Journal of the Korean Data & Information Science Society
2017, **28**(4), 927–936

http://dx.doi.org/10.7465/jkdi.2017.28.4.927
한국데이터정보과학회지

Poisson linear mixed models with ARMA random effects covariance matrix[†]

Jiin Choi[1] · Keunbaik Lee[2]

[12]Department of Statistics, Sungkyunkwan University

Received 24 June 2017, revised 17 July 2017, accepted 20 July 2017

Abstract

To analyze longitudinal count data, Poisson linear mixed models are commonly used. In the models the random effects covariance matrix explains both within-subject variation and serial correlation of repeated count outcomes. When the random effects covariance matrix is assumed to be misspecified, the estimates of covariates effects can be biased. Therefore, we propose reasonable and flexible structures of the covariance matrix using autoregressive and moving average Cholesky decomposition (ARMACD). The ARMACD factors the covariance matrix into generalized autoregressive parameters (GARPs), generalized moving average parameters (GMAPs) and innovation variances (IVs). Positive IVs guarantee the positive-definiteness of the covariance matrix. In this paper, we use the ARMACD to model the random effects covariance matrix in Poisson loglinear mixed models. We analyze epileptic seizure data using our proposed model.

Keywords: Cholesky decomposition, general linear mixed model, high dimensionality, longitudinal count data, positive-definite.

1. Introduction

In longitudinal studies repeated outcomes from each subject are measured over time and the outcomes are not independent (Diggle *et al.*, 2002). Therefore, the serial correlation of the outcomes should be properly considered in order to reduce bias of covariates effects. When subject-specific effects of covariates are of interest in analysis of longitudinal categorical data, generalized linear mixed models (GLMMs) are often used (Breslow and Clayton, 1993). Random effects in the GLMMs are used to explain both within-subject variation and serial correlation of repeated categorical data using a random effects covariance matrix. The estimation of the covariance matrix is not easy because the matrix is high-dimensional and should be positive-definite (Jeon and Lee, 2014; Lee *et al.*, 2012; Lee *et al.*, 2017; Lee and Kim, 2016). Therefore, a specific structure of the covariance matrix is assumed such

[†] This project was supported by Basic Science Research Program through the National Research Foundation of Korea (KRF) funded by the Ministry of Education, Science and Technology (NRF-2016R1D1A1B03930343).

[1] Graduate student, Department of Statistics, Sungkyunkwan University, Seoul 110-745, Korea.
[2] Corresponding author: Associate professor, Department of Statistics, Sungkyunkwan University, Seoul 110-745, Korea. E-mail: keunbaik@skku.edu.

as autoregressive or compound symmetry. In real data analysis the assumption of the specific structure may be too strong. In addition, when the random effects covariance matrix is assumed to be misspecified, the estimates of covariates effects can be biased (Heagerty and Kurland, 2001; Lee *et al.*, 2017). To relieve the strong assumption, more general structure of the covariance matrix using three Cholesky decompositions were proposed: modified Cholesky decomposition (MCD) (Pourahmadi, 1999, 2000; Daniels and Pourahmadi, 2002; Daniels and Zhao, 2003; Lee *et al.*, 2012; Lee, 2013; Lee and Sung, 2014), moving average Cholesky decomposition (MACD) (Zhang and Leng, 2012; Lee and Yoo, 2014; Kim *et al.*, 2017), and autoregressive moving-average Cholesky decomposition (ARMACD) (Han and Lee, 2016; Lee *et al.*, 2017; Nam and Lee, 2017).

The MCD decomposes the inverse of the covariance matrix (precision matrix) into generalized autoregressive parameters (GARPs) and innovation variances (IVs). The GARPs are coefficients of previous outcomes and the IVs are prediction error variances. The MACD decomposes the covariance matrix into generalized moving average parameters (GMAPs) and IVs. The GMAPs are moving average coefficients of previous errors. In both decompositions, the estimated covariance matrix is positive definite when all IVs are positive. The ARMACD was recently developed to model the covariance matrix with the heterogeneous autoregressive moving average (ARMA) structure (Lee *et al.*, 2017). It combines MCD and MACD to create a more flexible decomposition of the marginal covariance matrix in linear models. The ARMACD factors the heterogeneous marginal covariance matrix into GARPs, GMAPs, and IVs. The positive IVs guarantee the positive-definiteness of the covariance matrix. In this paper, we use the ARMACD to model random effects covariance matrix in Poisson loglinear mixed models for longitudinal count data.

