Statistics and Its Interface Volume 12 (2019) 479–487

# Analysis of longitudinal data by combining multiple dynamic covariance models

Lin Xu, Man-Lai Tang, and Ziqi Chen*

In longitudinal data analysis, it is crucial to understand the dynamic of the covariance matrix of repeated measurements and correctly model it in order to achieve efficient estimators of the mean regression parameters. It is well known that any incorrect covariance matrices can result in inefficient estimators of the mean regression parameters. In this article, we propose an empirical likelihood based method which combines the advantages of different dynamic covariance modeling approaches. The effectiveness of the proposed approach is demonstrated by an anesthesiology dataset and some simulation studies.

Keywords and phrases: Empirical likelihood, Longitudinal data analysis, Maximum likelihood, Modified Cholesky decomposition, Multiple covariance models.

## 1. INTRODUCTION

In longitudinal studies, repeated measurements are observed on subjects over time and responses from the same subject are very likely to be correlated (Liang and Zeger, 1986; Diggle et al., 2002). The within-subject correlation must be incorporated into the estimation of mean regression parameters; otherwise, the resultant estimators may be inefficient (Qu et al., 2000; Daniels and Zhao, 2003). Although one can specify the within-subject correlation structure (e.g., AR, MA, or exchangeable structure) with some index parameters (Diggle et al., 2002; Qu et al., 2000; Leung et al., 2009; Li and Pan, 2013), more general forms of the correlation structure are not allowed, and covariates, which may help to explain the dynamic of the covariance matrices for different subjects, can not be flexibly incorporated. To get efficient maximum likelihood or GEE estimators of the mean parameters, dynamic covariance matrices changing in response to some subject-dependent covariates could be adopted (Pourahmadi, 2000; Ye and Pan, 2006; Chen et al., 2011b).

Recently, parsimonious models for characterising the dependence structure among repeated measurements has attracted increasing attention. Pourahmadi (1999) proposed to dynamically model the covariance matrices by using the modified Cholesky decomposition (see also, Pourahmadi (2007); Pan and Mackenzie (2003); Leng et al. (2010)). An attractive aspect of such a decomposition is that the entries in this decomposition have autoregressive and log innovation interpretations. Zhang and Leng (2012) considered an alternative decomposition in which the entries have moving average and log innovation interpretations. By applying hyperspherical coordinates, Zhang et al. (2015) proposed a novel dynamic variance-correlation modeling approach. It is noteworthy that these parsimonious covariance models are more flexible and adaptive than those only specifying the correlation structure (e.g., AR, MA, or exchangeable structure). Thus, it is expected that more efficient maximum likelihood or GEE estimators of the mean regression parameters could be obtained based on those estimated dynamic covariance matrices (see, Pourahmadi (2000); Ye and Pan (2006)).

Different dynamic covariance modeling approaches may result in different estimators of the mean regression parameters but selection of the correct covariance modeling method has not been developed to the best of our knowledge. In this article, we propose an empirical likelihood (Owen, 1988, 2001) based method which combines the advantages of the existing dynamic covariance modeling approaches and thus yields efficient mean coefficient estimators.

The rest of the paper is organized as follows. In Section 2, we propose a combined multiple likelihood (*CML*) estimating procedure based on three well-known dynamic covariance models. In Section 3, the spinal anesthesiology data is analyzed, and comprehensive simulation studies are conducted to evaluate the finite-sample performance of the proposed method. Finally, some concluding remarks are briefly summarized in Section 4. The detailed calculation techniques of the combined multiple likelihood estimation and an constrained iterative profile optimization algorithm are given in the Supplementary Materials http://intlpress.com/site/pub/files/_supp/sii/2019/0012/0003/SII-2019-0012-0003-s003.pdf.

## 2. METHODOLOGY

Suppose we have $n$ independent subjects, and the $ith$ subject has $m_i$ generic repeated measurements $\mathbf{y}_i =$

*Corresponding author.

$(y_{i1}, y_{i2}, \ldots, y_{im_i})^\top$ which are observed at irregular time points $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{im_i})^\top$. The corresponding design data are recorded in a $m_i \times p$ matrix, denoted as $\mathbf{x}_i$, which includes a column of units if an intercept term is desired. By allowing the number of repeated measurements $m_i$ to be subject specific, our framework is valid for the unbalanced longitudinal data. Assume the response vector $\mathbf{y}_i$ follows a multivariate normal distribution $N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ with

$$\boldsymbol{\mu}_i = E(\mathbf{y}_i|\mathbf{x}_i) = \boldsymbol{\mu}(\mathbf{x}_i\boldsymbol{\beta}), \quad i = 1, \ldots, n,$$

where $\boldsymbol{\mu}(\cdot)$ is a differentiable link function and $\boldsymbol{\Sigma}_i$ is the $m_i \times m_i$ positive definite covariance matrix. We are interested in the efficient estimation of $\boldsymbol{\beta}$ in this paper.

The traditional static covariation patterns (e.g., AR, MA, or exchangeable structure) could not depict the dynamic dependence structure. Dynamic covariance modelling approaches provides some unconstrained parametrizations by modelling the dependence structure and innovation variance in a time series context (Kou and Pan, 2010; Fan and Wu, 2008). An appropriate modelling of covariance structure will improve efficiency of mean parameter estimators (Wang and Carey, 2003; Lin and Carroll, 2006; Carey and Wang, 2011).

## 2.1 Existing dynamic covariance modeling approaches

The modified Cholesky decomposition factor based covariance model originated from Pourahmadi (1999, 2000) is a commonly used dynamic covariance modelling method, which provides a unconstrained parametrization for the covariance matrix via modeling the autoregressive parameters and the innovation variances using covariates. They proposed to decompose the covariance matrix $\boldsymbol{\Sigma}_i$ as $\mathbf{T}_i\boldsymbol{\Sigma}_i\mathbf{T}_i^\top = \mathbf{D}_i$ where $\mathbf{T}_i$ is a lower unitriangular matrix with the $(j, k)$-th below diagonal entry being $-\phi_{ijk}$ and $\mathbf{D}_i = diag(\sigma_{i1}^2, \ldots, \sigma_{im_i}^2)$. The below diagonal entries of $\mathbf{T}_i$ are the negatives of the autoregressive coefficients $\phi_{ijk}$ in the following autoregressive model:

$$(1) \quad y_{ij} - \mu_{ij} = \sum_{k=1}^{j-1} \phi_{ijk}(y_{ik} - \mu_{ik}) + \epsilon_{ij}, \quad j = 1, 2, \ldots, m_i.$$

Here, the notation $\Sigma_{k=1}^0$ means zero throughout this paper. The diagonal elements $\sigma_{ij}^2$ of $\mathbf{D}_i$ are the innovation variances $\sigma_{ij}^2 = var(\epsilon_{ij})$. The autoregressive parameters $\phi_{ijk}$ and innovation variances could be modeled dynamically as

$$(2) \qquad \phi_{ijk} = \mathbf{z}_{ijk}^\top\boldsymbol{\gamma}, \qquad \ln(\sigma_{ij}^2) = \mathbf{h}_{ij}^\top\boldsymbol{\lambda},$$

where $\mathbf{z}_{ijk}$ and $\mathbf{h}_{ij}$ are $q \times 1$ and $d \times 1$ vectors of covariates, respectively. The covariates $\mathbf{z}_{ijk}$ usually could be taken as a polynomial of time difference $t_{ik} - t_{ij}$ for $k > j$. Up to constants, the minus twice log-likelihood function can be written as

$$\begin{aligned} -2l(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = &\sum_{i=1}^n \ln|\boldsymbol{\Sigma}_i| \\ &+ \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu}(\mathbf{x}_i\boldsymbol{\beta}))^T\boldsymbol{\Sigma}_i(\boldsymbol{\gamma}, \boldsymbol{\lambda})^{-1} \\ &\times (\mathbf{y}_i - \boldsymbol{\mu}(\mathbf{x}_i\boldsymbol{\beta})). \end{aligned}$$

We minimize $-2l(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda})$ to obtain the maximum likelihood estimators of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda}$ (i.e., $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\lambda}}$). The detailed algorithms could be found in the paper of Pourahmadi (2000). As a result, we can get the estimates $\hat{\mathbf{T}}_i$ and $\hat{\mathbf{D}}_i$ of $\mathbf{T}_i$ and $\mathbf{D}_i$, respectively. The estimate of $\boldsymbol{\Sigma}_i$ could be simply computed by $\hat{\boldsymbol{\Sigma}}_i^{MCDF} = \hat{\mathbf{T}}_i^{-1}\hat{\mathbf{D}}_i\hat{\mathbf{T}}_i^{T-1}$, which is positive definite. This is the so-called the modified Cholesky decomposition factor estimating procedure (*MCDF*).

