Supplementary Materials for "A Class of Markov Models for Longitudinal Ordinal Data"

Keunbaik Lee*

Ph.D. candidate, Department of Statistics, University of Florida, Gainesville FL 32611, U.S.A.

and

Michael J. Daniels†

Associate Professor, Department of Statistics, University of Florida, Gainesville FL 32611, U.S.A.

October 20, 2006

* *email:* lee@stat.ufl.edu
† *email:* mdaniels@stat.ufl.edu

Appendix A

Detailed calculations for Fisher-scoring for OMTM(1)

Let θ_1 and θ_1^* represent the parameter vector containing β_j and β for $j = 1, \cdots, K-1$. The forms of the derivatives for Fisher-scoring algorithm follow

$$
\begin{aligned}
\frac{\partial \log L^{(2)}}{\partial \theta_1} &= \sum_{i=1}^{N}\sum_{t=2}^{n_i}\sum_{k=1}^{K-1}\left(y_{itk} - P_{itk}^{c}\right)\frac{\partial \triangle_{itk}}{\partial \theta_1}, \\
\frac{\partial \log L^{(1)}}{\partial \theta_1} &= \sum_{i=1}^{N}\sum_{k=1}^{K-1}\left(R_{i1k} - R_{i1k+1}\frac{e^{\phi_{i1k}}}{1+e^{\phi_{i1k}}}\right)\left\{\frac{\partial \phi_{i1k}}{\partial P_{i1k}^{M}}\frac{\partial P_{i1k}^{M}}{\partial \theta_1} + \frac{\partial \phi_{i1k}}{\partial P_{i1k+1}^{M}}\frac{\partial P_{i1k+1}^{M}}{\partial \theta_1}\right\}, \\
\frac{\partial \log L^{(2)}}{\partial \alpha_{1b}^{(d)}} &= \sum_{i=1}^{N}\sum_{t=2}^{n_i}\left\{\sum_{k=1}^{K-1}\left(y_{itk} - P_{itk}^{c}\right)\frac{\partial \triangle_{itk}}{\partial \alpha_{1b}^{(d)}} + \left(y_{itd} - P_{itd}^{c}\right)y_{it-1b}z_{it}\right\},
\end{aligned}
$$

for $a = 1, \cdots, K-1$, and $d = 1, \cdots, K-1$.

$$
\begin{aligned}
E_y\left(-\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_1^{*T}}\right) &= E_y\left(\sum_{i=1}^{N}\sum_{t=2}^{n_i}\sum_{k=1}^{K-1}\sum_{g=1}^{K-1}\frac{\partial P_{itk}^{c}}{\partial \triangle_{itg}}\frac{\partial \triangle_{itk}}{\partial \theta_1}\frac{\partial \triangle_{itg}}{\partial \theta_1^{*T}}\right) + E_y\left(-\frac{\partial^2 \log L^{(1)}}{\partial \theta_1 \partial \theta_1^{*T}}\right), \\
E_y\left(-\frac{\partial^2 \log L}{\partial \alpha_{1b}^{(d)} \partial \alpha_{1e}^{(f)T}}\right) &= E_y\left[\sum_{i=1}^{N}\sum_{t=2}^{n_i}\left\{\sum_{k=1}^{K-1}\left(\sum_{g=1}^{K-1}\frac{\partial P_{itk}^{c}}{\partial \triangle_{itg}}\frac{\partial \triangle_{itg}}{\partial \alpha_{1b}^{(d)}}\frac{\partial \triangle_{itk}}{\partial \alpha_{1e}^{(f)T}} + \frac{\partial P_{itk}^{c}}{\partial \gamma_{itb}^{(d)}}\frac{\partial \gamma_{itb}^{(d)}}{\partial \alpha_{1b}^{(d)}}\frac{\partial \triangle_{itk}}{\partial \alpha_{1e}^{(f)T}}\right)\right.\right. \\
&\qquad \left.\left. + \sum_{g=1}^{K-1}\frac{\partial P_{itf}^{c}}{\partial \triangle_{itg}}\frac{\partial \triangle_{itg}}{\partial \alpha_{1b}^{(d)}}y_{it-1b}z_{it}^{T} + \frac{\partial P_{itf}^{c}}{\partial \gamma_{itb}^{(d)}}\frac{\partial \gamma_{itb}^{(d)}}{\partial \alpha_{1b}^{(d)}}y_{it-1f}z_{it}^{T}\right\}\right], \\
E_y\left(-\frac{\partial^2 \log L}{\partial \alpha_{1b}^{(d)} \partial \theta_1{}^{T}}\right) &= E_y\left[\sum_{i=1}^{N}\sum_{t=2}^{n_i}\sum_{k=1}^{K-1}\left\{\sum_{g=1}^{K-1}\frac{\partial P_{itk}^{c}}{\partial \triangle_{itg}}\frac{\partial \triangle_{itg}}{\partial \alpha_{1b}^{(d)}}\frac{\partial \triangle_{itk}}{\partial \theta_1{}^{T}} + \frac{\partial P_{itk}^{c}}{\partial \gamma_{itb}^{(d)}}\frac{\partial \gamma_{itb}^{(d)}}{\partial \alpha_{1b}^{(d)}}\frac{\partial \triangle_{itk}}{\partial \theta_1{}^{T}}\right\}\right],
\end{aligned}
\tag{A.1}
$$