In the analysis of count data, Poisson loglinear models are typically used. In the models, the log function is a natural link function for Poisson distribution and the log link is the canonical link for the Poisson distribution (Agresti, 2002). The log link accounts for effects of covariates to be multiplicative rather than additive. In the analysis of longitudinal Poisson data, Zeger (1988) introduced log linear model with random effect in order to account for time series count data using an estimating equation approach. Markov regression using a quasi-likelihood estimating approach is proposed by Zeger and Qaqish (1988) for time series count data. For Poisson data having transition patterns and non-ignorable missing values, Li *et al.* (2007) proposed first-order Markov transition models with random intercept. Different approaches for modeling longitudinal count data with dropouts are illustrated by Alosh (2010). The approaches include generalized estimating equations (GEE), weighted GEE methods, likelihood-based GLMMs and transition models which are extension of the Poisson autoregressive model with order one. In this paper, we consider Poisson linear mixed model with a general random effect covariance matrix. We decompose the random effects covariance matrix to several parameters using the ARMACD. Our proposed models were easily used to analyze longitudinal count data which commonly occur in longitudinal studies.

The paper is organized as follows. In Section 2, we propose Poisson regression model with ARMA structured random effects covariance matrix. We analyze seizure data using the proposed model in Section 3. Finally conclusions are included in Section 4.

2. Poisson linear mixed model

In this section, we propose Poisson loglinear mixed models with an ARMA structured random effects covariance matrix. The random effects covariance matrix in the models is modeled using the ARMACD.

2.1. Proposed models

Let $y_i = (y_{i1}, y_{i2}, ..., y_{in_i})^T$ be response vector of longitudinal Poisson data where y_{ij} is the Poisson response for subject i $(i = 1, ..., N)$ at time j $(j = 1, ..., n_i)$. We assume that the responses for each subject are conditionally independent given random effects. Then Poisson loglinear mixed model is given by

$$\begin{aligned} &y_{ij} \sim \text{Poisson}(\mu_{ij}(b_{ij})), \\ &\log(\mu_{ij}(b_{ij})) = x_{ij}^T\beta + b_{ij}, \\ &b_i \ \sim^{indep.} \ N(0, \Sigma_i), \end{aligned}$$

where $\mu_{ij}(b_{ij})$ is the conditional mean of y_i given random effects b_{ij}, x_{ij} is the $p\times 1$ covariates, β is the $p \times 1$ coefficient of covariate vector, and $b_i = (b_{i1}, b_{i2}, ..., b_{in_i})^T$ is the vector of the random effects.

Now we assume the ARMA structure of the random effects covariance matrix using the ARMACD

$$b_{ij} = \sum_{k=1}^{j-1} \phi_{ijk} b_{ik} + \sum_{k=1}^{j-1} l_{ijk} e_{ik} + e_{ij}, \tag{2.1}$$

where $e_i = (e_{i1}, e_{i2}, ..., e_{in_i})^T \ \sim \ N(0, D_i)$ with $D_i = diag(\sigma_{i1}^2, \cdots, \sigma_{in_i}^2)$, ϕ_{ijk} are generalized autoregressive parameters (GARPs), l_{ijk} are generalized moving average parameters (GMAPs), and σ_{ij}^2 are innovation variances (IVs). A similar structure was already used in logistic random effects models (Lee *et al.*, 2017a)

We rewrite (2.1) in matrix form as

$$T_i b_i \ = \ L_i e_i, \tag{2.2}$$

where

$$T_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\phi_{i21} & 1 & \cdots & 0 \\ -\phi_{i31} & -\phi_{i32} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\phi_{in_i1} & -\phi_{in_i2} & \cdots & 1 \end{pmatrix}, \quad b_i = \begin{pmatrix} b_{i1} \\ \\ \vdots \\ \\ b_{in_i} \end{pmatrix}, \quad L_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{i21} & 1 & \cdots & 0 \\ l_{i31} & l_{i32} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{in_i1} & l_{in_i2} & \cdots & 1 \end{pmatrix}.$$

