Journal of Mathematical Research with Applications
Mar., 2020, Vol. 40, No. 2, pp. 209–220
DOI:10.3770/j.issn:2095-2651.2020.02.009
Http://jmre.dlut.edu.cn

Orthogonality Based Empirical Likelihood Inferences for Linear Mixed Effects Models

Changqing LIU[1], **Peixin ZHAO**[2,3,*], **Yiping YANG**[2]

1. *College of Mathematics and Statistics, Baise University, Guangxi* 533000*, P. R. China;*
2. *College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing* 400067*, P. R. China;*
3. *Chongqing Key Laboratory of Social Economy and Applied Statistics, Chongqing* 400067*, P. R. China*

Abstract Based on empirical likelihood method and QR decomposition technique, an orthogonality empirical likelihood based estimation method for the fixed effects in linear mixed effects models is proposed. Under some regularity conditions, the proposed empirical log-likelihood ratio is proved to be asymptotically chi-squared, and then the confidence intervals for the fixed effects are constructed. The proposed estimation procedure is not affected by the random effects, and then the resulting estimator is more effective. Some simulations and a real data application are conducted for further illustrating the performances of the proposed method.

Keywords Linear mixed effects model; orthogonality empirical likelihood; QR decomposition; random effects

MR(2010) Subject Classification 62G05; 62G15; 62G20

1. Introduction

Suppose we have a random sample of n subjects with n_i observations for the ith subject. Let Y_{ij} and $(X_{ij},\ Z_{ij})$ be the response variable and the covariates, respectively, where X_{ij} is a p-dimensional covariate vector, and Z_{ij} is a q-dimensional covariate vector. Then the linear mixed effects model has the following structure

$$Y_{ij} = X_{ij}^T\beta + Z_{ij}^T b_i + \varepsilon_{ij},\quad i=1,\ldots,n,\ j=1,\ldots,n_i, \tag{1.1}$$

where $\beta = (\beta_1,\ldots,\beta_p)^T$ is a $p\times 1$ vector of fixed effects and $b_i = (b_{i1},\ldots,b_{iq})^T$ is a $q\times 1$ vector of random effects of the ith subject, ε_{ij} is a zero-mean model error. We assume that $\{b_i, i=1,\ldots,n\}$ and $\{\varepsilon_{ij}, i=1,\ldots,n, j=1,\ldots,n_i\}$ are both independent and identically distributed random series. Further, for the identifiability of the fixed effects β, we assume that the expectations of the random effects and errors are all zeros.

When the random effects and model errors are all normally distributed, the maximum likelihood estimation and restricted maximum likelihood estimation are popular to use for model (1.1) (see Hartley and Rao [1], Miller [2] and Jiang [3]). When only moments of the random

Received September 6, 2018; Accepted September 4, 2019
Supported by the National Social Science Foundation of China (Grant No. 18BTJ035).
* Corresponding author
E-mail address: zpx81@163.com (Peixin ZHAO)

effects and model errors are assumed to exist, the quasi-likelihood method is explored to handle the inferences of model (1.1) (see Heyde [4], Richardson and Welsh [5] and Jiang [6]). In addition, Goldstein [7] proposed an iterative generalized least-squares method to estimate the model parameters in model (1.1). Cui et al. [8] used the method of moments to estimate the model parameters, and Wu and Zhu [9] proposed an orthogonality estimation method of higher-order moments of the random effects and errors.

However, in general, the maximum likelihood estimation procedure and the restricted maximum likelihood estimation procedure should specify the distributions of random effects and model errors. Then taking this issue into account, we rely on the empirical likelihood method as a flexible estimation tool, and proposed a new estimation procedure for the fixed effects in model (1.1). The empirical likelihood estimation procedure is a more flexible nonparametric estimation method, which does not need any assumption about the distributions of random effects and model errors, and the confidence interval construction of resulting estimator does not need any asymptotic variance estimation. Compared with the existing empirical likelihood methods for linear mixed effects models such as Chen et al. [10], our orthogonality based estimation procedure can individually estimate the fixed effects without any influence of random effects. This is a positive improvement of Chen et al. [10]. In addition, although Wu and Zhu [9] considered the estimation for linear mixed effects models by using orthogonality estimation technique, our empirical likelihood based estimation method is different from the estimating equation based estimation procedure proposed by Wu and Zhu [9]. Hence, this paper provides a positive result of the orthogonality based estimation technology, and extends the application literature of the empirical likelihood method.

Recently, the QR decomposition based estimation technology also has been considered by some authors. For example, Huang and Zhao [11] considered the orthogonality estimation for longitudinal partially linear models based on the QR decomposition technique. Zhao and Zhou [12] proposed a robust empirical likelihood estimation procedure for partially linear models by combining the QR decomposition technology and weighted composite quantile regression method. The work of this paper is an additional positive result for the QR decomposition based orthogonality estimation technology. More works of the orthogonality estimation technology can be found in Zhao et al. [13], Huang and Zhao [14], Yang and Yang [15], and among others.

