M. Katsikatsou, Irini Moustaki, F. Yang-Wallentin, and Karl G. Jöreskog

Pair wise likelihood estimation for factor analysis models with ordinal data

Article (Accepted version) (Unrefereed)

Original citation:
Katsikatsou, Myrsini and Moustaki, Irini and Yang-Wallentin, Fan and Jöreskog, Karl G. (2012) *Pair wise likelihood estimation for factor analysis models with ordinal data. Computational statistics and data analysis*, 56 (12). pp. 4243-4258. ISSN 0167-9473
DOI: 10.1016/j.csda.2012.04.010

To appear in *Computational Statistics and Data Analysis* in 2012

Pairwise likelihood estimation for factor analysis models with ordinal data

Myrsini Katsikatsou[a,*], Irini Moustaki[b], Fan Yang-Wallentin[a], Karl G. Jöreskog[a,c]

[a] *Department of Statistics, University of Uppsala, S-75120, Sweden*

[b] *Department of Statistics, London School of Economics, WC2A 2AE, United Kingdom*

[c] *Department of Economics, Norwegian School of Management, 0442 Oslo, Norway*

Abstract

Pairwise maximum likelihood (PML) estimation method is developed for factor analysis models with ordinal data and fitted both in an exploratory and confirmatory set-up. The performance of the method is studied via simulations and comparisons with full information maximum likelihood (FIML) and three-stage limited information estimation methods, namely the robust unweighted least squares (3S-RULS) and robust diagonally weighted least squares (3S-RDWLS). The advantage of PML over FIML is mainly computational. Unlike PML estimation, the computational complexity of FIML estimation increases either with the number of factors or with the number of observed variables depending on the model formulation. Contrary to 3S-RULS and 3S-RDWLS estimation, PML estimates of all model parameters are obtained simultaneously and the PML method does not require the estimation of a weight matrix for the computation of correct standard errors. The simulation study on the performance of PML estimates and estimated asymptotic standard errors investigates the effect of different model and sample sizes. The bias and mean squared error of PML estimates and their standard errors are found to be small in all experimental conditions and decreasing with increasing sample size. Moreover, the PML estimates and their standard errors are found to be very close to those of FIML.

Keywords: composite maximum likelihood; factor analysis; ordinal data; pairwise likelihood; three-stage estimation; item response theory approach.

1. Introduction

Factor analysis is frequently employed in social sciences where the main interest lies in measuring and relating unobserved constructs, such as emotions, attitudes, beliefs and behavior. The main idea behind the analysis is that the latent variables (referred to also as factors) account for the dependencies among the observed variables (referred to also as items or indicators) in the sense that if the factors are held fixed, the observed variables would be independent. Theoretically, factor analysis can be distinguishable between exploratory and confirmatory analysis, but in practice the analysis always lie between the two. In exploratory factor analysis the goal is the following: for a given set of observed variables $x_1, \ldots, x_p$ one wants to find a set of latent factors $\xi_1, \ldots, \xi_k$, fewer in number than the observed variables ($k < p$), that contain essentially the same information. In confirmatory factor analysis, the objective is to verify a social theory. Hence, a factor model is specified in advance and its fit to the empirical data is tested.

The data usually encountered in social sciences is of categorical nature (ordinal or nominal). In the literature, there are two main approaches for the analysis of ordinal variables with factor models. The Underlying Response Variable (URV) approach (e.g. Jöreskog 1990, 1994; Lee et al. 1990, 1992; Muthén, 1984), and the Item Response Theory (IRT) approach (e.g. Bartholomew et al., 2011; Bock & Moustaki, 2007; van der Linden & Hambleton, 1997). In the URV approach, the ordinal variables are assumed to be generated by underlying continuous variables, which are

*Corresponding author. Tel.: +46 18 471 1038
Email addresses: myrsini.katsikatsou@statistik.uu.se (Myrsini Katsikatsou), i.moustaki@lse.ac.uk (Irini Moustaki), fan.yang@statistik.uu.se (Fan Yang-Wallentin), karl.joreskog@statistik.uu.se (Karl G. Jöreskog)

partially observed through their ordinal counterparts. In the IRT approach, ordinal indicators are treated as they are. In both approaches, one must specify the probability of each response pattern as a function of $\xi_1, \xi_2, \ldots, \xi_k$:

$$Pr(x_1 = c_1,\ x_2 = c_2, \ldots, x_p = c_p \mid \xi_1, \ldots, \xi_k) = f(\xi_1, \xi_2, \ldots, \xi_k)\ , \tag{1}$$

where $c_1, c_2, \ldots, c_p$ represent the different response categories of $x_1, x_2, \ldots, x_p$, respectively.

Marginal maximum likelihood, also known as full information maximum likelihood (FIML) (Muraki & Carlson, 1995), is used for estimating the model under the IRT framework. Whereas three-stage limited information robust unweighted least squares (3S-RULS) and robust diagonally weighted least squares (3S-RDWLS) methods (see e.g. Forero et al., 2009; Yang-Wallentin et al., 2010 for a comparison) are employed for estimating the model under the URV framework. In the latter framework, FIML is infeasible for problems with more than five observed variables. In both approaches, FIML involves high-dimensional integrations, the dimensions of which increase either with the number of observed variables under the URV approach or with the number of factors under the IRT approach. Under the URV approach, robust maximum likelihood (RML) estimation is also used. This is the standard maximum likelihood estimation employed with continuous observed variables but with standard errors and chi-square tests estimated under non-normality (see e.g. Browne, 1984; Satorra, 1989; Satorra & Bentler, 1988; Yang-Wallentin et al., 2010). Recently Bollen & Maydeu-Olivares (2007) developed the polychoric instrumental variable estimator (PIV) for structural equation models with categorical data, which is robust to structural misspecifications. According to our knowledge, the PIV estimator has not yet been incorporated into any available software.

The restricted applicability of FIML in both URV and IRT approaches, along with knowledge of the theoretical developments of Composite Maximum Likelihood (CML) estimation (Lindsay, 1988; Varin, 2008; Varin et. al., 2011), motivated us to consider CML as an alternative estimation method. CML estimators have the desired properties of being asymptotically unbiased, consistent, and normally distributed. Additionally, CML can be applied to any of the two aforementioned model formulations; although the computational gain is only for the URV approach, the computational complexity of which can be kept low regardless of the number of observed variables or factors. To our knowledge, there are very few studies that investigate the applicability and performance of CML approaches within the context of factor analysis with ordinal data. De Leon (2005) used the pairwise maximum likelihood approach to estimate thresholds and polychoric correlations of ordinal data. His simulation study indicates that the estimates are quite accurate, yielding minimal bias and small root mean squared errors. Jöreskog & Moustaki (2001) suggest the use of the *underlying bivariate normal* (UBN) method within the URV approach, which is found to yield estimates close to those of the FIML approach. UBN can be seen as a composite maximum likelihood method, involving both univariate and bivariate marginal distributions. However, Jöreskog & Moustaki (2001) do not incorporate their approach within the CML framework, nor within any other general framework. Thus, they do not provide any discussion about the standard errors and the properties of the UBN estimator. Liu (2007) proposes a multistage estimation method for structural equation models, an alternative to the commonly used three-stage methods. In particular, in a first stage, thresholds, polychoric, and polyserial correlations are estimated simultaneously by using the pairwise maximum likelihood approach. Given these estimates, structural parameters, such as loadings and factor correlations, are estimated in a second stage using generalized least squares. The simulation studies in Liu (2007) show that the proposed methodology performs equally well as the conventional three-stage methods. Finally, Vasdekis et al. (2012) have developed a pairwise estimation for longitudinal ordinal variables under the IRT approach.

Based on the results of the above studies, we develop a pairwise maximum likelihood (PML) estimation method for factor analysis models with ordinal variables under the URV approach. We compare the method's performance, in terms of bias and mean square error (MSE), with 3S-RULS, 3S-RDWLS, and FIML used with the IRT model formulation. The structure of the paper is as follows: Section 2 provides a brief presentation of the URV and IRT approaches, with a focus on the estimation. We discuss the computational issues arising in the case of FIML and the advantages and disadvantages of the three-stage limited information estimators. Section 3 presents the proposed methodology, namely the pairwise maximum likelihood estimation. It is followed by the simulation study on the performance of PML in Section 4 and the simulation study on the comparison of FIML, PML, 3S-RULS, and 3S-RDWLS in Section 5. The PML approach is applied to empirical data, both in the case of exploratory and confirmatory factor analysis, and a comparison of the estimates with those obtained by 3S-RULS, 3S-RDWLS, and FIML is discussed in Section 6. Discussion and conclusions are provided at the end.

2. Factor analysis models with ordinal observed variables

2.1. Basic framework and notation

Let $\mathbf{x}' = (x_1, x_2, \ldots, x_p)$ denote the vector of p ordinal observed variables, where x_i has m_i ordered categories, $i = 1, \ldots, p$. Thus, there are $R = \prod_{i=1}^{p} m_i$ possible response patterns of the form $\mathbf{x}_r' = (c_1, c_2, \ldots, c_p)$, where $c_i = 1, \ldots, m_i$. For a random sample of size n the log-likelihood is:

$$\ln L(\boldsymbol{\theta}; \mathbf{x}) = \sum_{r=1}^{R} n_r \ln \pi_r(\boldsymbol{\theta}) \ , \tag{2}$$

where $\boldsymbol{\theta}$ is a parameter vector, n_r and $\pi_r(\boldsymbol{\theta})$ are the observed frequency and the probability under the model, respectively, for the response pattern r, $\pi_r(\boldsymbol{\theta}) > 0$, $\sum_{r=1}^{R} n_r = n$, and $\sum_{r=1}^{R} \pi_r(\boldsymbol{\theta}) = 1$. Each approach imposes a different model on the probability $\pi_r(\boldsymbol{\theta})$, but both URV and IRT methods assume the presence of a k-dimensional vector of continuous latent variables $\boldsymbol{\xi}' = (\xi_1, \ldots, \xi_k)$, where $k < p$.