for $b = 1, \cdots, K-1$, and $d = 1, \cdots, K-1$. Note that (A.1) is zero if there are no missing observations.

To compute the score vector and information matrix, we also need derivatives of $\triangle_{it}$ with respect to β_0, β, and α. They can be obtained as the solution to the following system of linear equations,

$$\frac{\partial \triangle_{it1}}{\partial \theta_1}\sum_{g=1}^{K}\frac{\partial h_{ikg}^{(t)}}{\partial \triangle_{it1}}\pi_{ig}^{(t-1)} + \cdots + \frac{\partial \triangle_{itK-1}}{\partial \theta_1}\sum_{g=1}^{K}\frac{\partial h_{ikg}^{(t)}}{\partial \triangle_{itK-1}}\pi_{ig}^{(t-1)} = \frac{\partial \pi_{ik}^{(t)}}{\partial \theta_1} - \sum_{g=1}^{K} h_{ikg}^{(t)}\frac{\partial \pi_{ig}^{(t-1)}}{\partial \theta_1},$$

$$\frac{\partial \triangle_{it1}}{\partial \alpha_{1b}^{(d)}}\sum_{g=1}^{K}\frac{\partial h_{ikg}^{(t)}}{\partial \triangle_{it1}}\pi_{ig}^{(t-1)} + \cdots + \frac{\partial \triangle_{itK-1}}{\partial \alpha_{1b}^{(d)}}\sum_{g=1}^{K}\frac{\partial h_{ikg}^{(t)}}{\partial \triangle_{itK-1}}\pi_{ig}^{(t-1)} = -\sum_{g=1}^{K}\frac{\partial h_{ikg}^{(t)}}{\partial \gamma_{itb}^{(d)}}\frac{\partial \gamma_{itb}^{(d)}}{\partial \alpha_{1b}^{(d)}}\pi_{ig}^{(t-1)},$$

where $h_{itkg}^{(t)} = P(Y_{it} = k | Y_{it-1} = g, x_{it})$.

