Journal of Multivariate Analysis 108 (2012) 53–66

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Estimation of parameters in the growth curve model via an outer product least squares approach for covariance

Jianhua Hu [a,*], Fuxiang Liu [a,b], S. Ejaz Ahmed [c]

[a] School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, PR China
[b] Science College and Institute of Intelligent Vision and Image Information, China Three Gorges University, Yichang, Hubei 443002, PR China
[c] Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada N9B 3P4

ARTICLE INFO

*Article history:*
Received 6 January 2011
Available online 17 February 2012

*AMS 2000 subject classifications:*
primary 62H12
secondary 62F12
62H10

*Keywords:*
Estimation
Growth curve model
Outer product
Outer product least squares for covariance
COPLS estimator
Two-stage generalized least squares

ABSTRACT

In this paper, we propose a framework of outer product least squares for covariance (COPLS) to directly estimate covariance in the growth curve model based on an analogy, between the outer product of a data vector and covariance of a random vector, and the ordinary least squares technique. The COPLS estimator of covariance has an explicit expression and is shown to have the following properties: (1) following a linear transformation of two independent Wishart distribution for a normal error matrix; (2) having asymptotic normality for a nonnormal error matrix; and (3) having unbiasedness and invariance under a linear transformation group. And, a corresponding two-stage generalized least squares (GLS) estimator for the regression coefficient matrix in the model is obtained and its asymptotic normality is investigated. Simulation studies confirm that the COPLS estimator and the two-stage GLS estimator of the regression coefficient matrix are satisfying competitors with some evident merits to the existing maximum likelihood estimator in finite samples.



## 1. Introduction

Observations that occur in social studies, biological science, economics and medical research are usually measured over multiple time points on a particular characteristic to investigate temporal pattern of change on the characteristic. The growth curve model is a useful tool for statisticians to analyze the observations of repeated measurements. The growth curve model without assumption of a normal distribution is the model in which we observe

$$\mathbf{Y}_{n\times p} = X_{n\times m}\Theta_{m\times q}Z'_{p\times q} + \mathcal{E}_{n\times p}, \qquad E(\mathcal{E}) = \mathbf{0} \quad \text{and} \quad \mathrm{Cov}(\mathcal{E}) = I_n \otimes \Sigma_{p\times p}, \tag{1}$$

where $\mathbf{Y}$ is the observation matrix of the response consisting of $p$ repeated measurements taken on $n$ individuals, $X$ is the treatment design matrix with order $n\times m$, $Z$ is the profile matrix with order $p\times q$, and $\Theta$ is the unknown regression coefficient matrix with order $m \times q$. Assume that observations on individuals are independent, so that the rows of the random error matrix $\mathcal{E}$ are independent and identically distributed (iid) by a general continuous type distribution $\mathcal{F}$ with mean zero and a common covariance matrix $\Sigma$ of order $p$. An interested reader can refer to Kollo and von Rosen [9] or Pan and Fang [15] for the details of the growth curve model.

The growth curve model was initiated by Potthoff and Roy [17] and widely studied by many researchers. For no restriction for the structure of covariance, Potthoff and Roy [17] originally derived a class of weighted estimator only for the regression

* Corresponding author.
*E-mail addresses:* frank.jianhuahu@gmail.com (J. Hu), liufuxiangst@gmail.com (F. Liu), seahmed@uwindsor.ca (S.E. Ahmed).

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.02.007

coefficients. Khatri [8] derived the maximum likelihood estimator (MLE) and showed that the MLE also is a weighted estimator. Grizzle and Allen [3] used the technique of analysis of covariance to obtain an estimator of the regression coefficient matrix and showed that it is identical to the MLE. Rao [18], Reinsel [20], and Lange and Laird [10] considered estimation for the random effects covariance structure of the model. Rao [19] and Lee [11] studied the prediction problem of the model without or with a special covariance structures. Recently, Ohlson and von Rosen [14] presented explicit estimators of parameters for covariance with linear structure based on a decomposition of the whole tensor space. If pursuing (simultaneous) estimation of a linear parametric function $\mathrm{tr}(C\Sigma)$ (or $\mathrm{tr}(C\Sigma) + \mathrm{tr}(D'\Theta)$), the interested readers may refer to a series of works by Xu and Yang [24], Yang and Xu [29], Yang [25–27], and Yang and Jiang [28].

It is well-known that inference on the regression coefficient matrix strongly relies on the preestimated covariance matrix. Generally speaking, any estimator of the regression coefficient matrix is a function of the preestimated covariance matrix. Naturally, the estimator of covariance is very important to estimation of the regression coefficients. With the help of computers, although the maximum likelihood, the restricted maximum likelihood and iterative techniques can be used to obtain the estimators of the parameters of interest, non-iterative estimating methods still are worth being studied, specially for very large data sets or for non-normal errors.