2.2. Underlying Response Variable (URV) approach

Under the URV approach, the observed ordinal variables are taken to be manifestations of underlying continuous variables partially observed through their ordinal counterparts. The connection between an observed ordinal variable x_i and the underlying continuous variable $x_i^\star$ is

$$x_i = c_i \Longleftrightarrow \tau_{c_i-1}^{(x_i)} < x_i^\star < \tau_{c_i}^{(x_i)} \ , \tag{3}$$

where $\tau_{c_i}^{(x_i)}$ is the c_i^{th} threshold of variable x_i and $-\infty = \tau_0^{(x_i)} < \tau_1^{(x_i)} < \ldots < \tau_{m_i-1}^{(x_i)} < \tau_{m_i}^{(x_i)} = +\infty$. Since only ordinal information is available, the distribution of $x_i^\star$ is determined only up to a monotonic transformation. In practice it is convenient to assume a standard normal distribution. In the case that the mean and the variance of $x_i^\star$ are of interest, Jöreskog (2002) discusses an alternative parametrization.

The factor model is of the form

$$\mathbf{x}^\star = \Lambda\boldsymbol{\xi} + \boldsymbol{\delta} \ , \tag{4}$$

where $\mathbf{x}^\star$ is the p-dimensional vector of the underlying variables, Λ is the $p \times k$ matrix of loadings, and $\boldsymbol{\delta}$ is the p-dimensional vector of unique variables. In addition, it is assumed that $\boldsymbol{\xi} \sim N_k(\mathbf{0}, \Phi)$ where Φ has 1's on its main diagonal being this way, the correlation matrix of latent factors, $\boldsymbol{\delta} \sim N_p(\mathbf{0}, \Theta)$ with Θ a diagonal matrix, $\Theta = I - \text{diag}(\Lambda\Phi\Lambda')$, and $Cov(\boldsymbol{\xi}, \boldsymbol{\delta}) = \mathbf{0}$. The parameter vector $\boldsymbol{\theta}' = (\boldsymbol{\lambda}, \boldsymbol{\varphi}, \boldsymbol{\tau})$ contains $\boldsymbol{\lambda}$ and $\boldsymbol{\varphi}$ which are the vectors of the free non-redundant parameters in matrices Λ and Φ, respectively, and $\boldsymbol{\tau}$ which is the vector of all free thresholds.

Under the model, the probability of a response pattern r is

$$\pi_r(\boldsymbol{\theta}) = \pi\left(x_1 = c_1, x_2 = c_2, \ldots, x_p = c_p; \boldsymbol{\theta}\right) = \int_{\tau_{c_1-1}^{(x_1)}}^{\tau_{c_1}^{(x_1)}} \cdots \int_{\tau_{c_p-1}^{(x_p)}}^{\tau_{c_p}^{(x_p)}} \phi_p(\mathbf{x}^\star; \Sigma_{\mathbf{x}^\star}) d\mathbf{x}^\star \ , \tag{5}$$

where $\phi_p(\mathbf{x}^\star; \Sigma_{\mathbf{x}^\star})$ is a p-dimensional normal density with zero mean, and correlation matrix $\Sigma_{\mathbf{x}^\star} = \Lambda\Phi\Lambda' + \Theta$.

The maximization of log-likelihood defined in (2) over the parameter vector $\boldsymbol{\theta}$ requires the evaluation of the p-dimensional integral given in (5), which cannot be written in a closed form. Lee et al. (1990) discuss FIML estimation in the case of the URV approach, but restrict their example to the case of four ordinal observed variables. As a consequence, limited information estimation methods have been proposed and added to the analytical tools of commercial software, the most widely used ones being the three-stage estimation methods (Jöreskog, 1990, 1994; Muthén, 1984). In the application of these methods, thresholds are first estimated by maximizing the univariate marginal likelihoods separately. Then, given the estimated thresholds, polychoric correlations are estimated by maximizing the bivariate marginal likelihoods separately. In the third stage, the factor analysis model given in (4) is fitted to the estimated polychoric correlation matrix using a version of generalized least squares (GLS), such as unweighted least squares (ULS), diagonally weighted least squares (DWLS), and weighted least squares (WLS) (e.g. Jöreskog, 1990, 1994; Jöreskog & Sörbom, 1996, pp. 23-24; Muthén, 1984; Muthén et al., 1997). In WLS, the weight matrix is an estimate of the inverse of the asymptotic covariance matrix of polychoric correlations, while

DWLS involves only the diagonal elements of that weight matrix. Several studies have been carried out to compare these three least square methods (e.g. Forero et al., 2009; Yang-Wallentin et al., 2010) and all have led to similar conclusions. The WLS estimator converges very slowly to its asymptotic properties and therefore does not perform well in small sample sizes. DWLS and ULS are preferable to WLS and they seem to perform similarly well in finite samples. However, in order to compute correct standard errors, the full weight matrix is needed. The methods are then called robust, which is applied to the beginning of the acronyms (hence: RULS, RDWLS). The advantages of the three-stage RDWLS and RULS estimators are that they are computationally less demanding than FIML. However, the estimate of the weight matrix is relatively unstable in small sample sizes.

2.3. Item Response Theory (IRT) approach

Under the IRT approach where conditional independence is assumed, the probability $\pi_r(\boldsymbol{\theta})$ is written as:

$$\pi_r(\boldsymbol{\theta}) = \int_{R_\xi} \pi_r(\boldsymbol{\theta}|\boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi} = \int_{R_\xi} \prod_{i=1}^{p} \pi(x_i = c_i; \boldsymbol{\theta}|\boldsymbol{\xi}) f(\boldsymbol{\xi}) d\boldsymbol{\xi} \ , \tag{6}$$

where $f(\boldsymbol{\xi})$ is the joint distribution of latent variables, usually assumed to be the k-dimensional standard normal density function, $\pi(x_i = c_i; \boldsymbol{\theta}|\boldsymbol{\xi})$ is the conditional response category probability which is given by

$$\pi(x_i = c_i; \boldsymbol{\theta}|\boldsymbol{\xi}) = \gamma(x_i \leq c_i; \boldsymbol{\theta}|\boldsymbol{\xi}) - \gamma(x_i \leq c_i - 1; \boldsymbol{\theta}|\boldsymbol{\xi}) \ , \tag{7}$$

and $\gamma(x_i \leq c_i; \boldsymbol{\theta}|\boldsymbol{\xi})$ is the cumulative probability of a response in category c_i, or below, for variable x_i. The cumulative probability is modeled as follows:

$$\text{link}\left(\gamma(x_i \leq c_i; \boldsymbol{\theta}|\boldsymbol{\xi})\right) = \alpha_{c_i}^{(x_i)} - \sum_{j=1}^{k} \beta_{ij} \xi_j \ , \tag{8}$$

where the $\alpha_{c_i}^{(x_i)}$'s are thresholds ($-\infty = \alpha_0^{(x_i)} < \alpha_1^{(x_i)} < \ldots < \alpha_{m_i-1}^{(x_i)} < \alpha_{m_i}^{(x_i)} = +\infty$), and the β_{ij}'s are factor loadings. The link function can be any monotonically increasing function mapping $(0, 1)$ onto $(-\infty, \infty)$, such as the logit (Samejima, 1969) or the inverse normal (also called probit). The log-likelihood is maximized using the E-M algorithm (Bartholomew et al., 2011; Muraki, 1990; Muraki & Carlson, 1995). FIML requires the evaluation of $k-$dimensional integrals as defined in (6). The integrals cannot be written in a closed form, but there are several numerical methods that can be used to approximate them (see Schilling & Bock, 2005, for a discussion of various methods). However, for all these methods, the computational burden increases rapidly with the number of factors k, rendering FIML quite impractical or even infeasible beyond a certain number of factors.

3. Proposed methodology

The promising results of Jöreskog & Moustaki (2001), de Leon (2005), and Liu (2007) motivated us to develop the pairwise maximum likelihood (PML) approach for factor analysis models with ordinal data. Using the results of composite likelihood theory (see e.g. Varin, 2008; Varin et al., 2011), the computation of standard errors becomes straightforward for PML. Furthermore, test statistics for inference and model selection criteria are also available (see e.g. Maydeu-Olivares & Joe, 2005, 2006; Joe & Maydeu-Olivares, 2010; Varin & Vidoni, 2005). A main difference between the work of de Leon (2005), and Liu (2007) and that of Jöreskog & Moustaki (2001) is that the former define their composite log-likelihood as the sum of the bivariate log-likelihoods, while the latter as the sum of both the bivariate and the univariate log-likelihoods. Hence, a first thing to investigate is the role of the sum of univariate log-likelihoods in determining the level of accuracy and efficiency of estimation. In factor analysis with ordinal data, the univariate log-likelihoods only contain information for thresholds, while the bivariate log-likelihoods contain information for all URV model parameters, i.e. $\boldsymbol{\lambda}$, $\boldsymbol{\varphi}$, and $\boldsymbol{\tau}$. Thus we ask the question: can this information overlap be balanced out by an "optimal" weighting scheme which improves the efficiency of estimation? Cox & Reid (2004) provide a general framework that can be used in order to deal with such a question. Their proposed log-likelihood is very similar to the UBN log-likelihood of Jöreskog & Moustaki (2001), with the only difference being that they put a weight on the sum of the univariate log-likelihoods. Following the suggestion of Cox & Reid (2004), the UBN log-likelihood could be modified as follows:

$$l(\boldsymbol{\theta}; \mathbf{x}) = \sum_{i<j} \ln L\left(\boldsymbol{\theta}; (x_i, x_j)\right) - ap \sum_{i} \ln L\left(\boldsymbol{\theta}; x_i\right) \ ,$$

where a is a constant to be chosen for optimal efficiency, p is the number of observed variables, $\ln L(\boldsymbol{\theta}; (x_i, x_j))$ is the bivariate marginal log-likelihood of variables x_i and x_j, and $\ln L(\boldsymbol{\theta}; x_i)$ is the univariate marginal log-likelihood of variable x_i. Cox & Reid (2004) point out that if the univariate likelihoods are independent of $\boldsymbol{\theta}$ then the choice of $a = 0$ is appropriate; taking $a = \frac{1}{2}$ corresponds to the situation where all possible conditional distributions of one variable, given another, are considered. In general, they suggest that a non-negative value of a is appropriate. Trying different values of a so that the value of ap ranges from 0 to 1, and conducting some small scale simulation studies, our results indicate that, practically, the sum of univariate log-likelihoods affect neither the accuracy nor the efficiency of estimation. Therefore, we conclude that the most appropriate choice of a is zero in our case. Subsequently, we suggest that one could consider the composite pairwise log-likelihood, $pl(\boldsymbol{\theta}; \mathbf{x})$, to estimate the URV parameter $\boldsymbol{\theta}' = (\boldsymbol{\lambda}, \boldsymbol{\varphi}, \boldsymbol{\tau})$. That is of the form:

