

Linear and Nonlinear Multiset Canonical Correlations Analysis

Klaus Baggesen HILGER, *Technical University of Denmark*

Allan Aasbjerg NIELSEN, Rasmus LARSEN, Knut CONRADSEN, *Technical University of Denmark*

This paper deals with decomposition of multiset data. Friedman's alternating conditional expectations (ACE) algorithm is extended to handle multiple sets of variables of different mixtures. The new algorithm finds estimates of the optimal transformations of the involved variables that maximize the sum of the pair-wise correlations over all sets. The new algorithm is termed multiset ACE (MACE) and can find multiple orthogonal eigensolutions. MACE is a generalization of the linear multiset correlations analysis (MCCA). It handles multivariate multisets of arbitrary mixtures of both continuous and categorical variables by applying only bivariate scatterplot smoothers for which the data analyst may specify appropriate restrictions when performing an exploratory analysis of the data.

[Klaus Baggesen HILGER, Technical University of Denmark, Informatics and Mathematical Modelling, 2800 Kgs. Lyngby, Denmark; kbh@imm.dtu.dk]