The motivation of this paper is to exploit both the linear structure of mean in the growth curve model and the analogy between the outer product of data vectors and covariance and formulate a framework or an outer product least squares approach to directly do least squares to the unknown covariance, exactly as the ordinary least squares method that directly does least squares to regression coefficients in the Gauss–Markov models. The resulting estimator by the outer product least squares approach is named an outer product least squares estimator for covariance (COPLSE). After the COPLSE of the unknown covariance matrix has been obtained, the corresponding two-stage generalized least squares (GLS) estimator of the regression coefficient matrix will be derived and its properties in the large sample will be discussed.

The outer product least squares approach is formulated by combining the following basic ideas or techniques: (a) the analogy between the outer product of data vectors and covariance, (b) the complete set of error contracts, (c) an auxiliary least squares model, and (d) the ordinary least squares technique. The complete set of error contracts was used by Patterson and Thompson [16] to develop the restricted maximum likelihood technique which was modified by Harville [4]. The auxiliary least squares model for the growth curve model was used by Yang and Xu [29] in which it was said to be an induced model. In the literature, the main focus was on estimation of the trace of the linear transformation of covariance.

The outer product least squares approach seems to be useful and effective for estimating unknown parameters in covariance for a class of linear models with independent and identically distributed errors. The class of linear models includes many famous statistical models. The growth curve model is one member in the class. Two working papers are already to address estimation for a generalized GMANOVA model and the extended growth curve model via the proposed outer product least squares approach.

The organization of the paper is as follows. An outer product least squares approach is formulated or a framework for directly doing least squares to covariance in the growth curve model is established in Section 2. An outer product least squares estimator for covariance (COPLSE) in the growth curve model is represented. For normal errors, the exact distribution of the COPLSE is obtained in Section 3. The strong consistency and asymptotic normality without assumption of normal errors are studied in Section 4. A corresponding two-stage GLS estimator to the regression coefficient matrix is derived and its consistency and asymptotic normality are investigated in Section 5. Simulation studies are provided in Section 6 to demonstrate that the COPLSE and the resulting two-stage GLS estimator are alternative competitors with some evident merits, for example, more efficiency in the sense of bias or the mean squared error, to the existing maximum likelihood estimators. Brief concluding remarks are presented in Section 7.

## 2. Estimation of covariance based on an outer product least squares approach

In this paper, $\mathbb{M}_{n\times n}$ denotes the set of all $n \times n$ matrices over the real set $\mathbb{R}$ with the trace inner product $\langle\cdot,\cdot\rangle$. $\|\cdot\|$ denotes the trace norm on the set $\mathbb{M}_{n\times n}$. $\mathbb{N}_p$ denotes the set of all nonnegative definite matrices of order $p$. $A^-$ denotes the generalized inverse of a matrix $A$. $P_T = T(T'T)^-T'$ denotes the orthogonal projection matrix onto the column space $\mathscr{C}(T)$ of a matrix $T$ and $M_T = I - T(T'T)^-T'$ denotes the orthogonal projection matrix onto the orthogonal complement $\mathscr{C}(T)^\perp$ of $\mathscr{C}(T)$.

For the Gauss–Markov model $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\mathrm{Cov}(\boldsymbol{\epsilon}) = \sigma^2 I$, the ordinary least squares method is to find a point $\widehat{\boldsymbol{\beta}}(\mathbf{y})$ in the $m$-dimensional real space $\mathbb{R}^m$ such that

$$\widehat{\boldsymbol{\beta}}(\mathbf{y}) = \underset{\boldsymbol{\beta}\in\mathbb{R}^m}{\operatorname{argmin}} \|\mathbf{y} - X\boldsymbol{\beta}\|^2. \tag{2}$$

Equivalently, the ordinary least squares method takes the perpendicular projection $P_X\mathbf{y}$ of $\mathbf{y}$ as the least squares estimator of the expected value $E(\mathbf{y})$. As a by-product of the least squares problem (2), the following statistic

$$\widehat{\sigma}^2_{ols}(\mathbf{y}) = \frac{1}{n-r}(\mathbf{y} - X\widehat{\boldsymbol{\beta}}(\mathbf{y}))'(\mathbf{y} - X\widehat{\boldsymbol{\beta}}(\mathbf{y})) = \frac{1}{n-r}\mathbf{y}'(I - P_X)\mathbf{y} \tag{3}$$

is viewed as the ordinary least squares estimator of $\sigma^2$, where $r$ is the rank of the design matrix $X$. The major drawback of this indirect method (as a by-product) for estimating variance $\sigma^2$ is that the residuals cannot be explicitly expressed by the design matrix and observation $\mathbf{y}$ in many linear statistical models, e.g., the growth curve model. To overcome the drawback, an outer product least squares approach directly to covariance will be deliberately formulated in this paper.