The paper is organized as follows. In Section 2, we present the estimation procedure, and derive some asymptotic properties of the resulting estimator. In Section 3, we study the finite sample properties of the proposed estimation procedure by some simulations and a real data analysis. Finally, the technical proofs of all asymptotic results are provided in Section 4.

2. Methodology and main results

We denote $Y_i = (Y_{i1}, \ldots, Y_{in_i})^T$, $X_i = (X_{i1}, \ldots, X_{in_i})^T$, $Z_i = (Z_{i1}, \ldots, Z_{in_i})^T$ and $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})^T$. Then from model (1.1), we can get

$$Y_i = X_i\beta + Z_i b_i + \varepsilon_i, \quad i = 1, \ldots, n. \tag{2.1}$$

We further assume Z_i is a column full rank matrix. Then, based on the definition of the QR decomposition of a matrix, the matrix Z_i can be decomposed as

$$Z_i = Q_i \begin{pmatrix} R_i \\ \mathbf{0} \end{pmatrix},$$

where Q_i is an $n_i \times n_i$ orthogonal matrix, R_i is a $q \times q$ triangular matrix, and $\mathbf{0}$ is an $(n_i - q) \times q$ zero matrix. Furthermore, we denote $Q_i = (q_{i1}, q_{i2}, \ldots, q_{in_i})$, $Q_{i1} = (q_{i1}, \ldots, q_{iq})$ and $Q_{i2} = (q_{iq+1}, \ldots, q_{in_i})$. Then, it is easy to show that $Z_i = Q_{i1} R_i$ and $Q_{i2}^T Q_{i1} = \mathbf{0}$. Hence we have $Q_{i2}^T Z_i = Q_{i2}^T Q_{i1} R_i = \mathbf{0}$. By using this information, and left multiplying the both sides of equation (2.1) by Q_{i2}^T yields

$$Q_{i2}^T Y_i = Q_{i2}^T X_i \beta + Q_{i2}^T \varepsilon_i, \quad i = 1, \ldots, n. \tag{2.2}$$

Invoking the definition of Q_{i2}, model (2.2) can be rewritten as

$$q_{ij}^T Y_i = q_{ij}^T X_i \beta + q_{ij}^T \varepsilon_i, \quad i = 1, \ldots, n, \ j = q+1, \ldots, n_i. \tag{2.3}$$

Note that $Q_{i2}^T Q_{i2} = I_{n_i - q}$, where $I_{n_i - q}$ is an identity matrix, then we have $E\{q_{ij}^T \varepsilon_i \varepsilon_i^T q_{ij}\} = \sigma_\varepsilon^2$, which implies that model (2.3) is homogeneous. Hence, to construct an empirical log-likelihood ratio function for fixed effects β, invoking model (2.3), we define an auxiliary random vector as follows

$$\eta_i(\beta) = \sum_{i=q+1}^{n_i} X_i^T q_{ij} q_{ij}^T (Y_i - X_i \beta), \quad i = 1, \ldots, n. \tag{2.4}$$

Remark 2.1 Although linear mixed effects model contains fixed effects and random effects simultaneously, we construct model (2.3) in the orthogonal column space of Z_i, $i = 1, \ldots, n$, and make sure the auxiliary random vector (2.4) does not depend on the random effects. Then, based on (2.4), we can separately make statistical inferences for fixed effects without any affection from the random effects.

From (2.3), it can be shown that $E\{\eta_i(\beta)\} = 0$ when β is the true parameter. Hence, we can construct an empirical log-likelihood ratio function for β by using $\eta_i(\beta)$. More specifically, the empirical log-likelihood ratio for fixed effects β is defined as

$$R(\beta) = -2 \sup \Big\{ \sum_{i=1}^n \log(np_i) \big| p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \eta_i(\beta) = 0 \Big\},$$

where p_i means the probability of $\eta_i(\beta)$ occurrence. We assume that zero is inside the convex hull of the point $(\eta_1(\beta), \ldots, \eta_n(\beta))$, then a unique value for $R(\beta)$ exists. By the Lagrange multiplier method and using the same arguments as in Owen [16], $R(\beta)$ can be represented as

$$R(\beta) = 2 \sum_{i=1}^n \log\{1 + \lambda^T \eta_i(\beta)\}, \tag{2.5}$$

where λ is the Lagrange multiplier, which satisfies

$$\sum_{i=1}^n \frac{\eta_i(\beta)}{1 + \lambda^T \eta_i(\beta)} = 0. \tag{2.6}$$

Under some regularity conditions, Lemma 4.1 in Section 4 shows that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\xrightarrow{\mathcal{L}}N(0,\Phi),$$

where $\Phi=\sigma_\varepsilon^2\Gamma$, and Γ is defined in the following condition (C4). Based on this result, and using a regular proof method such as in Owen [16], we can show that $R(\beta)$ is asymptotically chi-square distributed when β is the true parameter. For clear exposition of the asymptotic behavior of $R(\beta)$, some regularity conditions are listed as follows.

