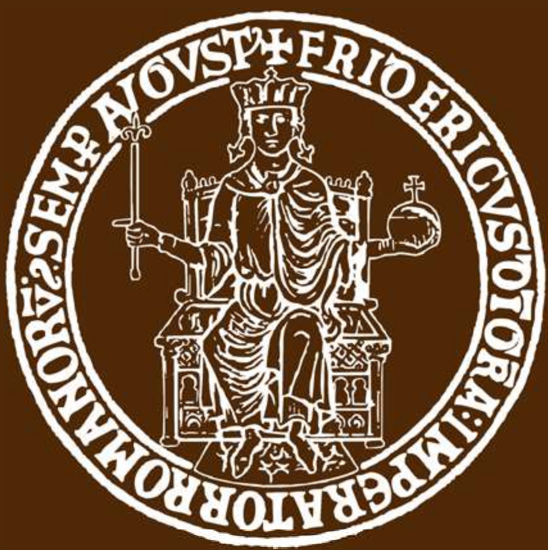
Università degli Studi di Napoli Federico II
Scuola delle Scienze Umane e Sociali
Quaderni
11

ASMOD 2018
Proceedings of the International Conference on Advances in Statistical Modelling of Ordinal Data

Naples, 24-26 October 2018

Editors
Stefania Capecchi, Francesca Di Iorio, Rosaria Simone



Federico II University Press

fedOA

Università degli Studi di Napoli Federico II
Scuola delle Scienze Umane e Sociali
Quaderni
11

ASMOD 2018

Proceedings of the International Conference on Advances in Statistical Modelling of Ordinal Data

Naples, 24-26 October 2018

Editors

Stefania Capecchi, Francesca Di Iorio, Rosaria Simone

Federico II University Press

fedOA Press

ASMOD 2018 : Proceedings of the Advanced Statistical Modelling for Ordinal Data Conference : Naples, 24-26 October 2018 / editors Stefania Capecchi, Francesca Di Iorio, Rosaria Simone. – Napoli : FedOAPress, 2018. – (Scuola di Scienze Umane e Sociali. Quaderni ; 11).

Accesso alla versione elettronica:
http://www.fedoabooks.unina.it

ISBN: 978-88-6887-042-3
DOI: 10.6093/978-88-6887-042-3
ISSN Collana: 2499-4774

Università degli Studi di Napoli Federico II
Centro di Ateneo per le Biblioteche “Roberto Pettorino”
Piazza Bellini 59-60
80138 Napoli, Italy
http://www.fedoapress.unina.it/

Published in Italy

Contents

Foreword ix

Invited papers

A. Agresti, *Simple ordinal model effect measures* 1

B. Francis, *Latent class models for multiple ordinal items* 3

B. Grün, G. Malsiner-Walli, *Bayesian latent class analysis with shrinkage priors: an application to the Hungarian heart disease data* 13

M. Kateri, *Modelling Ordinal Data: A ϕ-divergence based approach* 25

E. Ronchetti, *Robust statistical analysis of ordinal data* 27

G. Tutz, *Uncertainty, dispersion and response styles in ordinal regression* 33

Accepted papers

A. Barbiero, *Inducing a desired value of correlation between two point-scale variables* 45

A. Bonanomi, M.N. Ruscone, S.A. Osmetti, *Dissimilarity measure for ranking data via mixture of copulae* 53

G. Bove, E. Nuzzo, A. Serafini, *Measurement of interrater agreement for the assessment of language proficiency* 61

E. Brentari, M. Manisera, P. Zuccolotto, *Modelling perceived variety in a choice process with nonlinear CUB* 69

R. Colombi, S. Giordano, *A flexible distribution to handle response styles when modelling rating scale data* 77

M. Corduas, *Joint modelling of ordinal data: a copula-based method* 85

C. Davino, T. Naes, R. Romano, D. Vistocco, *Modeling preferences: beyond the average effects* 93

C. Davino, R. Simone, D. Vistocco, *Exploring synergy between CUB models and quantile regression: a comparative analysis through continuousized data* 101

E. di Bella, L. Leporatti, F. Maggino, L. Gandullia, *A Poset based indicator of gender equality at sub-national level* 109

M. Fattore, A. Arcagni, *Using mutual ranking probabilities for dimensionality reduction and ranking extraction in multidimensional systems of ordinal variables* 117

