



วิธีกำลังสองน้อยที่สุดสองชั้นแบบปรับปรุงสำหรับการประมาณค่าพารามิเตอร์ในตัวแบบการถดถอยไม่เชิงเส้นเมื่อความคลาดเคลื่อนมีสหสัมพันธ์

Modified Two-stage Least Squares Methods for Estimating Parameters in Nonlinear Regression Models with Correlated Errors

วรรณภา ภัคดี*

Wannapa Pukdee*

ภาควิชาวิทยาศาสตร์ทั่วไป คณะวิทยาศาสตร์และวิศวกรรมศาสตร์

มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตเฉลิมพระเกียรติ จังหวัดสกลนคร

Department of General Science, Faculty of Sciences and Engineering,

Kasetsart University Chalermphrakiat Sakon Nakhon Province Campus

Received : 9 February 2020

Revised : 14 April 2020

Accepted : 24 July 2020

บทคัดย่อ

วิธีกำลังสองน้อยที่สุดสองชั้นถูกใช้เพื่อหาค่าเหมาะสมของตัวแบบการถดถอยไม่เชิงเส้นเมื่อความคลาดเคลื่อนมีสหสัมพันธ์ภายใต้กระบวนการถดถอยในตัวคงที่อันดับหนึ่ง วิธีนี้ทำให้ได้ค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์ที่ต่ำเกินไป ดังนั้นงานวิจัยนี้จึงนำเสนอวิธีกำลังสองน้อยที่สุดสองชั้นแบบปรับปรุงโดยการใช้ค่าส่วนเหลือจากตัวแบบแอนโนวาทางเดียวและสัมประสิทธิ์สหสัมพันธ์จากกระบวนการกำลังสองน้อยที่สุดแบบมีเงื่อนไขเพื่อสร้างเมทริกซ์ถ่วงน้ำหนัก วิธีเหล่านี้ถูกใช้ประมาณค่าพารามิเตอร์ทุกตัวของตัวแบบการถดถอยไม่เชิงเส้น จากการศึกษาโดยการจำลองตัวแบบที่ระดับความสัมพันธ์ต่าง ๆ ผลการศึกษาพบว่า วิธีกำลังสองน้อยที่สุดสองชั้นแบบปรับปรุงทำให้การอนุมานเชิงสถิติมีประสิทธิภาพเนื่องจากตัวประมาณค่าพารามิเตอร์ไม่มีความเอนเอียงและลดความเอนเอียงในการประมาณความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์

คำสำคัญ : การถดถอยไม่เชิงเส้น ; ความคลาดเคลื่อนที่มีสหสัมพันธ์ ; การประมาณกำลังสองน้อยที่สุด ; วิธีกำลังสองน้อยที่สุดสองชั้น



Abstract

The two-stage least squares method is used to fit nonlinear regression models with correlated errors based on a stationary autoregressive process of order one. This method tends to underestimate the standard error of parameter estimates. Therefore, this paper presents modified two-stage least squares methods by using residuals from the one-way ANOVA model and estimating the correlation coefficient from the conditional least squares procedure to construct a weight matrix. These methods are used to estimate all parameters of nonlinear regression models. A simulation study covers a wide range of correlation levels on the models. The study result shows that the modified two-stage least squares methods can improve the efficient statistical inference since they produce unbiased estimators of parameters and reduce bias in estimating standard errors for parameter estimates.

Keywords : nonlinear regression ; correlated errors ; least squares estimation ;
the two-stage least squares method



Introduction

Nonlinear regression models are widely used by many researchers in order to describe data measured over time. For example, a curve model is applied to analyze pharmacodynamics data obtained in studying blood glucose after injection of insulin. By using the glucose-clamp technique, the responses are glucose infusion rates along with times (Bender & Heinemann, 1995). Moreover, there are many application areas in gene expression data studied on circadian rhythms. The rhythms can control many biological processes and help to predict the effects of particular drugs and assist in the design of protocols for drug administration. Sinusoidal models are used to describe cyclic rhythms components of biological organisms (Izumo *et al.*, 2003; Maier *et al.*, 2009). Responses arising from all the experimental studies are measured over the course of time, and these responses give serially correlated observations.