Zhang and Leng (2012) provided the moving average Cholesky factor modeling for parameterizing covariance structures. Briefly, they decomposed $\boldsymbol{\Sigma}_i$ as $\boldsymbol{\Sigma}_i = \mathbf{L}_i\mathbf{D}_i\mathbf{L}_i^\top$, where $\mathbf{L}_i$ is a lower unitriangular matrix. The entries of $\mathbf{L}_i = (\phi_{ijk})$ could be interpreted as the moving average parameters in

$$(3) \qquad y_{ij} - \mu_{ij} = \sum_{k=1}^{j-1} \phi_{ijk}\varepsilon_{ik} + \varepsilon_{ij}, \quad j = 1, \ldots, m_i,$$

where $\varepsilon_{i1} = y_{i1} - \mu_{i1}$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{im_i}) \sim N(\mathbf{0}, \mathbf{D}_i)$. The elements $\sigma_{ij}^2$ of the diagonal matrix $\mathbf{D}_i$ are the innovation variances, $\sigma_{ij}^2 = var(\varepsilon_{ij})$. The entries in this decomposition have a moving average and log innovation interpretation and are modeled as linear functions of covariates. Specifically, the models for $\phi_{ijk}$ and $\sigma_{ij}^2$ can be found in Equation (2). Based on Equation (2) or (3), we could get the maximum likelihood estimators of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\lambda}$ (i.e., $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\lambda}}$, for details, see Zhang and Leng (2012)). Thus, we can get the estimates $\hat{\mathbf{L}}_i$ and $\hat{\mathbf{D}}_i$ of $\mathbf{L}_i$ and $\mathbf{D}_i$, respectively. The estimate of $\boldsymbol{\Sigma}_i$ could then be calculated as $\hat{\boldsymbol{\Sigma}}_i^{MACF} = \hat{\mathbf{L}}_i\hat{\mathbf{D}}_i\hat{\mathbf{L}}_i^T$. This is the moving average Cholesky factor estimating procedure (*MACF*).

Zhang et al. (2015) proposed an unconstrained parametrization for the correlation matrix by using the hyperspherical coordinates. Specially, write $\boldsymbol{\Sigma}_i = \boldsymbol{\Gamma}_i\mathbf{R}_i\boldsymbol{\Gamma}_i$, where $\boldsymbol{\Gamma}_i = diag(\sigma_{i1}, \ldots, \sigma_{im_i})$ with $\sigma_{ij}$ being the standard deviation of $y_{ij}$ and $\mathbf{R}_i = (\rho_{ijk})_{j,k=1}^{m_i}$ is the correlation matrix of $\mathbf{y}_i$. Using hyperspherical coordinates and trigonometric functions, the correlation matrices $\mathbf{R}_i$ can be parameterized as $\mathbf{R}_i = \mathbf{C}_i\mathbf{C}_i^T$, where $\mathbf{C}_i$ is a lower triangular matrix:

$$\mathbf{C}_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ c_{i21} & s_{i21} & \cdots & 0 \\ c_{i31} & c_{i32}s_{i31} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{im_i1} & c_{im_i2}s_{im_i1} & \cdots & \prod_{l=1}^{m_i-1} s_{im_il} \end{pmatrix},$$

with $c_{ijk} = \cos(\phi_{ijk})$ and $s_{ijk} = \sin(\phi_{ijk})$. Here, $\phi_{ijk}$ is a function of the correlation parameters and could be interpreted as an angle (for details, see, Zhang et al. (2015)). $\phi_{ijk}$ and $\sigma_{ij}^2$ are modeled as in Equation (2). Using the algorithms proposed by Zhang et al. (2015), we could obtain the maximum likelihood estimates of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\lambda}$ (i.e., $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\lambda}}$). As a result, we get the estimators $\hat{\boldsymbol{\Gamma}}_i$ and $\hat{\mathbf{R}}_i$, and the estimator of $\boldsymbol{\Sigma}_i$ can be obtained by $\hat{\boldsymbol{\Sigma}}_i^{HSCF} = \hat{\boldsymbol{\Gamma}}_i \hat{\mathbf{R}}_i \hat{\boldsymbol{\Gamma}}_i$. We call this the hyperspherical coordinates factor estimating procedure *HSCF*.

## 2.2 The combined multiple likelihood (CML) estimating procedure

Suppose that $s$ dynamic covariance candidate models are available. For example, we could consider the $s = 3$ covariance modeling methods (i.e., *MCDF*, *MACF* and *HSCF*) introduced in the previous section. We denote the true covariance matrix for the $ith$ subject as $\boldsymbol{\Sigma}_i^{(k)}$ with $k = 1, \cdots, s$ and $i = 1, \ldots, n$. Ideally, if one of the $s$ dynamic covariance models is consistent with that of the true covariance structure, efficient and reliable estimator for $\boldsymbol{\beta}$ can be achieved. Unfortunately, ones seldom know the true dynamic covariance structures and misspecification often occurs, which will lead to considerable efficiency loss in parameter estimation. We employ the empirical likelihood to assemble advantages of all the candidate covariance models. The estimated covariance matrix for the $i$th subject is denoted as $\hat{\boldsymbol{\Sigma}}_i^{(k)}$ when the $k$th covariance modeling method is used, for $k = 1, \ldots, s$; $i = 1, \ldots, n$. The $i$th subject's information can be available for $s$ independent estimation functions expressed as a vectors of length $s \times p$,

$$\mathbf{g}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta}) = (g_1(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})^\top, \ldots, g_s(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})^\top)^\top,$$

for $i = 1, \ldots, n$. The component $g_k(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})$ is the $i$th subject's score function based on the $k$th candidate covariance modeling and can be termed as

$$g_k(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta}) = \mathbf{x}_i^\top \Delta_i \hat{\boldsymbol{\Sigma}}_i^{(k)-1}(\mathbf{y}_i - \boldsymbol{\mu}(\mathbf{x}_i \boldsymbol{\beta})), \quad k = 1, \ldots, s.$$

where $\Delta_i = \mathrm{diag}(\boldsymbol{\mu}^{(1)}(x_{ij}^\top \beta))$, where $\boldsymbol{\mu}^{(1)}(\cdot)$ is the first derivative of $\boldsymbol{\mu}(\cdot)$.