S. Golia, M. Carpita, *On classifiers to predict soccer match results* 125

L. Grilli, M.F. Marino, O. Paccagnella, C. Rampichini, *Multiple imputation and selection of ordinal level-2 predictors in multilevel models* 133

M. Iannario, A.C. Monti, P. Scalera, *Why the number of response categories in rating scales should be large* 139

A. Lorenzo-Arribas, M.J. Brewer, A.M. Overstall, *Solutions to issues with partial proportional odds models* 147

F. Pennoni, M. Nakai, *A latent variable model for a derived ordinal response accounting for sampling weights, missing values and covariates* 155

F. Pesarin, L. Salmaso, H. Huang, R. Arboretti, R. Ceccato, *Permutation tests for stochastic ordering with ordinal data* 163

A. Plaia, M. Sciandra, S. Buscemi, *Consensus measures for preference rankings with ties: an approach based on position weighted Kemeny distance* 171

M. Ranalli, R. Rocci, *Simultaneous clustering and dimensional reduction of mixed-type data* 179

V. Sansivieri, M. Matteucci, S. Mignani, *Bi-Factor MIRT observed-score equating under the NEAT design for tests with several content areas* 187

R. Simone, F. Di Iorio, C. Cappelli, *On the choice of splitting rules for model-based trees for ordinal responses* 195

F. Torti, S. Salini, M. Riani, *Robustness issues for categorical data* 203

M. Vives-Mestres, J. Antoni Martín-Fernández, S. Thió-Henestrosa, R.S. Kenett, *Applications and theoretical results of association rules and compositional data analysis: a contingency table perspective* 211

Simultaneous clustering and dimensional reduction of mixed-type data

Monia Ranalli*, Roberto Rocci**

Abstract: In real applications, it is very common to have the true clustering structure masked by the presence of noise variables and/or dimensions. A mixture model is proposed for simultaneous clustering and dimensionality reduction of mixed-type data: the continuous and the ordinal variables are assumed to follow a Gaussian mixture model, where, as regards the ordinal variables, it is only partially observed. To recognize discriminative and noise dimensions, the variables are considered to be linear combinations of two independent sets of latent factors where only one contains the information about the cluster structure while the other one contains noise dimensions. In order to overcome computational issues, the parameter estimation is carried out through an EM-like algorithm maximizing a composite log-likelihood based on low-dimensional margins.

Keywords: Mixture models, Composite likelihood, EM algorithm.

1. Introduction

The aim of cluster analysis is to partition the data into meaningful groups which should differ considerably from each other. The cluster analysis is made more difficult by the presence of mixed-type data (ordinal and continuous variables) combined by the presence of dimensions (named noise) that are uninformative for recovering the groups and could obscure the true cluster structure. It follows that there are two main points to be addressed: combining continuous with ordinal variables; taking into account the presence of noise variables/dimensions. As regards the first point, the literature on clustering for continuous data is rich and wide; the most commonly clustering model-based used is the finite mixture of Gaussians (McLachlan et al., 2016). Differently, that one developed for categorical data is still limited. Models used for ordinal data mainly adopt two approaches developed in the factor

*University of Tor Vergata, monia.ranalli@uniroma2.it
**University of Tor Vergata, roberto.rocci@uniroma2.it

