS 6154 - University of Caen, Basse-Normandie
Esplanade de la paix, F-14032 Caen, France
rafik.abdesselam@unicaen.fr

Abstract. The processing of mixed data - quantitative and qualitative variables cannot be carry out directly by classical methods of data analysis. In this work, a factorial method which analyze simultaneously quantitative and qualitative data is described. The proposed Mixed Principal Components Analysis is a standardized principal component analysis of both quantitative variables and the transformation of the dummy variables associated to qualitative variables on quantitative variables through orthogonal projections of configurations of statistical units in the individual-space with a relational inner product. An example resulting from real data illustrates the results of this method, which are also compared with those of the Mixed Data Factorial Analysis proposed by Pagès (2004).

References

PAGÈS, J. (2004): Analyse factorielle de données mixtes. *Revue de Statistique Appliquée, LII(4), 93-111.*

SAPORTA, G. (1990): Simultaneous analysis of qualitative and quantitative data. Società italiana di statistica, *Atti della XXXV riunione scientifica*, 63–72.

SCHEKTMAN, Y. and ABDESSELAM, R. (2000): A Geometrical Relational Model for Data Analyses. In: W.Gaul, O.Opitz and M.Schader (Eds.): *Data Analysis*. Springer, Berlin, 359–368.

TENENHAUS, M. (1977): Analyse en composantes principales d'un ensemble de variables nominales ou numériques. *Revue de Statistique Appliquée, XXV(2), 39-56.*

YOUNG, F.W. (1981): Quantitative analysis of qualitative data. *Psychometrika, 46(4), 357-388.*

keywords

PRINCIPAL COMPONENTS ANALYSIS, MULTIVARIATE ANALYSIS OF VARIANCE, GENERALIZED CORRELATION RATIO, RELATIONAL INNER PRODUCT.