Taking variance in (2.2), we have

$$T_i \Sigma_i T_i^T = L_i D_i L_i^T.$$

Then we have following equation

$$\Sigma_i = T_i^{-1} L_i D_i L_i^T T_i^{-T}. \tag{2.3}$$

Note that Σ_i is factored to the GARPs, GMAPs, and IVs. However, it is still high-dimensional and should be positive definite. To release the limitations, linear and loglinear models for the GARPs/GMAPs and IVs are respectively proposed as follows,

$$\phi_{ijk} = w_{ijk}^T\alpha, \quad l_{ijk} = z_{ijk}^T\gamma, \quad \log\sigma_{ij}^2 = h_{ij}^T\lambda, \tag{2.4}$$

where α and γ are vectors of unknown dependence parameters, λ is the vector of unknown variance parameters, and w_{ijk}, z_{ijk} and h_{ij} are time and/or subject-specific design vectors. Note that w_{ijk} and z_{ijk} are time covariate design vectors which play roles of controlling the time order of the model capturing the correlation between responses, and that h_{ij} is subject-specific design matrix. Therefore, the random effects covariance matrix can be structured nonstationary and heteroscedastic. We note that Σ_i is positive definite since all IVs are positive by the loglinear model (Lee *et al.*, 2017). We also note that the matrices T and L in (2.3) are identifiable when they have specific structure of the covariance matrix such as ARMA structure (Lee *et al.*, 2017).

The ARMA model can work when the series is partly autoregressive and partly moving average. This approach provides parsimonious models compared to the AR or the MA models (Judge *et al.*, 1980). Similarly, the parsimony of parametrization in the ARMACD enables reasonable interpretation, easy computation, and stable estimation of parameters (Lee *et al.*, 2017).

2.2. Maximum likelihood estimation

In this subsection, we present derivation of the maximum likelihood estimates for the proposed model. Let $\theta = (\beta, \alpha, \gamma, \lambda)$ be the vector of parameters. Then the marginal likelihood function is the integral over joint distribution of y_i and the random effects b_i, and is given by,

$$L(\theta; y) \;=\; \prod_{i=1}^{N}\int\prod_{j=1}^{n_i}\frac{\mu_{ij}(b_{ij})^{y_{ij}}e^{-\mu_{ij}(b_{ij})}}{y_{ij}!}\; f(b_i)\; db_i, \tag{2.5}$$

where $f(b_i)$ is the normal probability density function with mean vector 0 and covariance matrix Σ_i.

Using (2.3), $f(b_i)$ is re-expressed by

$$f(b_i) = (2\pi)^{-\frac{n_i}{2}}\left(\prod_{j=1}^{n_i}\sigma_{ij}^2\right)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}b_i^TT_i^TL_i^{-T}D_i^{-1}L_i^{-1}T_ib_i\right).$$

Then the log of marginal likelihood function is given by

$$\begin{aligned}\log L(\theta; y) \;=\; & \sum_{i=1}^{N}\log\int\left\{\prod_{j=1}^{n_i}\frac{\mu_{ij}(b_{ij})^{y_{ij}}e^{-\mu_{ij}(b_{ij})}}{y_{ij}!}\right\}\\ & \times(2\pi)^{-\frac{n_i}{2}}\left(\prod_{j=1}^{n_i}\sigma_{ij}^2\right)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}b_i^TT_i^TL_i^{-T}D_i^{-1}L_i^{-1}T_ib_i\right)db_i.\end{aligned}$$

