Estimating the variance of estimator of the latent factor linear mixed model using supplemented expectation-maximization algorithm

Angraini, Yenni; Notodiputro, Khairil Anwar; Folmer, Henk; Saefuddin, Asep; Toharudin, Toni

Published in:
Symmetry

DOI:
10.3390/sym13071286

Publication date:
2021

symmetry

MDPI

Article

Estimating the Variance of Estimator of the Latent Factor Linear Mixed Model Using Supplemented Expectation-Maximization Algorithm

Yenni Angraini [1,2,*], Khairil Anwar Notodiputro [1,*], Henk Folmer [2], Asep Saefuddin [1] and Toni Toharudin [3]

[1] Department of Statistics, IPB University, Bogor 16680, Indonesia; asaefuddin@apps.ipb.ac.id
[2] Faculty of Spatial Sciences, University of Groningen, 9747 Groningen, The Netherlands; h.folmer@rug.nl
[3] Department of Statistics, University of Padjadjaran, Bandung 16426, Indonesia; toni.toharudin@unpad.ac.id
* Correspondence: y_angraini@apps.ipb.ac.id (Y.A.); khairil@apps.ipb.ac.id (K.A.N.)

Abstract: This paper deals with symmetrical data that can be modelled based on Gaussian distribution, such as linear mixed models for longitudinal data. The latent factor linear mixed model (LFLMM) is a method generally used for analysing changes in high-dimensional longitudinal data. It is usual that the model estimates are based on the expectation-maximization (EM) algorithm, but unfortunately, the algorithm does not produce the standard errors of the regression coefficients, which then hampers testing procedures. To fill in the gap, the Supplemented EM (SEM) algorithm for the case of fixed variables is proposed in this paper. The computational aspects of the SEM algorithm have been investigated by means of simulation. We also calculate the variance matrix of beta using the second moment as a benchmark to compare with the asymptotic variance matrix of beta of SEM. Both the second moment and SEM produce symmetrical results, the variance estimates of beta are getting smaller when number of subjects in the simulation increases. In addition, the practical usefulness of this work was illustrated using real data on political attitudes and behaviour in Flanders-Belgium.

Keywords: latent factor linear mixed model (LFLMM); expectation-maximization (EM) algorithm; supplemented EM algorithm; longitudinal data analysis

Citation: Angraini, Y.; Notodiputro, K.A.; Folmer, H.; Saefuddin, A.; Toharudin, T. Estimating the Variance of Estimator of the Latent Factor Linear Mixed Model Using Supplemented Expectation-Maximization Algorithm. *Symmetry* **2021**, *13*, 1286. https://doi.org/10.3390/sym13071286

Academic Editors: Christophe Chesneau and José Carlos R. Alcantud

Received: 3 June 2021
Accepted: 15 July 2021
Published: 17 July 2021

1. Introduction

The latent factor multivariate linear mixed model (LFLMM) is a combination between the Factor Analysis (FA) and the Linear Mixed Model (LMM), as proposed by [1]. The model aims to analyze longitudinal data sets with large numbers of multivariate responses, i.e., high-dimensional longitudinal data. The authors proposed estimation of the LFLMM by means of the EM algorithm, which is a closed-form solution. They showed by way of simulation that EM estimation of the LFLMM provides accurate parameter estimates and is more efficient in terms of adding variables other than time variables in the model than alternatives like the structural equation model. As shown by [2,3], the combination of fixed and random effects and the interaction of covariates with time can be straightforwardly handled by the LFLMM estimated by EM.

The LFLMM assumes that the responses are continuous and that the number of latent variables is known [1]. Moreover, convergence of the EM algorithm is sometimes slow. The main disadvantage, however, is that the EM algorithm does not produce standard errors of the estimator of the regression coefficients because it does not calculate the derivatives of the likelihood function, which are often complicated and tedious to derive [4,5]. Thus, it is difficult to study the effects of different covariates or fixed variables for different latent factors simultaneously.

The general Supplemented EM algorithm was proposed by [6] to obtain the standard errors by calculating the complete information matrix as the base of the variance-covariance

matrix of the estimator. The Supplemented EM algorithm has been applied to various kinds of models, notably item response models [6–9]. However, its suitability and features in the case of application to the LFLMM have not been investigated yet. In this study, we extend the work of [1] by employing the Supplemented EM algorithm as a by-product of the EM estimator for the case of fixed variables. We used simulation studies to investigate the computational aspects of the Supplemented EM algorithm and used a real data example to illustrate the practical usefulness of this work.

The remainder of this study is organized as follows. In Section 2, we specify the LFLMM and summarize the EM algorithm to estimate it. Section 3 presents the Supplemented EM algorithm. Sections 4 and 5 present the results of the simulation and real data example. Conclusions follow in Section 6.

2. The LFLMM and the EM Algorithm

Following [1], an LFLMM can be composed of two parts. The first is the factor analysis model, which represents the relationships between the observed and latent variables. This is similar to the structural equation model, which explains the relationship of the latent variables and the measurement indicators carried out through factor analysis [10]. This part can be written as:

$$\boldsymbol{Y}_{it} = \boldsymbol{\Lambda}\boldsymbol{\eta}_{it} + \boldsymbol{\epsilon}_{it} \tag{1}$$

Specifically, for the i-th of N individuals, we observe $j = 1, \ldots, J$ responses which characterize d latent factors $\left(\boldsymbol{\eta}_{it} = \left(\boldsymbol{\eta}_{it}^{1}, \ldots, \boldsymbol{\eta}_{it}^{d}\right),\ d < J\right)$ at time t, $t = 1, \ldots, T_i$, where T_i is the number of time periods for subject i. $\boldsymbol{\Lambda}$ is the matrix of factor loadings and $\boldsymbol{\epsilon}_{it} = \left(\boldsymbol{\epsilon}_{it1}, \ldots, \boldsymbol{\epsilon}_{itJ}\right)$ the vector of measurement errors for subject i at time t. It is assumed that $\boldsymbol{\epsilon}_{itJ} \sim N\left(0, \boldsymbol{\tau}_j^2\right)$ and $\boldsymbol{\epsilon}_{itj} \perp \boldsymbol{\epsilon}_{ith}$, $j \neq h$. In matrix notation, Equation (1) reads:

$$\boldsymbol{Y}_i = \left(\boldsymbol{I}_{T_i} \bigotimes \boldsymbol{\Lambda}\right)\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i \tag{2}$$

where

$$\boldsymbol{Y}_i = \left(y'_{i1}, \ldots, y'_{iT_i}\right)'_{[J \times T_i, 1]}$$
$$\boldsymbol{\eta}_i = \left(\boldsymbol{\eta}'_{i1}, \ldots, \boldsymbol{\eta}'_{iT_i}\right)'_{[d \times T_i, 1]}$$
$$\boldsymbol{\Lambda}_{[J \times d]} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_J \end{pmatrix}$$

The second part of the LFLMM is a multivariate linear mixed model containing the fixed and random effects for each latent variable ($\boldsymbol{\eta}_{it}$). For individual i, $i = 1, 2, \ldots N$, at time t, $t = 1, 2, \ldots, T_i$, and latent variable l, $l = 1, 2, \ldots, d$, we thus have:

$$\boldsymbol{\eta}_{it}^{l} = x_{it}^{l}\boldsymbol{\beta}^{l} + z_{it}^{l}\boldsymbol{a}_i^{l} + \boldsymbol{\varepsilon}_{it}^{l} \tag{3}$$

where x_{it}^{l} and z_{it}^{l} are the elements of design matrices of the p fixed variables and q random effects, respectively. $\boldsymbol{\beta}^{l}$ is an unknown coefficient, $\boldsymbol{a}_i^{l} = \left(\boldsymbol{a}_{i1}^{l}, \ldots, \boldsymbol{a}_{iq}^{l}\right)$ and $\boldsymbol{\varepsilon}_i^{l} = \left(\boldsymbol{\varepsilon}_{i1}^{l}, \ldots, \boldsymbol{\varepsilon}_{iTi}^{l}\right)$, $l = 1, 2, \ldots d$, are the random effects and errors for subject i and factor l, respectively. The random effects are assumed to be normally distributed with mean 0 and variance-covariance matrix $\boldsymbol{V}(\boldsymbol{a}) = \boldsymbol{\Sigma}_{\boldsymbol{a}}$. It is assumed that $\boldsymbol{\Sigma}_{\boldsymbol{a}}$ captures the changes among the latent variables [1]. For example, a positive covariance between the random effects for the latent variables 1 and 2 means that if for a given individual i the latent variable 1 increases over time, the latent variable 2 also increases for that individual. Note that in this setting, the covariates are included in the multivariate linear mixed model (MLMM) of Equation (3) but not in the factor analysis model of Equation (1).