analysis framework: Item Response Theory (IRT) (see e.g. Bartholomew et al. (2011), Bock and Moustaki (2007)), and the Underlying Response Variable (URV) (see e.g. Jöreskog, 1990; Lee et al., 1990; Muthén, 1984). In the URV approach, the ordinal variables are seen as a discretization of continuous latent variables jointly distributed as a finite mixture; examples are: Everitt (1988), Lubke and Neale (2008), Ranalli and Rocci (2016a, 2017a, 2017b). However, this makes the maximum likelihood estimation rather complex because it requires the computation of many high dimensional integrals. The problem is usually solved by approximating the likelihood function. In this regard we mention some useful surrogate functions, such as the variational likelihood (Gollini and Murphy, 2014) or the composite likelihood (Ranalli and Rocci (2016a, 2017a, 2017b)). Although it is possible to cluster via a model based approach continuous or ordinal variables separately, combining both into a common framework may raise some issues. Following the URV approach, Everitt (1988) and Ranalli-Rocci (2017a) proposed a model according to which both the continuous and the categorical ordinal variables follow a Gaussian mixture model, where the ordinal variables are only partially observed through their ordinal counterparts. This satisfies the two main requirements: dealing with ordinal data properly and modelling dependencies between ordinal and continuous variables. As regards the presence of noise variables, different approaches exist in literature. Several techniques for simultaneous clustering and dimensionality reduction (SCR) have been proposed in a non-model based framework for quantitative (e.g.: Rocci et al., 2011; Vichi and Kiers, 2001) or categorical data (e.g.: Hwang et al., 2006; Van Buuren et al., 1989). There are also approaches based on a family of mixture models which fits the data into a common discriminative subspace (see e.g. Bouveyron and Brunet, 2012; Kumar and Andreou, 1998; Ranalli and Rocci, 2017b). The key idea is to assume a common latent subspace to all groups that is the most discriminative one. This allows to project the data into a lower dimensional space preserving the clustering characteristics in order to improve visualization and interpretation of the underlying structure of the data. The model can be formulated as a finite mixture of Gaussians with a particular set of constraints on the parameters. Combining all pieces together, following the URV approach, in our proposal the continuous and the ordinal variables are assumed to follow a heteroscedastic Gaussian mixture model, where, as regards the ordinal variables, it is only partially observed. To recognize discriminative and noise dimensions, these variables are considered to be linear combinations of two independent sets of latent factors where only one contains the information about the cluster structure, defining a discriminative subspace, distributed as a finite mixture of Gaussians. The other one contains noise dimensions distributed as a multivariate normal. The model specification is parsimonious and is able to identify a reduced set of discriminative latent factors/dimensions even when there are no noise variables to be detected. The main drawback of this model is that, in practice, it cannot be estimated through a full maximum likelihood approach, due to the presence of multidimensional integrals making the estimation time consuming. To overcome this issue, we propose to replace this cumbersome likelihood with a surrogate objective function, easier to maximize, that is the product of marginal likelihoods. It is a composite likelihood method (Lindsay, 1988; Varin et al., 2011) where surrogate functions are defined as the product of marginal or conditional events. In particular, our proposals is based on the existing results within a mixture model framework Ranalli-Rocci (2016a, 2017a, 2017b). It consists of replacing the joint likelihood with all possible marginals, like bivariate marginal distributions of ordinal variables and the marginal distributions of one ordinal variable and all continuous variables.

2. Model specification

Let $\mathbf{x} = [x_1, \ldots, x_O]'$ and $\mathbf{y}^{\bar{O}} = [y_{O+1}, \ldots, y_P]'$ be O ordinal and $\bar{O} = P - O$ continuous variables, respectively. The associated categories for each ordinal variable are denoted by $c_i = 1, 2, \ldots, C_i$ with $i = 1, 2, \ldots, O$. Following the underlying response variable approach (URV) developed within the SEM framework (see e.g. Jöreskog, 1990; Lee et al., 1990; Muthén, 1984), the ordinal variables $\mathbf{x}$ are considered as a categorization of a continuous multivariate latent variable $\mathbf{y}^O = [y_1, \ldots, y_O]'$. According to the URV, the joint distribution of $\mathbf{x}$ and $\mathbf{y}^{\bar{O}}$ can be constructed as follows. The latent relationship between $\mathbf{x}$ and $\mathbf{y}^O$ is explained by the threshold model, $x_i = c_i \Leftrightarrow \gamma^{(i)}_{c_i-1} \leq y_i < \gamma^{(i)}_{c_i}$, where $-\infty = \gamma^{(i)}_0 < \gamma^{(i)}_1 < \ldots < \gamma^{(i)}_{C_i-1} < \gamma^{(i)}_{C_i} = +\infty$ are the thresholds defining the C_i categories collected in a set $\mathbf{\Gamma}$ whose elements are given by the vectors $\boldsymbol{\gamma}^{(i)}$. To accommodate both cluster structure and dependence within the groups, we assume that $\mathbf{y} = [\mathbf{y}^{O\prime}, \mathbf{y}^{\bar{O}\prime}]'$ follows a heteroscedastic Gaussian mixture, $f(\mathbf{y}) = \sum_{g=1}^{G} p_g \phi_p(\mathbf{y}; \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)$, where the p_g's are the mixing weights and $\phi_p(\mathbf{y}; \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)$ is the density of a P-variate normal distribution with mean vector $\boldsymbol{\mu}_g$ and covariance matrix $\boldsymbol{\Sigma}_g$. Let us set $\boldsymbol{\psi} = \{p_1, \ldots, p_G, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_G, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_G, \boldsymbol{\Gamma}\} \in \boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ is the parameter space. For a random i.i.d. sample of size N, $(\mathbf{x}_1, \mathbf{y}_1^{\bar{Q}}), \ldots, (\mathbf{x}_N, \mathbf{y}_N^{\bar{Q}})$, the log-likelihood is