The efficient methods used to fit nonlinear regression models with correlated errors in the case of a stationary autoregressive process of order one, AR(1), are conditional least squares (CLS) estimation (Bates & Watts, 1988) and a two-stage least squares (TS) method (Seber & Wild, 2003). Pukdee *et al.*, (2018) presented TS and CLS methods based on the least squares procedures when assumptions on the joint distributions of the errors are not made for fitting sinusoidal models where the errors are correlated. The TS approach produced underestimated standard errors for period parameter estimates. In order to improve inferential statistics, Pukdee *et al.*, (2020) proposed a modified two-stage least squares (MTS_{ANOVA}) method by firstly using residuals from the one-way ANOVA model to modify the correlation coefficient in a weight matrix, and secondly applying the modified weight matrix to transform nonlinear regression models. The transformed models are estimated by the ordinary least squares approach. The MTS_{ANOVA} method was applied to estimate a period parameter of sinusoidal models and its standard error was more accurate. In this paper, the TS method modified by using the CLS procedure to estimate the correlation coefficient in a weight matrix is proposed, and this is here called the MTS_{CLS} method. Moreover, these modified two-stage MTS_{CLS} and MTS_{ANOVA} methods are used to estimate all the parameters of nonlinear regression models with correlated errors in literatures, i. e. , exponential growth, log-normal as well as sinusoidal models, and then compared them with the TS method.

Methods

This section provides analysis methods for studying correlated data. The observed data is presented as a general nonlinear regression model with correlated errors,

$$\mathbf{y}_i = \mathbf{f}(\mathbf{t}_i; \boldsymbol{\theta}) + \boldsymbol{\varepsilon}_i; \quad i = 1, \dots, r, \quad (1)$$



where $\mathbf{y}_i = (y_{i,1}, \dots, y_{i,n})'$ is a response variable vector for the i^{th} dataset at n time points, and each observed data with sample size n is measured for r replicates, $\mathbf{t}_i = (t_{i,1}, \dots, t_{i,n})'$ is a time vector as an independent variable, $\mathbf{f}(\mathbf{t}_i; \boldsymbol{\theta}) = (f(t_{i,1}; \boldsymbol{\theta}), \dots, f(t_{i,n}; \boldsymbol{\theta}))'$ is any nonlinear function of t with an unknown parameter vector $\boldsymbol{\theta}$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i,1}, \dots, \varepsilon_{i,n})'$ is an error under assuming a stationary autoregressive process of order one, AR(1), as given by

$$\varepsilon_{i,j} = \rho \varepsilon_{i,j-1} + \delta_{i,j}; \quad j = 1, \dots, n, \quad (2)$$

where $|\rho| < 1$ is the simple correlation between errors $\varepsilon_{i,j}$ and $\varepsilon_{i,j-1}$, $\delta_{i,j}$ are random errors assumed to be independent and identically distributed (i.i.d.) variables with zero mean and constant variance.

Nonlinear functions to model the correlated data are used and found in the literatures as follows:

$$f(t; \boldsymbol{\theta}) = \theta_1 (1 - \theta_2 \exp(-\theta_3 t)) \quad (3)$$

is an exponential growth function with an unknown parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)'$ where $\theta_1, \theta_2, \theta_3$ are the maximum observed data, an initial value of the response at the reaction time point and a rate constant, as presented by Bates & Watts (1988). Next, shown in Bender & Heinemann (1995) the log-normal function is

$$f(t; \boldsymbol{\theta}) = \frac{A}{t} \exp\left(-B [\log(t) - C]^2\right), \quad (4)$$

where the parameter vector of the function is $\boldsymbol{\theta} = (A, B, C)'$ and A, B, C are the regression coefficients of the log-normal function (Bender, 2000). Finally, the one-sine function adapted from Izumo *et al.*, (2003) by adding a linear trend is

$$f(t; \boldsymbol{\theta}) = \alpha + \beta t + a \exp(-dt) \sin\left(\frac{2\pi t}{\tau} + \Phi\right), \quad (5)$$

where the vector of unknown parameters is $\boldsymbol{\theta} = (\alpha, \beta, a, d, \tau, \Phi)'$ and let α, β be intercept and slope of the regression line, a, d, τ, Φ are the amplitude, a damping parameter, the period and the phase of the sine wave, respectively.