The estimate of $\boldsymbol{\beta}$ is obtained by maximizing the following empirical likelihood (Owen, 1988, 2001; Qin and Lawless, 1994):

$$L_{CML}(\boldsymbol{\beta}) = \prod_{i=1}^{n} p_i,$$

subject to the constrains $p_i \geq 0$, $\sum_{i=1}^n p_i = 1$ and

$$\sum_{i=1}^{n} p_i \mathbf{g}(\mathbf{y}_i, \mathbf{x}_i; \boldsymbol{\beta}) = \mathbf{0}.$$

likelihood estimating equations (*CML*). The potential advantage of this method is that it can get the information of estimating a $p$-dimensional parameter $\boldsymbol{\beta}$ by an efficient combination of the $s \times p$-dimensional zero-mean estimating functions $\mathbf{g}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})$, for $i = 1, \ldots, n$. Specifically, we consider the class of $p$-dimensional estimating functions

$$\Xi = \{\psi(\beta) | \psi(\beta) = \tau(\beta) \sum_{i=1}^{n} \mathbf{g}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})\},$$

where $\tau(\beta)$ is a $p \times s \times p$ matrix of real functions. Based on the estimating function theory (Godambe and Heyde, 1987), we call an estimating function $\psi^*(\beta) \in \Xi$ optimum in $\Xi$ if the estimator $\hat{\beta}$ from $\psi^*(\beta) = 0$ has minimum asymptotical variance. By this understanding, our empirical likelihood based estimator $\hat{\boldsymbol{\beta}}_{CML}$ is fully efficient because of Corollary 2 in Qin and Lawless (1994). That is, our proposed method combines the advantages of the $s$ estimating functions $g_1(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta}), \ldots, g_s(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})$ and thus efficient mean coefficient estimator could be yielded. readily derive the asymptotical normality property of the *CML* estimator (i.e., $\hat{\boldsymbol{\beta}}_{CML}$). The detailed estimation procedure and calculations can be found in the Supplementary Materials. Using similar arguments of Theorem 1 in Qin and Lawless (1994), we can readily derive the asymptotical normality property of the *CML* estimator (i.e., $\hat{\boldsymbol{\beta}}_{CML}$). Specifically, we give the following Theorem. Let $|| \cdot ||_F$ be the Frobenius norm.

**Theorem 1.** *Assume*

- *(i)* $\boldsymbol{\mu}$ *has continuous second order derivative;*
- *(ii) there exists a matrix* $\boldsymbol{\Sigma}_i^{(k)}$, *such that* $||\hat{\boldsymbol{\Sigma}}_i^{(k)} - \boldsymbol{\Sigma}_i^{(k)}||_F = o_p(1)$ *and* $\boldsymbol{\Sigma}_i^{(k)}$ *is positive definite;*
- *(iii) the largest eigenvalues of* $\boldsymbol{\Sigma}_i^{(k)}$ *and* $\Sigma_i = var(\mathbf{y}_i)$ *both can be bounded by some integrable functions, for* $k = 1, \cdots, s$.

*Define*

$$\bar{\mathbf{g}}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta}) = (\bar{g_1}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})^\top, \cdots, \bar{g_s}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta})^\top)^\top,$$

*for* $i = 1, \ldots, n$, *where*

$$\bar{g_k}(\mathbf{y}_i; \mathbf{x}_i, \boldsymbol{\beta}) = \mathbf{x}_i^\top \Delta_i \boldsymbol{\Sigma}_i^{(k)-1}(\mathbf{y}_i - \boldsymbol{\mu}(\mathbf{x}_i \boldsymbol{\beta})), \quad k = 1, \ldots, s.$$

*Assume* $E(\bar{\mathbf{g}}\bar{\mathbf{g}}^\top | \boldsymbol{\beta} = \boldsymbol{\beta}_0)$ *is positive definite. We have*

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{CML} - \boldsymbol{\beta}_0) \rightarrow_L N(\mathbf{0}, \mathbf{V}),$$

*where* $\mathbf{V} = (\mathbf{V}_{12}^\top \mathbf{V}_{22}^{-1} \mathbf{V}_{12})^{-1}$, *with*

$$\mathbf{V}_{12} = E\left(\frac{\partial \bar{\mathbf{g}}}{\partial \boldsymbol{\beta}} | \boldsymbol{\beta} = \boldsymbol{\beta}_0\right) \qquad \text{and} \qquad \mathbf{V}_{22} = E(\bar{\mathbf{g}}\bar{\mathbf{g}}^\top | \boldsymbol{\beta} = \boldsymbol{\beta}_0).$$

**Remark 1.** Let $s = 1$. If the covariance model is correctly specified, i.e, $\boldsymbol{\Sigma}_i^{(1)} = \boldsymbol{\Sigma}_i$, we can have that, our proposed estimator is fully efficient. Otherwise, the estimator is not efficient.

**Remark 2.** By Corollary 1 or the same arguments in page 318 in Qin and Lawless (1994), dropping $g_k(\mathbf{y}_i;\mathbf{x}_i,\boldsymbol{\beta})$ (i.e., one candidate covariance model), the asymptotic variance of $\sqrt{n}(\hat{\boldsymbol{\beta}}_{CML}-\boldsymbol{\beta}_0)$ cannot decrease. Therefore, if one of the $s$ candidate covariance models is correctly specified, i.e, there exists $k$ such that, $\boldsymbol{\Sigma}_i^{(k)}=\boldsymbol{\Sigma}_i$, $\hat{\boldsymbol{\beta}}_{CML}$ can be fully efficient.

**Remark 3.** We use the candidate covariance models: *MCDF*, *MACF* and *HSCF* and assume the link function $\mu(\cdot)$ is an identity function. $\mathbf{g}(\mathbf{y}_i;\mathbf{x}_i,\boldsymbol{\beta})$ can be written as

$$\mathbf{g}(\mathbf{y}_i,\mathbf{x}_i;\boldsymbol{\beta})=\begin{pmatrix}\mathbf{x}_i^\top\hat{\boldsymbol{\Sigma}}_i^{MCDF-1}(\mathbf{y}_i-\mathbf{x}_i\boldsymbol{\beta})\\ \mathbf{x}_i^\top\hat{\boldsymbol{\Sigma}}_i^{MACF-1}(\mathbf{y}_i-\mathbf{x}_i\boldsymbol{\beta})\\ \mathbf{x}_i^\top\hat{\boldsymbol{\Sigma}}_i^{HSCF-1}(\mathbf{y}_i-\mathbf{x}_i\boldsymbol{\beta})\end{pmatrix},\quad i=1,\ldots,n.$$

The asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}_{CML}$ can be estimated by

$$\frac{1}{n}(\tilde{\mathbf{V}}_{12}^\top\tilde{\mathbf{V}}_{22}^{-1}\tilde{\mathbf{V}}_{12})^{-1}, \tag{4}$$

where

$$\tilde{\mathbf{V}}_{12}=\begin{pmatrix}-\frac{1}{n}\sum_{i=1}^n\mathbf{x}_i^\top\hat{\boldsymbol{\Sigma}}_i^{MCDF-1}\mathbf{x}_i\\ -\frac{1}{n}\sum_{i=1}^n\mathbf{x}_i^\top\hat{\boldsymbol{\Sigma}}_i^{MACF-1}\mathbf{x}_i\\ -\frac{1}{n}\sum_{i=1}^n\mathbf{x}_i^\top\hat{\boldsymbol{\Sigma}}_i^{HSCF-1}\mathbf{x}_i\end{pmatrix}^\top$$

and $\tilde{\mathbf{V}}_{22}=\frac{1}{n}\sum_{i=1}^n\mathbf{g}(\mathbf{y}_i,\mathbf{x}_i;\hat{\boldsymbol{\beta}}_{CML})\mathbf{g}^\top(\mathbf{y}_i,\mathbf{x}_i;\hat{\boldsymbol{\beta}}_{CML})$.

# 3. NUMERICAL STUDIES

## 3.1 Real data analysis

Spinal anesthesia is a common clinic anesthetic technique used in surgery and may cause hypotension during operation. Diastolic blood pressure (*DBP*) is an important index in measuring a patient's physical status during a surgical operation. Therefore, investigating the relationship between DBP and particular risk factors (e.g., age, gender, and anesthesia drug doses) is valuable to anesthesiologists Sharma et al. (1997); Lin et al. (2008); Samur et al. (2014). In the following analysis, all subjects that were admitted to the Akdeniz University Hospital Anesthesiology and Reanimation Department during the period of January 2008 to January 2011 were evaluated retrospectively. There are 375 patients (with 210 males and 165 females) and the diastolic blood pressures (DBP) were observed 9 times for each individual, which were measured every 5 minutes during the surgery.