$$pl(\boldsymbol{\theta}; \mathbf{x}) = \sum_{i<j} \ln L(\boldsymbol{\theta}; (x_i, x_j)) = \sum_{i<j} \sum_{c_i=1}^{m_i} \sum_{c_j=1}^{m_j} n_{c_i c_j}^{(x_i x_j)} \ln \pi_{c_i c_j}^{(x_i x_j)}(\boldsymbol{\theta}) , \tag{9}$$

where $n_{c_i c_j}^{(x_i x_j)}$ is the observed frequency of a response in category c_i and c_j for variables x_i and x_j, respectively, and $\pi_{c_i c_j}^{(x_i x_j)}(\boldsymbol{\theta})$ is the corresponding probability under the model. Based on equation (5), the latter is of the form:

$$\begin{aligned} \pi_{c_i c_j}^{(x_i x_j)}(\boldsymbol{\theta}) &= \pi(x_i = c_i, x_j = c_j; \boldsymbol{\theta}) = \\ &= \Phi_2\left(\tau_{c_i}^{(x_i)}, \tau_{c_j}^{(x_j)}; \rho_{x_i x_j}\right) - \Phi_2\left(\tau_{c_i}^{(x_i)}, \tau_{c_j-1}^{(x_j)}; \rho_{x_i x_j}\right) - \\ &\quad - \Phi_2\left(\tau_{c_i-1}^{(x_i)}, \tau_{c_j}^{(x_j)}; \rho_{x_i x_j}\right) + \Phi_2\left(\tau_{c_i-1}^{(x_i)}, \tau_{c_j-1}^{(x_j)}; \rho_{x_i x_j}\right) , \end{aligned} \tag{10}$$

where $\Phi_2(a, b; \rho)$ is the bivariate cumulative normal distribution with correlation ρ evaluated at the point (a, b),

$$\rho_{x_i x_j}(\boldsymbol{\theta}) = \boldsymbol{\lambda}_{i\cdot} \Phi \boldsymbol{\lambda}_{j\cdot}^{'} ,$$

and $\boldsymbol{\lambda}_{i\cdot}$ is a $1 \times k$ row vector containing the elements of the i^{th} row of matrix Λ, $i = 1, \ldots, p-1$, $j = i+1, \ldots, p$. Maximizing the log-likelihood function in (9) over the parameter $\boldsymbol{\theta}$ we get the composite pairwise maximum likelihood estimator $\hat{\boldsymbol{\theta}}_{PML}$. The gradient of the pairwise log-likelihood $\nabla pl(\boldsymbol{\theta}; \mathbf{x})$ is equal to the sum of the gradients of the bivariate log-likelihood components $\nabla \ln L(\boldsymbol{\theta}; (x_i, x_j))$. The explicit form of the latter is given in Appendix 1. Under regularity conditions upon the component likelihoods, the central limit theorem for the composite likelihood score statistic can be applied, leading to the result

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_{PML} - \boldsymbol{\theta}\right) \xrightarrow{d} N_q\left(0, G^{-1}(\boldsymbol{\theta})\right) ,$$

where q is the dimension of $\boldsymbol{\theta}$, and $G(\boldsymbol{\theta})$ is the Godambe information (also known as sandwich information) matrix of a single observation (Varin, 2008; Varin et al., 2011). In particular,

$$G(\boldsymbol{\theta}) = H(\boldsymbol{\theta}) J^{-1}(\boldsymbol{\theta}) H(\boldsymbol{\theta}) ,$$

where $H(\boldsymbol{\theta})$ is the sensitivity matrix, $H(\boldsymbol{\theta}) = E\left\{-\nabla^2 pl(\boldsymbol{\theta}; \mathbf{x})\right\}$, and $J(\boldsymbol{\theta})$ is the variability matrix, $J(\boldsymbol{\theta}) = Var\{\nabla pl(\boldsymbol{\theta}; \mathbf{x})\}$. In general, the identity $H(\boldsymbol{\theta}) = -J(\boldsymbol{\theta})$ does not hold in the case of composite likelihoods. The assumed independence among the likelihood components forming the composite likelihood is usually not valid when the full likelihood is considered. The sample estimates of $H(\boldsymbol{\theta})$ and $J(\boldsymbol{\theta})$ are

$$\hat{H}(\hat{\boldsymbol{\theta}}_{PML}) = \nabla^2 pl(\hat{\boldsymbol{\theta}}_{PML}; \mathbf{x}) , \text{ and}$$

$$\hat{J}(\hat{\boldsymbol{\theta}}_{PML}) = \frac{1}{n} \sum_{h=1}^{n} \left(\nabla pl(\boldsymbol{\theta}; \mathbf{x}_h)\right) \left(\nabla pl(\boldsymbol{\theta}; \mathbf{x}_h)\right)^T .$$

The obvious advantage of PML over FIML estimation is that it only requires the evaluation of up to two-dimensional normal probabilities, regardless of the number of observed or latent variables. In this way, it is always computationally feasible. The advantage of the PML approach when compared with the three-stage limited information estimation methods is that the estimation of all parameters is carried out simultaneously. Moreover, the standard errors of the estimates can be obtained without the usage of any weight matrix. Varin et al. (2011) discuss some other qualities of the composite likelihood approach. Composite likelihood can be seen as a robust alternative in terms of modeling. In some cases it is easier and more straightforward to model lower order dimensional distributions, while modeling uncertainty increases with dimensionality. By applying composite likelihood, possible misspecification of the higher order dimensional distributions can be avoided. In addition, a model assumed for lower order distributions can be compatible with more than one of the possible model options available for higher dimensional distributions. Moreover, in some settings, there are no obvious high dimensional distributions.

In the current study, the maximization of $pl(\boldsymbol{\theta};\mathbf{x})$ has been carried out by using the "maxLik" command in the "maxLik" package of the 2.10.1 version of R software. As an input to the command we have specified two "function objects" (according to R terminology), which we wrote. One function is for the composite pairwise log-likelihood as expressed in (9), and the other is for the gradient given in Appendix 1. The "maxLik" command offers several options regarding the maximization algorithm. We have used one of the group of quasi-newton methods, namely the Broyden-Fletcher-Goldfarb-Shanno method (denoted as BFGS in R software). Regarding the estimate $\hat{H}(\hat{\boldsymbol{\theta}}_{PML})$, it is part of the output of "maxLik" command. Concerning $\hat{J}(\hat{\boldsymbol{\theta}}_{PML})$, we wrote our own function object.

4. Simulation study on the performance of PML estimator

4.1. Simulation study set-up

A simulation study has been conducted to evaluate the performance of the PML estimator within the framework of confirmatory factor analysis with ordinal data. Eight experimental conditions have been investigated by combining four sample sizes, namely 100, 200, 500, and 1000, with two model sizes, which are referred to as Model I and Model II, and are detailed below. For each condition, 1000 replications have been carried out.

Model I is regarded as a small size model with 6 observed variables ($p = 6$) and 2 factors ($k = 2$). The number of indicators for each factor is relatively small: 3 or 4 indicators for each factor. The true values of matrices Λ and Φ are:

$$\Lambda = \begin{pmatrix} 0.9 & \\ 0.8 & \\ 0.7 & \\ 0.5 & 0.6 \\ & 0.7 \\ & 0.8 \end{pmatrix}, \; \Phi = \begin{pmatrix} 1 & \\ 0.5 & 1 \end{pmatrix}.$$

As it can be seen, the loadings range from high (0.9) to relatively low (0.5), and there is one multidimensional indicator, the loadings of which are relatively small and close in value (0.5 and 0.6). The factor correlation is of moderate size (0.5). All observed variables are assumed to have 4 response categories, a case quite often met in applications, and the same thresholds, namely $\tau_1^{(x_i)} = -1.2$, $\tau_2^{(x_i)} = 0$, $\tau_3^{(x_i)} = 1.2$, $i = 1, \ldots, 6$. Therefore, there are 26 free parameters to be estimated for Model I.

Model II is of larger size, with 15 observed variables ($p = 15$) and 3 factors ($k = 3$). The number of indicators per factor is moderate (5 or 6). The true values of matrices Λ and Φ are:

$$\Lambda = \begin{pmatrix} 0.4 & & \\ 0.5 & & \\ 0.6 & & \\ 0.7 & & \\ 0.8 & & \\ 0.3 & 0.8 & \\ & 0.7 & \\ & 0.6 & \\ & 0.5 & \\ & 0.4 & 0.5 \\ & & 0.6 \\ & & 0.7 \\ & & 0.8 \\ & & 0.9 \\ & & 0.4 \end{pmatrix}, \; \Phi = \begin{pmatrix} 1 & & \\ 0.2 & 1 & \\ 0.5 & 0.8 & 1 \end{pmatrix}.$$

The loadings now range from 0.3 to 0.9. In this model, there are two multidimensional indicators; the first of which has one high loading (0.8), and one low (0.3), while the second has loadings that are both low and close in value (0.4 and 0.5), thus presenting a similar pattern as in Model I. The factor correlations range from low to high (0.2, 0.5, and 0.8 respectively). Again all observed variables are assumed to have 4 response categories and the same thresholds as in Model I. There are 65 free parameters to be estimated in the larger model.

4.2. Data generation

Within each replication, the data are generated as follows:

1. A random vector $\boldsymbol{\xi}$ and a random vector $\boldsymbol{\delta}$ are generated from $N_k(\mathbf{0}, \Phi)$ and $N_p(\mathbf{0}, \Theta)$, respectively.
2. A random vector of underlying variables $\mathbf{x}^\star$ is generated by applying the assumed model $\mathbf{x}^\star = \Lambda\boldsymbol{\xi} + \boldsymbol{\delta}$.
3. A random vector of ordinal variables $\mathbf{x}$ is obtained from $\mathbf{x}^\star$ by applying the relationship in (3), which connects the continuous underlying variables with the ordinal observed variables. Thus, the values of thresholds $\tau_{c_i}^{(x_i)}$ are used in this step.
4. Steps 1-3 are repeated n times to get a sample of size n.