In matrix notation, Equation (3) reads:

$$\boldsymbol{\eta}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{a}_i + \boldsymbol{\varepsilon}_i \tag{4}$$

where

$$\boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}_{i1} \\ \vdots \\ \boldsymbol{x}_{iT_i} \end{pmatrix}_{[d\times T_i, p\times d]}, \; \boldsymbol{x}_{it} = \begin{pmatrix} x_{it}^1 & 0 & \cdots & 0 \\ 0 & x_{it}^2 & \cdots & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{it}^d \end{pmatrix}_{[d,\; p\times d]}$$

$$\boldsymbol{Z}_i = \begin{pmatrix} \boldsymbol{z}_{i1} \\ \vdots \\ \boldsymbol{z}_{iT_i} \end{pmatrix}_{[d\times T_i, q\times d]}, \; \boldsymbol{z}_{it} = \begin{pmatrix} z_{it}^1 & 0 & \cdots & 0 \\ 0 & z_{it}^2 & \cdots & 0 \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_{it}^d \end{pmatrix}_{[d,\; q\times d]}$$

$$\boldsymbol{\beta} = \left(\boldsymbol{\beta}^{1\prime}, \ldots, \boldsymbol{\beta}^{d\prime}\right)\prime_{[p\times d,1]}$$

$$\boldsymbol{a}_i = \left(\boldsymbol{a}_i^{1\prime}, \ldots, \boldsymbol{a}_i^{d\prime}\right)\prime_{[q\times d,1]} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_a)$$

$$\boldsymbol{\varepsilon}_i = \left(\boldsymbol{\varepsilon}_{i1}^{\prime}, \ldots, \boldsymbol{\varepsilon}_{iT_i}^{\prime}\right)\prime_{[d\times T_i,1]}, \; \boldsymbol{\varepsilon}_{it} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$$

The marginal distribution of $\boldsymbol{Y}_i$ is assumed multivariate normal with mean:

$$\boldsymbol{E}\left(\boldsymbol{Y}_i\right) = \left(\boldsymbol{I}_{T_i} \bigotimes \boldsymbol{\Lambda}\right)\boldsymbol{X}_i\boldsymbol{\beta}$$

and variance-covariance matrix

$$\boldsymbol{V}(\boldsymbol{Y}_i) = \left(\boldsymbol{I}_{T_i} \otimes \boldsymbol{\Lambda}\right)\boldsymbol{V}(\boldsymbol{\eta}_i)\left(\boldsymbol{I}_{T_i} \otimes \boldsymbol{\Lambda}\right)^{\prime} + \boldsymbol{I}_{T_i} \otimes diag\left(\boldsymbol{\tau}_1^2, \ldots, \boldsymbol{\tau}_J^2\right)$$

The first term in $\boldsymbol{V}(\boldsymbol{Y}_i)$ denotes the variances and covariance of the latent factors and the last term the variances of the error term, $\boldsymbol{\epsilon}_{it}$. The mean and variance-covariance matrix of $\boldsymbol{\eta}_i$ are $\boldsymbol{E}(\boldsymbol{\eta}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$ and $\boldsymbol{V}(\boldsymbol{\eta}_i) = \boldsymbol{Z}_i\boldsymbol{\Sigma}_a\boldsymbol{Z}_i^{\prime} + \boldsymbol{I}_{T_i} \otimes \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$, respectively.

To estimate the LFLMM by EM, we summarize it below, as proposed by [1]. Before going into detail, we observe that $\{\boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}$ is treated as missing data. Hence, the complete dataset is $\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}$ whereas the observed data is $\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}$. It follows that the complete data likelihood is:

$$L = \prod_{i=1}^{N} \boldsymbol{P}\left(\boldsymbol{Y}_i \middle| \boldsymbol{\eta}_i, \boldsymbol{\Lambda},\ \boldsymbol{\tau}^2\right)\boldsymbol{P}(\boldsymbol{\eta}_i|\boldsymbol{X}_i, \boldsymbol{Z}_i,\ \boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\boldsymbol{P}(\boldsymbol{a}_i|\boldsymbol{\Sigma}_{\boldsymbol{a}}) \tag{5}$$

The corresponding complete data loglikelihood is:

$$\log L = \sum_{i=1}^{N}\left[\log \boldsymbol{P}\left(\boldsymbol{Y}_i \middle| \boldsymbol{\eta}_i, \boldsymbol{\Lambda},\ \boldsymbol{\tau}^2\right) + \log \boldsymbol{P}(\boldsymbol{\eta}_i|\boldsymbol{X}_i, \boldsymbol{Z}_i,\ \boldsymbol{a}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) + \log \boldsymbol{P}(\boldsymbol{a}_i|\boldsymbol{\Sigma}_{\boldsymbol{a}})\right] \tag{6}$$

where

$$\begin{aligned} \sum_{i=1}^{N} \log \boldsymbol{P}\left(\boldsymbol{Y}_i \middle| \boldsymbol{\eta}_i, \boldsymbol{\Lambda},\ \boldsymbol{\tau}^2\right) &= \sum_{i=1}^{N}\sum_{j=1}^{J} \log \boldsymbol{P}\left(\boldsymbol{Y}_{ij} \middle| \boldsymbol{\eta}_i, \boldsymbol{\Lambda}_j,\ \boldsymbol{\tau}_j^2\right) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{J}\left[-\tfrac{n_i}{2}\log \boldsymbol{\tau}_j^2 - \tfrac{1}{2\boldsymbol{\tau}_j^2}\left(\boldsymbol{Y}_{ij} - \boldsymbol{\eta}_i^{\prime}\boldsymbol{\Lambda}_j\right)^{\prime}\left(\boldsymbol{Y}_{ij} - \boldsymbol{\eta}_i^{\prime}\boldsymbol{\Lambda}_j\right)\right] \end{aligned} \tag{7}$$

$$\begin{aligned} \sum_{i=1}^{N} \log \boldsymbol{P}(\boldsymbol{\eta}_i|\boldsymbol{X}_i, \boldsymbol{Z}_i,\ \boldsymbol{\beta}, \boldsymbol{a}_i, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) &= \sum_{i=1}^{N}\sum_{t=1}^{T_i} \log \boldsymbol{P}(\boldsymbol{\eta}_{it}|\boldsymbol{X}_i, \boldsymbol{Z}_i,\ \boldsymbol{a}_i, \boldsymbol{\beta},\ \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T_i}\left[-\tfrac{1}{2}\log|\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}| - \tfrac{1}{2}(\boldsymbol{\eta}_{it} - \boldsymbol{X}_{it}\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{a}_i)^{\prime}\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}(\boldsymbol{\eta}_{it} - \boldsymbol{X}_{it}\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{a}_i)\ \right] \end{aligned} \tag{8}$$

$$\sum_{i=1}^{N} \log \boldsymbol{P}(\boldsymbol{a}_i|\boldsymbol{\Sigma}_{\boldsymbol{a}}) = -\frac{1}{2}\sum_{i=1}^{N}\left[\log|\boldsymbol{\Sigma}_{\boldsymbol{a}}| - \frac{1}{2}\boldsymbol{a}\prime_i\boldsymbol{\Sigma}_{\boldsymbol{a}}^{-1}\boldsymbol{a}_i\right] \tag{9}$$