These three nonlinear regression functions with the AR(1) errors are fitted by a two-stage least squares method and modified two-stage least squares estimation. These methods are described below and are based on the least squares procedure.

The Two-stage least squares method

In order to estimate parameters of a nonlinear regression with the AR(1) error structure, the two-stage least squares (TS) method has two least squares procedures. Firstly, ignoring the correlation structure, the model (1) is fitted by ordinary least squares (OLS) estimation to produce parameter estimates $\hat{\theta}_{OLS}$ and fitted values $f(\mathbf{t}_i; \hat{\theta}_{OLS})$. The residual vector for the i^{th} dataset

$$\hat{\mathbf{e}}_i = \mathbf{y}_i - f(\mathbf{t}_i; \hat{\theta}_{OLS}), \quad (6)$$

is calculated to produce an estimate of the correlation ρ_i given by Park & Mitchell (1980) as

$$\hat{\rho}_i = \frac{\sum_{j=2}^n \hat{\mathbf{e}}_{i,j} \hat{\mathbf{e}}_{i,j-1}}{\sum_{j=2}^n \hat{\mathbf{e}}_{i,j}^2}. \quad (7)$$

The mean of the r estimates obtained by $\hat{\rho} = \frac{\hat{\rho}_1 + \dots + \hat{\rho}_r}{r}$, is used to estimate the simple correlation coefficient ρ , and to construct the $n \times n$ weight matrix $\hat{\mathbf{R}}_i$ of elements $\hat{\rho}$ as

$$\hat{\mathbf{R}}_i = \begin{bmatrix} \sqrt{1 - \hat{\rho}^2} & 0 & 0 & \dots & 0 & 0 \\ -\hat{\rho} & 1 & 0 & \dots & 0 & 0 \\ 0 & -\hat{\rho} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\hat{\rho} & 1 \end{bmatrix}. \quad (8)$$

Secondly, to transform the model (1) to an OLS model, the model multiplied by $\hat{\mathbf{R}}_i$ is



$$\mathbf{z}_i = \mathbf{g}(t_i; \boldsymbol{\theta}) + \boldsymbol{\delta}_i; \quad i = 1, \dots, r, \tag{9}$$

where $\mathbf{z}_i = \hat{\mathbf{R}}_i \mathbf{y}_i$, $\mathbf{g}(t_i; \boldsymbol{\theta}) = \hat{\mathbf{R}}_i f(t_i; \boldsymbol{\theta})$, $\boldsymbol{\delta}_i = \hat{\mathbf{R}}_i \boldsymbol{\varepsilon}_i$. The TS model (9), as transformed under i.i.d. errors $\boldsymbol{\delta}_i \sim (\mathbf{0}, \sigma^2 \mathbf{I}_i)$, can be fitted by using OLS (Seber & Wild, 2003). The TS method produces parameter estimates $\hat{\boldsymbol{\theta}}_{TS}$ with asymptotic properties as $\hat{\boldsymbol{\theta}}_{OLS}$ (Gallant & Goebel, 1976).

Modified Two-stage least squares estimation

This estimation develops a modification to the previous TS method. The first stage of the TS method constructs the weight matrix $\hat{\mathbf{R}}_i$ (8) of elements $\hat{\rho}$ by using the mean of (7) from the residuals obtained by (6) based on the OLS estimation, but the matrix $\hat{\mathbf{R}}_i$ in the modified two-stage least squares (MTS) estimation can use different correlation coefficients (Asikgil & Erar, 2009).

The correlation coefficient based on the conditional least squares (CLS) approach with correlated errors under a stationary AR(1) process, $\hat{\rho}_{CLS}$, is estimated by minimizing

$$S(\boldsymbol{\theta}, \rho) = \sum_{i=1}^r \sum_{j=2}^n (y_{i,j} - \rho y_{i,j-1} - f(t_{i,j}; \boldsymbol{\theta}) + \rho f(t_{i,j-1}; \boldsymbol{\theta}))^2, \tag{10}$$

with respect to $\boldsymbol{\theta}$ and ρ simultaneously.