Estimates of $\triangle_{it}$ can be obtained using Newton-Raphson. Let

$$f(\triangle_{it}) = (f_1(\triangle_{it}), \cdots, f_{K-1}(\triangle_{it}))),$$

where $f_k(\triangle_{it}) = \sum_{g=1}^{K} h_{ikg}^{(t)}\pi_{ig}^{(t)} - \pi_{ik}^{(t)}$ and $h_{itkg}^{(t)} = P(Y_{it} = k | Y_{it-1} = g, x_{it-1})$. We obtain

$$\triangle_{it}^{(n+1)} = \triangle_{it}^{(n)} - \left(\frac{\partial f(\triangle_{it})}{\partial \triangle_{it}}\right)^{-1} f(\triangle_{it}),$$

where

$$\frac{\partial f(\triangle_{it})}{\partial \triangle_{itj}} = \begin{cases} \sum_{g=1}^{K} h_{ikg}^{(t)}\left(1 - h_{ikg}^{(t)}\right)\pi_{ig}^{(t-1)}, & \text{if } j = k; \\ -\sum_{g=1}^{K} h_{ikg}^{(t)} h_{ijg}^{(t)}\pi_{ig}^{(t-1)}, & \text{if } j \neq k. \end{cases}$$

Appendix B

Maximum likelihood for the second order models

The maximum likelihood algorithm for the OMTM(2) is given below. The OMTM(2) likelihood function consists of three components, conditional probabilities for $t > 2$, $t = 2$, and marginal probabilities for $t = 1$. The log likelihood is given as

$$\begin{aligned}
&\log L(\theta; y) \\
=\ & \sum_{i=1}^{N}\sum_{t=3}^{n_i}\left\{\sum_{k=1}^{K-1} y_{itk}\left(\triangle_{itk} + \sum_{m=1}^{2}\gamma_{itm1}^{(k)} y_{it-m1} + \cdots + \gamma_{itmK-1}^{(k)} y_{it-mK-1}\right) + \log P_{itK}^{c}\right\} \\
& + \sum_{i=1}^{N}\left\{\sum_{k=1}^{K-1} y_{i2k}\left(\triangle_{i2k} + \gamma_{i211}^{(k)} y_{i11} + \cdots + \gamma_{i21K-1}^{(k)} y_{i1K-1}\right) + \log P_{i2K}^{c}\right\} \\
& + \sum_{i=1}^{N}\sum_{k=1}^{K-1}\{R_{i1k}\phi_{i1k} - R_{i1k+1} g(\phi_{i1k})\} \\
\overset{\text{let}}{=}\ & \log L^{(3)}(\theta; y) + \log L^{(2)}(\theta; y) + \log L^{(1)}(\theta; y). \qquad \text{(B.2)}
\end{aligned}$$

where

$$P_{itk}^{c} = \begin{cases} P(Y_{it} = k | y_{it-1}, y_{it-2}), & \text{if } t \geq 3; \\ P(Y_{it} = k | y_{it-1}), & \text{if } t = 2. \end{cases}$$

To evaluate the likelihood, we need to evaluate both the contributions from the initial state, $L^{(1)}$ and $L^{(2)}$, and the subsequent contribution from each transition probability, $L^{(3)}$. Let $\theta = (\beta_0, \beta, \alpha)$ be the vector of parameters. Then,

$$\frac{\partial \log L}{\partial \theta} = \frac{\partial \log L^{(3)}}{\partial \theta} + \frac{\partial \log L^{(2)}}{\partial \theta} + \frac{\partial \log L^{(1)}}{\partial \theta}.$$

The likelihood equations are given in the web appendix. The Fisher-scoring method can be used to solve the likelihood equations. The expectations of the negative second derivatives of log likelihood are given by

$$E\left(-\frac{\partial^2 \log L}{\partial \theta \partial \theta^T}\right) = E\left(-\frac{\partial^2 \log L^{(3)}}{\partial \theta \partial \theta^T}\right) + E\left(-\frac{\partial^2 \log L^{(2)}}{\partial \theta \partial \theta^T}\right) + E\left(-\frac{\partial^2 \log L^{(1)}}{\partial \theta \partial \theta^T}\right).$$

For the explicit forms of these expectations, see the web appendix. The derivatives $\frac{\partial \triangle_{itk}}{\partial \beta_{0b}}$, $\frac{\partial \triangle_{itk}}{\partial \beta}$, and $\frac{\partial \triangle_{itg}}{\partial \alpha_b^{(d)}}$ are given in the web appendix.