Let the $\boldsymbol{\theta}$ denote the parameter vector $(\boldsymbol{\Lambda},\ \boldsymbol{\tau}^2,\ \boldsymbol{\beta},$ and $\boldsymbol{\Sigma}_{\varepsilon})$, $\boldsymbol{\theta}^{(w)}$ be the ML estimate of $\boldsymbol{\theta}$ at the wth iteration for $w = 0, 1, \ldots$, and $Q\left(\boldsymbol{\theta}|\,\boldsymbol{\theta}^{(w)}\right)$ the expectation of the joint loglikelihood for the complete data $\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}$ conditional on the observed data $\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}$:

$$Q\left(\boldsymbol{\theta}|\,\boldsymbol{\theta}^{(w)}\right) = \boldsymbol{E}\left\{\log L(\,\boldsymbol{\theta}|\,\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i)\middle|\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i, \boldsymbol{\theta}^{(w)}\right\} \tag{10}$$

Then the $(w+1)$th iteration of the EM algorithm consists of (i) the E–step, which is the expectation of the joint loglikelihood computed according to (10) and (ii) the M-step, which maximizes $Q\left(\boldsymbol{\theta}|\,\boldsymbol{\theta}^{(w)}\right)$ to yield $\boldsymbol{\theta}^{(w+1)}$. Further details on EM estimation of the LFLMM can be found in [1].

3. The Supplemented EM

Below we discuss the Supplemented EM algorithm, denoted as SEM. Before going into detail, we observe that the main purpose of this study is to estimate the standard errors of the fixed effects, $\boldsymbol{\beta}$.

Consider the mapping M defined by iteration w of the EM algorithm:

$$\boldsymbol{\beta}^{(w+1)} = M\left(\boldsymbol{\beta}^{(w)}\right), \text{ for } w = \mathbf{0}, \mathbf{1}, \ldots$$

when the parameter vector converges to $\boldsymbol{\beta}^*$, we obtain $\boldsymbol{\beta}^* = M(\boldsymbol{\beta}^*)$. For $M(\boldsymbol{\beta})$ continuous we have by Taylor expansion in the neighbourhood of $\boldsymbol{\beta}^*$

$$\boldsymbol{\beta}^{(\boldsymbol{w}+1)} = M\left(\boldsymbol{\beta}^{(\boldsymbol{w})}\right) \approx M(\boldsymbol{\beta}^*) + \boldsymbol{DM}\left(\boldsymbol{\beta}^{(\boldsymbol{w})} - \boldsymbol{\beta}^*\right) = \boldsymbol{\beta}^* + \boldsymbol{DM}\left(\boldsymbol{\beta}^{(\boldsymbol{w})} - \boldsymbol{\beta}^*\right) \tag{11}$$

where

$$\boldsymbol{DM} = \left(\frac{\partial M_h(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_g}\right)\Bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*} \tag{12}$$

$g = 1, 2, \ldots, k$ and $h = 1, 2, \ldots, k$ is the $k \times k$ Jacobian matrix of $M(\boldsymbol{\beta}) = (M_1(\boldsymbol{\beta}), \ldots, M_k(\boldsymbol{\beta}))$ evaluated at the ML estimate of $\boldsymbol{\beta}$ with $k = p \times d$. $\boldsymbol{DM}$ is known as the rate matrix. To obtain the loglikelihood of $\boldsymbol{\beta}$, we consider the complete data density of the LFLMM:

$$\begin{aligned} &f(\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\boldsymbol{\theta}) \\ &= f(\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}|\boldsymbol{\theta}) f(\{\boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}, \boldsymbol{\theta}) \end{aligned}$$

where $f(\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}|\boldsymbol{\theta})$ is the density of the observed data and $f(\{\boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}, \boldsymbol{\theta})$ the density of missing data, given the observed data. Thus, the loglikelihood of $\boldsymbol{\beta}$ given the complete data is:

$$\log L(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}) = \log L(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}) + \log f(\{\boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}, \boldsymbol{\beta}) \tag{13}$$

where $\log L(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\})$ is the observed-data loglikelihood and $\log L(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\})$ is the complete data loglikelihood.

The asymptotic variance-covariance matrix of $\boldsymbol{\beta}$, $\boldsymbol{V}(\boldsymbol{\beta})$, is the inverse of the observed information matrix $(\boldsymbol{I}_o)$. In the case of the LFLMM, the observed data is $\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}$ so that $\boldsymbol{V}(\boldsymbol{\beta})$ is:

$$\boldsymbol{V}(\boldsymbol{\beta}) = \boldsymbol{I}_o^{-1}(\boldsymbol{\beta}\Big|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}) \tag{14}$$

where $\boldsymbol{I}_o(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\})$ is the information matrix of the observed data loglikelihood (which is assumed to exist). That is [6,11]:

$$\boldsymbol{I}_o(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}) = -E\left[\frac{\partial^2 \log L(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\})}{\partial\boldsymbol{\beta}\cdot\partial\boldsymbol{\beta}}\right] \tag{15}$$

Equation (15) is difficult to evaluate directly using the EM algorithm [6,11]. As a way out, [7] suggested to evaluate the complete data information matrix:

$$\boldsymbol{I}_o(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}) = -E\left[\frac{\partial^2 \log L(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\})}{\partial\boldsymbol{\beta}\cdot\partial\boldsymbol{\beta}}\right] \tag{16}$$

The conditional complete data information, given the observed data evaluated at $\boldsymbol{\beta} = \boldsymbol{\beta}^*$, is:

$$\boldsymbol{I}_{oc} = E[\boldsymbol{I}_o(\boldsymbol{\beta}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i,\ \boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}, \boldsymbol{\beta}^*\ ] \tag{17}$$

After taking second derivatives, averaging over $f(\{\boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}, \boldsymbol{\beta})$, and evaluating at $\boldsymbol{\beta} = \boldsymbol{\beta}^*$, Equation (13) implies:

$$\boldsymbol{I}_o(\boldsymbol{\beta}^*|(\ \{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}) = \boldsymbol{I}_{oc} - \boldsymbol{I}_{om} \tag{18}$$

where the missing information matrix $(\boldsymbol{I}_{om})$ is

$$\boldsymbol{I}_{om} = E\left[-\frac{\partial^2 \log f\ (\{\boldsymbol{\eta}_i,\ \boldsymbol{a}_i\}|\{\boldsymbol{Y}_i, \boldsymbol{X}_i, \boldsymbol{Z}_i\}, \boldsymbol{\beta}^*\ )}{\partial\boldsymbol{\beta}\cdot\partial\boldsymbol{\beta}}\right] \tag{19}$$

ref [12] interpreted Equation (18) as

$$\textit{observed information} = \textit{complete information} - \textit{missing information}$$

and called it the "missing information principle". Equation (18) can be written as:

$$\boldsymbol{I}_o(\boldsymbol{\beta}^*|(\ \{\boldsymbol{Y}_i,\ \boldsymbol{X}_i,\ \boldsymbol{Z}_i\}) = \left(\boldsymbol{I} - \boldsymbol{I}_{om}\boldsymbol{I}_{oc}^{-1}\right)\boldsymbol{I}_{oc}, \tag{20}$$

where $\boldsymbol{I}$ is the $k \times k$ identity matrix and $\boldsymbol{I}_{om}\boldsymbol{I}_{oc}^{-1}$ is the matrix of the fraction of missing information [7,11]. According to [13], the rate of convergence of the EM algorithm is determined by the fraction of missing information in the neighborhood of $\boldsymbol{\beta}^*$:

$$\boldsymbol{DM} = \boldsymbol{I}_{om}\boldsymbol{I}_{oc}^{-1} \tag{21}$$

Substituting $\boldsymbol{DM} = \boldsymbol{I}_{om}\boldsymbol{I}_{oc}^{-1}$ into Equation (20) and inverting, the asymptotic variance-covariance matrix of $\boldsymbol{\beta}^*$, $\boldsymbol{V}(\boldsymbol{\beta}^*)$ is:

$$\boldsymbol{V}(\boldsymbol{\beta}^*) = \boldsymbol{I}_{oc}^{-1}(\boldsymbol{I} - \boldsymbol{DM})^{-1} \tag{22}$$