(C1) The random effects b_i, $i=1,\ldots,n$, are independent and identically distributed, and satisfy $E(b_i|X_{ij},Z_{ij})=0$ and $E(\|b_i\|^4)<\infty$.

(C2) Let $\varepsilon_i=(\varepsilon_{i1},\ldots,\varepsilon_{in_i})$ and $V_i=E(\varepsilon_i\varepsilon_i^T)$. Then V_i satisfies $\sup_i\|V_i\|<\infty$. In addition, there exists a positive constant δ such that $E(\|\varepsilon_i\|^{2+\delta})<\infty$.

(C3) The covariates X_{ij} and Z_{ij} satisfy $\sup_{ij}E\{\|X_{ij}\|^4\}<\infty$ and $\sup_{ij}E\{\|Z_{ij}\|^4\}<\infty$, $i=1,\ldots,n$, $j=1,\ldots,n_i$. In addition, the covariate matrices Z_i, $i=1,\ldots,n$, defined in model (2.1), are all column full rank matrices.

(C4) Denote $P_{z_i}=Q_{i2}Q_{i2}^T$, then we assume that

$$\frac{1}{n}\sum_{i=1}^{n}E\{X_i^TP_{z_i}X_i\}\longrightarrow\Gamma,$$

where Γ is an invertible matrix.

Under these regularity conditions, we give the following theorem that states the asymptotic distribution of $R(\beta)$.

Theorem 2.2 *Suppose that the conditions (C1)–(C4) hold. Then, if β is the true value of the parameter, we have*

$$R(\beta)\xrightarrow{\mathcal{L}}\chi_p^2,$$

where "$\xrightarrow{\mathcal{L}}$" means the convergence in distribution, and χ_p^2 denotes the chi-square distribution with p degrees of freedom.

As a consequence of the Theorem 2.2, the confidence region for the fixed effects β can be constructed. More specifically, for any given α with $0<\alpha<1$, let c_α satisfy $P(\chi_p^2\le c_\alpha)=1-\alpha$, then the approximate $1-\alpha$ confidence region for β can be given as

$$C_\alpha(\beta)=\{\beta|R(\beta)\le c_\alpha\}.$$

Furthermore, we also can maximize $-R(\beta)$ to obtain the maximum empirical likelihood estimator $\hat\beta$ of β. In addition, from the proof procedure of Theorem 2.2 in Section 4, we have that

$$R(\beta)=\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\Big\}^T\Phi_n^{-1}\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\Big\}+o_p(1),\tag{2.7}$$

where $\Phi_n=n^{-1}\sum_{i=1}^{n}\eta_i(\beta)\eta_i^T(\beta)$. Clearly, this is a quadratic form, then maximizing $-R(\beta)$ to obtain the estimator $\hat{\beta}$ is asymptotic equivalent to solving the following estimating equation

$$\sum_{i=1}^{n}\eta_i(\beta)=\sum_{i=1}^{n}\sum_{i=q+1}^{n_i}X_i^Tq_{ij}q_{ij}^T(Y_i-X_i\beta)=0. \tag{2.8}$$

Note that $\sum_{i=q+1}^{n_i}q_{ij}q_{ij}^T=Q_{i2}Q_{i2}^T$, then (2.8) can be rewritten as

$$\sum_{i=1}^{n}X_i^TQ_{i2}Q_{i2}^T(Y_i-X_i\beta)=0. \tag{2.9}$$

Let $\Gamma_n=n^{-1}\sum_{i=1}^{n}X_i^TQ_{i2}Q_{i2}^TX_i$. Then it follows from (2.9) that

$$\hat{\beta}=\Gamma_n^{-1}\frac{1}{n}\sum_{i=1}^{n}X_i^TQ_{i2}Q_{i2}^TY_i. \tag{2.10}$$

Obviously, the estimator $\hat{\beta}$ defined by (2.10) is the same as the estimator obtained by Wu and Zhu [9]. Then, using the similar arguments to Wu and Zhu [9], we can prove that, under some mild regularity conditions, the estimator $\hat{\beta}$ obtained by maximizing $-R(\beta)$ has asymptotic normality, which is stated in the following Theorem 2.3.