4.3. Performance Criteria

We compute the bias and mean squared error (MSE) for each parameter as follows:

$$Bias = \frac{1}{R}\sum_{i=1}^{R}\left(\hat{\theta}_i - \theta\right) ,$$

and

$$MSE = \frac{1}{R}\sum_{i=1}^{R}\left(\hat{\theta}_i - \theta\right)^2 ,$$

where R here is the number of valid replicates, $\hat{\theta}_i$ is the estimate of a parameter or of its asymptotic standard error at the i^{th} valid replication, and θ is the corresponding true value. In the case of standard errors, where the true value θ is unknown, the standard deviation of parameter estimates across valid replications is used.

4.4. Results

First we report in Table 1 the percentage of replications per condition that produced a proper solution, i.e. estimated loadings and factor correlations between -1 and 1. For the smaller model (Model I), the percentage is 100% for all sample sizes except for $n = 100$. However, for Model II, the percentage ranges from 89.4% to 99.3%, increasing as the sample size increases. It is worthwhile to note that, for the sample sizes 200, 500, and 1000, improper solutions only occurred for the loadings of the second multidimensional indicator, $\lambda_{10,2}$ and $\lambda_{10,3}$, and in some cases for the correlation between the 2nd and 3rd factor, ϕ_{23}. Interestingly, Forero et al. (2009) comment that “convergence and estimation accuracy problems are aggravated in the presence of multidimensional indicators”. In our study, the related indicator has, in addition, relatively low and close loadings on the two dimensions. For sample size 100, improper solutions occurred for other parameters as well. The final results only consider replications with proper solutions as in Forero & Maydeu-Olivares (2009) and Forero et al. (2009).

Sample size n	Model I	Model II
100	96.6%	89.4%
200	100%	94.6%
500	100%	98.2%
1000	100%	99.3%

Table 1: Percentage of valid replications per condition, PML

We present the results for the factor loadings and factor correlations, although our conclusions are valid for the thresholds as well. Tables 2 - 4 give for each parameter: the true value, the average parameter estimate across valid replications, the bias, the MSE, the standard deviation of the parameter estimate across valid replications, the average standard error of a parameter across valid replications and its corresponding bias and MSE. Figures 1 and 2 depict the bias and MSE of both estimates and their standard errors for all sample sizes for Model I and Model II, respectively. On the horizontal axis of the graphs, the parameters are denoted with an index, and are presented in the same order as in Tables 2 - 4. From the tables and the figures, we conclude that for both models, the PML parameter estimates and their estimated asymptotic standard errors have bias and MSE close to zero, decreasing with the increase of the sample size. Concerning the estimated asymptotic standard errors, the bias is negative in most cases, i.e. the standard deviations computed from the replications are slightly bigger than the average of estimated asymptotic standard errors. In Model I, the MSE of both estimates and standard errors are relatively higher for the loadings of the multidimensional indicator, λ_{41} and λ_{42} (parameter indices 4 and 5 respectively, Figure 1). A similar pattern occurs in Model II regarding the loadings of the second multidimensional indicator, $\lambda_{10,2}$ and $\lambda_{10,3}$, (parameter indices 11 and 12 respectively, Figure 2).

		Estimate				Estimated standard error		
	True	Mean	Bias	MSE	Standard Deviation	Mean	Bias	MSE
$n = 100$								
λ_{11}	0.9	0.897	-0.003	0.002	0.049	0.050	0.001	0.000114
λ_{21}	0.8	0.803	0.003	0.003	0.058	0.058	0.000	0.000132
λ_{31}	0.7	0.703	0.003	0.005	0.072	0.070	-0.002	0.000166
λ_{41}	0.5	0.497	-0.003	0.012	0.109	0.106	-0.003	0.001183
λ_{42}	0.6	0.605	0.005	0.013	0.112	0.109	-0.004	0.001175
λ_{52}	0.7	0.703	0.003	0.007	0.085	0.083	-0.002	0.000261
λ_{62}	0.8	0.804	0.004	0.007	0.085	0.082	-0.002	0.000379
ϕ_{12}	0.5	0.509	0.009	0.013	0.113	0.113	0.000	0.000281
$n = 200$								
λ_{11}	0.9	0.901	0.001	0.001	0.035	0.034	-0.000	0.000026
λ_{21}	0.8	0.803	0.003	0.002	0.044	0.041	-0.003	0.000041
λ_{31}	0.7	0.701	0.001	0.003	0.051	0.050	-0.001	0.000042
λ_{41}	0.5	0.500	-0.000	0.006	0.076	0.072	-0.004	0.000244
λ_{42}	0.6	0.600	0.000	0.006	0.076	0.073	-0.003	0.000225
λ_{52}	0.7	0.704	0.004	0.004	0.059	0.059	-0.001	0.000059
λ_{62}	0.8	0.803	0.003	0.004	0.059	0.057	-0.002	0.000079
ϕ_{12}	0.5	0.504	0.004	0.007	0.083	0.080	-0.002	0.000073
$n = 500$								
λ_{11}	0.9	0.900	0.000	0.001	0.022	0.022	0.000	0.000004
λ_{21}	0.8	0.801	0.001	0.001	0.026	0.026	0.000	0.000005
λ_{31}	0.7	0.701	0.001	0.001	0.031	0.032	0.001	0.000007
λ_{41}	0.5	0.497	-0.003	0.002	0.044	0.045	0.001	0.000023
λ_{42}	0.6	0.602	0.002	0.002	0.046	0.045	-0.001	0.000022
λ_{52}	0.7	0.699	-0.001	0.002	0.038	0.037	-0.001	0.000010
λ_{62}	0.8	0.801	0.001	0.001	0.036	0.036	0.000	0.000011
ϕ_{12}	0.5	0.503	0.003	0.003	0.052	0.051	0.001	0.000012
$n = 1000$								
λ_{11}	0.9	0.901	0.001	0.000	0.016	0.016	-0.001	0.000001
λ_{21}	0.8	0.800	0.000	0.000	0.018	0.019	0.000	0.000001
λ_{31}	0.7	0.700	-0.000	0.001	0.022	0.023	0.001	0.000002
λ_{41}	0.5	0.500	-0.001	0.001	0.032	0.031	-0.000	0.000006
λ_{42}	0.6	0.601	0.001	0.001	0.032	0.031	-0.001	0.000006
λ_{52}	0.7	0.700	-0.001	0.001	0.026	0.026	0.000	0.000002
λ_{62}	0.8	0.801	0.001	0.001	0.027	0.026	-0.001	0.000004
ϕ_{12}	0.5	0.500	0.000	0.001	0.038	0.036	-0.001	0.000005

Table 2: Simulation results for Model I, PML

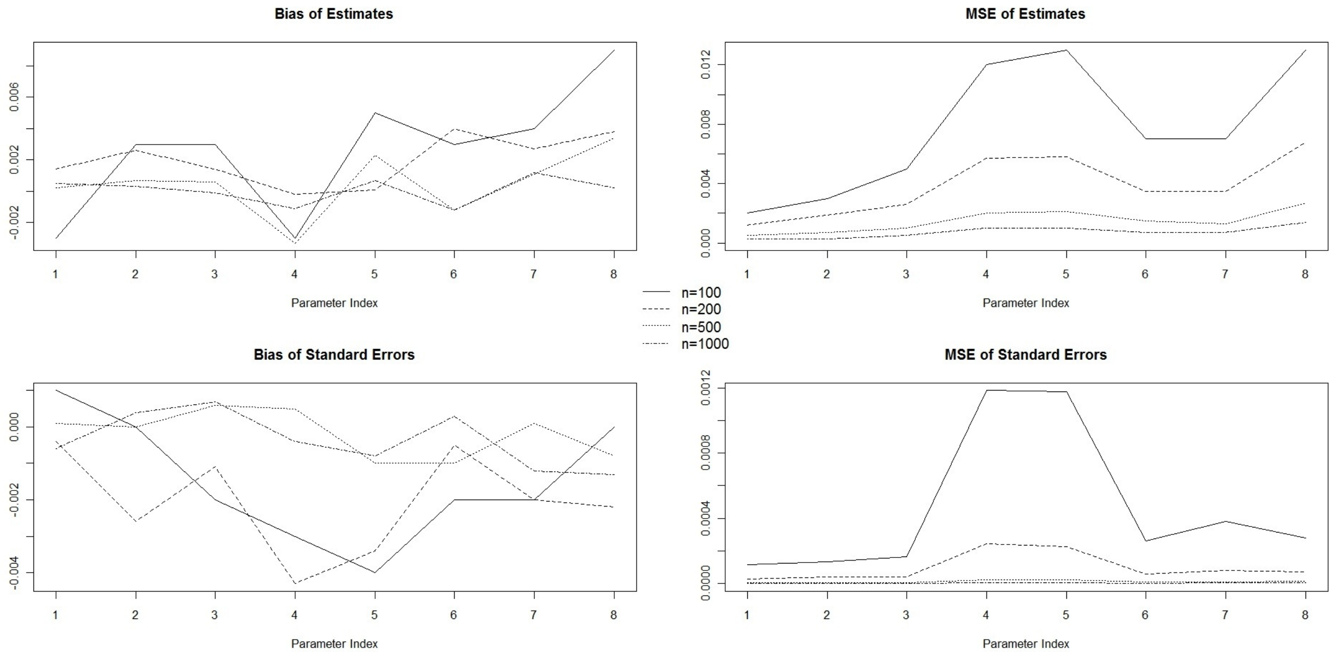

Figure 1: Bias and MSE of estimates and standard errors for all sample sizes, Model I