From the equality $(\boldsymbol{I} - \boldsymbol{P})^{-1} = (\boldsymbol{I} - \boldsymbol{P} + \boldsymbol{P})(\boldsymbol{I} - \boldsymbol{P})^{-1} = \boldsymbol{I} + \boldsymbol{P}(\boldsymbol{I} - \boldsymbol{P})^{-1}$ it follows that:

$$\boldsymbol{V}(\boldsymbol{\beta}^*) = \boldsymbol{I}_{oc}^{-1}\left\{\boldsymbol{I} + \boldsymbol{DM}(\boldsymbol{I} - \boldsymbol{DM})^{-1}\right\} = \boldsymbol{I}_{oc}^{-1} + \boldsymbol{I}_{oc}^{-1}\boldsymbol{DM}(\boldsymbol{I} - \boldsymbol{DM})^{-1} \tag{23}$$

or

$$\boldsymbol{V}(\boldsymbol{\beta}^*) = \boldsymbol{I}_{oc}^{-1} + \Delta\boldsymbol{V}(\boldsymbol{\beta}^*) \tag{24}$$

where $\Delta\boldsymbol{V}(\boldsymbol{\beta}^*)$ is the increase of the diagonal elements of $\boldsymbol{V}(\boldsymbol{\beta}^*)$ related to missing information.

Calculation of the $\boldsymbol{DM}$ matrix can be done using the code and output of the original EM algorithm as follows [6,7]. The $\boldsymbol{DM}$ matrix represents the differential of the parameter mappings during the EM algorithm. Hence, each element of the $\boldsymbol{DM}$ matrix represents

a component-wise increase of the rate of convergence per iteration of the EM algorithm. Let r_{gh} be the (g,h) th element of the ***DM*** matrix. From Equation (13), we have:

$$\begin{aligned} r_{gh} = \frac{\partial M_h(\boldsymbol{\beta}^*)}{\partial \boldsymbol{\beta}_g} &= \lim_{\beta_g \to \beta_g^*} \frac{M_h\left(\beta_1^*, \ldots, \beta_{g-1}^*, \beta_g, \beta_{g+1}^*, \ldots, \beta_k^*\right) - M_h(\boldsymbol{\beta}^*)}{\beta_g - \beta_g^*} \\ &= \lim_{w \to \infty} \frac{M_h\left(\beta^{(w)}(g)\right) - M_h(\beta^*)}{\beta_g^{(w)} - \beta_g^*} \equiv \lim_{w \to \infty} r_{gh}^{(w)} \end{aligned} \tag{25}$$

$g = 1, 2, \ldots, k$ and $h = 1, 2, \ldots, k$
where $\boldsymbol{\beta}^{(w)}(g)$ is called the semi-active parameter set

$$\boldsymbol{\beta}^{(w)}(g) = \left(\beta_1^*, \ldots, \beta_{g-1}^*, \beta_g^{(w)}, \beta_{g+1}^*, \ldots, \beta_k^*\right), \; w = 1, 2, \ldots \tag{26}$$

which converges to β_g^*. Note that only the gth component in $\boldsymbol{\beta}^{(w)}(g)$ takes a value different from its maximum likelihood estimate.

To calculate r_{gh}, the Supplemented EM algorithm requires $\boldsymbol{\theta}^* = \left\{\boldsymbol{\Lambda}^*, \boldsymbol{\tau}^{2*}, \boldsymbol{\beta}^* \text{and } \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}{}^*\right\}$ and $\boldsymbol{\theta}^{(\mathrm{w})} = \left\{\boldsymbol{\Lambda}^{(w)}, \boldsymbol{\tau}^{2(w)}, \boldsymbol{\beta}^{(w)} \text{and } \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}{}^{(w)}\right\}$ for $w = 1, 2, \ldots$ as input. $\boldsymbol{\theta}^*$ can be obtained by the EM algorithm using a set of arbitrarily chosen initial parameters $\boldsymbol{\theta}^{init}$ including $\boldsymbol{\theta}^{(w)}$ for $w = 1$, i.e., $\boldsymbol{\theta}^{(1)}$. The starting point $\boldsymbol{\theta}^{(1)}$ may, but need not, be close $\boldsymbol{\theta}^*$. The algorithm below closely follows [14,15].

1. Select input: $\boldsymbol{\theta}^{(\mathrm{w})}$ and $\boldsymbol{\theta}^*$
2. Set $\boldsymbol{\theta}^{(\mathrm{w})}$ for $w = 1, 2, \ldots$. Then take the E step and M step of the LFLMM EM algorithm to produce $\boldsymbol{\theta}^{(w+1)}$.
3. For rows $= 1, 2, \ldots, k$:
   (i) Set $\widetilde{\boldsymbol{\beta}}^{(w)}(g)$ be equal to $\boldsymbol{\beta}^*$, except for the gth element: $\left(\widetilde{\boldsymbol{\beta}}^{(w)}(g) = (\beta_1^*, \ldots, \beta_{g-1}^*, \beta_g^{(w)}, \beta_{g+1}^*, \ldots, \beta_k^*)\right)$
   (ii) Run the LFLMM EM algorithm with $\widetilde{\boldsymbol{\beta}}^{(w)}(g)$ as the current estimate of $\boldsymbol{\beta}$ to obtain $\widetilde{\boldsymbol{\beta}}^{(w+1)}(g)$.
   (iii) Calculate the gth row of $r_{gh}^{(w)}$ as

$$r_{gh}^{(w)} = \frac{\widetilde{\beta}_h{}^{(w+1)}(g) - \beta_h^*}{\beta_g^{(w)} - \beta_g^*}, \text{ for } h = 1, 2, \ldots, k$$

The output after a single run of the Supplemented EM algorithm (Step 1 and 2) are $\boldsymbol{\beta}^{w+1}$ and $r_{gh}^{(w)}$ $g = 1, 2, \ldots, k$ and $h = 1, 2, \ldots, k$. Based on the final estimates of DM, $\boldsymbol{V}(\boldsymbol{\beta}^*)$ is calculated using (24). The diagonal elements of $\boldsymbol{V}$ are the variance of $\boldsymbol{\beta}^*$.

4. Simulation

To evaluate the statistical properties and computational aspects of the SEM, we set up a simulation study. The number of subjects (N) is set at 500, 1000, and 1500 with six time periods. The number of simulations (S) is set at 50 and 250. The other set-up of the simulations is adopted from [1]. Particularly, we use the same initial values of the parameters of the LFLMM model (12 items, 2 latent factors, and a simple structure to model the relationship between the items and the latent factors). It is done to check if the bias resulting from the Supplemented EM algorithm on LFLMM is in line with the results presented in [1]. Table 1 presents the absolute difference for the true parameters, and the averages of the SEM estimates are calculated as a measure of performance.