Theorem 2.3 *Suppose that the conditions (C1)–(C4) hold. Then, if β is the true value of the parameter, we have*

$$\sqrt{n}(\hat{\beta}-\beta)\xrightarrow{\mathcal{L}}N(0,\Sigma),$$

where $\Sigma=\sigma_\varepsilon^2\Gamma^{-1}$, and Γ is defined in conditions (C4).

We also can construct the confidence region for β based on Theorem 2.3. But the asymptotic variance Σ should be estimated. Here, we propose a consistent estimation procedure of Σ based on plug-in method. More specifically, from the proof procedure of Lemma 4.1 in Section 4, we have that Γ_n is a consistent estimator of Γ, where Γ_n is defined by (2.10). In addition, from the proof procedure of Theorem 2.2 in Section 4, we have that Φ_n is a consistent estimator of Φ, where $\Phi_n=n^{-1}\sum_{i=1}^{n}\eta_i(\beta)\eta_i(\beta)^T$ and $\Phi=\sigma_\varepsilon^2\Gamma$. Hence invoking $\Sigma=\Gamma^{-1}\Phi\Gamma^{-1}$ and by the plug-in method, a consistent estimator of Σ is given by $\hat{\Sigma}=\Gamma_n^{-1}\hat{\Phi}_n\Gamma_n^{-1}$, where $\hat{\Phi}_n=n^{-1}\sum_{i=1}^{n}\eta_i(\hat{\beta})\eta_i(\hat{\beta})^T$. In addition, for random effect b, an interesting topic is to estimate the variance component of b. Invoking the maximum empirical likelihood estimator $\hat{\beta}$, and using the same arguments as in Wu and Zhu [9], can easily obtain the estimation of the moments of random effect b. Then, we omit the estimation details.

3. Numerical results

In this section, we present some simulation experiments to illustrate the finite sample performance of the proposed method, and consider a real data set analysis for further illustration.

3.1. Simulation studies

We first present some simulation studies to evaluate the performance of the proposed method. The data are generated from the following model

$$Y_{ij}=X_{ij}^T\beta+Z_{ij}^Tb_i+\varepsilon_{ij}, \tag{3.1}$$

where the fixed effects are taken as $\beta = (\beta_1, \beta_2)^T = (1, 2)^T$. The covariates $X_{ij} = (X_{1ij}, X_{2ij})^T$, $i = 1, \dots, n$, $j = 1, \dots, n_i$ are independently generated from the normal distribution $N(\mu_x, \Sigma_x)$ with $\mu_x = (1, 1)^T$ and $\Sigma_x = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$. The covariate Z_{ij} is taken as $Z_{ij} = (Z_{1ij}, Z_{2ij})^T$, where Z_{1ij} and Z_{2ij}, $i = 1, \dots, n$, $j = 1, \dots, n_i$ are independently generated from the normal distribution $N(0, 1)$. Furthermore, the response Y_{ij} is generated according to the model (3.1), where the model error ε_{ij} and random effects $b_i = (b_{1i}, b_{2i})^T$ are separately generated from the following three combinations:

(i) $\varepsilon_{ij} \sim 0.5N(0, 1)$, $b_{1i} \sim 0.5N(0, 1)$ and $b_{2i} \sim 0.5N(0, 1)$;

(ii) $\varepsilon_{ij} \sim 0.5N(0, 1)$, $b_{1i} \sim 0.5N(0, 1)$ and $b_{2i} \sim 0.5t(3)$;

(iii) $\varepsilon_{ij} \sim 0.5N(0, 1)$, $b_{1i} \sim 0.5t(3)$ and $b_{2i} \sim 0.5t(3)$.

To perform the simulation, the sample size is taken as $n = 100$, 200 and 300, respectively, and for the ith subject, the number of repeated measurements is randomly drawn from a Poisson distribution with mean $\lambda = 8$.

		$n = 100$		$n = 200$		$n = 300$	
Model	Method	Bias	SD	Bias	SD	Bias	SD
(i)	OEL	-0.0032	0.0168	-0.0019	0.0113	0.0007	0.0097
	OBE	0.0035	0.0172	0.0018	0.0116	-0.0009	0.0097
	LSE	-0.0041	0.025	0.0021	0.0201	-0.0013	0.0184
(ii)	OEL	0.0035	0.0166	0.0016	0.0116	0.0008	0.0104
	OBE	-0.0039	0.0169	-0.0017	0.0118	0.0009	0.0107
	LSE	-0.0045	0.0272	0.0023	0.0205	0.0011	0.0196
(iii)	OEL	-0.0035	0.0169	0.0017	0.0118	-0.0008	0.0103
	OBE	0.0038	0.0172	-0.0016	0.0116	0.0008	0.0102
	LSE	0.0041	0.0279	-0.0027	0.0203	-0.0009	0.0198