Maximizing the log of marginal likelihood with respect to parameter vector $\theta = (\beta, \alpha, \gamma, \lambda)^T$ yields the estimation equations

$$\sum_{i=1}^{N} \frac{\partial log L(\theta; y_i)}{\partial \theta} \;=\; 0, \tag{2.6}$$

where

$$\frac{\partial \log L(\theta; y_i)}{\partial \beta} = \sum_{i=1}^{N} \frac{1}{L(\theta; y)} \int L(\theta; y, b_i) \sum_{j=1}^{n_i} \{y_{ij} - \mu_{ij}(b_{ij})\} \;\; x_{ij} \;\; f(b_i) db_i, \tag{2.7}$$

$$\begin{aligned}\frac{\partial \log L(\theta; y_i)}{\partial \alpha_a} =& \sum_{i=1}^{N} \frac{1}{L(\theta; y)} \int L(\theta; y, b_i) \\ & \times \left\{ -\frac{1}{2} b_i^T \left(\frac{\partial T_i^T}{\partial \alpha_a} L_i^{-T} D_i^{-1} L_i^{-1} T_i + T_i^T L_i^{-T} D_i^{-1} L_i^{-1} \frac{\partial T_i}{\partial \alpha_a} \right) \right\} f(b_i) db_i,\end{aligned} \tag{2.8}$$

$$\begin{aligned}\frac{\partial \log L(\theta; y_i)}{\partial \gamma_a} =& \sum_{i=1}^{N} \frac{1}{L(\theta; y)} \int L(\theta; y, b_i) \\ & \times \left\{ -\frac{1}{2} b_i^T \left(T_i^T \frac{\partial L_i^{-T}}{\partial \gamma_a} D_i^{-1} L_i^{-1} T_i + T_i^T L_i^{-T} D_i^{-1} \frac{\partial L_i^{-1}}{\partial \gamma_a} T_i \right) b_i \right\} f(b_i) db_i,\end{aligned} \tag{2.9}$$

$$\begin{aligned}\frac{\partial \log L(\theta; y_i)}{\partial \lambda_a} =& \sum_{i=1}^{N} \frac{1}{L(\theta; y)} \int L(\theta; y, b_i) \\ & \times \left\{ -\frac{1}{2} \sum_{j=1}^{n_i} h_{ija} - \frac{1}{2} b_i^T T_i^T L_i^{-T} \frac{\partial D_i^{-1}}{\partial \lambda_l} L_i^{-1} T_i b_i \right\} f(b_i) db_i,\end{aligned} \tag{2.10}$$

with

$$\frac{\partial L_i^{-1}}{\partial \gamma_a} = -L_i \frac{\partial L_i}{\partial \gamma_a} L_i^T,$$

$$\frac{\partial T_i}{\partial \alpha_a} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ -w_{i21a} & 0 & 0 & \cdots & 0 \\ -w_{i31a} & -w_{i32a} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -w_{in_i1a} & -w_{in_i2a} & -w_{in_i3a} & \cdots & 0 \end{pmatrix},$$

$$\frac{\partial L_i}{\partial \gamma_a} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ z_{i21a} & 0 & 0 & \cdots & 0 \\ z_{i31a} & z_{i32a} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{in_i1a} & z_{in_i2a} & z_{in_i3a} & \cdots & 0 \end{pmatrix},$$

$$\frac{\partial D_i^{-1}}{\partial \lambda_l} = diag \left\{ -\frac{1}{\sigma_{i1}^2} h_{i1l}, \cdots, -\frac{1}{\sigma_{in_i}^2} h_{in_i l} \right\}.$$

Since the solution of the equations (2.7)-(2.10) is computationally intensive, we used Quasi-Monte Carlo (QMC) approximation to compute integrations inside of the the log-likelihood in (2.5). So we have the following result:

$$L(\theta; y) \approx \prod_{i=1}^{N} \frac{1}{M} \sum_{l=1}^{M} \prod_{j=1}^{n_i} \frac{\mu_{ij}(b_{ij}^{(l)})^{y_{ij}} e^{-\mu_{ij}(b_{ij}^{(l)})}}{y_{ij}!}, \tag{2.11}$$

where the set $(b_i^{(1)}, \cdots, b_i^{(M)})$ is a subsequence of a low-discrepancy sequence. In Section 3, we use a function, **rnorm.sobol()**, in the library fOptions (Wuertz, 2005) of R to get the set. The QMC in (2.7)-(2.10) are similar to the approximation in (2.11).