Table 1 shows that the absolute difference of $\sigma_{\boldsymbol{a},11}$ has a range from 0 to 0.0444 $(N = 500$ and $S = 250)$. The results are in line with [1] the parameters of the measurement model (factor loadings and error variances) are estimated more precisely than those of the latent mixed regression model. Overall, these results indicate that with the increasing number of

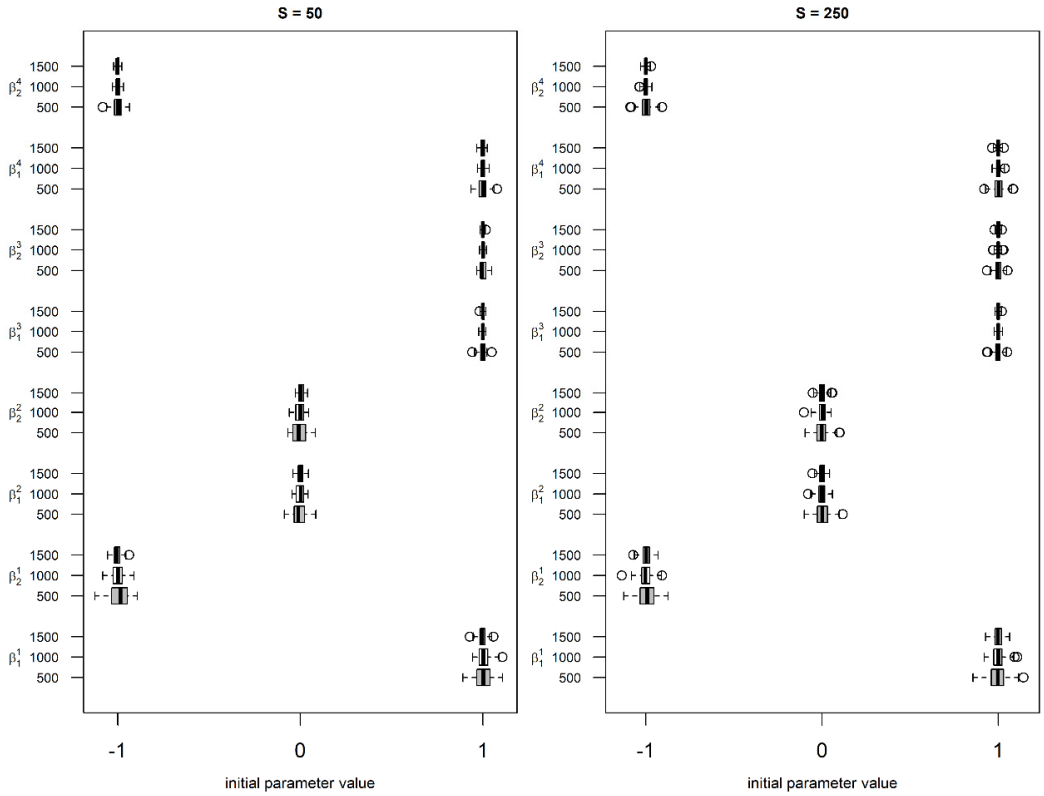
subjects in the simulation, the absolute difference between the actual parameters and the SEM average is getting smaller. This means the SEM can estimate the model parameter very well.

Although the results in Table 1 indicate that the accuracy of the estimates in the latent mixed regression model (No. 23–43) is not as good as in the measurement model part (No. 1–22), through Figure 1a–c, it can be shown that the median of boxplots (which generally is close to the true parameters) are all at the same level. This means that the parameters of the latent mixed regression models (β, σ_a, σ_ε) are estimated more precisely, especially for the number of simulations $S = 250$. Furthermore, all the boxplots are also shown to have different distributions of views with an increasing number of subjects in the simulation. This is indicated by the smaller size of the boxplot as the number of subjects increases for both numbers of simulations.

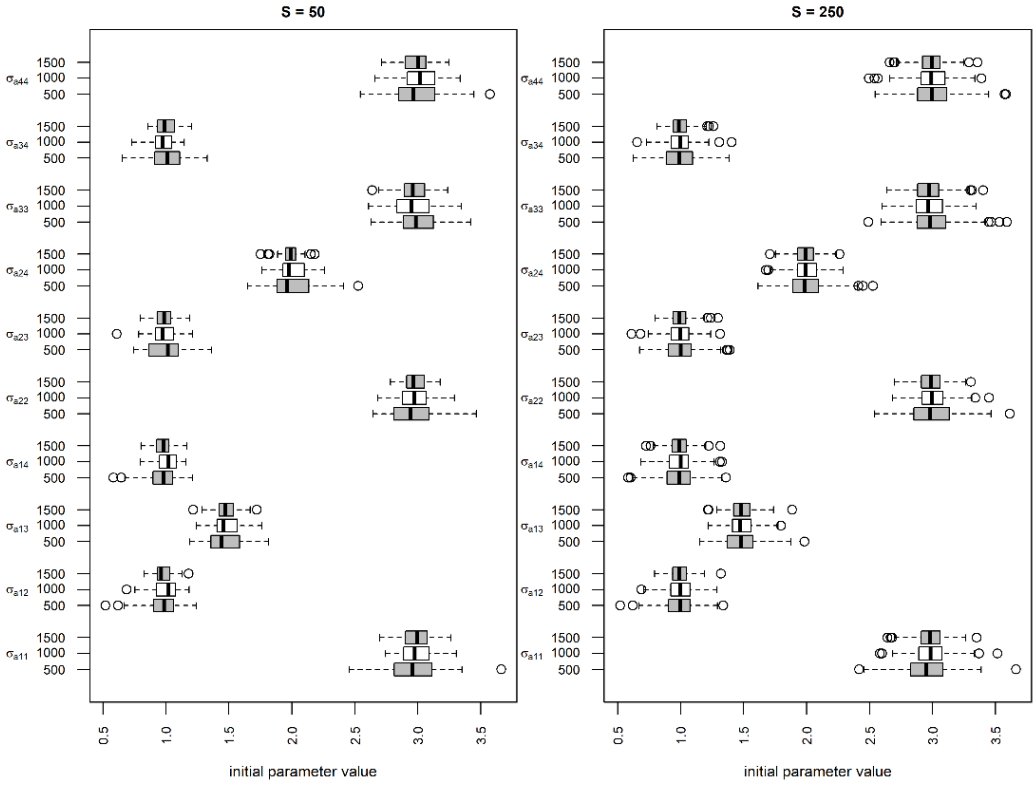
The results from the simulations of the Supplemented EM algorithm in estimating the asymptotic variance matrix of beta is summarized as a standard deviation of beta in Table 2. We also calculate the standard deviation of beta using the 2nd moment, $\sqrt{V\left(\hat{\beta}\right)} = \sqrt{\frac{1}{S-1}\sum_{s=1}^{S}\left(\hat{\beta}_s - \overline{\hat{\beta}}\right)^2}$ as a benchmark to compare with the standard deviation of beta of SEM. Both the 2nd moment and SEM produce symmetrical results the parameter estimate for the standard deviation of beta is getting smaller with the increasing number of subjects in the simulation. Overall, it can be concluded that by using SEM, changes in the parameter estimate for the standard deviation of beta are not too different for all number of subjects (Figure 2). Therefore, the simulation results suggest that the asymptotic variance of beta from the Supplemented EM Algorithm can be used to estimate the asymptotic variance of beta in real data analysis.

Table 1. The absolute difference between the true parameters and the SEM averages.