$$\ell(\boldsymbol{\psi}) = \sum_{n=1}^{N} \log \left[\sum_{g=1}^{G} p_g \phi_{\bar{O}}(\mathbf{y}_n^{\bar{O}}; \boldsymbol{\mu}_g^{\bar{O}}, \boldsymbol{\Sigma}_g^{\bar{O}}) \pi_n \left(\boldsymbol{\mu}_{n;g}^{O|\bar{O}}, \boldsymbol{\Sigma}_g^{O|\bar{O}}, \boldsymbol{\Gamma} \right) \right], \qquad (1)$$

where, with obvious notation

$$\begin{aligned}
\pi_n \left(\boldsymbol{\mu}_{n;g}^{O|\bar{O}}, \boldsymbol{\Sigma}_g^{O|\bar{O}}, \boldsymbol{\Gamma} \right) &= \int_{\gamma_{c_1-1}^{(1)}}^{\gamma_{c_1}^{(1)}} \cdots \int_{\gamma_{c_O-1}^{(O)}}^{\gamma_{c_O}^{(O)}} \phi_O(\mathbf{u}; \boldsymbol{\mu}_{n;g}^{O|\bar{O}}, \boldsymbol{\Sigma}_g^{O|\bar{O}}) d\mathbf{u} \\
\boldsymbol{\mu}_{n;g}^{O|\bar{O}} &= \boldsymbol{\mu}_g^{O} + \boldsymbol{\Sigma}_g^{O\bar{O}} (\boldsymbol{\Sigma}_g^{\bar{O}\bar{O}})^{-1} (\mathbf{y}_n^{\bar{O}} - \boldsymbol{\mu}_g^{\bar{O}}) \\
\boldsymbol{\Sigma}_g^{O|\bar{O}} &= \boldsymbol{\Sigma}_g^{OO} - \boldsymbol{\Sigma}_g^{O\bar{O}} (\boldsymbol{\Sigma}_g^{\bar{O}\bar{O}})^{-1} \boldsymbol{\Sigma}_g^{\bar{O}O},
\end{aligned}$$

$\pi_n \left(\boldsymbol{\mu}_{n;g}^{O|\bar{O}}, \boldsymbol{\Sigma}_g^{O|\bar{O}}, \boldsymbol{\gamma} \right)$ is the conditional joint probability of response pattern $\mathbf{x}_n = (c_1^{(1)}, \ldots, c_O^{(O)})$ given the cluster g and the continuous variables $\mathbf{y}_n^{\bar{O}}$. Finally p_g is the probability of belonging to group g subject to $p_g > 0$ and $\sum_{g=1}^{G} p_g = 1$. In order to identify the discriminative dimensions, it is assumed that there is a set of P latent factors $\tilde{\mathbf{y}}$, formed of two independent subsets. In the first one, there are Q (with $Q \leq P$) factors that have some clustering information distributed as a mixture of Gaussians with class conditional means and variances equal to $E(\tilde{\mathbf{y}}^{Q} \mid g) = \boldsymbol{\eta}_g$ and $\text{Cov}(\tilde{\mathbf{y}}^{Q} \mid g) = \boldsymbol{\Omega}_g$, respectively. In the second set there are $\bar{Q} = P - Q$ noise factors defining the so-called noise dimensions, that are independent of $\tilde{\mathbf{y}}^{Q}$ and their distribution does not vary from one class to another: $E(\tilde{\mathbf{y}}^{\bar{Q}} \mid g) = \boldsymbol{\eta}_0$ and $\text{Cov}(\tilde{\mathbf{y}}^{\bar{Q}} \mid g) = \boldsymbol{\Omega}_0$. The link between $\tilde{\mathbf{y}}$ and $\mathbf{y}$ is given by a non-singular $P \times P$ matrix $\mathbf{A}$, as $\mathbf{y} = \mathbf{A}\tilde{\mathbf{y}}$. The final step is to identify the variables that could be considered as noise. Intuitively y_p is a noise variable if it is well explained by $\tilde{\mathbf{y}}^{\bar{Q}}$. Exploiting the independence between $\tilde{\mathbf{y}}^{Q}$ and $\tilde{\mathbf{y}}^{\bar{Q}}$, it is possible to compute proportions of each variable's variance that can be explained by the noise factors, and by one's complement, the proportions of each variable's variance that can be explained by the discriminative factors.