The matrix of second derivatives of the observed data log likelihood has a very complex form. Fortunately, the sample empirical covariance matrix of the individual scores in any correctly specified model is a consistent estimator of the information and involves only the first derivatives (Lee and Daniels, 2008). Therefore, the quasi-Newton method can be used to solve the likelihood equations, using

$$\theta^{(c+1)} = \theta^{(c)} + [H(\theta^{(c)}; y)]^{-1} \frac{\partial \log L(\theta; y_i)}{\partial \theta^{(c)}},$$

where $H(\theta^{(c)}; y)$ is empirically estimated information matrix at step c, which is given by

$$H(\theta^{(c)}; y) = \sum_{i=1}^{N} \frac{\partial log L(\theta; y_i)}{\partial \theta^{(c)}} \frac{\partial log L(\theta; y_i)}{\partial \theta^{(c)T}}.$$

At convergence, the large-sample covariance matrix of the estimated parameters is the inverse of $H(\theta^{(c)}; y)$.

3. Data analysis

3.1. Data description

Epilepsy is a part of neurological diseases which can be characterized by epileptic seizures. Symptoms of epilepsy include episodes that can vary from brief and nearly undetectable to long periods of vigorous shaking and can result in physical injuries such as occasionally broken bones. We used medical data including information of various patients with epilepsy and anti-epileptic drug (AED). The data were collected from a randomized, double-blind, parallel group multi-center study in purpose of comparing placebo with a new anti-epileptic drug (AED), in combination with one or two other AEDs (Faught *et al.*, 1996).

The effect of AED was stabilized for 12-week baseline period during which the number of seizures was collected. The randomization of patients with epilepsy took place after the above period, 45 patients were categorized to placebo group and rest of 44 patients to the AED treatment group. Then the number of seizures of each patient was counted weekly. During 16 weeks patients were followed based on double-blind rule after which they entered to a long-term open-extension study. Several patients were followed for 27 weeks. The objective of this research was to decide whether or not this additional new treatment had effect on reducing the number of epileptic seizures.

Table 3.1 Models for covariance matrix with GARPs (ϕ_{ijk}), GMAPs (l_{ijk}) and IVs ($\log\sigma_{ij}^2$) in the epileptic seizure data

Model	GARP	GAMP	IV
Model 1	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)}$		$\log\sigma_{ij}^2 = \lambda_0$
Model 2	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)} + \alpha_1 I_{(\lvert j-k\rvert=2)}$		$\log\sigma_{ij}^2 = \lambda_0$
Model 3	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)}$	$l_{ijk} = \gamma_0 I_{(\lvert j-k\rvert=1)}$	$\log\sigma_{ij}^2 = \lambda_0$
Model 4	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)}$	$l_{ijk} = \gamma_0 I_{(\lvert j-k\rvert=1)}$	$\log\sigma_{ij}^2 = \lambda_0 + \lambda_1 \times Trt_i$
Model 5	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)} + \alpha_1 I_{(\lvert j-k\rvert=1)} \times Trt_i$	$l_{ijk} = \gamma_0 I_{(\lvert j-k\rvert=1)} + \gamma_1 I_{(\lvert j-k\rvert=1)} \times Trt_i$	$\log\sigma_{ij}^2 = \lambda_0$ $\log\sigma_{ij}^2 = \lambda_0$
Model 6	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)} + \alpha_1 I_{(\lvert j-k\rvert=1)} \times Trt_i$	$l_{ijk} = \gamma_0 I_{(\lvert j-k\rvert=1)} + \gamma_1 I_{(\lvert j-k\rvert=1)} \times Trt_i$	$\log\sigma_{ij}^2 = \lambda_0 + \lambda_1 \times Trt_i$
Model 7	$\phi_{ijk} = \alpha_0 I_{(\lvert j-k\rvert=1)} + \alpha_1 I_{(\lvert j-k\rvert=1)} \times Trt_i$	$l_{ijk} = \gamma_0 I_{(\lvert j-k\rvert=1)} + \gamma_1 I_{(\lvert j-k\rvert=1)} \times Trt_i + \gamma_2 I_{(\lvert j-k\rvert=2)} + \gamma_3 I_{(\lvert j-k\rvert=2)} \times Trt_i$	$\log\sigma_{ij}^2 = \lambda_0 + \lambda_1 \times Trt_i$

To summarize data, the number of seizures for one week was mostly from 0 to 10 times. However, the data were very skewed and the maximum number of seizures was up to 73 times for a one week. The median and average values of outcomes were affected by extreme values such as 73 and number of observations was some time-points such as 3 patients in week 27.