No	Para	True	SEM						No	Para	True	SEM					
	N		500		1000		1500			*N*		500		1000		1500	
	S		50	250	50	250	50	250		*S*		50	250	50	250	50	250
1	$\lambda_{2,1}$	1	0.0013	0.0012	0.0010	0.0009	0.0009	0.0009	23	β_1^1	1	0.0040	0.0036	0.0032	0.0004	0.0018	0.0007
2	$\lambda_{3,1}$	1	0.0019	0.0015	0.0013	0.0009	0.0012	0.0009	24	β_1^2	−1	0.0111	0.0079	0.0008	0.0008	0.0034	0.0017
3	$\lambda_{4,1}$	1	0.0015	0.0009	0.0005	0.0009	0.0012	0.0010	25	β_2^1	0	0.0038	0.0018	0.0012	0.0007	0.0016	0.0003
4	$\lambda_{5,1}$	1	0.0014	0.0016	0.0007	0.0009	0.0007	0.0009	26	β_2^2	0	0.0031	0.0039	0.0014	0.0021	0.0044	0.0006
5	$\lambda_{6,1}$	1	0.0016	0.0014	0.0011	0.0009	0.0013	0.0010	27	β_3^1	1	0.0012	0.0016	0.0008	0.0003	0.0007	0.0004
6	$\lambda_{7,2}$	1	0.0021	0.0018	0.0013	0.0014	0.0016	0.0013	28	β_3^2	1	0.0016	0.0010	0.0002	0.0004	0.0000	0.0005
7	$\lambda_{8,2}$	1	0.0015	0.0016	0.0017	0.0014	0.0010	0.0013	29	β_4^1	1	0.0025	0.0027	0.0002	0.0001	0.0007	0.0002
8	$\lambda_{9,2}$	1	0.0020	0.0015	0.0015	0.0014	0.0019	0.0013	30	β_4^2	−1	0.0018	0.0025	0.0004	0.0001	0.0003	0.0001
9	$\lambda_{10,2}$	1	0.0015	0.0015	0.0015	0.0014	0.0022	0.0017	31	$\sigma_{a,11}$	3	0.0332	0.0444	0.0152	0.0154	0.0039	0.0160
10	$\lambda_{11,2}$	1	0.0019	0.0012	0.0017	0.0017	0.0016	0.0016	32	$\sigma_{a,12}$	1	0.0410	0.0120	0.0024	0.0057	0.0265	0.0085
11	τ_1^2	0.5	0.0007	0.0001	0.0006	0.0004	0.0009	0.0004	33	$\sigma_{a,13}$	1.5	0.0289	0.0200	0.0218	0.0141	0.0229	0.0133
12	τ_2^2	0.5	0.0003	0.0006	0.0000	0.0011	0.0019	0.0006	34	$\sigma_{a,14}$	1	0.0286	0.0131	0.0077	0.0083	0.0169	0.0094
13	τ_3^2	0.5	0.0027	0.0001	0.0021	0.0001	0.0005	0.0007	35	$\sigma_{a,22}$	3	0.0233	0.0105	0.0247	0.0017	0.0276	0.0150
14	τ_4^2	0.5	0.0007	0.0009	0.0002	0.0002	0.0005	0.0008	36	$\sigma_{a,23}$	1	0.0003	0.0045	0.0158	0.0042	0.0124	0.0051
15	τ_5^2	0.5	0.0002	0.0006	0.0010	0.0013	0.0010	0.0006	37	$\sigma_{a,24}$	2	0.0073	0.0020	0.0057	0.0004	0.0126	0.0089
16	τ_6^2	0.5	0.0044	0.0010	0.0017	0.0005	0.0004	0.0004	38	$\sigma_{a,33}$	3	0.0014	0.0125	0.0383	0.0260	0.0387	0.0274
17	τ_7^2	0.5	0.0005	0.0008	0.0016	0.0021	0.0017	0.0004	39	$\sigma_{a,34}$	1	0.0080	0.0075	0.0236	0.0069	0.0043	0.0022
18	τ_8^2	0.5	0.0027	0.0012	0.0006	0.0006	0.0011	0.0009	40	$\sigma_{a,44}$	3	0.0021	0.0015	0.0215	0.0020	0.0148	0.0094
19	τ_9^2	0.5	0.0016	0.0010	0.0004	0.0009	0.0001	0.0004	41	$\sigma_{\varepsilon,11}$	0.5	0.0020	0.0013	0.0028	0.0005	0.0015	0.0013
20	τ_{10}^2	0.5	0.0009	0.0012	0.0012	0.0005	0.0003	0.0008	42	$\sigma_{\varepsilon,12}$	0.2	0.0017	0.0006	0.0006	0.0003	0.0006	0.0001
21	τ_{11}^2	0.5	0.0001	0.0001	0.0013	0.0000	0.0012	0.0001	43	$\sigma_{\varepsilon,22}$	0.5	0.0021	0.0033	0.0027	0.0029	0.0028	0.0023
22	τ_{12}^2	0.5	0.0011	0.0012	0.0001	0.0002	0.0011	0.0002									

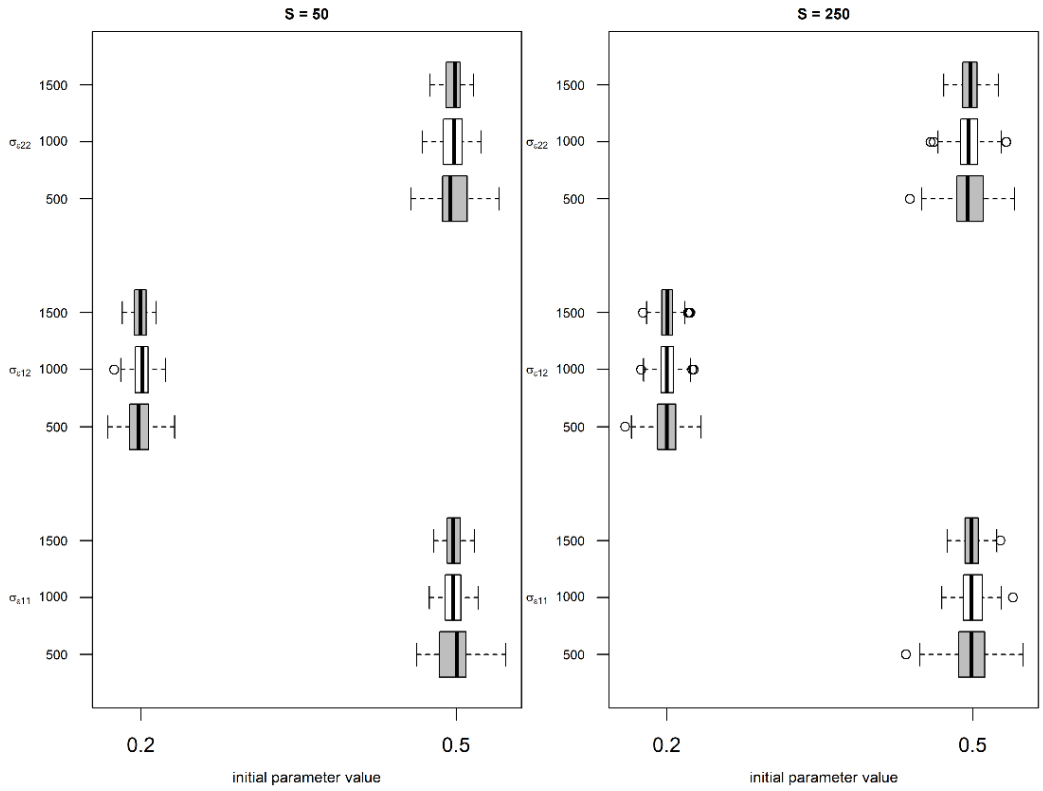


(**a**)



(**b**)

Figure 1. *Cont.*



(**c**)

Figure 1. (**a**) Boxplot the parameters estimate of the latent mixed regression models (β). (**b**) Boxplot the parameters estimate of the latent mixed regression models (σ_a). (**c**) Boxplot the parameters estimate of the latent mixed regression models (σ_ε).

Table 2. The parameter estimates for $\sqrt{V(\hat{\beta})}$.

Number of Subjects	Parameter	The 2nd Moment		SEM	
		50	250	50	250
500	β_1^1	0.0512	0.0528	0.0249	0.0251
	β_1^2	0.0587	0.0530	0.0187	0.0195
	β_2^1	0.0378	0.0392	0.0110	0.0126
	β_2^2	0.0391	0.0379	0.0164	0.0158
	β_3^1	0.0184	0.0190	0.0245	0.0268
	β_3^2	0.0184	0.0190	0.0212	0.0200
	β_4^1	0.0283	0.0283	0.0105	0.0105
	β_4^2	0.0288	0.0308	0.0145	0.0145
1000	β_1^1	0.0338	0.0327	0.0179	0.0184
	β_1^2	0.0342	0.0339	0.0138	0.0134
	β_2^1	0.0232	0.0232	0.0071	0.0077
	β_2^2	0.0253	0.0232	0.0105	0.0100
	β_3^1	0.0095	0.0100	0.0176	0.0184
	β_3^2	0.0089	0.0095	0.0130	0.0130
	β_4^1	0.0141	0.0130	0.0077	0.0077
	β_4^2	0.0130	0.0130	0.0110	0.0105