Alternatively, instead of directly using the correlation estimated from (10), the first step is to produce residuals from fitting a one-way ANOVA model of the replicate observations at each time point. That residual is

$$\hat{\varepsilon}_{ij} = y_{ij} - \hat{y}_{j,iid}, \tag{11}$$

where $\hat{y}_{j,iid}$ is the sample mean for the j^{th} time point, and the residual (11) is next used to estimate the correlation coefficient (7) for each replicate. This correlation coefficient is $\hat{\rho}_{ANOVA}$ (Pukdee *et al.*, 2020). With correlation coefficients $\hat{\rho}_{CLS}$ and $\hat{\rho}_{ANOVA}$ replaced in the weight matrix (8) respectively, the MTS methods are then used to fit nonlinear regression models in the second step as the same previous method.



Simulation study

In order to compare the performance of the above methods, each simulated dataset is generated from the CLS model in case of the AR(1) errors that the first observation is $y_{i,1} = f(t_{i,1}; \boldsymbol{\theta}) + \delta_{i,1}$, but the next one $y_{i,j}$ is generated as follows:

$$y_{i,j} = \rho y_{i,j-1} + f(t_{i,j}; \boldsymbol{\theta}) - \rho f(t_{i,j-1}; \boldsymbol{\theta}) + \delta_{i,j} \quad ; \quad j = 2, \dots, n, \quad (12)$$

where $f(t_{i,j}; \boldsymbol{\theta})$ is the function (3) with known parameters, as given by

$$f(t_{i,j}; \boldsymbol{\theta}) = 35 \left(1 - 0.91 \exp(-0.22 t_{i,j}) \right), \quad (13)$$

let $\delta_{i,j} \square N(0, 0.2^2)$, $t_{i,j} = 2.45, 2.55, \dots, 7.75$ and $n = 54$. Next, the function (4) with a known parameter vector is

$$f(t_{i,j}; \boldsymbol{\theta}) = \frac{1000}{t_{i,j}} \exp\left(-1 \left[\log(t_{i,j}) - 6 \right]^2\right), \quad (14)$$

with $\delta_{i,j} \square N(0, 0.5^2)$, $t_{i,j} = 10, 20, \dots, 1000$ and $n = 100$. Finally, the function (5) of values of a parameter vector is

$$f(t_{i,j}; \boldsymbol{\theta}) = 330 - 3t_{i,j} + 180 \exp(-0.07 t_{i,j}) \sin \left[\frac{2\pi t_{i,j}}{24} + 0.31 \right], \quad (15)$$

where $\delta_{i,j} \square N(0, 5^2)$, $t_{i,j} = 0, 1.5, \dots, 78$ and $n = 53$. In the simulation study, for four replicates, $r = 4$, repeated measures are simulated datasets, $y_{1,j}, \dots, y_{4,j}$, at each time point. Examples of simulated datasets under the three nonlinear models are shown in Figure 1.