The main question of interest in this experiment design was to know the effectiveness of the AED. For the purpose, we let the count response variable Y be the number of seizure. To examine effect of AED, we made type of treatment (Trt=0 for placebo group, Trt=1 for AED group). The visit number was standardized and the interaction effect between treatment and visit number was also included.

3.2. Model fit

We analyzed the epileptic seizure data using our proposed models. Models with the AR or ARMA structure of random effects covariance matrix were considered to compare. In particular, we considered seven models specified in Table 3.1

Models 1 and 2 indicate Poisson linear mixed models (PLMMs) with a homoscedastic AR(1) and AR(2) covariance matrix, respectively. Model 3 indicates the PLMM with a homoscedastic ARMA(1,1) random effects covariance matrix. Model 4 indicates the PLMM with a heteroscedastic ARMA(1,1) covariance matrix having IV depending on type of treatment. Models 5 and 6 indicate the PLMMs with homogeneous and heteroscedastic ARMA(1,1) covariance matrices having IV depending on type of treatment, respectively. Lastly, Model 7 indicates the PLMM with a heteroscedastic ARMA(1,2) covariance matrix having IVs depending on type of treatment.

We implemented a quasi-Newton algorithm to find MLEs of parameters for the models using quasi Monte Carlo integration. In our analysis below, we obtained convergence in 50 iterations using a fairly strict convergence criterion, $|(\theta^{new} - \theta^{old})/\theta^{new}| < 10^{-3}$ where θ^{new} and θ^{old} were current and previous fitted values of parameters, respectively. To compare the models, we considered maximized loglikelihood values and Akaike information criteria (AIC) (Akaike, 1974) (Table 3.2). Note that $AIC = -2\ \log L(\hat{\theta}) + 2\ p$ where $\hat{\theta}$ is the maximum likelihood estimate and p is the number of parameters.

The five ARMA models (Models 3-7) fit better than AR models (Models 1 and 2) according

Table 3.2 Real data analysis. Maximized loglikelihood and AIC for the models

	Max. loglikelihood	AIC
Model 1	-3967.810	7949.620
Model 2	-3947.215	7910.430
Model 3	-3925.025	7864.050
Model 4	-3917.827	7851.654
Model 5	-3918.196	7854.392
Model 6	-3913.314	7846.628
Model 7	-3870.293	7764.586

to AIC. The AIC of the Model 7 was the smallest among five ARMA models demonstrating that Model 7 fits most appropriately for this data. We also conducted likelihood ratio test (LRT) for the comparison of nested models (Table 3.3). Note that

$$\chi^2 = -2\left\{\log L(\hat{\theta}_{M1}) - \log L(\hat{\theta}_{M2})\right\},$$

where $\hat{\theta}_{M1}$ and $\hat{\theta}_{M2}$ are the maximum likelihood estimates under two nested models M_1 and M_2, respectively. Test result between Models 3 and 4 shows that there was significant difference. Similarly, the LRTs between Models 5 and 6, Models 3 and 6, Models 6 and 7 indicate that there were significant difference. Among models with ARMA(1,1) random effects covariance matrices, Model 6 was better than Models 3 and 5. We also compared Models 6 and 7 and the LRT indicates that Model 7 was better than Model 6.

Table 3.3 Likelihood ratio test

	χ^2	*df*	p-value
Model 3 vs Model 4	14.396	1	≤ 0.001
Model 5 vs Model 6	9.764	1	0.002
Model 3 vs Model 6	23.422	2	≤ 0.001
Model 6 vs Model 7	43.021	2	≤ 0.001