The outcome variable of interest is DBP, and the explanatory variables include age, gender, pulse and the doses of marcain-heavy, midazolam, chirocaine and fentanyl. There is no missing observations in either outcome or covariates. We compared *MCDF*, *MACF*, *HSCF* and *CML* in terms of estimated standard errors (SEs) and mean squared regression errors (MSRE). Let *MCDF*, *MACF*, *HSCF* share the same covariates for covariance modeling, i.e., $\mathbf{h}_{ij}=(1,pulse_{ij}/100)$ and $\mathbf{z}_{ijk}=(1,(t_{ij}-t_{ik}),(t_{ij}-t_{ik})^2)$, where $pulse_{ij}$ and $t_{ij}$ are the heart pulse value and observed time point for the $j-$th observation of subject $i$, respectively. The *CML* estimator is conducted based on *MCDF*, *MACF* and *HSCF* covariance models (for details, see Remark 3). We obain the *MCDF*, *MACF*, *HSCF* and *CML* regression coefficient estimates and the estimated standard errors (SEs), respectively. We calculate the mean squared regression error as ($MSRE=\frac{1}{\sum_{i=1}^{375}m_i}\sum_{i=1}^{375}\sum_{j=1}^{m_i}(y_{ij}-x_{ij}\hat{\boldsymbol{\beta}})^2$).

The results of the $GEE$ and $QIF$ (Qu et al., 2000) are also shown in Table 1 for comparison. The $GEE$ method is carried out with $AR(1)$ and *exchangeable* correlation structures, respectively. For the $QIF$ method, following Qu et al. (2000) and Leung et al. (2009), we combined the three basis matrices: $\mathbf{M}_0$ is the identity matrix, $\mathbf{M}_1$ is a matrix with 0 on the diagonal and 1 off the diagonal, and $\mathbf{M}_2$ is a matrix with 1 on the two main off-diagonals and 0 elsewhere.

The main difference among *MCDF*, *MACF* and *HSCF* lies in the covariance structure modelings. Table 1 shows that the combining multiple likelihood estimation (*CML*) has the generally lowest SE and MSRE values among all the approaches. This may indicate that the *CML* method combines the advantages of all the candidate covariance models and is the most efficient estimating procedure. We conclude *CML* estimation is the most preferred method for analyzing this dataset. Program codes prepared in R have been developed to implement the methodologies developed in this article and are available from the first author upon request.

## 3.2 Simulation studies

In this subsection, we carry out some simulation studies to investigate the finite sample performance of the *CML* estimation. The main purposes of the simulation studies are to evaluate:

($i$) the efficiency of $GEE$, $QIF$, *MCDF*,*MACF*, *HSCF* and *CML* under different scenarios; and

($ii$) the flexibility of *CML* estimation for the non-normal data generated from a mixture of same-mean and different-covariance multivariate normal distributions.

For all simulation studies, the datasets are generated from the model:

$$y_{ij}=\mathbf{x}_{ij}^\top\boldsymbol{\beta}+e_{ij},\quad j=1,\ldots,m_i;\quad i=1,\ldots,n, \tag{5}$$

where $\mathbf{x}_i=(\mathbf{x}_{i1},\ldots,\mathbf{x}_{im_i})$ and the zero-mean error $\mathbf{e}_i=(e_{i1},\ldots,e_{im_i})^\top$ is of a specified covariance structure (e.g., *MCDF*, *MACF* or *HSCF*). The mean regression parameters are $\boldsymbol{\beta}=(1,-0.5,0.5)$ and the corresponding covariate is $\mathbf{x}_{ij}=(1,x_{ij1},x_{ij2})^\top$, where $(x_{ij1},x_{ij2})^\top$ is generated from a multivariate normal distribution with mean zero, marginal

Table 1. The spinal anesthesia data analysis. The estimates of regression coefficients, the corresponding standard errors (SEs) and the mean squared regression error (MSRE).

| Covariate | $MCDF$ | | $MACF$ | | $HSCF$ | | $CML$ | | $GEE.AR(1)$ | | $GEE.EXch$ | | $QIF$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{\boldsymbol{\beta}}$ | $SE$ | $\hat{\boldsymbol{\beta}}$ | $SE$ | $\hat{\boldsymbol{\beta}}$ | $SE$ | $\hat{\boldsymbol{\beta}}$ | $SE$ | $\hat{\boldsymbol{\beta}}$ | $SE$ | $\hat{\boldsymbol{\beta}}$ | $SE$ | $\hat{\boldsymbol{\beta}}$ | $SE$ |
| **Intercept** | 39.65 | 2.23 | 40.18 | 2.23 | 42.85 | 7.97 | 42.76 | 2.31 | 36.95 | 3.06 | 27.16 | 3.07 | 42.75 | 3.89 |
| **Age** | 0.10 | 0.02 | 0.10 | 0.02 | 0.16 | 0.07 | 0.08 | 0.02 | 0.09 | 0.04 | 0.09 | 0.04 | 0.08 | 0.04 |
| **Gender** | 6.30 | 0.89 | 6.28 | 0.89 | 7.33 | 2.90 | 7.29 | 0.91 | 7.24 | 1.41 | 8.33 | 1.54 | 7.24 | 1.47 |
| **Pulse** | 0.34 | 0.02 | 0.33 | 0.02 | 0.29 | 0.07 | 0.25 | 0.02 | 0.33 | 0.02 | 0.41 | 0.02 | 0.25 | 0.03 |
| **Marcain-heavy** | -0.07 | 0.08 | -0.09 | 0.08 | 0.01 | 0.27 | -0.11 | 0.07 | -0.10 | 0.13 | -0.08 | 0.14 | -0.09 | 0.14 |
| **Midazolam** | -1.08 | 0.33 | -1.06 | 0.33 | -0.85 | 1.08 | -0.98 | 0.32 | -0.90 | 0.52 | -0.93 | 0.57 | -0.94 | 0.52 |
| **Chirocaine** | 0.04 | 0.06 | 0.05 | 0.06 | 0.16 | 0.21 | 0.07 | 0.05 | 0.06 | 0.10 | 0.08 | 0.11 | 0.06 | 0.09 |
| **Fentanyl** | -25.42 | 6.62 | -23.85 | 6.62 | 26.10 | 21.14 | -22.32 | 6.04 | -25.78 | 9.91 | -29.61 | 10.84 | -26.19 | 9.92 |
| **MSRE** | 162.17 | | 160.55 | | 228.91 | | 144.74 | | 147.35 | | 149.75 | | 152.83 | |

variance 1 and two dimensional exchange correlation structure with $\rho = 0.5$. Motivated by the anesthesia data analysis in the previous section, we take quadratic polynomials of time difference $\mathbf{z}_{ijk} = (1, (t_{ij} - t_{ik}), (t_{ij} - t_{ik})^2)^\top$ as the covariate and set $\boldsymbol{\gamma} = (0.3, -0.2, 0.3)^\top$. The covariate for log innovation structure is taken as $\mathbf{h}_{ij} = (1, h_{ij1}, h_{ij2}, h_{ij3})^\top$ with $(h_{ij1}, h_{ij2}, h_{ij3})^\top$ being generated from a multivariate normal distribution with mean zero, marginal variance 1 and three dimensional exchange correlation structure with $\rho = 0.5$. The true log innovation parameters $\boldsymbol{\lambda} = (-0.5, 0.5, -0.3, 0.1)^\top$. Each subject is measured $m_i$ times with $m_i \sim Binomial(10, 0.8)$ and the measurement time $t_{ij}$'s are independently generated from $U(0, 1)$. Hence, the repeated measurements for each subject are observed at irregular and unbalanced time points. For each setting, we generate 1000 datasets.