2.1. Construction of surrogate functions

The presence of multidimensional integrals makes the maximum likelihood estimation computationally demanding and infeasible as the number of observed ordinal variables increases. To overcome this, a composite likelihood approach is adopted (Lindsay, 1988). It allows us to simplify the problem by replacing the full likelihood with a surrogate function. As suggested in Ranalli-Rocci (2016a, 2017a, 2017b) within a similar context, the full log-likelihood could be replaced by the sum of two estimating-block functions: $O(O-1)/2$ bivariate marginals of ordinal variables; O marginal distributions each of them composed of one ordinal variable and the $\bar{O}$ continuous variables. This leads to the following surrogate function

$$c\ell(\boldsymbol{\psi}) = \sum_{n=1}^{N}\sum_{i=1}^{O-1}\sum_{j=i+1}^{O}\sum_{c_i=1}^{C_i}\sum_{c_j=1}^{C_j}\delta_{nc_ic_j}^{(ij)}\log\left[\sum_{g=1}^{G}p_g\pi_{c_ic_j}^{(ij)}(\boldsymbol{\mu}_g^{(ij)},\boldsymbol{\Sigma}_g^{(ij)},\boldsymbol{\Gamma}^{(ij)})\right]$$
$$+\quad \sum_{n=1}^{N}\sum_{j=1}^{O}\sum_{c_j=1}^{C_j}\delta_{nc_j}^{(j)}\log\left[\sum_{g=1}^{G}p_g\pi_{c_j}^{(j|\bar{O})}(\mu_{n;g}^{(j|\bar{O})},\sigma_g^{(j|\bar{O})},\boldsymbol{\Gamma}^{j})\phi_{\bar{O}}(\mathbf{y}_n^{\bar{O}};\boldsymbol{\mu}_g^{\bar{O}},\boldsymbol{\Sigma}_g^{\bar{O}\bar{O}})\right],$$

where now, after the reparameterization induced by the reduction model, the set of parameters is $\boldsymbol{\psi} = \{p_1,\dots,p_G,\boldsymbol{\eta}_0,\boldsymbol{\eta}_1,\dots,\boldsymbol{\eta}_G,\boldsymbol{\Omega}_0,\boldsymbol{\Omega}_1,\dots,\boldsymbol{\Omega}_G,\mathbf{A},\gamma\}$, $\delta_{nc_ic_j}^{(ij)}$ is a dummy variable assuming 1 if the n-th observation presents the combination of categories c_i and c_j for variables x_i and x_j respectively, 0 otherwise; similarly $\delta_{nc_j}^{(j)}$ is a dummy variable assuming 1 if the n-th observation presents category c_j for variable x_j, 0 otherwise; $\pi_{c_ic_j}^{(ij)}(\boldsymbol{\mu}_g^{(ij)},\boldsymbol{\Sigma}_g^{(ij)},\boldsymbol{\Gamma}^{(ij)})$ is the probability under the model obtained by integrating the density of a bivariate normal distribution with parameters $(\boldsymbol{\mu}_g^{(ij)},\boldsymbol{\Sigma}_g^{(ij)},\boldsymbol{\Gamma}^{(ij)})$ between the corresponding threshold parameters. On the other hand, $\pi_{c_j}^{(j|\bar{Q})}(\mu_{n;g}^{(j|\bar{Q})},\sigma_g^{(j|\bar{Q})},\boldsymbol{\Gamma}^{(j)})$ is the conditional probability of variable x_j of being in category c_j given all the continuous variables $\mathbf{y}^{\bar{Q}}$. Finally, $\boldsymbol{\mu}_g = E(\mathbf{y} \mid g) = \mathbf{A}E(\tilde{\mathbf{y}} \mid g)$, while $\boldsymbol{\Sigma}_g = \text{Cov}(\mathbf{y} \mid g) = \mathbf{A}\text{Cov}(\tilde{\mathbf{y}} \mid g)\mathbf{A}'$, as specified previously. The parameter estimates are carried out through an EM-like algorithm, that works in the same manner as the standard EM.