χ^2 =likelihood ratio test statistic

Table 3.4 presents the MLEs and SEs of parameters for Model 2 and Models 5-7. The MLEs of the fixed effect of covariates were similar in the models with ARMA (1,1) structure of random effect covariance matrix (Models 5 and 6). The models with different structure of random effect covariance matrix (Model 2 and Models 5-7) had slightly different estimates for fixed effect. Since Model 7 was the best among the models considered, we focus on Model 7. The estimate of the IVs, λ_0 and λ_1 were significant ($\hat{\lambda_0} = -1.355$, $SE = 0.147$, $\hat{\lambda_1} = 0.694$, $SE = 0.171$). This result indicates that the random effects covariance matrix differed by type of group. The parameters of GARPs and GMAPs were not significant except for intercept of AR(1) ($\hat{\alpha_0} = 1.010$, $SE = 0.033$). For the fixed effects, time effect and interaction between time and treatment effect were significant under 5% significant level ($\beta_2 = -0.678$, $SE = 0.189$, $\beta_3 = 0.966$, $SE = 0.209$). It indicates that the time effect was significantly different between placebo and treatment groups. The estimate of coefficient for Trt was not significant. For the group with AED, ($Trt = 1$), controlling the random effect, log of the estimated mean of epileptic seizure occurrence during one week increased by 0.288 ($= -0.678 + 0.966$) as visit time increased by one. For the placebo ($Trt = 0$), log of the estimated mean decreased by 0.678 as visit time increased by one controlling random effect.

Table 3.4 Maximum likelihood estimates for five models.

	Model 2	Model 5	Model 6	Model 7
Parameters: β				
Intercept	0.743*(0.081)	0.846*(0.090)	0.908*(0.071)	0.910*(0.089)
Treatment	0.269*(0.092)	0.345*(0.103)	0.229*(0.087)	0.090(0.091)
Time	-0.315(0.237)	-0.451(0.259)	-0.681*(0.161)	-0.678*(0.189)
Treatment×Time	-0.108(0.262)	0.579*(0.294)	0.781*(0.218)	0.966*(0.209)
GARPs and GMAPs: γ and η				
Intercept (AR(1)-1)	0.587*(0.055)	1.006*(0.026)	1.007*(0.026)	1.010*(0.033)
Treatment (AR(1)-2)		-0.007(0.032)	-0.006(0.032)	0.013(0.038)
Intercept (AR(2)-1)	-0.395*(0.061)			-0.355(0.196)
Treatment (AR(2)-2)				0.038(0.221)
Intercept (MA(1)-1)		-0.486*(0.096)	-0.407*(0.110)	-0.068(0.269)
Treatment (MA(1)-2)		0.200(0.119)	0.016(0.134)	-0.205(0.286)
IVs parameters: λ				
Intercept	-1.11*(0.064)	-1.015*(0.065)	-1.338*(0.119)	-1.355*(0.147)
Treatment			0.613*(0.148)	0.694*(0.171)

*Significance at 95% confidence level.

4. Conclusion

We proposed Poisson loglinear mixed models (PLMM) with the random effects covariance matrix for longitudinal Poisson data. The random effect covariance matrix was factored via ARMA Choleksy decomposition into generalized auto regressive parameters (GARPs), generalized moving average parameters (GMAPs) and innovation variance (IVs). When the IVs are positive, the random effects covariance matrix is positive definite. The number of parameters in the covariance matrix is reduced by considering linear and loglinear models for the GARPs/GMAPs and the IVs, respectively. It is well known that the ARMACD allows flexible and parsimonious modeling of the covariance matrix.

Estimation of parameters in the proposed model was implemented by a quasi-Newton algorithm. Integrations in the derivatives of loglikelihood were approximated using quasi Monte Carlo integration.

In analysis of epileptic seizure data, we fitted four PLMMs with several AR/ARMA structures of the random effects covariance matrix. The PLMMs with an ARMA(1,2) random effects covariance matrix having IV depending on treatment fitted better than the other three models. Time effect was significantly different between placebo and treatment groups under 5% significant level.

References

Agresti, A. (2002). *Categorical data analysis*, Wiley and Sons, New York.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions Automatic Control*, **19**, 716-723.

Alosh, M. (2010). Modeling longitudinal count data with dropouts. *Pharmaceutical Statistics*, **9**, 35-45.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, **88**, 125-134.

Daniels, M. J. and Pourahmadi, M. (2002). Bayesian analysis of covariance matrices and dynamic models for longitudinal data. *Biometrika*, **89**, 553-566.