### Study 1

The purpose of the first study is to compare the performances of CML with those of MCDF when the working log innovation variance structure is either correct or less specified. Here, three candidate $MCDF$ covariance models all employ the same auto-regression model $\phi_{ijk} = \mathbf{z}_{ijk}^\top \boldsymbol{\gamma}$, but model $\ln(\sigma_{ij}^2)$ as $\lambda_0 + \sum_{k=1}^{J} h_{ijk}\lambda_k, J = 1, 2, 3$, respectively.

With mean $\mathbf{x}_i^T \boldsymbol{\beta}$ and the covariance structure being one of the three aforementioned candidate models, the maximum likelihood estimates of the mean regression parameter $\boldsymbol{\beta}$ are obtained based on multivariate normal distribution, respectively. The $CML$ estimator of $\boldsymbol{\beta}$ is conducted based on the three candidate $MCDF$ models. We also report the $GEE$ and $QIF$ estimators in each simulation study for comparison. We consider $n = 100, 200, 400$ and summarized the simulation results in the Table 2.

For each method, we calculate the mean of the 1,000 absolute bias $(|\hat{\beta}_k - \beta_k|, k = 0, 1, 2)$ $(MAB)$ and the bias norm of the mean $(\|\hat{\boldsymbol{\mu}}_d\| = \frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x}_i^\top(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|)$. We calculate the sample standard deviation $(SD)$ of 1,000 parameter estimates and the sample average of 1,000 estimated standard errors $(SE)$. The 95% estimated confidence interval for $\beta_k$ is calculated by $\hat{\beta}_k \pm 1.96SE(\beta_k), k = 0, 1, 2$. The coverage probability is also reported.

### Study 2

We note that the statistical interpretations for $\phi_{ijk}$ and $\ln(\sigma_{ij}^2)$ are different for the three dynamic covariance models (e.g., *MCDF*, *MACF* and *HSCF*). However, we use the same models $\phi_{ijk} = \mathbf{z}_{ijk}^\top \boldsymbol{\gamma}$ and $\ln(\sigma_{ij}^2) = \mathbf{h}_{ij}^\top \boldsymbol{\lambda}$ for different covariance structure models in this study. Using Equation (5) with $\phi_{ijk}$ and $\ln(\sigma_{ij}^2)$ in the covariance structure, the datasets are generated from the *MCDF* covariance model, *MACF* covariance model, *HSCF* covariance model, respectively. With the covariance structure being the *MCDF* model, *MACF* model and *HSCF* model respectively, we obtain the maximum likelihood estimates of $\boldsymbol{\beta}$ based on the multivariate normal distribution. The *CML* estimates of $\boldsymbol{\beta}$ are obtained based on the three aforementioned covariance models. The sample size is set to be $n = 200$ and the performances of different methods are summarized in the Table 3.

### Study 3

In this study, we demonstrate the flexibility of *CML* estimation for the non-normal data. The datasets are generated from a mixture of three same mean multivariate normal distributions:

$$\pi_1 N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i^{(MCDF)}) + \pi_2 N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i^{(MACF)}) + \pi_3 N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i^{(HSCF)})$$

based on different ratios $\pi_1 : \pi_2 : \pi_3$ with $\sum_{i=1}^{3} \pi_i = 1$, where $\boldsymbol{\mu}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{im_i})^T \boldsymbol{\beta}$. Here, $\boldsymbol{\Sigma}_i^{(MCDF)}$, $\boldsymbol{\Sigma}_i^{(MACF)}$ and $\boldsymbol{\Sigma}_i^{(MCDF)}$ denote the covariance matrices for subject $i$ when the *MCDF* covariance model, *MACF* covariance model

*Table 2. Study 1. The $MABs$, $SEs$, $SDs$ and $\|\boldsymbol{\mu}_d\|$ based on 1000 parameter estimates. (all the results are multiplied by a factor $10^2$)*

| | | Sample Size | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $(n=100))$ | | | | $(n=200)$ | | | | $(n=400)$ | | | |
| Working Model | | $MAB$ | $SE$ | $SD$ | $CP$ | $MAB$ | $SE$ | $SD$ | $CP$ | $MAB$ | $SE$ | $SD$ | $CP$ |
| | $\beta_0$ | 2.90 | 3.28 | 3.65 | 91.9% | 2.09 | 2.31 | 2.63 | 91.5% | 1.53 | 1.64 | 1.90 | 90.9% |
| $\ln(\sigma_{ij}^2)=\sum_{k=0}^{1} h_{ijk}\lambda_k$ | $\beta_1$ | 2.15 | 2.53 | 2.66 | 93.5% | 1.57 | 1.79 | 1.97 | 93.1% | 1.06 | 1.27 | 1.33 | 94.3% |
| (less-specified model | $\beta_2$ | 1.97 | 2.53 | 2.50 | 95.3% | 1.42 | 1.79 | 1.76 | 96.4% | 1.03 | 1.27 | 1.31 | 94.0% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.234 | | | | 0.121 | | | | 0.062 | | | |
| | $\beta_0$ | 2.85 | 3.20 | 3.57 | 91.7% | 2.04 | 2.26 | 2.57 | 91.6% | 1.50 | 1.60 | 1.87 | 90.4% |
| $\ln(\sigma_{ij}^2)=\sum_{k=0}^{2} h_{ijk}\lambda_k$ | $\beta_1$ | 2.10 | 2.47 | 2.61 | 93.8% | 1.54 | 1.75 | 1.92 | 92.8% | 1.04 | 1.24 | 1.30 | 93.9% |
| (less-specified model) | $\beta_2$ | 1.94 | 2.47 | 2.47 | 94.1% | 1.38 | 1.75 | 1.72 | 96.1% | 1.02 | 1.24 | 1.29 | 93.8% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.224 | | | | 0.115 | | | | 0.060 | | | |
| | $\beta_0$ | 2.85 | 3.19 | 3.56 | 92.2% | 2.03 | 2.25 | 2.56 | 91.9% | 1.49 | 1.59 | 1.86 | 90.3% |
| $\ln(\sigma_{ij}^2)=\sum_{k=0}^{3} h_{ijk}\lambda_k$ | $\beta_1$ | 2.10 | 2.46 | 2.62 | 93.9% | 1.54 | 1.74 | 1.92 | 92.8% | 1.04 | 1.23 | 1.30 | 93.7% |
| (correct model) | $\beta_2$ | 1.95 | 2.46 | 2.47 | 94.1% | 1.39 | 1.74 | 1.72 | 96.0% | 1.02 | 1.23 | 1.29 | 93.5% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.223 | | | | 0.114 | | | | 0.059 | | | |
| | $\beta_0$ | 2.88 | 3.34 | 3.59 | 92.0% | 2.03 | 2.43 | 2.56 | 93.5% | 1.49 | 1.75 | 1.86 | 93.4% |
| $CML$ | $\beta_1$ | 2.11 | 2.42 | 2.63 | 93.3% | 1.54 | 1.75 | 1.92 | 93.2% | 1.04 | 1.25 | 1.30 | 94.7% |
| | $\beta_2$ | 1.95 | 2.40 | 2.49 | 92.8% | 1.38 | 1.75 | 1.71 | 96.3% | 1.02 | 1.25 | 1.29 | 94.6% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.227 | | | | 0.114 | | | | 0.059 | | | |
| | $\beta_0$ | 8.43 | 4.35 | 10.53 | 58.4% | 5.77 | 3.10 | 7.33 | 59.3% | 4.01 | 2.19 | 5.05 | 60.2% |
| $GEE.Ind$ | $\beta_1$ | 4.12 | 5.03 | 5.13 | 93.5% | 2.85 | 3.58 | 3.58 | 95.4% | 2.02 | 2.53 | 2.54 | 94.4% |
| | $\beta_2$ | 3.99 | 5.03 | 5.03 | 94.8% | 2.83 | 3.58 | 3.54 | 95.7% | 2.06 | 2.53 | 2.57 | 94.5% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 1.486 | | | | 0.727 | | | | 0.352 | | | |
| | $\beta_0$ | 8.85 | 7.78 | 11.05 | 82.4% | 6.08 | 5.55 | 7.73 | 84.8% | 4.25 | 3.94 | 5.34 | 85.1% |
| $GEE.AR(1)$ | $\beta_1$ | 2.65 | 3.46 | 3.32 | 96.5% | 1.92 | 2.45 | 2.40 | 95.1% | 1.33 | 1.73 | 1.66 | 96.6% |
| | $\beta_2$ | 2.59 | 3.46 | 3.24 | 97.2% | 1.82 | 2.45 | 2.27 | 96.5% | 1.31 | 1.73 | 1.66 | 95.5% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 1.383 | | | | 0.681 | | | | 0.327 | | | |
| | $\beta_0$ | 8.02 | 10.24 | 10.01 | 95.1% | 5.49 | 7.31 | 6.96 | 95.0% | 3.83 | 5.18 | 4.83 | 96.2% |
| $GEE.Exch$ | $\beta_1$ | 2.96 | 3.68 | 3.71 | 95.0% | 2.17 | 2.61 | 2.73 | 93.6% | 1.50 | 1.84 | 1.88 | 94.2% |
| | $\beta_2$ | 2.88 | 3.68 | 3.65 | 95.3% | 2.03 | 2.61 | 2.54 | 94.7% | 1.48 | 1.84 | 1.84 | 95.0% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 1.209 | | | | 0.590 | | | | 0.285 | | | |
| | $\beta_0$ | 6.90 | 8.81 | 8.78 | 93.0% | 4.78 | 5.91 | 5.99 | 93.8% | 3.38 | 4.23 | 4.21 | 95.6% |
| $QIF$ | $\beta_1$ | 2.80 | 3.22 | 3.51 | 92.8% | 1.99 | 2.35 | 2.50 | 93.9% | 1.40 | 1.69 | 1.73 | 95.4% |
| | $\beta_2$ | 2.69 | 3.22 | 3.38 | 93.1% | 1.88 | 2.35 | 2.36 | 94.9% | 1.38 | 1.69 | 1.73 | 93.8% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.953 | | | | 0.447 | | | | 0.222 | | | |