Table 1 The biases and standard deviations of $\hat{\beta}_1$ by different estimation methods

We first evaluate the efficiencies of the proposed method for the fixed effects estimator $\hat{\beta}$. Note that the response variables follow the normal distribution for case (i), then the maximum likelihood estimation (MLE) procedure is equivalent to the least squares estimation (LSE) method for case (i). Hence, in this simulation, three estimation methods are compared: the orthogonality empirical likelihood method (OEL) proposed by this paper, the orthogonality based estimation procedure (OBE) proposed by Wu and Zhu [9] and the least squares based estimation method (LSE) defined in Wu et al. [17]. With 1000 simulation runs, Tables 1 and 2 report the averages of the biases (Bias) and the standard deviations (SD) of $\hat{\beta}_1$ and $\hat{\beta}_2$, respectively. From Tables 1 and 2, we can obtain the following observations:

(i) For any given distribution of random effects, when the sample size increases, the biases obtained by the three methods all decrease uniformly. This implies that the OEL, OBE and LSE methods all can give a consistent estimator for fixed effects β.

(ii) The proposed OEL method outperforms the LSE method in terms of standard deviation (SD). The standard deviations obtained by the OEL method decrease significantly when the sample size increases. However, the LSE method still gives a relatively larger standard deviation when the sample size increases. This implies that the estimation procedure for fixed effects, based on the OEL method, can avoid the influence of random effects.

(iii) For any given n, the results obtained by the OEL method are very similar for different model designs, which implies that the proposed OEL estimation method is insensitive to the distribution of random effects.

(iv) The performances of OEL and OBE methods are very similar in terms of bias and standard deviation. This also implies the estimators obtained by OEL and OBE methods are asymptotic equivalent, which agrees with the theoretical results presented in Section 2.

		$n = 100$		$n = 200$		$n = 300$	
Model	Method	Bias	SD	Bias	SD	Bias	SD
(i)	OEL	-0.0048	0.0176	0.0026	0.0121	-0.0018	0.0105
	OBE	-0.0051	0.0182	-0.0025	0.0126	0.0019	0.0107
	LSE	-0.0056	0.0344	0.0033	0.0241	0.0019	0.0182
(ii)	OEL	0.0047	0.0178	-0.0027	0.0123	0.0018	0.0107
	OBE	0.0049	0.0179	-0.0028	0.0126	-0.0019	0.0106
	LSE	-0.0055	0.0372	0.0035	0.0305	-0.0021	0.0296
(iii)	OEL	0.0049	0.0175	-0.0025	0.0126	0.0019	0.0109
	OBE	-0.0051	0.0178	-0.0026	0.0125	-0.0018	0.0107
	LSE	0.0061	0.0379	0.0027	0.0353	-0.0019	0.0293

Table 2 The biases and standard deviations of $\hat{\beta}_2$ by different estimation methods

		$n = 100$		$n = 200$		$n = 300$	
Model	Method	Len	CP	Len	CP	Len	CP
(i)	OEL	0.3578	0.938	0.2409	0.942	0.1334	0.951
	NA	0.5043	0.936	0.4156	0.941	0.2238	0.948
(ii)	OEL	0.3589	0.934	0.2462	0.941	0.1365	0.950
	NA	0.5112	0.934	0.4285	0.942	0.2399	0.948
(iii)	OEL	0.3563	0.938	0.2497	0.943	0.1371	0.948
	NA	0.5123	0.936	0.4364	0.943	0.2420	0.948

Table 3 The average lengths and coverage probabilities of the 95% confidence intervals for β_1

Next, we evaluate the performances of the confidence intervals for fixed effects obtained by the OEL method proposed by this paper and the normal approximation method (NA) proposed by

Wu and Zhu [9]. In Tables 3 and 4, we show the average interval lengths (Len) and corresponding coverage probabilities (CP) of the 95% confidence intervals for β_1 and β_2, respectively, based on 1000 simulation runs. From Tables 3 and 4, we can see that, although the coverage probabilities obtained by OEL and NA methods are similar, the confidence intervals based on the OEL method have uniformly shorter average lengths than those obtained by the NA method. This implies that the OEL method proposed by this paper performs better than the NA method for confidence interval construction, which is mainly because the confidence intervals obtained by the OEL method do not need any asymptotic variance estimation.

