Editorial

Published online: 3 November 2010
© Springer-Verlag 2010

This last issue of volume 4 (2010) of the journal *Advances in Data Analysis and Classification (ADAC)* contains six scientific articles on data analysis methods. They deal with Procrustes analysis, robust classification, fuzzy clustering with cluster-specific GLM models, agreement measures for comparing nominal rankings, detecting atypical observations in finance data, and specification of outliers by mathematical programming techniques.

The introductory paper by *Casper J. Albers* and *John C. Gower* considers Procrustes analysis, i.e., the problem of transforming a set of given configurations (e.g., several clouds of data points) to an optimal compromise configuration, typically by suitable configuration-specific (e.g., orthogonal) linear transformations. The paper presents methods for estimating missing values in these configurations, with or without side constraints (centering, standardization). Essentially they proceed by minimizing quadratic performance criteria by an alternating least-squares algorithm. The novelty of the paper resides in the formulation and (sometimes even the explicit) solution of interesting constrained quadratic optimization problems.

The paper by *Mia Hubert* and *Stephan Van der Veen* considers (supervised) classification problems. It is well-known that classical discrimination rules (such as linear or quadratic ones related to normal or elliptical distributions) are sensitive to outliers. The authors propose a new robustification method for skewed unimodal distributions that proceeds by minimizing the adjusted outlyingness to each group. In the case of high dimensional data, it provides an extension of the “Soft Independent Modelling by Class Analogy” method that is both robustified and adjusted for skewness when combining it with a robust Principal Component approach for skewed data. Various simulation studies underline the usefulness of the proposed approach.

The following article by *Heungsun Hwang* and *Marc A. Tomiuk* considers a fuzzy clustering problem in which the observations from each class follow a class-specific generalized linear model (GLM). So, in contrast to classical GLM approaches, this model accounts for cluster-level heterogeneity. Estimation of clusters and parameters is performed by minimizing a regularized quasi-likelihood clustering criterion.

Monte-Carlo studies and a real-data example on the coupon usage behaviour of consumers illustrate the approach.

A major topic in psychometry and sociology is the problem of measuring the agreement between rankings obtained by several persons (raters) for the same set of objects. Various corresponding measures have been proposed in the literature, e.g., the classical kappa measure by Cohen and its modifications and extensions. In his paper ‘Inequalities between multi-rater kappas’ *Matthijs J. Warrens* proves that the values of some of these measures are strictly ordered, and other orderings hold under special data situations. These results generalize former publications by the author where similar results have been obtained for the case of two raters only.

In his paper on ‘Detecting atypical observations in financial data: the forward search for elliptical copulas’ *Tiziano Bellini* starts from the fact that in financial applications elliptical copulas are widely applied for describing the underlying probability distributions. He shows how in this framework multivariate outliers can be detected by an extension of the classical Forward Search algorithm based on subsets of contiguous observations, where squared Mahalanobis distances are computed and monitored along the search. Inferential issues for these distances are also addressed by means of simulation envelopes.

The problem of detecting multivariate outliers is considered by *Tri-Dzung Nguyen* and *Roy E. Welsch* from a mathematical programming point of view. Their proposal is to design a special mathematical programming model to find the optimal weights for all observations. At this optimal solution, outliers are given smaller weights and can then be detected. The suggested method produces a covariance estimator which is affine equivariant, computationally efficient and can incorporate prior beliefs.

On its last pages this issue contains a Call for Papers for a Special Issue of the journal ADAC that is devoted to methodological and applied aspects of the topic ‘Clustering of Time Series and Related Problems’. The readers of ADAC and other interested researchers are kindly invited to submit manuscripts to EditorADAC@stochastik.rwth-aachen.de before the deadline 15 January 2011.

Hans-Hermann Bock (Aachen)
Wolfgang Gaul (Karlsruhe)
Akinori Okada (Tokyo)
Maurizio Vichi (Rome)