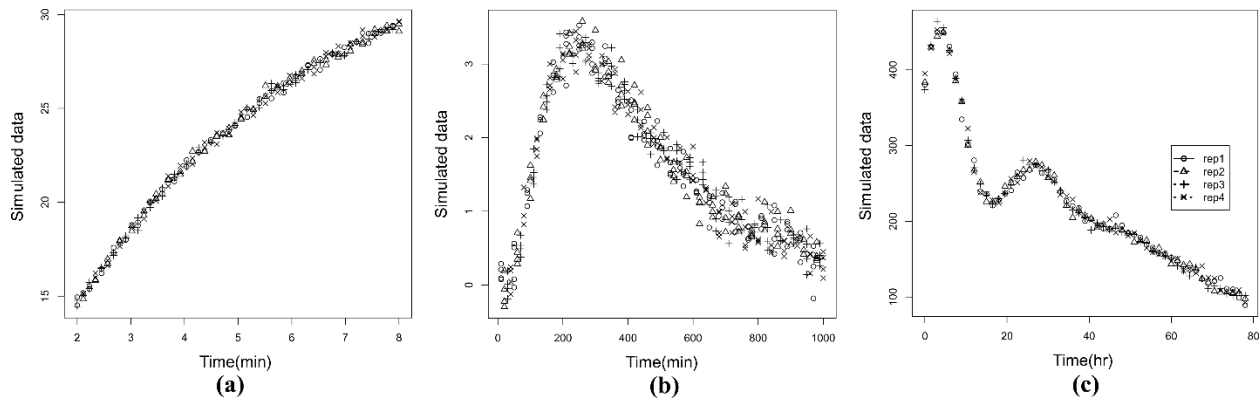


Figure 1 Examples of the simulated datasets for four replicates generated by the following functions:

(a) exponential growth, (b) log-normal and (c) one-sine with the AR(1) errors at $\rho = 0.25$.

At a level of ρ (0, 0.25, 0.75), each simulated dataset is run with a total of 20,000 replicate studies by using the R software (R Core Team, 2013) with the nls function based on a Gauss-Newton algorithm (Ritz & Streibig, 2008; Crawley, 2013). Bias of estimates, coverage probability and accuracy of the variance estimators from the three methods are next used to compare. The percentage bias of the estimator of the parameter is

$$\% \text{ Bias} = 100 \left(\frac{\hat{\theta} - \theta}{\theta} \right),$$

where $\hat{\theta}$ is the mean of $\hat{\theta}_m$ and $\hat{\theta}_m$ is the parameter estimate obtained from the m^{th} simulation run ($m = 1, 2, \dots, M$).

To evaluate the precision of the standard error for parameter estimates, the percentage relative difference between the standard error and the standard deviation for the estimate is

$$\% \text{ Diff} = 100 \left(\frac{\text{SE}(\hat{\theta}) - \text{SD}(\hat{\theta})}{\text{SD}(\hat{\theta})} \right),$$

where $\text{SD}(\hat{\theta}) = \sqrt{\frac{1}{M-1} \sum_{m=1}^M (\hat{\theta}_m - \hat{\theta})^2}$ and $\text{SE}(\hat{\theta}) = \frac{1}{M} \sum_{m=1}^M \text{SE}(\hat{\theta}_m)$, with $\text{SE}(\hat{\theta}_m)$ the estimated standard error of the estimate from the m^{th} simulated dataset. The estimated coverage probability (CP) is the proportion of times that the $100(1 - \alpha)\%$ confidence interval covers the true value of θ , as given by

$$\hat{\theta}_m \pm t_{\frac{\alpha}{2}, v} SE(\hat{\theta}_m),$$

where $t_{\frac{\alpha}{2}, v}$ is the critical value of student t distribution with the significance level α and degrees of freedom v .

In the simulation study, the simulated datasets (12) are analyzed by fitting the nonlinear regression, based on the Gauss-Newton iterative algorithm with initial values setting for both θ in (13)-(15) and the level of ρ , in which this setting is close to the parameter estimate with the total number of successful fits M .

Results

This section shows results of the simulation evaluation by means of the two-stage least squares method and modified two-stage least squares methods in fitting the nonlinear regression models described in the previous section. Evaluations are presented by using percentage bias (%Bias), percentage relative difference (%Diff) and coverage probability (CP) of 95% confidence interval for the parameter estimates for all ρ (0.00, 0.25 and 0.75).

Table 1 Percentage bias, percentage relative difference and coverage probability of the parameter estimates when the fitted model is the exponential growth function