		$n = 100$		$n = 200$		$n = 300$	
Model	Method	Len	CP	Len	CP	Len	CP
(i)	OEL	0.3682	0.939	0.2510	0.942	0.1469	0.951
	NA	0.5149	0.937	0.4175	0.943	0.2217	0.947
(ii)	OEL	0.3687	0.938	0.2571	0.944	0.1477	0.949
	NA	0.5217	0.938	0.4294	0.942	0.2598	0.948
(iii)	OEL	0.3689	0.937	0.2595	0.943	0.1476	0.947
	NA	0.5263	0.936	0.4362	0.941	0.2837	0.947

Table 4 The average lengths and coverage probabilities of the 95% confidence intervals for β_2

3.2. Application to CD4 Data

We now illustrate the proposed estimation method in this paper through analysis of a data set from the Multi-Center AIDS Cohort study. The data set contains the human immunodeficiency virus (HIV) status of 283 homosexual men who were infected with HIV during a follow-up period between 1984 and 1991. Although the original design was to collect the measurements for all individuals semiannually, some individuals missed scheduled visits, which resulted in unequal numbers of measurements. More details about the related design, methods and medical implications of the Multi-Center AIDS Cohort study have been described by Kaslow et al. [18].

This data set has been used by many authors under different models. For example, Xue and Zhu [19], Wang et al. [20] and Huang et al. [21] analysis this data by using varying coefficient models. Fan and Li [22] and Xue and Zhu [23] analyzed this data by using partially linear models. In addition, He et al. [24] [25] analyzed this data by using semi-varying coefficient models with fixed effects. In this paper, the objective of the study is to evaluate the effects of the pre-HIV infection CD4 percentage, the time after HIV infection, and age at HIV infection on the mean CD4 percentage after infection. Note that the effect of age at HIV infection may vary for different individuals, then we analyze this data by using the following mixed effects model

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + b_i Z_{ij} + \varepsilon_{ij},$$

where Y_{ij} is the individual's CD4 percentage, X_{1ij} is the time after HIV infection, X_{2ij} is the pre-HIV infection CD4 percentage, and Z_{ij} is the individual's age at HIV infection. β_0, β_1 and β_2 are fixed effects, and b_i is random effect, which reflects the individual variations.

By using the orthogonality empirical likelihood method proposed by this paper, the estimators and the 95% confidence intervals of fixed effects $\beta_k,\ k = 0, 1, 2$ are reported in Table 5. For comparison, Table 5 also presents the OBE estimators and the normal approximation based confidence intervals proposed by Wu and Zhu [9]. From Table 5, we can see that $\beta_1 < 0$, which implies that the CD4 percentage decreases as the time after HIV infection goes on. $\beta_2 > 0$ means that the pre-HIV infection CD4 percentage has a positive effect on the individual's CD4 percentage after HIV infection. Furthermore, the estimator of β_2 is 0.3621, which is little than $\beta_2 = 0.74$ obtained by Xue and Zhu [23]. This implies that the effect of pre-HIV infection CD4 percentage obtained by Xue and Zhu [23] might be overestimated. In addition, we also can see that the interval lengths obtained by the proposed orthogonality empirical likelihood method are uniformly shorter than those obtained by the normal approximation method, which basically agrees with what was discovered in simulation studies.

Method	Fixed effects	Estimator	Confidence interval	Interval length
OEL	β_0	19.3839	(18.9106, 19.8573)	0.9467
	β_1	-2.3747	(-2.5747, -2.1830)	0.3917
	β_2	0.3621	(0.3503, 0.3737)	0.0234
OBE	β_0	19.3841	(16.7775, 21.9907)	5.2132
	β_1	-2.3749	(-2.6814, -2.0684)	0.6130
	β_2	0.3622	(0.3051, 0.4193)	0.1142

Table 5 Application to CD4 data. The estimators and 95% confidence intervals for fixed effects

4. Proofs of Theorems

In this section, we present the technical proofs of Theorems 2.2 and 2.3. For convenience and simplicity, let c denote a positive constant which may be different values at each appearance throughout this paper. To facilitate the proof procedure, firstly, we list some lemmas.

Lemma 4.1 *Suppose that conditions (C1)–(C4) hold. If β is the true value of the parameter, then we have*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\stackrel{\mathcal{L}}{\longrightarrow}N(0,\Phi),$$

where $\Phi = \sigma_\varepsilon^2\Gamma$, and Γ is defined in condition (C4).