The algorithm for maximum likelihood estimation for the IOMTM(2) is similar.

Appendix C

Additional simulation results

We also conducted simulations with the same marginal mean models as in the simulations in the manuscript, but now generating data under the following IOMTM(2),

$$\log\left(\frac{P(Y_{it}=k|Y_{it-1},Y_{it-2},x_{it})}{P(Y_{it}=K|Y_{it-1},Y_{it-2},x_{it})}\right) = \triangle_{itk} + Y_{it-1}\gamma^{(k)}_{it11} + Y^2_{it-1}\gamma^{(k)}_{it12} + Y_{it-2}\gamma^{(k)}_{it21} + Y^2_{it-2}\gamma^{(k)}_{it22},$$

with $\gamma^{(k)}_{itmj} = \alpha^{(k)}_{mj0}$ for $m = 1,2$; $j = 1,2$ where $(\alpha^{(1)}_{110},\alpha^{(2)}_{110},\alpha^{(3)}_{110}) = (-0.1, 0.1, 0.2)$, $(\alpha^{(1)}_{120},\alpha^{(2)}_{120},\alpha^{(3)}_{120}) = (-0.2, 0.1, 0.1)$, $(\alpha^{(1)}_{210},\alpha^{(2)}_{210},\alpha^{(3)}_{210}) = (-0.5,-0.7,-0.5)$, and $(\alpha^{(1)}_{220},\alpha^{(2)}_{220},\alpha^{(3)}_{220}) = (-0.4, 0.5, 0.6)$. We then fit an IOMTM(1) with $z_{it} = 1$. The results with no dropout appear in Table 1 and those with (MAR) dropout appear in Table 2.

Table 1

Bias of IOMTM(2) maximum likelihood estimators when no data are missing. Displayed is the average regression coefficient estimates and the percent relative bias, $100 \times (\hat{\beta} - \beta)/\beta$

Unstructured Marginal Mean				Structured Marginal Mean			
para.	true	mean	bias(%)	para.	true	mean	bias(%)
int1	-1.00	-1.01	1.0	int1	-1.00	-1.00	0.0
int2	0.70	0.70	0.0	int2	0.70	0.70	0.0
int3	2.00	2.00	0.0	int3	2.00	2.00	0.0
grp1	0.10	0.10	0.0	grp1	0.10	0.10	0.0
grp2	-0.50	-0.50	0.0	grp2	-0.50	-0.50	0.0
vis1	0.10	0.10	0.0	vis	-0.50	-0.51	2.0
vis2	-1.50	-1.50	0.0				
vis3	1.20	1.21	0.8				
vis4	-0.40	-0.40	0.0				
vis5	0.80	0.80	0.0				
vis6	-0.30	-0.30	0.0				
vis7	-1.00	-1.00	0.0				

Table 2

Bias of IOMTM(2) maximum likelihood estimators when data are missing at random (MAR). Displayed is the average regression coefficient estimates and the percent relative bias, $100 \times (\hat{\beta} - \beta)/\beta$

Unstructured Marginal mean				Structured Marginal mean			
para.	truth	mean	bias(%)	truth	para.	mean	bias(%)
int1	-1.00	-0.94	-5.6	int1	-1.00	-1.02	1.8
int2	0.70	0.79	12.6	int2	0.70	0.71	1.0
int3	2.00	2.06	3.0	int3	2.00	1.97	-1.6
grp1	0.10	0.10	0.0	grp1	0.10	0.10	0.0
grp2	-0.50	-0.49	-1.6	grp2	-0.50	-0.49	-1.2
vis1	0.10	0.03	-74.0	vis	-0.50	-0.36	-28.2
vis2	-1.50	-1.57	4.9				
vis3	1.20	1.16	-3.3				
vis4	-0.40	-0.39	-2.5				
vis5	0.80	0.76	-5.1				
vis6	-0.30	-0.31	3.0				
vis7	-1.00	-1.06	6.0				