Editorial

Murtagh and Legendre take the lead in this issue with a painstaking analysis of two different, yet very close implementations of Ward's classic method of hierarchical clustering that can be found in major software packages. It all has to do with taking squares or square roots at the right time and in the right place. Ward's method is special due to its intricate connection with K -means, and I believe this comprehensive and authoritative paper is a must-read for anyone interested in the practice of sum of squares clustering.

We then have two papers about supervised classification of functional data, in which the covariates or predictors are correlated as a function of an ordered parameter such as time. In the first of these papers, Escabias, Aguilera, and Aguilera-Morillo consider a multinomial response model, in which functional principal components analysis is used to reduce the dimensionality of the predictor space, and they test their methodology extensively in both simulated and real data. In the second supervised classification paper, Alonso, Casado, Lopez-Pintado and Romo use the integrated periodogram to classify times series, where the periodograms are calculated in local sections of the series to accommodate non-stationarity. To make the method robust, they use the L1 distance between any individual curve and the reference curve of a group as the basis of classification, and determine the reference curve as the α -trimmed mean of the curves in the group. In a series of simulation experiments, they show that their method has the desired robustness properties and is suitable for classifying long time series.

The last two papers of this issue consider the analysis of proximity data which have some special property. In the first of these papers, the special property considered is Robinson form, that is the situation where the rows and columns of a table of proximities can be reordered so that the entries never decrease when going away from the diagonal. Préa and Fortin give an optimal quadratic time algorithm to determine if a given proximity table has Robinson form or not. In the second paper, the special property considered is structural asymmetry, that is the situation where there are differences in corresponding entries above and below the diagonal that cannot be attributed to noise. Vicari attacks this problem by first decomposing the proximity table in a symmetric part and a skew-symmetric part, which are uncorrelated. She then offers a family of clustering models which can be the same or different for the two parts, and which can also model within and between cluster effects. Alternating least squares algorithms are developed for all models, and are shown in a simulation study to perform well, even when error levels are relatively large.

Willem J. Heiser