$\hat{\theta}$	ρ	%Bias			%Diff			CP		
		TS	MTS _{CLS}	MTS _{ANOVA}	TS	MTS _{CLS}	MTS _{ANOVA}	TS	MTS _{CLS}	MTS _{ANOVA}
$\hat{\theta}_1$	0.00	0.0206	0.0205	0.0205	-1.3946	-1.3580	0.0000	0.9434	0.9436	0.9465
	0.25	0.0260	0.0262	0.0260	-1.4424	-1.2069	0.7065	0.9465	0.9471	0.9505
	0.75	0.1183	0.1180	0.1177	-4.8547	-2.4043	0.4031	0.9343	0.9398	0.9422
$\hat{\theta}_2$	0.00	0.0064	0.0063	0.0064	-0.8739	-0.8233	0.5252	0.9472	0.9466	0.9500
	0.25	0.0227	0.0226	0.0227	-1.4270	-1.1809	0.6752	0.9454	0.9456	0.9490
	0.75	0.1181	0.1177	0.1173	8.9984	11.7566	15.0356	0.9588	0.9619	0.9644
$\hat{\theta}_3$	0.00	0.0145	0.0128	0.0146	-2.2822	-1.0375	0.3730	0.9413	0.9444	0.9471
	0.25	0.0137	0.0133	0.0126	-3.4324	-0.6554	0.7288	0.9394	0.9452	0.9472
	0.75	0.1180	0.1126	0.1115	-12.8529	-0.0982	-1.3382	0.9067	0.9452	0.9407

It can be seen from Table 1 that %Bias for all three methods applied to fit simulated data generated by the exponential growth function are comparable and they are unbiased for all scenarios. However, %Diff of TS and MTS_{CLS} methods are negative and biased in estimating the standard error of parameter estimates. And also the coverage probability of the TS method is significantly less than 95%, but the one from the MTS_{CLS} method is quite good. On the other hand, %Diff of the MTS_{ANOVA} method is getting to zero, so its coverage probability is quite better and close to the expected value. Except for $\hat{\theta}_3$ the estimated standard error of the MTS_{ANOVA} method is slightly



underestimated at a high correlation, i.e. 0.75, and it makes the percentage coverage probability approximately 94%. Noticeably, since the standard error of $\hat{\theta}_2$ for the three methods are positive and large at $\rho = 0.75$, as shown in %Diff, their coverage probabilities are bigger than 0.95.

For all the methods for estimating parameters of the log-normal model, all the bias of parameter estimates are small, even when the correlation levels increase. Obviously, the bias of the estimated standard errors, as presented by %Diff, obtained from the TS method is much underestimated when the correlation is high, and its coverage probability is around 0.92. While the standard error of the MTS_{CLS} method is good, but it still produces the poorer coverage probability than that of the MTS_{ANOVA} method, as provided in Table 2.

Table 2 Percentage bias, percentage relative difference and coverage probability of the parameter estimates when the fitted model is the log-normal function

$\hat{\theta}$	ρ	%Bias			%Diff			CP		
		TS	MTS_{CLS}	MTS_{ANOVA}	TS	MTS_{CLS}	MTS_{ANOVA}	TS	MTS_{CLS}	MTS_{ANOVA}
\hat{A}	0.00	0.0373	0.0376	0.0373	-0.8332	-0.8433	-0.0002	0.9480	0.9480	0.9499
	0.25	0.1083	0.1083	0.1082	-1.1978	-1.0547	0.2172	0.9454	0.9454	0.9470
	0.75	0.6028	0.6004	0.5998	-6.8943	-4.6454	-2.4833	0.9246	0.9309	0.9353
\hat{B}	0.00	0.1177	0.1176	0.1175	-0.9510	-0.9567	-0.1269	0.9465	0.9468	0.9484
	0.25	0.3218	0.3218	0.3217	-1.2433	-1.1014	0.1418	0.9464	0.9468	0.9494
	0.75	1.4164	1.4127	1.4120	-8.2625	-6.4371	-4.7564	0.9228	0.9284	0.9314
\hat{C}	0.00	0.0018	0.0018	0.0019	-0.9606	-0.9473	-0.1276	0.9473	0.9474	0.9489
	0.25	0.0014	0.0014	0.0014	-0.5547	-0.4115	0.8631	0.9462	0.9463	0.9494
	0.75	0.0637	0.0631	0.0628	-8.8416	-6.7059	-4.6889	0.9247	0.9305	0.9353



Table 3 Percentage bias, percentage relative difference and coverage probability of the parameter estimates when the fitted model is the one-sine function