and $HSCF$ covariance model are employed, respectively. The estimates of $\boldsymbol{\beta}$ using the $CML$ approach are obtained based on the three candidate covariance models. The sample size is set to be n = 200 and the Table 4 lists the corresponding simulation results.

We have the following observations from Tables 2–4:

1). Covariance structure has remarkable impact on the estimation efficiency of the mean regression parameters. In general, the closer the candidate working covariance structure to the correct working covariance structure, the smaller the $MAB$, $SD$ or $\|\boldsymbol{\mu}_d\|$ is.

2). As expected, the most efficient estimators can be obtained when the correct working covariance model is specified. The regressive efficiency of the estimated $\boldsymbol{\beta}$ by $CML$ estimation is better than the $MLE$ based on incorrect covariance models. However, our proposed $CML$ estimation is only slightly inferior to the $MLE$ based on correct working covariance structures.

3). If the datasets are generated from a mixture of different dynamic covariance matrices models, $CML$ estimation is more efficient and flexible than the other regression methods since the $MABs$, $SDs$ and $\|\boldsymbol{\mu}_d s\|$ for the $CML$ estimation

*Table 3. Study 2. The $MAB$s, $SE$s, $SD$s and $\|\boldsymbol{\mu}_d\|$ based on 1000 parameter estimates (all the results are multiplied by a factor $10^2$)*

| | | True Model | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $MCDF$ | | | | $MACF$ | | | | $HSCF$ | | | |
| Working Model | | $MAB$ | $SE$ | $SD$ | $CP$ | $MAB$ | $SE$ | $SD$ | $CP$ | $MAB$ | $SE$ | $SD$ | $CP$ |
| $MCDF$ | $\beta_0$ | 2.03 | 2.25 | 2.56 | 91.9% | 2.98 | 3.24 | 3.84 | 89.8% | 4.14 | 1.78 | 5.13 | 90.4% |
| | $\beta_1$ | 1.54 | 1.74 | 1.92 | 92.8% | 1.51 | 1.85 | 1.89 | 94.2% | 0.70 | 0.99 | 0.87 | 97.1% |
| | $\beta_2$ | 1.39 | 1.74 | 1.72 | 96.0% | 1.53 | 1.85 | 1.93 | 94.0% | 0.70 | 0.99 | 0.88 | 97.3% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.114 | | | | 0.201 | | | | 0.275 | | |
| $MACF$ | $\beta_0$ | 2.61 | 2.98 | 3.29 | 92.5% | 2.80 | 3.65 | 3.62 | 95.1% | 4.08 | 2.16 | 5.06 | 58.4% |
| | $\beta_1$ | 1.64 | 1.82 | 2.03 | 93.7% | 1.47 | 1.80 | 1.85 | 94.3% | 0.52 | 0.71 | 0.65 | 96.7% |
| | $\beta_2$ | 1.48 | 1.82 | 1.84 | 96.5% | 1.50 | 1.80 | 1.88 | 94.0% | 0.55 | 0.71 | 0.69 | 95.6% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.165 | | | | 0.182 | | | | 0.263 | | |
| $HSCF$ | $\beta_0$ | 5.29 | 5.68 | 6.65 | 91.4% | 3.12 | 3.57 | 3.95 | 92.2% | 0.06 | 0.04 | 0.08 | 63.7% |
| | $\beta_1$ | 1.86 | 2.25 | 2.33 | 93.7% | 1.52 | 1.90 | 1.89 | 94.7% | 0.02 | 0.01 | 0.02 | 64.1% |
| | $\beta_2$ | 1.74 | 2.25 | 2.15 | 96.5% | 1.55 | 1.90 | 1.96 | 94.3% | 0.02 | 0.01 | 0.02 | 64.8% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.518 | | | | 0.210 | | | | 0.001 | | |
| $CML$ | $\beta_0$ | 2.02 | 2.33 | 2.52 | 92.6% | 2.86 | 3.52 | 3.69 | 94.3% | 0.06 | 0.06 | 0.08 | 85.3% |
| | $\beta_1$ | 1.56 | 1.74 | 1.94 | 91.5% | 1.52 | 1.76 | 1.92 | 92.6% | 0.02 | 0.02 | 0.02 | 85.2% |
| | $\beta_2$ | 1.40 | 1.74 | 1.75 | 95.1% | 1.53 | 1.76 | 1.93 | 93.0% | 0.02 | 0.02 | 0.02 | 87.2% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.115 | | | | 0.191 | | | | 0.001 | | |
| $GEE.Ind$ | $\beta_0$ | 5.77 | 3.10 | 7.33 | 59.3% | 3.47 | 2.17 | 4.44 | 66.1% | 4.63 | 1.86 | 5.79 | 45.9% |
| | $\beta_1$ | 2.85 | 3.58 | 3.58 | 95.4% | 1.94 | 2.50 | 2.45 | 95.3% | 1.63 | 2.15 | 2.05 | 95.9% |
| | $\beta_2$ | 2.83 | 3.58 | 3.54 | 95.7% | 2.04 | 2.50 | 2.58 | 94.6% | 1.76 | 2.15 | 2.19 | 94.3% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.727 | | | | 0.292 | | | | 0.401 | | |
| $GEE.AR(1)$ | $\beta_0$ | 6.08 | 5.88 | 7.73 | 84.8% | 3.47 | 3.22 | 4.41 | 84.7% | 4.69 | 5.17 | 5.79 | 93.0% |
| | $\beta_1$ | 1.92 | 2.45 | 2.40 | 95.1% | 1.70 | 2.09 | 2.14 | 94.8% | 0.50 | 0.55 | 0.63 | 92.2% |
| | $\beta_2$ | 1.82 | 2.45 | 2.27 | 96.5% | 1.76 | 2.09 | 2.22 | 93.7% | 0.53 | 0.55 | 0.64 | 91.5% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.681 | | | | 0.265 | | | | 0.342 | | |
| $GEE.Exch$ | $\beta_0$ | 5.49 | 7.73 | 6.96 | 95.0% | 3.42 | 4.45 | 4.36 | 95.5% | 4.63 | 5.65 | 5.76 | 95.2% |
| | $\beta_1$ | 2.17 | 2.61 | 2.73 | 93.6% | 1.66 | 2.08 | 2.09 | 95.1% | 0.43 | 0.57 | 0.54 | 94.4% |
| | $\beta_2$ | 2.03 | 2.61 | 2.54 | 94.7% | 1.67 | 2.08 | 2.13 | 93.9% | 0.45 | 0.57 | 0.56 | 95.4% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.590 | | | | 0.256 | | | | 0.337 | | |
| $QIF$ | $\beta_0$ | 4.78 | 5.90 | 5.99 | 93.8% | 3.50 | 4.30 | 4.46 | 93.6% | 4.76 | 5.44 | 5.93 | 92.5% |
| | $\beta_1$ | 1.99 | 2.35 | 2.50 | 93.9% | 1.65 | 2.01 | 2.10 | 93.4% | 0.46 | 0.59 | 0.57 | 94.4% |
| | $\beta_2$ | 1.88 | 2.35 | 2.36 | 94.9% | 1.69 | 2.01 | 2.13 | 93.6% | 0.50 | 0.59 | 0.61 | 95.4% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | | 0.447 | | | | 0.266 | | | | 0.356 | | |