Table 2. *Cont.*

Number of Subjects	Parameter	The 2nd Moment		SEM	
		50	250	50	250
1500	β_1^1	0.0276	0.0253	0.0152	0.0152
	β_1^2	0.0270	0.0265	0.0105	0.0110
	β_2^1	0.0187	0.0182	0.0063	0.0063
	β_2^2	0.0164	0.0192	0.0084	0.0084
	β_3^1	0.0071	0.0071	0.0145	0.0152
	β_3^2	0.0063	0.0071	0.0105	0.0105
	β_4^1	0.0110	0.0105	0.0063	0.0063
	β_4^2	0.0110	0.0105	0.0084	0.0084

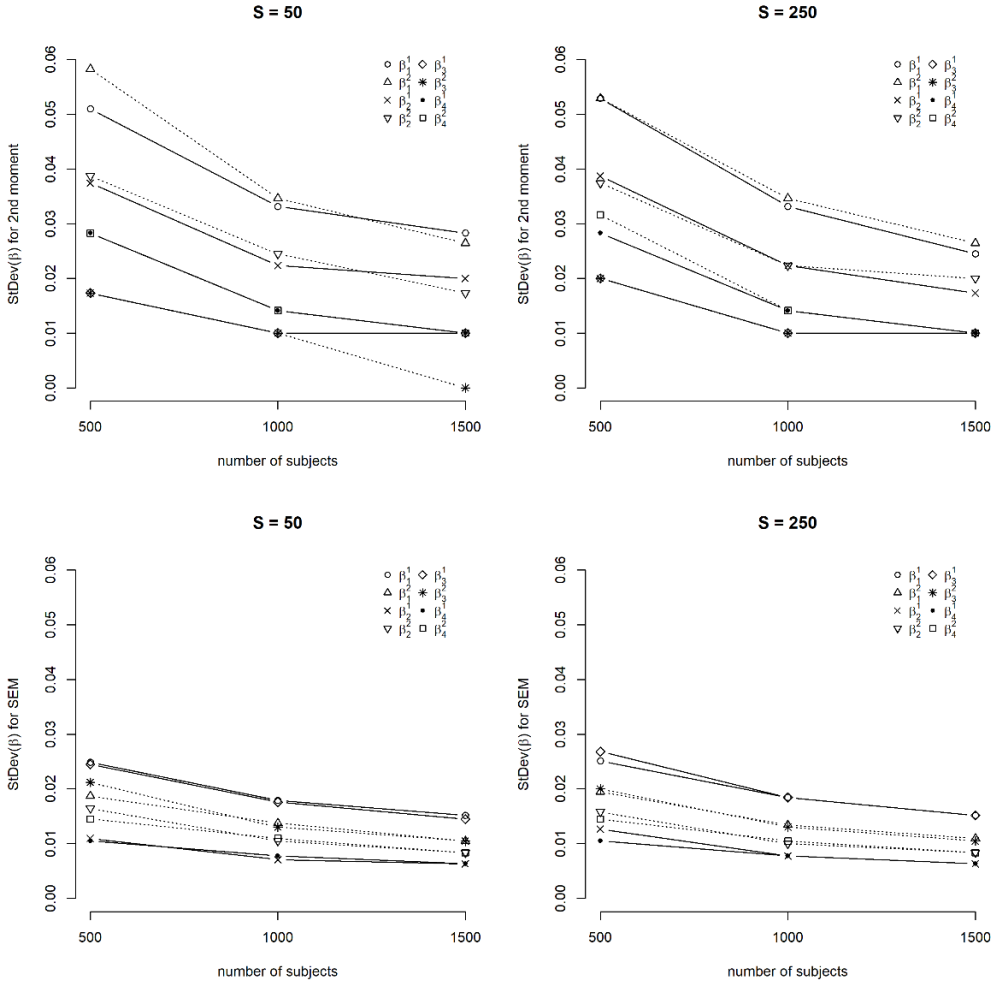


Figure 2. Line plot for the standard deviation of beta.

5. Real Data Example

The real data-set that we used to illustrate the development of the Supplemented EM algorithm is the political attitudes and behavior data of Flemish. The data was designed to include a representative sample of the target population under the Belgian electorate. The Flemish data set (Flemish and Dutch speaking respondents from Brussels Capital Region) consists of 1274 respondents, who have been interviewed three times (1991, 1995, and 1999) [16–18]. There are four latent factors measured on political attitudes of Flemish

used, i.e., Individualism, Nationalism, Ethnocentrism, and Authoritarianism. This data has been analyzed using various methods by several authors, including [19–23]. There are three interesting questions in this real data case, i.e., how Individualism, Nationalism, Ethnocentrism, and Authoritarianism of the Flemish develop over time; whether there is an association between these four developments, and whether the gender of the respondent affects the change patterns of latent developments.

I, N, E, and A in Table 3 correspond to Individualism, Nationalism, Ethnocentrism, and Authoritarianism, respectively. a_{11} and a_{12} are the random intercept and random slope for Individualism. a_{21} and a_{22} are the random intercept and random slope for Nationalism. a_{31} and a_{32} are the random intercept and random slope for Ethnocentrism. a_{41} and a_{42} are the random intercept and random slope for Authoritarianism. The positive correlation of random intercept between a_{11} and a_{21}, a_{11} and a_{31}, a_{11} and a_{41} suggests that the development of Individualism and other political attitudes is highly related, which highest correlated with Ethnocentrism. The results indicate that those who have a better sense of Individualism tend to have a better sense of Nationalism, Ethnocentrism, and Authoritarianism. The results find a positive correlation of random intercept between a_{21} and a_{31}, a_{21} and a_{41}. It suggests that those who have a better sense of Nationalism tend to have a better sense of Ethnocentrism and Authoritarianism, as well as those who have a better sense of Ethnocentrism tend to have a better sense of Authoritarianism. There is also a positive correlation of random slope between a_{12} and a_{22}. It means that if one subject's Individualism decreases over time, then it is reasonable to expect that his or her Nationalism will decrease over time and vice versa. This also holds between Individualism and Ethnocentrism and between Individualism and Authoritarianism. The positive correlation of random slope between a_{22} and a_{32}, meaning that if one subject's Nationalism decreases over time, then it is reasonable to expect that his or her Ethnocentrism will decrease over time. The correlation matrix of random effects confirms that all latent factors have a positive correlation over time.

Table 3. Correlation matrix of random effects.

Random Effects		**I**		**N**		**E**		**A**	
		a_{11}	a_{12}	a_{21}	a_{22}	a_{31}	a_{32}	a_{41}	a_{42}
I	a_{11}	1	0.892	0.832	0.903	0.966	0.930	0.956	0.880
	a_{12}	0.892	1	0.854	0.878	0.937	0.931	0.913	0.895
N	a_{21}	0.832	0.854	1	0.311	0.864	0.883	0.856	0.835
	a_{22}	0.903	0.878	0.311	1	0.918	0.893	0.899	0.861
E	a_{31}	0.966	0.937	0.864	0.918	1	0.946	0.973	0.919
	a_{32}	0.930	0.931	0.883	0.893	0.946	1	0.946	0.918
A	a_{41}	0.956	0.913	0.856	0.899	0.973	0.946	1	0.873
	a_{42}	0.880	0.895	0.835	0.861	0.919	0.918	0.873	1

The significance of parameter estimate of β is analyzed via the z–values. By using the Supplemented EM algorithm, the standard errors of β for all parameters can be calculated. The standard errors of β are listed in Table 4. Using a 95 percent confidence interval of β, almost all confidence intervals do not include the null value, except the slope of Male on Authoritarianism. Hence there are statistically significant differences in the parameter estimate of β. In other words, all latent factors of Flemish people decrease over time, with Ethnocentrism having the highest rate of decline over time (-0.252) and Nationalism the lowest (-0.177). On average, the Individualism and Nationalism of the male respondent are higher than that of the female. However, Ethnocentrism of the male respondent is lower than that of the female.

Table 4. Parameter estimates of β.