$\hat{\theta}$	ρ	%Bias			%Diff			CP		
		TS	MTS _{CLS}	MTS _{ANOVA}	TS	MTS _{CLS}	MTS _{ANOVA}	TS	MTS _{CLS}	MTS _{ANOVA}
$\hat{\alpha}$	0.00	0.0028	0.0025	0.0028	-2.6828	-2.7627	-0.0510	0.9430	0.9429	0.9481
	0.25	-0.0026	-0.0026	-0.0026	-3.6150	-3.3524	0.3675	0.9393	0.9402	0.9481
	0.75	0.0068	0.0071	0.0070	-10.9039	-7.7710	-3.3006	0.9080	0.9184	0.9299
$\hat{\beta}$	0.00	0.0033	0.0038	0.0034	-2.5209	-2.5396	0.1220	0.9423	0.9423	0.9482
	0.25	-0.0033	-0.0033	-0.0032	-3.8440	-3.6038	0.0264	0.9387	0.9392	0.9466
	0.75	-0.0090	-0.0083	-0.0079	-11.2689	-8.1938	-3.9476	0.9092	0.9187	0.9296
\hat{a}	0.00	-0.0014	-0.0009	-0.0012	-2.3057	-2.3751	0.0035	0.9418	0.9417	0.9476
	0.25	0.0256	0.0259	0.0261	-3.2248	-3.0078	-0.0838	0.9414	0.9425	0.9494
	0.75	0.0151	0.0126	0.0151	-6.9832	-5.6428	-3.9611	0.9310	0.9349	0.9385
\hat{d}	0.00	-0.0143	-0.0145	-0.0138	-3.3004	-3.3577	-0.9248	0.9408	0.9408	0.9456
	0.25	0.0054	0.0061	0.0062	-3.5905	-3.3531	-0.3109	0.9392	0.9403	0.9469
	0.75	-0.0693	-0.0719	-0.0678	-5.6647	-4.5103	-3.2195	0.9344	0.9380	0.9414
$\hat{\tau}$	0.00	0.0070	0.0072	0.0070	-1.9816	-2.0090	0.1259	0.9439	0.9434	0.9485
	0.25	0.0055	0.0054	0.0053	-2.5437	-2.3479	0.3989	0.9416	0.9422	0.9485
	0.75	0.0175	0.0168	0.0161	-3.8802	-2.4729	-0.6687	0.9371	0.9411	0.9453
$\hat{\phi}$	0.00	0.0377	0.0363	0.0375	-1.8400	-1.9353	-0.3185	0.9478	0.9473	0.9512
	0.25	0.0201	0.0204	0.0184	-0.7825	-0.6427	1.3088	0.9474	0.9478	0.9516
	0.75	0.0114	0.0094	0.0040	11.2228	12.5323	14.2836	0.9692	0.9712	0.9741

When the one-sine model is fitted, %Bias for all parameter estimates obtained by all the methods are similar for all correlation coefficients. Almost values of %Diff are negative, but except for the bias in estimating the standard error from the MTS_{ANOVA} method. As TS and MTS_{CLS} methods provide more biased variance estimates than the MTS_{ANOVA} method, the coverage probabilities of TS and MTS_{CLS} methods are quite poor. The standard error of $\hat{\phi}$ based on the MTS_{ANOVA} method is the best choice at $\rho = 0$ while the MTS_{CLS} method is the best one at $\rho = 0.25$. However, the three methods produce significantly inflated bias in the standard error of $\hat{\phi}$ for $\rho = 0.75$, as shown in %Diff, their coverage probabilities are considerably larger than 0.95, as proved in Table 3.

In summary, the MTS_{ANOVA} and MTS_{CLS} methods are good performances in terms of parameter estimates and standard errors. On the other hand, the TS method gives significantly underestimated standard errors of



parameter estimates and it also provides invalid coverage probabilities. However, the three methods produce overestimated standard errors for $\hat{\theta}_2$ and $\hat{\Phi}$ at a high correlation, which will be discussed in the next section.