are generally the smallest. This is due to the fact that our $CML$ estimation combines the advantages of all the candidate covariance models.

4). The dynamic covariance modelings (e.g., $MCDF$, $MACF$, $HSCF$ and $CML$) are more efficient and flexible than $GEE$ or $QIF$ when the within-subject covariation pattern is dynamic.

## 4. DISCUSSION

The $MCDF$, $MACF$ and $HSCF$ are the commonly used dynamic covariance models for longitudinal data. The advantages of the three covariance models are combined for estimating the mean regression parameters of longitudinal data by our proposed method. The proposed method generally yields efficient estimators.

We give several directions for future study. First, we can take into account the advantages of the nonparametric and semiparametric covariance modeling (Fan and Wu, 2008; Yin et al., 2010; Chen and Leng, 2016). Second, we are also interested in estimating parameters efficiently for binary longitudinal data by combining different dynamical log odds-ratio models (Carey et al., 1993; Fitzmaurice and Laird, 1993; Chen et al., 2011a). Third, it is necessary to pursue the variable selection problem and new method for missing data under our proposed framework. Finally, for the given candidate static correlation structures, the empirical likelihood based estimator and the QIF estimator (Qu et al., 2000) have the same asymptotic normal distribution and both can achieve an optimal linear combination of the given

*Table 4. Study 3. The $MABs$, $SEs$, $SDs$ and $\|\boldsymbol{\mu}_d\|$ based on 1000 parameter estimates (all the results are multiplied by a factor $10^2$)*

| | | Mixture Distribution: $\pi_1 : \pi_2 : \pi_3$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 50%:25%:25% | | | | 25%:50%:25% | | | | 25%:25%:50% | | | |
| Working Model | | $MAB$ | $SE$ | $SD$ | $CP$ | $MAB$ | $SE$ | $SD$ | $CP$ | $MAB$ | $SE$ | $SD$ | $CP$ |
| $MCDF$ | $\beta_0$ | 3.37 | 3.00 | 4.29 | 82.7% | 3.25 | 3.11 | 4.13 | 85.3% | 3.66 | 2.90 | 4.61 | 78.9% |
| | $\beta_1$ | 1.44 | 1.74 | 1.80 | 93.9% | 1.40 | 1.71 | 1.76 | 94.9% | 1.28 | 1.58 | 1.61 | 95.0% |
| | $\beta_2$ | 1.48 | 1.74 | 1.84 | 93.7% | 1.39 | 1.71 | 1.74 | 94.2% | 1.29 | 1.58 | 1.63 | 93.8% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.234 | | | | 0.216 | | | | 0.252 | | | |
| $MACF$ | $\beta_0$ | 3.77 | 4.32 | 4.70 | 92.1% | 3.60 | 4.41 | 4.53 | 94.4% | 3.55 | 4.21 | 4.42 | 94.1% |
| | $\beta_1$ | 1.40 | 1.66 | 1.76 | 94.1% | 1.38 | 1.63 | 1.79 | 94.2% | 1.22 | 1.44 | 1.53 | 93.6% |
| | $\beta_2$ | 1.50 | 1.66 | 1.87 | 92.5% | 1.43 | 1.63 | 1.79 | 91.7% | 1.25 | 1.44 | 1.58 | 93.2% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.271 | | | | 0.252 | | | | 0.232 | | | |
| $HSCF$ | $\beta_0$ | 4.35 | 4.55 | 5.47 | 90.2% | 3.74 | 3.92 | 4.70 | 90.4% | 3.67 | 3.63 | 4.60 | 88.2% |
| | $\beta_1$ | 1.55 | 1.94 | 1.93 | 95.7% | 1.44 | 1.82 | 1.81 | 95.0% | 1.24 | 1.55 | 1.57 | 94.4% |
| | $\beta_2$ | 1.55 | 1.94 | 1.94 | 95.2% | 1.45 | 1.82 | 1.81 | 92.9% | 1.25 | 1.55 | 1.59 | 94.7% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.356 | | | | 0.270 | | | | 0.249 | | | |
| $CML$ | $\beta_0$ | 3.24 | 3.60 | 4.21 | 92.1% | 3.15 | 3.64 | 4.00 | 92.9% | 3.40 | 3.73 | 4.30 | 91.1% |
| | $\beta_1$ | 1.37 | 1.62 | 1.72 | 93.0% | 1.34 | 1.58 | 1.69 | 92.4% | 1.21 | 1.38 | 1.51 | 93.4% |
| | $\beta_2$ | 1.46 | 1.62 | 1.80 | 91.8% | 1.36 | 1.58 | 1.71 | 92.5% | 1.21 | 1.38 | 1.52 | 91.5% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.225 | | | | 0.203 | | | | 0.220 | | | |
| $GEE.Ind$ | $\beta_0$ | 5.25 | 2.80 | 6.74 | 60.2% | 4.73 | 2.54 | 6.05 | 61.0% | 5.12 | 2.46 | 6.43 | 55.3% |
| | $\beta_1$ | 2.56 | 3.23 | 3.20 | 94.4% | 2.33 | 2.93 | 2.92 | 94.7% | 2.27 | 2.85 | 2.83 | 95.1% |
| | $\beta_2$ | 2.58 | 3.23 | 3.25 | 94.0% | 2.29 | 2.93 | 2.89 | 95.8% | 2.30 | 2.85 | 2.90 | 94.7% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.613 | | | | 0.494 | | | | 0.537 | | | |
| $GEE.AR(1)$ | $\beta_0$ | 5.41 | 5.01 | 6.94 | 83.9% | 4.80 | 4.36 | 6.13 | 84.1% | 5.25 | 4.63 | 6.57 | 83.5% |
| | $\beta_1$ | 1.75 | 2.22 | 2.16 | 96.3% | 1.70 | 2.11 | 2.14 | 94.4% | 1.51 | 1.83 | 1.89 | 94.5% |
| | $\beta_2$ | 1.76 | 2.22 | 2.22 | 95.3% | 1.70 | 2.11 | 2.16 | 94.0% | 1.54 | 1.83 | 1.93 | 93.9% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.556 | | | | 0.446 | | | | 0.488 | | | |
| $GEE.Exch$ | $\beta_0$ | 5.05 | 6.70 | 6.45 | 95.6% | 4.57 | 5.94 | 5.85 | 94.9% | 4.97 | 6.18 | 6.22 | 95.1% |
| | $\beta_1$ | 1.88 | 2.31 | 2.36 | 94.4% | 1.74 | 2.16 | 2.15 | 96.0% | 1.56 | 1.89 | 1.93 | 94.6% |
| | $\beta_2$ | 1.87 | 2.31 | 2.36 | 94.4% | 1.68 | 2.16 | 2.13 | 94.5% | 1.54 | 1.89 | 1.94 | 94.2% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.502 | | | | 0.412 | | | | 0.444 | | | |
| $QIF$ | $\beta_0$ | 4.85 | 5.94 | 6.17 | 93.5% | 4.53 | 5.52 | 5.78 | 93.5% | 4.89 | 5.75 | 6.08 | 93.1% |
| | $\beta_1$ | 1.79 | 2.12 | 2.23 | 95.0% | 1.67 | 2.02 | 2.10 | 94.4% | 1.47 | 1.76 | 1.84 | 93.6% |
| | $\beta_2$ | 1.77 | 2.13 | 2.23 | 93.1% | 1.66 | 2.02 | 2.09 | 93.5% | 1.47 | 1.76 | 1.85 | 92.6% |
| | $\|\hat{\boldsymbol{\mu}}_d\|$ | 0.457 | | | | 0.401 | | | | 0.422 | | | |