Parameter	Estimate	$SE\,(\hat{\beta})$	Lower	Upper
Individualism	−0.195	0.004	−0.203	−0.187
Nationalism	−0.177	0.028	−0.231	−0.123
Ethnocentrism	−0.252	0.011	−0.273	−0.231
Authoritarianism	−0.186	0.002	−0.190	−0.182
Slope of Male on Individualism	0.110	0.010	0.091	0.129
Slope of Male on Nationalism	0.219	0.017	0.185	0.253
Slope of Male on Ethnocentrism	−0.038	0.003	−0.045	−0.031
Slope of Male on Authoritarianism	0.022	0.017	−0.011	0.055

6. Conclusions

This paper proposed the Supplemented EM algorithm for LFLMM in estimating the asymptotic variance-covariance matrix as a by-product of the EM estimator for the case of fixed variables in the model. Results from simulation studies suggest that the Supplemented EM algorithm can estimate the model very close to the initial parameters.

As a result of the development of EM algorithm of LFLMM, the Supplemented EM algorithm is very slow to converge, as stated by [1], especially when the number of simulations is 250 times with 1500 subjects. For this reason, further research is needed to find techniques that can be used to accelerate the speed of the algorithm. Several approaches to speed the EM algorithm have been proposed and can be found in [24–26] (the ECM algorithm), [27] (the ECME algorithm), and [28] (the Parameter-Expanded EM algorithm).

Author Contributions: Conceptualization, Y.A.; methodology, Y.A.; software, Y.A.; validation, K.A.N. and H.F.; formal analysis, Y.A.; data curation, Y.A. and T.T.; writing—original draft preparation, Y.A.; writing—review and editing, K.A.N., H.F., A.S. and T.T.; visualization, T.T.; supervision, K.A.N., H.F. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by RUG and IPB University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from ISPO which were used under license, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of ISPO.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. An, X.; Yang, Q.; Bentler, P.M. A latent factor linear mixed model for high-dimensional longitudinal data analysis. *Stat. Med.* **2013**, *32*, 4229–4239. [CrossRef] [PubMed]
2. Kondaurova, M.V.; Bergeson, R.R.; Xu, H.; Kitamura, C. Affective Properties of Mothers' Speech to Infants with Hearing Impairment and Cochlear Implants. *J. Speech Lang. Hear. Res.* **2015**, *58*, 590–600. [CrossRef]
3. Wang, J.; Luo, S. Multidimensional latent trait linear mixed model: An application in clinical studies with multivariate longitudinal outcomes. *Stat. Med.* **2017**, *36*, 3244–3256. [CrossRef]
4. Ng, S.K.; Krishnan, T.; McLachlan, G. The EM algorithm. In *Handbook of Computational Statistics*; Springer: Berlin, Germany, 2004; pp. 137–168, ISBN 9783642215513.
5. Mclachlan, G.J.; Krishnan, T. *The EM Algorithm and Extensions Second Edition*, 2nd ed.; Wiley: New York, NY, USA, 2007; ISBN 9780471201700.
6. Meng, A.X.; Rubin, D.B. Using EM to Obtain Asymptotic Variance-Covariance Matrices: The SEM Algorithm. *J. Am. Stat. Assoc.* **1991**, *86*, 899–909. [CrossRef]
7. Cai, L. SEM of another flavour: Two new applications of the supplemented EM algorithm. *Br. J. Math. Stat. Psychol.* **2008**, *61*, 309–329. [CrossRef] [PubMed]
8. Cai, L.; Lee, T.; Lee, T. Covariance Structure Model Fit Testing Under Missing Data: An Application of the Supplemented EM Algorithm Covariance Structure Model Fit Testing Under Missing Data: An Application of the Supplemented EM Algorithm. *Multivar. Behav. Res.* **2009**, *44*, 281–304. [CrossRef] [PubMed]

9. Tian, W.; Cai, L.; Thissen, D.; Xin, T. Numerical Differentiation Methods for Computing Error Covariance Matrices in Item Response Theory Modeling: An Evaluation and a New Proposal. *Educ. Psychol. Meas* **2012**, *73*, 412–439. [CrossRef]
10. Caraka, R.E.; Noh, M.; Chen, R.C.; Lee, Y.; Gio, P.U.; Pardamean, B. Connecting climate and communicable disease to penta helix using hierarchical likelihood structural equation modelling. *Symmetry* **2021**, *13*, 657. [CrossRef]
11. Pritikin, J.N. A comparison of parameter covariance estimation methods for item response models in an expectation-maximization framework. *Cogent Psychol.* **2017**, *4*, 1–11. [CrossRef]
12. Orchard, T.; Woodbury, M. A Missing Information Principle: Theory and Applications. In *Theory of Statistics*; University of California Press: Berkeley, CA, USA, 1972; Volume 1, pp. 697–715. Available online: https://projecteuclid.org/download/pdf_1/euclid.bsmsp/1200514117 (accessed on 1 June 2021).
13. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm A. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–38.
14. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing*; Wiley: New York, NY, USA, 2002; ISBN 3175723993.
15. Abel, G.J. International Migration Flow Table Estimation. Ph.D. Thesis, University of Southampton, Southampton, UK, 2009.
16. Interuniversitair Steunpunt Politieke-Opinieonderzoek. *General Election Study: Codebook and Questionnaire*; ISPO: Leuven, Belgium, 1991; ISBN 9067841161.
17. Interuniversitair Steunpunt Politieke-Opinieonderzoek. *General Election Study: Codebook and Questionnaire*; ISPO: Leuven, Belgium, 1995; ISBN 9067841366.
18. Interuniversitair Steunpunt Politieke-Opinieonderzoek. *General Election Study: Codebook and Questionnaire*; ISPO: Leuven, Belgium, 1999.
19. Billiet, J. Church Involvement, Individualism, and Ethnic Prejudice among Flemish Roman Catholics: New Evidence of a Moderating Effect. *J. Sci. Study Relig.* **1995**, *34*, 224–233. [CrossRef]
20. Billiet, J.; Coffe, H.; Maddens, B. Een Vlaams-nationale identiteit en de houding tegenover allochtonen in een longitudinaal perspectief. In *Proceedings of the Paper Presented at the Marktdag Sociologie*; Universitaire Pers Leuven: Leuven, Belgium, 2005.
21. Toharudin, T.; Oud, J.H.L.; Billiet, J.B. Assessing the relationships between Nationalism, Ethnocentrism, and Individualism in Flanders using Bergstrom's approximate discrete model. *Stat. Neerl.* **2008**, *62*, 83–103. [CrossRef]
22. Toharudin, T.; Oud, J.H.L.; Billiet, J.; Folmer, H. Measuring Authoritarianism with Different Sets of Items in a Longitudinal Study. In *Methods, Theories, Andempirical Applications in the Social Sciences*; Salzborn, S., Davidov, E., Reinecke, J., Eds.; Springer: Heidelberg, Germany, 2012; pp. 193–200, ISBN 9783531188980.
23. Angraini, Y.; Toharudin, T.; Folmer, H.; Oud, J.H.L. The Relationships between Individualism, Nationalism, Ethnocentrism, and Authoritarianism in Flanders: A Continuous Time-Structural Equation Modeling Approach. *Multivar. Behav. Res.* **2014**, *49*, 41–53. [CrossRef] [PubMed]
24. Meng, X.; Rubin, D.B. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* **1993**, *80*, 267–278. [CrossRef]
25. Van Dyk, D.A.; Meng, X.; Rubin, D.B. Maximum Likelihood Estimation via the ECM Algorithm: Computing The Asymptotic Variance. *Stat. Sin.* **1995**, *5*, 55–75.
26. Li, H.; Tian, W. Slashed lomax distribution and regression model. *Symmetry* **2020**, *12*, 1877. [CrossRef]
27. Liu, B.Y.C.; Rubin, D.B. The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. *Biometrika* **1994**, *81*, 633–648. [CrossRef]
28. Liu, C.; Rubin, D.B.; Wu, Y.N. Parameter Expansion to Accelerate EM: The PX-EM Algorithm. *Biometrika* **1998**, *85*, 755–770. [CrossRef]