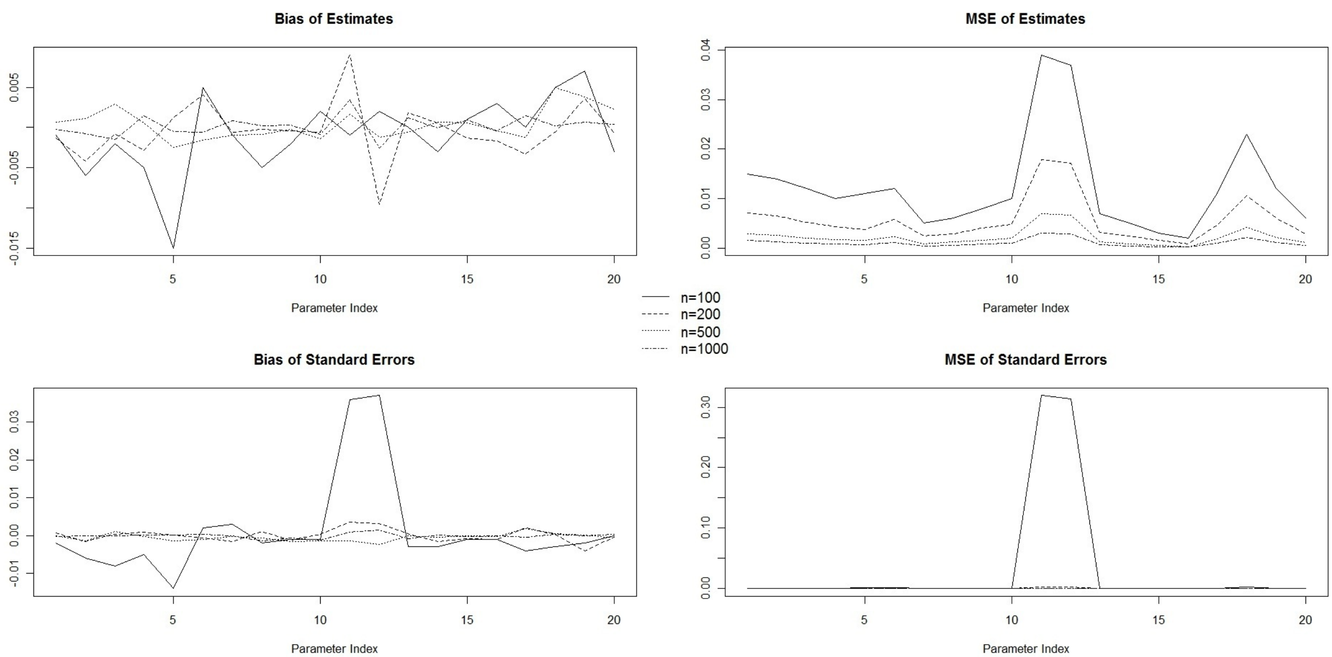

Figure 2: Bias and MSE of estimates and standard errors for all sample sizes, Model II

		Estimate				Estimated standard error		
	True	Mean	Bias	MSE	Standard Deviation	Mean	Bias	MSE
$n = 100$								
$\lambda_{1,1}$	0.4	0.399	-0.001	0.015	0.122	0.121	-0.002	0.000359
$\lambda_{2,1}$	0.5	0.494	-0.006	0.014	0.119	0.113	-0.006	0.000454
$\lambda_{3,1}$	0.6	0.598	-0.002	0.012	0.111	0.103	-0.008	0.000435
$\lambda_{4,1}$	0.7	0.695	-0.005	0.010	0.101	0.096	-0.005	0.000489
$\lambda_{5,1}$	0.8	0.785	-0.015	0.011	0.106	0.092	-0.014	0.000979
$\lambda_{6,1}$	0.3	0.305	0.005	0.012	0.110	0.112	0.002	0.001887
$\lambda_{6,2}$	0.8	0.799	-0.001	0.005	0.068	0.071	0.003	0.000280
$\lambda_{7,2}$	0.7	0.695	-0.005	0.006	0.079	0.078	-0.002	0.000371
$\lambda_{8,2}$	0.6	0.598	-0.002	0.008	0.089	0.088	-0.001	0.000281
$\lambda_{9,2}$	0.5	0.502	0.002	0.010	0.099	0.098	-0.001	0.000226
$\lambda_{10,2}$	0.4	0.399	-0.001	0.039	0.199	0.234	0.036	0.319323
$\lambda_{10,3}$	0.5	0.502	0.002	0.037	0.192	0.229	0.037	0.313953
$\lambda_{11,3}$	0.6	0.600	0.000	0.007	0.082	0.079	-0.003	0.000162
$\lambda_{12,3}$	0.7	0.697	-0.003	0.005	0.070	0.067	-0.003	0.000154
$\lambda_{13,3}$	0.8	0.801	0.001	0.003	0.054	0.053	-0.001	0.000126
$\lambda_{14,3}$	0.9	0.903	0.003	0.002	0.041	0.039	-0.001	0.000082
$\lambda_{15,3}$	0.4	0.400	0.000	0.011	0.103	0.099	-0.004	0.000162
ϕ_{12}	0.2	0.205	0.005	0.023	0.153	0.150	-0.003	0.002229
ϕ_{13}	0.5	0.507	0.007	0.012	0.107	0.105	-0.002	0.000356
ϕ_{23}	0.8	0.797	-0.003	0.006	0.075	0.076	0.000	0.000474
$n = 200$								
$\lambda_{1,1}$	0.4	0.399	-0.001	0.007	0.084	0.085	0.001	0.000074
$\lambda_{2,1}$	0.5	0.496	-0.004	0.007	0.081	0.079	-0.002	0.000080
$\lambda_{3,1}$	0.6	0.599	-0.001	0.005	0.072	0.072	0.000	0.000079
$\lambda_{4,1}$	0.7	0.697	-0.003	0.004	0.065	0.066	0.001	0.000077
$\lambda_{5,1}$	0.8	0.801	0.001	0.004	0.061	0.062	0.000	0.000077
$\lambda_{6,1}$	0.3	0.304	0.004	0.006	0.076	0.075	-0.001	0.000094
$\lambda_{6,2}$	0.8	0.799	-0.001	0.002	0.049	0.047	-0.002	0.000040
$\lambda_{7,2}$	0.7	0.700	0.000	0.003	0.053	0.054	0.001	0.000051
$\lambda_{8,2}$	0.6	0.600	0.000	0.004	0.063	0.062	-0.001	0.000054
$\lambda_{9,2}$	0.5	0.499	-0.001	0.005	0.069	0.069	0.000	0.000055
$\lambda_{10,2}$	0.4	0.409	0.009	0.018	0.134	0.137	0.004	0.002527
$\lambda_{10,3}$	0.5	0.490	-0.010	0.017	0.131	0.134	0.003	0.002486
$\lambda_{11,3}$	0.6	0.602	0.002	0.003	0.056	0.056	0.001	0.000040
$\lambda_{12,3}$	0.7	0.701	0.001	0.002	0.049	0.047	-0.002	0.000039
$\lambda_{13,3}$	0.8	0.799	-0.001	0.002	0.038	0.037	-0.001	0.000029
$\lambda_{14,3}$	0.9	0.898	-0.002	0.001	0.029	0.028	-0.001	0.000019
$\lambda_{15,3}$	0.4	0.397	-0.003	0.005	0.068	0.070	0.002	0.000038
ϕ_{12}	0.2	0.199	-0.001	0.011	0.103	0.104	0.000	0.000091
ϕ_{13}	0.5	0.504	0.004	0.006	0.077	0.073	-0.004	0.000077
ϕ_{23}	0.8	0.799	-0.001	0.003	0.053	0.052	-0.000	0.000066

Table 3: Simulation results for Model II, $n = 100$ & $n = 200$, PML

	Estimate					Estimated standard error		
	True	Mean	Bias	MSE	Standard Deviation	Mean	Bias	MSE
$n = 500$								
$\lambda_{1,1}$	0.4	0.401	0.001	0.003	0.054	0.054	0.000	0.000011
$\lambda_{2,1}$	0.5	0.501	0.001	0.003	0.051	0.050	-0.001	0.000014
$\lambda_{3,1}$	0.6	0.603	0.003	0.002	0.044	0.045	0.001	0.000014
$\lambda_{4,1}$	0.7	0.701	0.001	0.002	0.042	0.041	-0.000	0.000012
$\lambda_{5,1}$	0.8	0.798	-0.003	0.002	0.040	0.039	-0.001	0.000014
$\lambda_{6,1}$	0.3	0.298	-0.002	0.002	0.048	0.047	-0.001	0.000014
$\lambda_{6,2}$	0.8	0.799	-0.001	0.001	0.029	0.029	-0.000	0.000005
$\lambda_{7,2}$	0.7	0.699	-0.001	0.001	0.040	0.034	-0.007	0.000009
$\lambda_{8,2}$	0.6	0.600	0.000	0.002	0.041	0.039	-0.002	0.000011
$\lambda_{9,2}$	0.5	0.499	-0.001	0.002	0.045	0.044	-0.001	0.000011
$\lambda_{10,2}$	0.4	0.402	0.002	0.007	0.083	0.082	-0.001	0.000149
$\lambda_{10,3}$	0.5	0.499	-0.001	0.007	0.082	0.080	-0.002	0.000157
$\lambda_{11,3}$	0.6	0.599	-0.001	0.001	0.036	0.036	-0.000	0.000006
$\lambda_{12,3}$	0.7	0.701	0.001	0.001	0.030	0.030	-0.001	0.000006
$\lambda_{13,3}$	0.8	0.801	0.001	0.001	0.024	0.024	-0.000	0.000004
$\lambda_{14,3}$	0.9	0.900	0.000	0.000	0.018	0.018	-0.000	0.000003
$\lambda_{15,3}$	0.4	0.399	-0.001	0.002	0.043	0.044	0.002	0.000009
ϕ_{12}	0.2	0.205	0.005	0.004	0.065	0.065	0.001	0.000014
ϕ_{13}	0.5	0.504	0.004	0.002	0.046	0.046	0.000	0.000009
ϕ_{23}	0.8	0.802	0.002	0.001	0.033	0.033	-0.001	0.000010
$n = 1000$								
$\lambda_{1,1}$	0.4	0.400	-0.000	0.002	0.038	0.038	-0.000	0.000003
$\lambda_{2,1}$	0.5	0.499	-0.001	0.001	0.035	0.035	-0.000	0.000003
$\lambda_{3,1}$	0.6	0.599	-0.002	0.001	0.032	0.032	0.000	0.000003
$\lambda_{4,1}$	0.7	0.702	0.002	0.001	0.029	0.029	0.000	0.000003
$\lambda_{5,1}$	0.8	0.800	-0.001	0.001	0.027	0.027	0.000	0.000003
$\lambda_{6,1}$	0.3	0.299	-0.001	0.001	0.033	0.033	0.000	0.000003
$\lambda_{6,2}$	0.8	0.801	0.001	0.000	0.021	0.020	-0.000	0.000001
$\lambda_{7,2}$	0.7	0.700	0.000	0.001	0.025	0.024	-0.001	0.000004
$\lambda_{8,2}$	0.6	0.600	0.000	0.001	0.028	0.028	-0.001	0.000003
$\lambda_{9,2}$	0.5	0.499	-0.001	0.001	0.032	0.031	-0.001	0.000003
$\lambda_{10,2}$	0.4	0.403	0.003	0.003	0.055	0.056	0.001	0.000028
$\lambda_{10,3}$	0.5	0.497	-0.003	0.003	0.053	0.055	0.001	0.000030
$\lambda_{11,3}$	0.6	0.601	0.001	0.001	0.026	0.025	-0.001	0.000002
$\lambda_{12,3}$	0.7	0.700	0.000	0.000	0.021	0.021	0.000	0.000001
$\lambda_{13,3}$	0.8	0.801	0.001	0.000	0.017	0.017	-0.000	0.000001
$\lambda_{14,3}$	0.9	0.900	0.000	0.000	0.012	0.012	0.000	0.000001
$\lambda_{15,3}$	0.4	0.402	0.002	0.001	0.032	0.031	-0.001	0.000002
ϕ_{12}	0.2	0.200	0.000	0.002	0.046	0.046	0.000	0.000004
ϕ_{13}	0.5	0.501	0.001	0.001	0.033	0.033	-0.000	0.000002
ϕ_{23}	0.8	0.800	0.000	0.001	0.023	0.023	0.000	0.000003