estimating functions (Li and Pan, 2013). It is of interest to investigate the properties of QIF method in the framework of dynamic covariance structures.


## ACKNOWLEDGEMENT

We thank the two reviewers and the Joint Editor for their constructive comments. The research of Dr. Chen was supported by NSFC grant 11871477 and Hunan Provincial Natural Science Foundation of China 2016JJ3138. Dr. Xu's research was supported by Zhejiang Provincial NSF of China grant LY19A010006, Zhejiang Federation of Humanities and Social Sciences Circles grant 16NDJC154YB and First Class Discipline of Zhejiang–A (Zhejiang University of Finance and Economics- Statistics). Dr. Tang's research was supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region (Project No. UGC/FDS14/P01/16), a grant from National Natural Science Foundation of China (Grant No. 11871124), and the Big Data and Artificial Intelligence Group of School of Decision Sciences.

*Received 17 March 2018*


## REFERENCES


Carey, V. and Wang, Y. G. (2011). Working covariance model selection for generalized estimating equations. *Statistics in Medicine* **30**, 3117–3124. MR2845681

Carey, V. J., Zeger, S. L., and Diggle, P. J. (1993). Modelling multivariate binary data with alternating logistic regressions. *Biometrika* **80**, 517–526.

Chen, Z. and Leng, C. (2016). Dynamic covariance models. *Journal of the American Statistical Association* **111**, 1196–1207. MR3561942

Chen, Z., Shi, N. Z., and Gao, W. (2011). Nonparametric estimation of the log odds ratio for sparse data by kernel smoothing. *Statistics and Probability Letters* **81**, 1802–1807. MR2845891

Chen, Z., Shi, N. Z., Gao, W., and Tang, M. L. (2011). Efficient semiparametric estimation via cholesky decomposition for longitudinal data. *Computational Statistics and Data Analysis* **55**, 3344–3354. MR2825416

Daniels, M. J. and Zhao, Y. D. (2003). Modeling 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. L. (2002). *Analysis of Longitudinal Data*. Oxford University Press, Oxford, UK. MR2049007

Fan, J. and Wu, Y. (2008). Semiparametric estimation of covariance matrices for longitudinal data. *Journal of the American Statistical Association* **105**, 1520–1533. MR2504201

Fitzmaurice, G. M. and Laird, N. M. (1993). A likelihood-based method for analysing longitudinal binary responses. *Biometrika* **80**, 141–151.

Godambe, V. P. and Heyde, C. C. (1987). Quasi-likelihood and optimal estimation. *International Statistical Review* **55**, 231–244. MR0963141

Kou, C. F. and Pan, J. X. (2010). Variable selection for joint mean and covariance models via penalized likelihood. Tech. rep., School of Mathematics, The University of Manchester.

Leng, C., Zhang, W., and Pan, J. (2010). Semiparametric mean-covariance regression analysis for longitudinal data. *Journal of the American Statistical Association* **105**, 181–193. MR2656048

Leung, D. H., Wang, Y. G., and Zhu, M. (2009). Efficient parameter estimation in longitudinal data analysis using a hybrid gee method. *Biostatistics* **10**, 436–445.

Li, D. and Pan, J. (2013). Empirical likelihood for generalized linear models with longitudinal data. *Journal of Multivariate Analysis* **114**, 63–73. MR2993874

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22. MR0836430

Lin, C. S., Chiu, J. S., Hsieh, M. H., Mok, M. S., Li, Y. C., and Chiu, H. W. (2008). Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks. *Computer Methods and Programs in Biomedicine* **92**, 193–197.

Lin, X. and Carroll, R. J. (2006). Semiparametric estimation in general repeated measures problems. *Journal of the Royal Statistical Society* **68**, 69–88. MR2212575

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–249. MR0946049

Owen, A. (2001). *Empirical Likelihood*. Chapman and Hall, Boca Raton, FL.

Pan, J. and Mackenzie, G. (2003). Model selection for joint mean-covariance struc-tures in longitudinal studies. *Biometrika* **90**, 239–244. MR1966564

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrika* **86**, 677–690. MR1723786

Pourahmadi, M. (2000). Maximum likelihood estimation of generalised linear models for multivariate normal covariance matrix. *Biometrika* **87**, 425–435. MR1782488

Pourahmadi, M. (2007). Cholesky decompositions and estimation of a covariance matrix: orthogonality of variance-correlation parameters. *Biometrika* **94**, 1006–1013. MR2376812

Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Annals of Statistics* **22**, 300–325. MR1272085

Qu, A., Lindsay, B. G., and Li, B. (2000). Improving estimating equations using quadratic inference functions. *Biometrika* **87**, 823–836. MR1813977

Samur, A. A., Coskunfirat, N., and Saka, O. (2014). Comparison of predictor approaches for longitudinal binary outcomes application to anesthesiology data. *PeerJ* **2**, DOI:10.7717/peerj.648.

Sharma, S. K., Gajraj, N. M., and Sidawi, J. E. (1997). Prevention of hypotension during spinal anesthesia: acomparison of intravascular administration of hetastarch versus lactated ringer's solution. *Anesthesia and Analgesia* **84**, 111–114.

Wang, Y. G. and Carey, V. (2003). Working correlation structure misspecification, estimation and covariate design: Implications for generalised estimating equations performance. *Biometrika* **90**, 29–41. MR1966548

Ye, H. and Pan, J. (2006). Modelling covariance structures in generalized estimating equations for longitudinal data. *Biometrika* **93**, 927–941. MR2285080

Yin, J., Geng, Z., Li, R., and Wang, H. (2010). Nonparametric covariance model. *Statistica Sinica* **20**, 469–479. MR2640671

Zhang, W. and Leng, C. (2012). A moving average cholesky factor model in covariance modeling for longitudinal data. *Biometrika* **99**, 141–150. MR2899669

Zhang, W., Leng, C., and Tang, C. Y. (2015). A joint modeling approach for longitudinal studies. *Journal of the Royal Statistical Society* **77**, 219–238. MR3299406

Lin Xu
School of Data Sciences
Zhejiang University of Finance and Economics
Hangzhou
P.R. China
E-mail address: xulin@zufe.edu.cn

Man-Lai Tang
Department of Mathematics and Statistics
Hang Seng University of Hong Kong
Hong Kong
E-mail address: mltang@hsu.edu.hk

Ziqi Chen
School of Mathematics and Statistics
Central South University
Changsha
P.R. China
E-mail address: chenzq453@gmail.com