Proof From (2.3), a simple calculation yields

$$\eta_i(\beta)=\sum_{j=q+1}^{n_i}X_i^Tq_{ij}(q_{ij}^TY_i-q_{ij}^TX_i\beta)=\sum_{j=q+1}^{n_i}X_i^Tq_{ij}q_{ij}^T\varepsilon_i=X_i^TQ_{i2}Q_{i2}^T\varepsilon_i. \tag{4.1}$$

Then we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i^TQ_{i2}Q_{i2}^T\varepsilon_i\equiv\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i. \tag{4.2}$$

Invoking $E\{\varepsilon_i|X_i,Z_i\}=0$, and some calculations yield $E\{n^{-1/2}\sum_{i=1}^n\xi_i\}=0$ and

$$\mathrm{Var}\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i\Big\}=\frac{1}{n}\sum_{i=1}^{n}E\{\xi_i\xi_i^T\}=\frac{1}{n}\sum_{i=1}^{n}\sigma_\varepsilon^2E\{X_i^TQ_{i2}Q_{i2}^TX_i\}\longrightarrow\sigma_\varepsilon^2\Gamma. \tag{4.3}$$

Hence, invoking (4.2) and (4.3), and using the central limits theorem, we complete the proof of this lemma. □

Proof of Theorem 2.2 Invoking the proof of Lemma 4.1, and using the same arguments as in Owen [16], we can obtain

$$\|\lambda\|=O_p(n^{-1/2}). \tag{4.4}$$

In addition, from Lemma 4.1, it is easy to show that

$$\max_{1\le i\le n}\|\eta_i(\beta)\|=o_p(n^{1/2}). \tag{4.5}$$

Then, invoking (4.4) and (4.5), and using the Taylor expansion to (2.5), we obtain that

$$R(\beta)=2\sum_{i=1}^{n}\{\lambda^T\eta_i(\beta)-(\lambda^T\eta_i(\beta))^2/2\}+o_p(1). \tag{4.6}$$

Furthermore, it follows from (2.6) that

$$\sum_{i=1}^{n}\eta_i(\beta)-\sum_{i=1}^{n}\eta_i(\beta)\eta_i(\beta)^T\lambda+\sum_{i=1}^{n}\frac{\eta_i(\beta)[\lambda^T\eta_i(\beta)]^2}{1+\lambda^T\eta_i(\beta)}=0. \tag{4.7}$$

Invoking (4.4) and (4.5), we can prove that

$$\sum_{i=1}^{n}[\lambda^T\eta_i(\beta)]^2=\sum_{i=1}^{n}\lambda^T\eta_i(\beta)+o_p(1), \tag{4.8}$$

$$\lambda=\Big\{\sum_{i=1}^{n}\eta_i(\beta)\eta_i(\beta)^T\Big\}^{-1}\sum_{i=1}^{n}\eta_i(\beta)+o_p(n^{-1/2}). \tag{4.9}$$

Invoking (4.6)–(4.9), and using the same arguments as in Owen [16], we can obtain

$$R(\beta)=\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\Big\}^T\Phi_n^{-1}\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i(\beta)\Big\}+o_p(1), \tag{4.10}$$

where $\Phi_n=n^{-1}\sum_{i=1}^n\eta_i(\beta)\eta_i(\beta)^T$. In addition, from the proof of Lemma 4.1, we can obtain that

$$\Phi_n=\frac{1}{n}\sum_{i=1}^{n}\eta_i(\beta)\eta_i(\beta)^T=\frac{1}{n}\sum_{i=1}^{n}\xi_i\xi^T\xrightarrow{\mathcal{P}}\Phi. \tag{4.11}$$

Invoking (4.10), (4.11) and Lemma 4.1, we complete the proof of Theorem 2.2. □

Proof of Theorem 2.3 From (2.8), we have that $\hat\beta$ is the solution of estimating equation

$\sum_{i=1}^n \eta_i(\beta) = 0$. Then a simple calculation yields

$$\begin{aligned} 0 = \sum_{i=1}^n \eta_i(\hat{\beta}) &= \sum_{i=1}^n X_i^T Q_{i2} Q_{i2}^T (Y_i - X_i \hat{\beta}) \\ &= \sum_{i=1}^n X_i^T Q_{i2} Q_{i2}^T (Y_i - X_i \beta) + \sum_{i=1}^n X_i^T Q_{i2} Q_{i2}^T X_i (\beta - \hat{\beta}). \end{aligned} \tag{4.12}$$

From (4.12), we obtain that

$$\begin{aligned} \sqrt{n}(\hat{\beta} - \beta) &= \Gamma_n^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n X_i^T Q_{i2} Q_{i2}^T (Y_i - X_i \beta) \\ &= \Gamma_n^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n \eta_i(\beta), \end{aligned} \tag{4.13}$$

where Γ_n is defined by (2.10). In addition, invoking condition (C4), and by the law of large numbers, we can prove that $\Gamma_n \xrightarrow{\mathcal{P}} \Gamma$. Hence, invoking (4.13) and Lemma 4.1, and using the Slutsky's theorem, we obtain

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{L}} N(0, \Sigma),$$

where $\Sigma = \Gamma^{-1} \Phi \Gamma^{-1} = \sigma_\varepsilon^2 \Gamma^{-1}$. This completes the proof of Theorem 2.3. □

References

[1] H. O. HARTLEY, J. N. K. RAO. *Maximum likelihood estimation for the mixed analysis of variance model*. Biometrika, 1967, **54**: 93–108.