2.2. Classification model selection and identifiability

When we adopt a composite likelihood approach, since we do not compute the joint distribution for each observation, it is not possible anymore to assign the observation to the component with the maximum a posteriori probability (MAP criterion) without further computations. To solve the problem we follow the CMAP criterion (Ranalli-Rocci, 2017a, 2017b), according to which the observation is assigned to the component with the maximum scaled composite fit (scaled by the corresponding mixing weight). As regards model selection, the best model is chosen by minimizing the composite version of penalized likelihood selection criteria like BIC or CLC (see Ranalli-Rocci, 2016b and references therein). Finally, as regards identifiability, within a full maximum likelihood approach, it is well known that a sufficient condition for local identifiability is given by the non singularity of the information matrix; while a necessary condition is that the number of parameters must be less than or equal to the number of canonical parameters. Adopting a composite likelihood approach, the sufficient condition should be reformulated by investigating the Godambe information matrix, that is, the analogous of the information matrix in composite likelihood estimation. However, as far as we know, such modification has not been formally investigate yet. About the necessary condition, we note that the number of essential parameters in the block of ordinal variables equals the number of parameters of a log linear model with only two factor interaction terms. Thus it means that we can estimate a lower number of parameters compared to a full maximum likelihood approach. Furthermore, under the underlying response variable approach, the means and the variances of the latent variables are set to 0 and 1, respectively, because they are not identified. This identification constraint individualizes uniquely the mixture components (ignoring the label switching problem), as well described in Millsap and Yun-Tein (2004). This is sufficient to estimate both thresholds and component parameters if all the observed variables have three categories at least and when groups are known. Given the particular structure of the mean vectors and covariance matrices, it is preferable to adopt an alternative, but equivalent, parametrization. This is analogous to that one used by Jöreskog and Sörbom (1996); it consists in setting the first two thresholds to 0 and 1, respectively. This means that there is a one-to-one correspondence between the two sets of parameters. If there is a binary variable, then the variance of the corresponding latent variable is set equal to 1 (while its mean should be still kept free). Finally, we note that the model has the same rotational freedom that characterizes the classical factor analysis model. In other words, writing $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2]$ according to $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}^{Q\prime}, \tilde{\mathbf{y}}^{\bar{Q}\prime}]'$, only the subspaces generated by the columns of $\mathbf{A}_1$ and $\mathbf{A}_2$ are identified. In order to estimate such subspaces, we impose some constraints on the model parameters, in complete analogy with what is usually done in the factor analysis model. In this way, we select a particular solution, one which is convenient to find, and leave the experimenter to apply whatever rotation he thinks desirable, as suggested by Lawley and Maxwell (1962). In particular, we require a spherical distribution for the noise factors, i.e. $\mathbf{\Omega}_0 = \mathbf{I}$, and informative factors in the first cluster, i.e. $\mathbf{\Omega}_1 = \mathbf{I}$. Such constraints still allow a rotational freedom by orthonormal matrices. This can be eliminated by requiring a "lower" triangular form for the two loading matrices. In general, $\mathbf{A}_1$ and $\mathbf{A}_2$ have a lower triangular matrix in the first Q and $(P - Q)$ rows, respectively. Of course, after the estimation the parameter matrices can be rotated to enhance the interpretation.

Further details will be given in the extended version of the paper along with simulation and real data results to show the effectiveness of the proposal.

References

Bartholomew D.J., Knott M., Moustaki I. (2011) *Latent Variable Models and Factor Analysis: A Unified Approach.* Wiley Series in Probability and Statistics. Wiley, third edition, 2011.

Bock D., Moustaki I. (2007) *Handbook of Statistics on Psychonometrics*, chapter: Item response theory in a general framework. Elsevier.

Bouveyron C., Brunet C. (2012) Simultaneous model-based clustering and visualization in the fisher discriminative subspace, *Statistics and Computing*, 22, 301-324.

Everitt B.S. (1988) A finite mixture model for the clustering of mixed-mode data, *Statistics & Probability Letters*, 6, 305-309.

Gollini I., Murphy T.B. (2014) Mixture of latent trait analyzers for model-based clustering of categorical data, *Statistics and Computing*, 24, 569-588.