Daniels, M. J. and Zhao, Y. D. (2003). Modelling the random effects covariance matrix in longitudinal data. *Statistics in Medicine*, **22**, 1631-1647

Diggle, P. J., Heagerty, P., Liang, K. Y. and Zeger, S. (2002). *Analysis of longitudinal data*, Oxford University Press, Oxford.

Faught, E., Wilder, B. J., Ramsay, R. E., Reife, R. A., Kramer, L. D., Pledger, G. W. and Karim, R. M. (1996). Topiramate placebo-controlled dose-ranging trial in refractory partial epilepsy using 200-, 400-, and 600-mg daily dosages. *Neurology*, **46**, 1684-1690.

Han, E. J. and Lee, K. (2016). Dynamic linear mixed models with ARMA covariance matrix. *Communications for Statistical Applications and Methods*, **23**, 575-585.

Heagerty, P. J. and Kurland, B. F. (2001). Misspecified maximum likelihood estimates and generalised linear mixed models. *Biometrika*, **88**, 973-985.

Jeon, J. and Lee, K. (2014). Review and discussion of marginalized random effects models. *Journal of the Korean Data & Information Science Society*, **25**, 1263-1272.

Judge, G. G., Griffiths, W. E., Hill, R. C., and Lee, T. C. (1980). *The Theory and practice of ecomometrics*, Wiley, New York.

Kim, J., Sohn, I., and Lee, K. (2017). Bayesian modeling of random effects precision/covariance matrix in cumulative logit random effects models. *Communications for Statistical Applications and Methods*, **24**, 81-96.

Lee, K., Baek, C. and Daniels, M. J. (2017). ARMA Cholesky factor models for the covariance matrix of linear models. *Computational Statistics and Data Analysis*, In press.

Lee, K. and Daniels, M. (2008). Marginalized models for longitudinal ordinal data with application to quality of life studies. *Statistics in Medicine*, **27**, 4359-4380.

Lee, K., Jung, H. and Yoo, J. K. (2017a). Modeling of the ARMA random effects covariance matrix in logistic random effects models. *Working paper*.

Lee, K. and Kim, S. (2016). Modeling of random effects covariance matrix in marginalized random effects models. *Journal of the Korean Data & Information Science Society*, **27**, 815-825.

Lee, K. (2013). Bayesian modeling of random effects covariance matrix for generalized linear mixed models. *Communications for Statistical Applications and Methods*, **20**, 235-240.

Lee, K., Lee, J., Hagan, J. and Yoo, J. K. (2012). Modeling the random effects covariance matrix for the generalized linear mixed models. *Computational Statistics and Data Analysis*, **56**, 1545-1551.

Lee, K. and Sung S. (2014). Autoregressive Cholesky factor model for marginalized random effects model. *Communications for Statistical Applications and Methods*, **21**, 169-181.

Lee, K. and Yoo, J. K. (2014). Bayesian Cholesky factor models in random effects covariance matrix for generalized linear mixed models. *Computational Statistics and Data Analysis*, **80**, 111-116.

Li, J., Yang, X., Wu, Y., and Shoptaw, S. (2007). A random-effects Markov transition model for Poisson-distributed repeated measures with nonignorable missing values. *Statistics in Medicine*, **26**, 2519-2532.

Nam, S. and Lee, K. (2017). Comparison of the covariance matrix for general linear model. *The Korean Journal of Applied Statistics*, **30**, 103-117.

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation. *Biometrika*, **86**, 677-690.

Pourahmadi, M. (2000). Maximum likelihood estimation of generalised linear models for multivariate normal covariance matrix. *Biometrika*, **87**, 425-435.

Wuertz, D. (2005). *fOptions: Financial software collection-fOptions*, R package version 220.10063. (`http://www.rmetrics.org`).

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika*, **75**, 621-629.

Zeger, S. L. and Qaqish, B. (1988). Markov regression models for time series: A quasi-likelihood approach. *Biometrics*, **44**, 1019-1031.

Zhang, W. and Leng, C. (2012). A moving average Cholesky factor model in covariance modelling for longitudinal data. *Biometrika*, **99**, 141-150.