Discussion

The results indicate that the methods, the two-stage least squares (TS) method and modified two-stage least squares estimation (MTS_{CLS} and MTS_{ANOVA} methods) for fitting nonlinear regression models, exponential growth, log-normal and one-sine functions, with correlated errors in the case of AR(1) produce small estimates, as shown by %Bias in Tables 1-3. And also, standard error estimates are slightly biased when correlation coefficients are low, which corresponded with Bender & Heinemann (1995), as dictated in Table 2. However, the TS method produces %Diff worst, which leads to invalid coverage probabilities (CP). Moreover, for a high correlation, the expected values of CP from the TS method are more inappropriate than the ones of MTS_{CLS} and MTS_{ANOVA} methods, as given in Tables 1-3. While the estimated standard errors for parameter estimates $\hat{\theta}_2$ of the exponential growth function and $\hat{\Phi}$ for the fitted one-sine model, as respectively, are overestimated, these three methods based on the least squares estimation give wrong conclusions of the coverage probabilities of the parameters, so for further work the author will propose methods exploring more accurate variance estimation by using maximum likelihood procedures.

Conclusions

In order to improve inferential statistics, the modified two-stage least squares methods that explicitly account for the correlation are proposed. The methods are developed by using the correlation estimates based on conditional least squares (the MTS_{CLS} method) and applying pure errors from the one-way ANOVA model (the MTS_{ANOVA} method) to compute the correlation coefficient in the weight matrix. The proposed methods are compared to the two-stage least squares (TS) estimation for fitting nonlinear regression models. Simulation results suggest even when these methods produce unbiased estimators of the parameters, almost all outcomes of the TS method are poorer for both standard errors of parameter estimates and confidence intervals than those of the MTS methods.

Acknowledgements

The author would like to thank Kasetsart University, Chalermphrakiat Sakon Nakhon Province Campus (KUCSC) Province Campus for the financial support, and also extremely thanks to the reviewers for valuable comments and suggestions toward the completion of this paper.



References

- Asikgil, B., & Erar, A. (2009). Modified two-stage least squares method. In *Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis*. (pp.124-128). Lithuania: Vilnius.
- Bates, D.M, & Watts, D.G. (1988). *Nonlinear regression analysis and its applications*. New York: John Wiley & Sons.
- Bender, R., & Heinemann, L. (1995). Fitting nonlinear regression models with correlated errors to individual pharmacodynamics data using SAS software. *Journal of Pharmacokinetics and Biopharmaceutics*, 23(1), 87-100.
- Bender, R. (1995). Calculating summary measures of unimodal response curve by means of nonlinear regression models. *Journal of Theoretical Medicine*, 2(1), 73-80.
- Crawley, M.J. (2013). *The R book*. West Sussex: John Wiley & Sons.
- Gallant, A.R., & Goebel, J.J. (1976). Nonlinear regression with autoregressive errors. *Journal of the American Statistical Association*, 71(356), 961-967.
- Izumo, M., Johnson, C.H., & Yamazaki, S. (2003). Circadian gene expression in mammalian fibroblasts revealed by real-time luminescence reporting: Temperature compensation and damping. *The National Academy of Sciences of the USA*, 100(26), 16089-16094.
- Maier, B., Wendt, S., Vanselow, J.T., Wallach, T., Reischl, S., Oehmke, S., Schlosser, A., & Kramer, A. (2009). A large-scale functional RNAi screen reveals a role for CK2 in the mammalian circadian clock. *Genes & Development*, 23(1), 708-718.
- Park, R.E., & Mitchell, B.M. (1980). Estimating the autocorrelated error model with trended data. *Journal of Econometrics*, 13(2), 185-201.



Pukdee, W., Polsen, O., & Baksh, M.F. (2018) Improved methods for the analysis of circadian rhythms in correlated gene expression data. *Songklanakarin Journal of Science and Technology*, 40(3), 692-700.

Pukdee, W., Polsen, O., & Baksh, M.F. (2020). A Modified Two-Stage Method for Parameter Estimation in Sinusoidal Models of Correlated Gene Expression Profiles. *Thailand Statistician*, 18(1), 77-89.

R Core Team. (2013). *A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved March 3, 2015, from: <http://www.R-project.org/>

Ritz, C., & Streibig, J.C. (2008). *Nonlinear regression with R*. New York: Springer.

Seber, G.A.F., & Wild, C.J. (2003). *Nonlinear regression*. New York: Wiley.