Hwang H., Montréal H., Dillon W.R., Takane Y. (2006) An extension of multiple correspondence analysis for identifying heterogeneous subgroups of respondents, *Psychometrika*, 71, 161-171.

Jöreskog K.G. (1990) New developments in Lisrel: analysis of ordinal variables using polychoric correlations and weighted least squares, *Quality and Quantity*, 24, 387-404.

Jöreskog K.G., Sörbom D. (1996) *LISREL 8: User's Reference Guide*. Scientific Software.

Kumar N., Andreou A.G. (1998) Heteroscedastic discriminant analysis and reduced rank HMMs for improved speech recognition, *Speech Communication*, 26, 283 - 297.

Lawley D.N., Maxwell A.E. (1962) Factor analysis as a statistical method, *Journal of the Royal Statistical Society. Series D (The Statistician)*, 12, 209-229.

Lee S-Y, Poon W-Y, Bentler P.M. (1990) Full maximum likelihood analysis of structural equation models with polytomous variables, *Statistics & Probability Letters*, 9, 91-97.

Lindsay B.G. (1988) Composite likelihood methods, *Contemporary Mathematics*, 80, 221-239.

Lubke G., Neale M. (2008) Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models, *Multivariate Behavioral Research*, 43, 592-620.

McLachlan G., Rathnayake S.I. (2016) Mixture models for standard p-dimensional Euclidean data. *Handbook of cluster analysis*, CRC Press, 145-172.

Millsap R.E., Yun-Tein J. (2004) Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39, 479-515.

Muthén B. (1984) A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators, *Psychometrika*, 49, 115-132.

Ranalli M., Rocci R. (2016a) Mixture models for ordinal data: a pairwise likelihood approach, *Statistics and Computing*, 26, 529-547.

Ranalli M., Rocci R. (2016b) Standard and novel model selection criteria in the pairwise likelihood estimation of a mixture model for ordinal data. *Analysis of Large and Complex Data. Studies in Classification,Data Analysis and Knowledge Organization. Editors: Adalbert F.X. Wilhelm Hans A. Kestler.*.

Ranalli M., Rocci R. (2017a) Mixture models for mixed-type data through a composite likelihood approach, *Computational Statistics & Data Analysis*, 110, 87-102.

Ranalli M., Rocci R. (2017b) A model-based approach to simultaneous clustering and dimensional reduction of ordinal data, *Psychometrika*, 82, 1007-1034.

Rocci R., Gattone S.A., Vichi M. (2011) A new dimension reduction method: Factor discriminant k-means. *Journal of classification*, 28, 210-226.

Van Buuren S. and Heiser W.J. (1989) Clustering objects into *k* groups under optimal scaling of variables, *Psychometrika*, 54, 699-706.

Varin C., Reid N., Firth D. (2011) An overview of composite likelihood methods, *Statistica Sinica*, 21, 1-41.

Vichi M., Kiers H.A.L. (2001) Factorial k-means analysis for two-way data, *Computational Statistics & Data Analysis*, 37, 49-64.

Università degli Studi di Napoli Federico II
Scuola delle Scienze Umane e Sociali

This volume collects the peer-reviewed contributions presented at the 2nd International Conference on "Advances in Statistical Modelling of Ordinal Data" - ASMOD 2018 - held at the Department of Political Sciences of the University of Naples Federico II (24-26 October 2018). The Conference brought together theoretical and applied statisticians to share the latest studies and developments in the field. In addition to the fundamental topic of latent structure analysis and modelling, the contributions in this volume cover a broad range of topics including measuring dissimilarity, clustering, robustness, CUB models, multivariate models, and permutation tests. The Conference featured six distinguished keynote speakers: Alan Agresti (University of Florida, USA), Brian Francis (Lancaster University, UK), Bettina Gruen (Johannes Kepler University Linz, Austria), Maria Kateri (RWTH Aachen, Germany), Elvezio Ronchetti (University of Geneva, Switzerland), Gerhard Tutz (Ludwig-Maximilians University of Munich, Germany). The volume includes 22 contributions from scholars that were accepted as full papers for inclusion in this edited volume after a blind review process of two anonymous referees.

ISBN: 978-88-6887-042-3
DOI: 10.6093/978-88-6887-042-3
Online ISSN: 2532-4608

9 788868 870423