[2] J. J. MILLER. *Asymptotic properties of maximum likelihood estimates in the mixed model of analysis of variance*. Ann. Statist., 1977, **5**(4): 746–762.

[3] Jiming JIANG. *REML estimation: asymptotic behavior and related topics*. Ann. Statist., 1996, **24**(1): 255–286.

[4] C. C. HEYDE. *Quasi-Likelihood and Its Application*. Springer-Verlag, New York, 1997.

[5] A. M. RICHARDSON, A. H. WELSH. *Asymptotic properties of restricted maximum likelihood (REML) estimates for hierarchical mixed linear models*. Austral. J. Statist., 1994, **36**(1): 31–43.

[6] Jiming JIANG. *Wald consistency and the method of sieves in REML estimation*. Ann. Statist., 1997, **25**(4): 1781–1803.

[7] H. GOLDSTEIN. *Multilevel mixed linear model analysis using iterative generalized least squares*. Biometrika, 1986, **73**(1): 43–56.

[8] Hengjian CUI, K. W. NG, Lixing ZHU. *Estimation in mixed effects model with errors in variables*. J. Multivariate Anal., 2004, **91**(1): 53–73.

[9] Ping WU, Lixing ZHU. *An orthogonality-based estimation of moments for linear mixed models*. Scand. J. Stat., 2010, **37**(2): 253–263.

[10] Qiuhua CHEN, Pingshou ZHONG, Hengjian CUI. *Empirical likelihood for mixed-effects error-in-variables model*. Acta Math. Appl. Sin. Engl. Ser., 2009, **25**(4): 561–578.

[11] Jiting HUANG, Peixin ZHAO. *QR decomposition based orthogonality estimation for partially linear models with longitudinal data*. J. Comput. Appl. Math., 2017, **321**: 406–415.

[12] Peixin ZHAO, Xiaoshuang ZHOU. *Robust empirical likelihood for partially linear models via weighted composite quantile regression*. Comput. Statist., 2018, **33**(2): 659–674.

[13] Yanyong ZHAO, Jinguan LIN, Peirong XU, et al. *Orthogonality-projection-based estimation for semi-varying coefficient models with heteroscedastic errors*. Comput. Statist. Data Anal., 2015, **89**: 204–221.

[14] Jiting HUANG, Peixin ZHAO. *Orthogonal weighted empirical likelihoodbased variable selection for semiparametric instrumental variable models*. Communications in Statistics-Theory and Methods, 2018, **47**: 4375–4388.

[15] Jing YANG, Hu YANG. *Smooth-threshold estimating equations for varying coefficient partially nonlinear models based on orthogonality-projection method*. J. Comput. Appl. Math., 2016, **302**: 24–37.

[16] A. B. OWEN. *Empirical likelihood ratio confidence regions*. Ann. Statist., 1990, **18**(1): 90–120.

[17] Ping WU, Yun FANG, Lixing ZHU. *Estimating moments in linear mixed models*. Comm. Statist. Theory Methods, 2008, **37**(16-17): 2582–2594.

[18] R. A. KASLOW, D. G. OSTROW, R. DETELS, et al. *The multicenter AIDS cohort study: rationale, organization and selected characteristics of the participants*. Am. J. Epidemiol, 1987, **126**: 310–318.

[19] Liugen XUE, Lixing ZHU. *Empirical likelihood for a varying coefficient model with longitudinal data*. J. Amer. Statist. Assoc., 2007, **102**(478): 642–654.

[20] Lifeng WANG, Hongzhe LI, Jianhua Z. HUANG. *Variable selection in nonparametric varying-coefficient models for analysis of repeated measurements*. J. Amer. Statist. Assoc., 2008, **103**(484): 1556–1569.

[21] Jianhua Z. HUANG, Colin O. WU, Lan ZHOU. *Varying coefficient models and basis function approximations for the analysis of repeated measurements*. Biometrika, 2002, **89**(11): 111–128.

[22] Jianqing FAN, Runze LI. *New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis*. J. Amer. Statist. Assoc., 2004, **99**(467): 710–723.

[23] Liugen XUE, Lixing ZHU. *Empirical likelihood semiparametric regression analysis for longitudinal data*. Biometrika, 2007, **94**(4): 921–937.

[24] Bangqiang HE, Xingjian HONG, Guoliang FAN. *Empirical likelihood for semi-varying coefficient models for panel data with fixed effects*. J. Korean Statist. Soc., 2016, **45**(3): 395–408.

[25] Bangqiang HE, Xingjian HONG, Guoliang FAN. *Block empirical likelihood for partially linear panel data models with fixed effects*. Statist. Probab. Lett., 2017, **123**: 128–138.