Table 4: Simulation results for Model II, $n = 500$ & $n = 1000$, PML

5. A comparison of FIML, PML, 3S-RULS, 3S-RDWLS estimators with a simulation study

In this section, the performances of PML, FIML, 3S-RULS, and 3S-RDWLS estimators are compared using the same simulated samples as in the previous simulation study and under the same experimental conditions. The FIML estimates and standard errors are derived under the IRT approach with the probit link. Although the parameters in this case are the $\alpha_{c_i}^{(x_i)}$ and β_{ij} given in equation (8), the corresponding estimates of λ_{ij} and $\tau_{c_i}^{(x_i)}$ can be obtained easily. This is because the two sets of parameters are proved to be equivalent (see Takane & de Leeuw, 1987, and Bartholomew et al., 2011), and the related formulas can be found in Jöreskog & Moustaki (2001). All statistical packages provide both sets of parameters. The package Mplus 6.1 has been used to derive the FIML estimates and standard errors. The numerical algorithm used to evaluate the integral in (6) was the default of Mplus, i.e. adaptive quadrature with 15 integration points for each dimension. To get the 3S-RULS and 3S-RDWLS estimates and standard errors either LISREL 8.80 or Mplus can be used. Both packages give very similar results. The 3S-RULS estimation is denoted as ULS in LISREL and ULSMV in Mplus, while 3S-RDWLS is denoted as DWLS in LISREL and WLSMV in Mplus.

Table 5 gives the percentage of valid replications per condition for each estimator. For Model I, replication with improper solutions occur for all estimation methods when the sample size is 100, the lowest percentage being that of 3S-RULS, 95.3%. For the rest sample sizes (200, 500, and 1000) PML has 100% valid replications, while the other three methods have a very minor percentage of improper solutions (of around 0.3%) for $n = 200$ and 100% valid replications for sizes 500 and 1000. For Model II, all methods give quite a few improper solutions when $n = 100$, indicating that this is a relatively small sample size compared to the model size. The lowest percentage is that of PML, 89.4%. The percentage of improper solutions decreases with the increase of the sample size. All methods reach 100% of valid replications for sizes 500 and 1000 except for PML, which reaches 99.3% when $n = 1000$. For the performance comparisons, only the replications for which all four methods give valid solutions have been used.

	Model I				Model II			
Sample size n	FIML	PML	3S-RULS	3S-RDWLS	FIML	PML	3S-RULS	3S-RDWLS
100	98.7%	96.6%	95.3%	96.6%	95.9%	89.4%	92.7%	94.5%
200	99.6%	100%	99.7%	99.8%	99.7%	94.6%	99.5%	99.7%
500	99.9%	100%	100%	100%	100%	98.2%	100%	100%
1000	100%	100%	100%	100%	100%	99.3%	100%	100%

Table 5: Percentage of valid replications per condition for FIML, PML, 3S-RULS, 3S-RDWLS

The performance of the estimators is compared on the basis of the following criteria: bias, MSE and standard deviation of parameter estimates, as well as average, bias and MSE of estimated standard errors. To facilitate the comparisons, instead of presenting the results for each model parameter as in Section 4, we calculated the average, over all λ and ϕ parameters, of the aforementioned performance criteria. The results are reported in Table 6. All methods give very close results. As expected, FIML performs slightly better with respect to almost all criteria with almost all experimental conditions, and of course, it presents the smaller average standard deviation of estimates and average of estimated standard errors. The only case where FIML presents slightly bigger average mean of estimated standard errors is within Model I when $n = 100$. FIML presents slightly worse average MSE for the standard errors for the small sample sizes (100 and 200). This is due to the relatively worse performance for the standard errors of the loadings of the multidimensional indicators. Concerning the other three methods, the results of PML and 3S-RDWLS are, on average, slightly closer to those of the FIML approach than the results of the 3S-RULS.

	Estimate			Estimated standard error		
	Average Bias	Average MSE	Av. Standard Deviation	Average Mean	Average Bias	Average MSE
Model I						
$n = 100$ - Results based on 937 (93.7%) valid replicates						
FIML	0.003	0.007	0.080	0.088	0.009	0.015090
PML	0.002	0.008	0.085	0.084	-0.001	0.000464
3S-RULS	0.002	0.008	0.088	0.088	-0.001	0.000472
3S-RDWLS	0.003	0.008	0.086	0.084	-0.002	0.000509
$n = 200$ - Results based on 993 (99.3%) valid replicates						
FIML	0.002	0.004	0.057	0.057	-0.001	0.001095
PML	0.002	0.004	0.060	0.058	-0.002	0.000098
3S-RULS	0.002	0.004	0.062	0.061	-0.002	0.000095
3S-RDWLS	0.003	0.004	0.060	0.058	-0.002	0.000101
$n = 500$ - Results based on 999 (99.9%) valid replicates						
FIML	0.001	0.001	0.036	0.035	-0.000	0.000031
PML	0.001	0.002	0.037	0.037	-0.000	0.000012
3S-RULS	0.001	0.002	0.039	0.038	-0.000	0.000012
3S-RDWLS	0.001	0.002	0.037	0.037	-0.000	0.000012
$n = 1000$ - Results based on 1000 (100%) valid replicates						
FIML	0.000	0.001	0.026	0.025	-0.000	0.000003
PML	0.000	0.001	0.026	0.026	-0.000	0.000003
3S-RULS	0.000	0.001	0.027	0.027	-0.000	0.000004
3S-RDWLS	0.000	0.001	0.026	0.026	-0.000	0.000004
Model II						
$n = 100$ - Results based on 864 (86.4%) valid replicates						
FIML	0.000	0.011	0.097	0.099	0.002	0.004104
PML	-0.001	0.012	0.103	0.102	-0.001	0.001150
3S-RULS	-0.000	0.012	0.103	0.099	-0.004	0.000758
3S-RDWLS	0.003	0.012	0.101	0.094	-0.007	0.000686
$n = 200$ - Results based on 942 (94.2%) valid replicates						
FIML	0.000	0.005	0.067	0.067	0.000	0.000181
PML	-0.000	0.006	0.070	0.070	0.000	0.000284
3S-RULS	-0.000	0.006	0.073	0.071	-0.002	0.000192
3S-RDWLS	0.001	0.006	0.071	0.068	-0.003	0.000147
$n = 500$ - Results based on 982 (98.2%) valid replicates						
FIML	0.001	0.002	0.042	0.042	-0.000	0.000020
PML	0.000	0.002	0.044	0.044	-0.001	0.000024
3S-RULS	0.000	0.003	0.046	0.045	-0.001	0.000028
3S-RDWLS	0.001	0.002	0.045	0.043	-0.001	0.000022
$n = 1000$ - Results based on 993 (99.3%) valid replicates						
FIML	0.000	0.001	0.030	0.029	0.000	0.000005
PML	0.000	0.001	0.031	0.031	-0.000	0.000005
3S-RULS	0.000	0.001	0.032	0.032	-0.000	0.000006
3S-RDWLS	0.001	0.001	0.031	0.031	-0.000	0.000005

Table 6: Average, over λ and ϕ parameters, results for FIML, PML, 3S-RULS, 3S-RDWLS

6. Comparison of PML, FIML, 3S-RULS, 3S-RDWLS using empirical data

In this section, the PML estimation is demonstrated by applying the method in exploratory and confirmatory factor analysis, and using some empirical data. Moreover, the PML estimates and standard errors are compared with those obtained by FIML, 3S-RULS, and 3S-RDWLS approaches. Again, the FIML estimates and standard errors are derived under the IRT approach with probit as the link function where an adaptive quadrature with 15 integration points for each dimension has been employed to evaluate the integral in (6). However, in the example of the four-factor model, this numerical method is remarkably slow and a Monte Carlo method with the default settings of Mplus has been used. In all the examples, we focus again on the estimates of factor loadings and correlations.

6.1. Exploratory factor analysis - Science & Technology (S&T) data

The data used in this example come from the Consumer Protection and Perceptions of Science and Technology section of the 1992 Eurobarometer Survey (Karlheinz & Melich, 1992) and particularly, it is based on a sample from Great Britain. Seven indicators are used in the current analysis, all presented in Appendix 2. All the indicators were measured on a four-point scale with response categories "strongly disagree", "disagree to some extent", "agree to some extent", and "strongly agree". Almost all indicators present a considerable amount of skewness. The sample size is 392 after eliminating the cases with missing values in any of the indicators (listwise deletion). Exploratory factor analysis with one and two factors have been carried out where the factors are assumed to follow standard normal distribution. For the two-factor exploratory analysis, the loading of the first indicator on the second factor, namely λ_{12}, has been fixed to 0 for identification reasons. The estimates of factor loadings obtained by the four estimation methods in the case of one factor analysis are given in Table 7. As it can be seen, all methods give fairly close parameter estimates and standard errors. However, the PML estimates are on average slightly closer to those of the FIML than the 3S-RULS and 3S-RDWLS estimates. An interesting pattern regarding the standard errors is that 3S-RULS and 3S-RDWLS give smaller standard errors for all parameters compared with those of FIML and PML approaches.

Parameter	FIML	PML	3S-RULS	3S-RDWLS
λ_1	0.491	0.536	0.568	0.580
	(0.102)	(0.121)	(0.072)	(0.069)
λ_2	-0.018	0.044	0.069	0.196
	(0.128)	(0.215)	(0.081)	(0.077)
λ_3	0.548	0.503	0.464	0.448
	(0.070)	(0.119)	(0.069)	(0.067)
λ_4	0.795	0.752	0.713	0.670
	(0.084)	(0.135)	(0.059)	(0.058)
λ_5	-0.023	0.042	0.068	0.203
	(0.123)	(0.214)	(0.076)	(0.071)
λ_6	0.139	0.192	0.217	0.331
	(0.122)	(0.209)	(0.077)	(0.071)
λ_7	0.511	0.538	0.562	0.530
	(0.085)	(0.083)	(0.065)	(0.064)

Table 7: Estimated loadings and standard errors (in brackets) for S&T data, one-factor exploratory analysis

Table 8 reports the results of the exploratory analysis with two factors. As with one factor analysis, all four estimation methods give very similar estimates. However, the PML estimates and standard errors are slightly closer to those of the FIML than those of the 3S-RULS and 3S-RDWLS approaches. Again, the 3S-RULS and 3S-RDWLS approaches give the smaller standard errors among the four methods for almost all parameter estimates.

Parameter	FIML	PML	3S-RULS	3S-RDWLS	Parameter	FIML	PML	3S-RULS	3S-RDWLS
λ_{11}	0.529	0.545	0.556	0.561	λ_{12}	0* (fixed)	0* (fixed)	0* (fixed)	0* (fixed)
	(0.081)	(0.076)	(0.069)	(0.068)		—	—	—	—
λ_{21}	0.197	0.183	0.174	0.176	λ_{22}	0.637	0.633	0.634	0.634
	(0.121)	(0.116)	(0.111)	(0.109)		(0.074)	(0.074)	(0.074)	(0.073)
λ_{31}	0.479	0.464	0.449	0.469	λ_{32}	-0.312	-0.297	-0.283	-0.292
	(0.079)	(0.080)	(0.079)	(0.079)		(0.101)	(0.100)	(0.087)	(0.088)
λ_{41}	0.731	0.719	0.707	0.715	λ_{42}	-0.285	-0.279	-0.261	-0.267
	(0.070)	(0.069)	(0.063)	(0.063)		(0.131)	(0.127)	(0.112)	(0.112)
λ_{51}	0.206	0.189	0.179	0.181	λ_{52}	0.675	0.674	0.675	0.675
	(0.120)	(0.116)	(0.111)	(0.110)		(0.073)	(0.071)	(0.070)	(0.070)
λ_{61}	0.369	0.353	0.345	0.344	λ_{62}	0.554	0.549	0.552	0.552
	(0.102)	(0.100)	(0.098)	(0.096)		(0.085)	(0.083)	(0.082)	(0.082)
λ_{71}	0.493	0.510	0.532	0.524	λ_{72}	-0.153	-0.135	-0.140	-0.134
	(0.083)	(0.077)	(0.066)	(0.065)		(0.080)	(0.077)	(0.087)	(0.086)

Table 8: Estimated loadings and their standard errors (in brackets) for S&T data, two-factor exploratory analysis

6.2. Confirmatory factor analysis - Relationship Learning (RL) data

In this example, we use part of the data gathered by Selnes & Sallis (2003) who aimed to study whether specific factors affect the learning capabilities of targeted customer-supplier relationships. We focus on the data coming from suppliers and referring to 18 specific indicators, which measure four factors: collaborative commitment (ξ_1), internal complexity (ξ_2), relational trust (ξ_3), and environmental uncertainty (ξ_4) as named by Selnes & Sallis (2003). The indicators used to measure each factor are presented in Appendix 3. All indicators were measured on a seven-point scale; with 1 referring to "strongly disagree" or "low" and 7 to "strongly agree" or "high" depending on the form of the question. Quite a number of the indicators present rather skewed observed distribution, while the rest present a more symmetric distribution. The sample size is 286 after listwise deletion. The structure of matrices Λ and Φ to be estimated is of the form:

$$\Lambda = \begin{pmatrix} \lambda_{1,1} & & & \\ \lambda_{2,1} & & & \\ \lambda_{3,1} & & & \\ \lambda_{4,1} & & & \\ \lambda_{5,1} & & & \\ & \lambda_{6,2} & & \\ & \lambda_{7,2} & & \\ & \lambda_{8,2} & & \\ & & \lambda_{9,3} & \\ & & \lambda_{10,3} & \\ & & \lambda_{11,3} & \\ & & \lambda_{12,3} & \\ & & \lambda_{13,3} & \\ & & & \lambda_{14,4} \\ & & & \lambda_{15,4} \\ & & & \lambda_{16,4} \\ & & & \lambda_{17,4} \\ & & & \lambda_{18,4} \end{pmatrix}, \ \Phi = \begin{pmatrix} 1 & & & \\ \phi_{21} & 1 & & \\ \phi_{31} & \phi_{32} & 1 & \\ \phi_{41} & \phi_{42} & \phi_{43} & 1 \end{pmatrix}.$$

Along with the thresholds which are six for each indicator, there are a total of 132 free parameters to be estimated. The estimates of factor loadings and correlations are reported in Table 10. All four methods give very similar estimates. The PML and 3S-DWLS estimates are slightly closer to those of the FIML than those of the 3S-RULS approach. For the standard errors, it is those of the PML that are closer to those of the FIML approach.

Parameter	FIML	PML	3S-RULS	3S-RDWLS	Parameter	FIML	PML	3S-RULS	3S-RDWLS
$\lambda_{1,1}$	0.880	0.882	0.905	0.888	$\lambda_{14,4}$	0.775	0.767	0.774	0.768
	(0.022)	(0.023)	(0.027)	(0.015)		(0.031)	(0.033)	(0.035)	(0.027)
$\lambda_{2,1}$	0.894	0.891	0.897	0.893	$\lambda_{15,4}$	0.802	0.854	0.873	0.865
	(0.016)	(0.017)	(0.023)	(0.015)		(0.032)	(0.027)	(0.028)	(0.023)
$\lambda_{3,1}$	0.900	0.884	0.867	0.889	$\lambda_{16,4}$	0.783	0.752	0.740	0.752
	(0.015)	(0.018)	(0.023)	(0.015)		(0.035)	(0.040)	(0.042)	(0.029)
$\lambda_{4,1}$	0.905	0.897	0.907	0.909	$\lambda_{17,4}$	0.743	0.697	0.697	0.725
	(0.015)	(0.016)	(0.018)	(0.012)		(0.044)	(0.044)	(0.043)	(0.029)
$\lambda_{5,1}$	0.886	0.875	0.850	0.871	$\lambda_{18,4}$	0.724	0.705	0.710	0.739
	(0.018)	(0.020)	(0.025)	(0.015)		(0.042)	(0.042)	(0.042)	(0.030)
$\lambda_{6,2}$	0.520	0.622	0.775	0.661	ϕ_{21}	0.208	0.255	0.285	0.261
	(0.064)	(0.079)	(0.113)	(0.049)		(0.084)	(0.083)	(0.080)	(0.053)
$\lambda_{7,2}$	0.849	0.821	0.711	0.838	ϕ_{31}	0.627	0.627	0.630	0.626
	(0.051)	(0.065)	(0.090)	(0.041)		(0.042)	(0.044)	(0.043)	(0.035)
$\lambda_{8,2}$	0.834	0.784	0.669	0.780	ϕ_{41}	0.659	0.658	0.658	0.645
	(0.050)	(0.069)	(0.093)	(0.042)		(0.049)	(0.047)	(0.047)	(0.036)
$\lambda_{9,3}$	0.801	0.808	0.826	0.813	ϕ_{32}	0.113	0.125	0.135	0.126
	(0.027)	(0.027)	(0.031)	(0.024)		(0.076)	(0.073)	(0.074)	(0.056)
$\lambda_{10,3}$	0.873	0.866	0.857	0.865	ϕ_{42}	0.147	0.197	0.220	0.198
	(0.022)	(0.023)	(0.026)	(0.018)		(0.079)	(0.079)	(0.077)	(0.056)
$\lambda_{11,3}$	0.884	0.867	0.842	0.867	ϕ_{43}	0.641	0.651	0.638	0.645
	(0.019)	(0.024)	(0.028)	(0.018)		(0.052)	(0.048)	(0.047)	(0.037)
$\lambda_{12,3}$	0.913	0.908	0.903	0.910					
	(0.015)	(0.016)	(0.019)	(0.013)					
$\lambda_{13,3}$	0.865	0.871	0.888	0.873					
	(0.020)	(0.020)	(0.023)	(0.017)					

Table 9: Estimated loadings, correlations and their standard errors (in brackets) for RL data, confirmatory factor analysis

7. Discussion and Conclusions

Within the context of factor analysis models there are two main approaches for the analysis of ordinal variables, namely the item response theory (IRT) and the underlying response variable (URV) approaches. In both cases, full information maximum likelihood (FIML) estimation cannot be considered as a practical general method, since the level of computational complexity rises greatly with increases in the model size. In particular, the numerical evaluation of multidimensional integrals are necessary, the dimensionality of which depend either on the number of factors in the case of IRT approach, or on the number of observed variables in the case of URV approach. In the latter case, as the number of observed variables is often large, FIML is not used at all in practice. Instead, three-stage limited information estimation methods have been developed.

In this paper, we propose a pairwise maximum likelihood (PML) method that operates within the URV approach. As such, it is considered to be a computationally general method since it involves the evaluation of up to two-dimensional integrals written in a closed form, regardless of the number of observed variables or factors. Moreover, the PML estimator is asymptotically unbiased, consistent, and normally distributed. The main advantages of our proposed method over the commonly used three-stage limited information estimators (3S-RULS, 3S-RDWLS, 3S-WLS) are that all model parameters are estimated in one single step, and it removes the need to estimate a weight matrix to obtain the correct standard errors.

To investigate the performance of PML estimator, and the associated standard error, we conducted a simulation study on the effect of model and sample sizes within the confirmatory factor analysis framework. The main result of the study is that the PML parameter estimates and their estimated asymptotic standard errors are found to have very small bias and mean squared error, both of which decrease with the sample size. To compare the estimates and standard errors provided by the FIML, PML, 3S-RULS, and 3S-RDWLS approaches we have both conducted a simulation study and used some real data examples. The simulation study has been based on the same simulated samples used to investigate the performance of the PML approach. In the case of real data, the comparisons are made in an exploratory and confirmatory set-up, with one and two factors in the case of exploratory analysis, and four factors in the case of confirmatory analysis. The main conclusion is that the estimates and standard errors of all four methods are fairly close to each other. However, there is a tendency for the PML and 3S-RDWLS estimates and standard errors to be slightly closer to those of FIML than those of 3S-RULS approach.

Our study indicates that PML can be considered as a competitive estimation method for estimating factor analysis models with ordinal data. Subsequently, further study is advised to examine the efficiency of the proposed method relatively to the FIML approach. Additionally, PML estimation can be readily extended to the case of full structural equation models with a mixed type of data.

Appendix 1 - The gradient of the bivariate log-likelihood $\ln L(\boldsymbol{\theta};(x_i,x_j))$

The gradient $\nabla \ln L(\boldsymbol{\theta};(x_i,x_j))$, $i,j=1,\ldots,p$, $j\neq i$, can be distinguished in three main blocks as follows:

$$\nabla \ln L(\boldsymbol{\theta};(x_i,x_j)) = \begin{pmatrix} \frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \boldsymbol{\lambda}} \\ \frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \boldsymbol{\varphi}} \\ \frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \boldsymbol{\tau}} \end{pmatrix} .$$

The elements of the subvector $\frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \boldsymbol{\tau}}$ are the first derivatives with respect to thresholds $\frac{\partial \ln L\left(\boldsymbol{\theta};(x_i^*,x_j^*)\right)}{\partial \tau_{c_i}^{(x_i)}}$ and are given in Olsson (1979, eq. (13)). In particular,

$$\frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \tau_{c_i}^{(x_i)}} = \sum_{c_j=1}^{m_j} \left(\frac{n_{c_ic_j}^{(x_ix_j)}}{\pi_{c_ic_j}^{(x_ix_j)}} - \frac{n_{(c_i+1)c_j}^{(x_ix_j)}}{\pi_{(c_i+1)c_j}^{(x_ix_j)}} \right) \frac{\partial \pi_{c_ic_j}^{(x_ix_j)}}{\partial \tau_{c_i}^{(x_i)}} ,$$

where

$$\frac{\partial \pi_{c_ic_j}^{(x_ix_j)}}{\partial \tau_{c_i}^{(x_i)}} = \phi_1(\tau_{c_i}^{(x_i)}) \left[\Phi_1 \left(\frac{\tau_{c_j}^{(x_j)} - \rho_{x_ix_j}\tau_{c_i}^{(x_i)}}{\sqrt{1-\rho_{x_ix_j}^2}} \right) - \Phi_1 \left(\frac{\tau_{c_j-1}^{(x_j)} - \rho_{x_ix_j}\tau_{c_i}^{(x_i)}}{\sqrt{1-\rho_{x_ix_j}^2}} \right) \right] ,$$

and ϕ_1 and Φ_1 are the standard univariate normal density and distribution respectively.

To find the partial derivatives with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\varphi}$ we use the chain rule, i.e.

$$\frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \boldsymbol{\lambda}} = \frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \rho_{x_ix_j}} \frac{\partial \rho_{x_ix_j}}{\partial \boldsymbol{\lambda}} \text{ and}$$

$$\frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \boldsymbol{\varphi}} = \frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \rho_{x_ix_j}} \frac{\partial \rho_{x_ix_j}}{\partial \boldsymbol{\varphi}} .$$

The partial derivative with respect to $\rho_{x_ix_j}$ is

$$\frac{\partial \ln L(\boldsymbol{\theta};(x_i,x_j))}{\partial \rho_{x_ix_j}} = \sum_{c_i=1}^{m_i} \sum_{c_j=1}^{m_j} \frac{n_{c_ic_j}^{(x_ix_j)}}{\pi_{c_ic_j}^{(x_ix_j)}} \frac{\partial \pi_{c_ic_j}^{(x_ix_j)}}{\partial \rho_{x_ix_j}} ,$$

where $\frac{\partial \pi_{c_ic_j}^{(x_ix_j)}}{\partial \rho_{x_ix_j}}$ is given in Olsson (1979, eqs (8)) and it is

$$\frac{\partial \pi_{c_ic_j}^{(x_ix_j)}}{\partial \rho_{x_ix_j}} = \phi\left(\tau_{c_i}^{(x_i)}, \tau_{c_j}^{(x_j)}; \rho_{x_ix_j}\right) - \phi\left(\tau_{c_i}^{(x_i)}, \tau_{c_j-1}^{(x_j)}; \rho_{x_ix_j}\right) - \phi\left(\tau_{c_i-1}^{(x_i)}, \tau_{c_j}^{(x_j)}; \rho_{x_ix_j}\right) + \phi\left(\tau_{c_i-1}^{(x_i)}, \tau_{c_j-1}^{(x_j)}; \rho_{x_ix_j}\right) .$$

The partial derivative of $\rho_{x_ix_j}$ with respect to $\boldsymbol{\lambda}$ in vector - matrix form is as follows:

$$\frac{\partial \rho_{x_ix_j}}{\partial \boldsymbol{\lambda}} = \frac{\partial \rho_{x_ix_j}}{\partial \boldsymbol{\lambda}_{i\cdot}} \frac{\partial \boldsymbol{\lambda}_{i\cdot}}{\partial \boldsymbol{\lambda}} + \frac{\partial \rho_{x_ix_j}}{\partial \boldsymbol{\lambda}_{j\cdot}} \frac{\partial \boldsymbol{\lambda}_{j\cdot}}{\partial \boldsymbol{\lambda}} = \boldsymbol{\lambda}_{j\cdot} \Phi \frac{\partial \boldsymbol{\lambda}_{i\cdot}}{\partial \boldsymbol{\lambda}} + \boldsymbol{\lambda}_{i\cdot}^{(x)} \Phi \frac{\partial \boldsymbol{\lambda}_{j\cdot}}{\partial \boldsymbol{\lambda}} ,$$

where the terms of type $\frac{\partial \boldsymbol{\lambda}_{i\cdot}}{\partial \boldsymbol{\lambda}}$ are matrices of zeroes and ones with k rows (as many as the number of columns of matrix Λ) and as many columns as the size of $\boldsymbol{\lambda}$ (the number of free parameters in matrix Λ).

The partial derivative of $\rho_{x_ix_j}$ with respect to $\boldsymbol{\varphi}$ in vector - matrix form is

$$\frac{\partial \rho_{x_ix_j}}{\partial \boldsymbol{\varphi}} = \frac{\partial \rho_{x_ix_j}}{\partial \Phi} \frac{\partial \Phi}{\partial \boldsymbol{\varphi}} = \left(\boldsymbol{\lambda}_{i\cdot}^{(x)}\right)' \boldsymbol{\lambda}_{j\cdot}^{(x)} \frac{\partial \Phi}{\partial \boldsymbol{\varphi}} ,$$

where $\frac{\partial \Phi}{\partial \boldsymbol{\varphi}}$ is a matrix of zeroes and ones with k^2 rows (as many as the total number of elements of matrix Φ) and as many columns as the size of vector $\boldsymbol{\varphi}$ (the number of free non-redundant parameters in Φ). Note that $\frac{\partial \Phi}{\partial \boldsymbol{\varphi}}$ is sometimes (more appropriately) denoted as $\frac{\partial vec(\Phi)}{\partial \boldsymbol{\varphi}'}$, where vec is the function transforming a mxn matrix to a mnx1 vector by stacking its columns one underneath the other.

Appendix 2 - The indicators of Science & Technology data

1. Science and technology are making our lives healthier, easier and more comfortable.
2. Scientific and technological research cannot play an important role in protecting the environment and repairing it.
3. The application of science and new technology will make work more interesting.
4. Thanks to science and technology, there will be more opportunities for the future generations.
5. New technology does not depend on basic scientific research.
6. Scientific and technological research do not play an important role in industrial development.
7. The benefits of science are greater than any harmful effects it may have.

Appendix 3 - The indicators of Relationship Learning data

Collaborative Commitment

1. To what degree do you discuss company goals with the other party in this relationship?

2. To what degree are these goals developed through joint analysis of potentials?

3. To what degree are these goals formalized in a joint agreement or contract?

4. To what degree are these goals implemented in day-to-day work?

5. To what degree have you developed measures that capture performance related to these goals?

Internal Complexity

6. The products we exchange are generally very complex.

7. There are many operating units involved from both organizations.

8. There are many contract points between different departments or professions between the two organizations.

Relational Trust

9. I believe the other organization will respond with understanding in the event of problems.

10. I trust that the other organization is able to fulfill contractual agreements.

11. We trust that the other organization is competent at what they are doing.

12. There is a general agreement in my organization that the other organization is trustworthy.

13. There is a general agreement in my organization that the contact people on the other organization are trustworthy.

Environmental Uncertainty

14. End-users needs and preferences change rapidly in our industry.

15. The competitors in our industry frequently make aggressive moves to capture market share.

16. Crises have caused some of our competitors to shut down or radically change the way they operate.

17. It is very difficult to forecast where the technology will be in the next 2-3 years in our industry.

18. In recent years, a large number of new product ideas have been made possible through technological breakthroughs in our industry.