The *outer product* over the $np$-dimensional real space $\mathbb{R}^{np}$ is defined as

$$\mathbf{a}\Box\mathbf{a} = \mathbf{a}\mathbf{a}' = \left(a_i a_j\right)_{np\times np}$$

for any $\mathbf{a} = (a_1, \ldots, a_{np})' \in \mathbb{R}^{np}$. The covariance of a $np$-dimensional random vectors $\mathbf{y}$ is the expectation of outer product of $\mathbf{y}$, namely,

$$\mathrm{Cov}(\mathbf{y}) = E(\mathbf{y}\Box\mathbf{y}) = \left(\mathrm{Cov}(y_i, y_j)\right)_{np\times np}.$$

Exactly as using the moments of a random sample to estimate the moments of its population as well exactly as using the quantile of a random sample to estimate the quantile of its population, it is a natural and reasonable thing for us to use the outer products of a random sample to estimate covariance of its population. The outer product of the data vector in a random sample should contain the information of the behavior of the unknown parameters in variance or covariance of the population. Therefore, the framework to be developed is motivated from the above mentioned residuals $M_X\mathbf{y}$, noticed by Patterson and Thompson [16], and the analogy between outer product and covariance.

### 2.1. Least squares problem to covariance

Under the vec operator, the growth curve model without assumption of normality can be written as

$$\mathrm{vec}(\mathbf{Y}) = T\boldsymbol{\beta} + \boldsymbol{\zeta}, \qquad E(\boldsymbol{\zeta}) = \mathbf{0} \quad \text{and} \quad \mathrm{Cov}(\boldsymbol{\zeta}) = I_n \otimes \Sigma, \tag{4}$$

where $\boldsymbol{\beta} = \mathrm{vec}(\Theta)$, $T = X \otimes Z$ and $\boldsymbol{\zeta} = \mathrm{vec}(\mathcal{E})$. Here the *vec operator* transforms a matrix into a vector by *stacking the rows of* $\mathbf{Y}$ *one under another.*

An error contrast is defined as any linear combination of the response vector $\mathrm{vec}(\mathbf{Y})$, which has zero expectation. Error contrasts form an $np - r$ dimension linear space. A set of $np - r$ linearly independent error contrasts is said to be a complete set of error contrasts. The concept about the complete set of error contracts was first proposed and used by Patterson and Thompson [16] to develop the restricted maximum likelihood technique which was modified by Harville [4]. The columns of the orthogonal projection $M_T$ of the orthogonal complement $\mathscr{C}(T)^{\perp}$, forms a complete set of error contrasts.

A framework, which we shall establish, for directly doing least squares to the unknown covariance associates with considerations from the following four aspects.

The first step is to use the complete set of error contracts constructed by $M_T$ for the response $\mathrm{vec}(\mathbf{Y})$ to do the outer product. In other words, only the outer product of $M_T\mathrm{vec}(\mathbf{Y})$, or the residuals $\mathrm{vec}(\mathbf{Y}) - P_T\mathrm{vec}(\mathbf{Y})$, is considered.

Based on the opinion that the covariance matrix of two random vectors can be viewed as a special outer product of the two random vectors, the second step is to use the outer product of the residual vector to estimate unknown covariance of random errors. To be more precise, we shall use the outer product $M_T\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})'M_T$ to estimate covariance as same as using $\mathrm{vec}(\mathbf{Y})$ to estimate $\mathrm{vec}(\Theta)$ in the ordinary least squares approach.

The third step is to construct an auxiliary linear model. Let $Q(\mathbf{Y}) = M_T\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})'M_T$ (hereafter $M$ replaces $M_T$ for brevity). Then $Q(\mathbf{Y})$ is the outer product of the orthogonal projection vector of the random vector $\mathrm{vec}(\mathbf{Y})$ onto the error space $\mathscr{C}(T)^{\perp}$. All $Q(\mathbf{Y})$ form a subset of $\mathbb{M}_{np\times np}$, which is spanned by the columns of $M_T$. In essence, $Q(\mathbf{Y})$ is a random matrix with the mean structure of the form

$$\boldsymbol{\mu} = E(Q(\mathbf{Y})) = M(I \otimes \Sigma)M. \tag{5}$$

Naturally, *an auxiliary least squares model, called an outer product least squares model*, is defined as

$$Q(\mathbf{Y}) = M(I \otimes \Sigma)M + \boldsymbol{\xi} \tag{6}$$

where $E(\boldsymbol{\xi}) = \mathbf{0}$ and $\mathrm{Cov}(\boldsymbol{\xi}) = (M \otimes M)E\left((\mathcal{E} \otimes \mathcal{E})(\mathcal{E}' \otimes \mathcal{E}')\right)(M \otimes M)$.

The fourth step is to define the trace distance of the difference of the matrix $Q(\mathbf{Y})$ and its expected values $M(I \otimes \Sigma)M$ as

$$D(\Sigma, \mathbf{Y}) = \|Q(\mathbf{Y}) - M(I \otimes \Sigma)M\|^2.$$

*A least squares problem to covariance for the growth curve model* (4) is to find a nonnegative definite matrix $\widehat{\Sigma}_{ls}(\mathbf{Y})$ such that the trace distance function $D(\Sigma, \mathbf{Y})$ is minimized at $\widehat{\Sigma}_{ls}(\mathbf{Y})$, namely,

$$\widehat{\Sigma}_{ls}(\mathbf{Y}) = \underset{\Sigma\in\mathbb{N}_p}{\mathrm{argmin}}\, D(\Sigma, \mathbf{Y}). \tag{7}$$

In other words, the least squares problem to covariance for the model (1) or (4) is an ordinary least squares problem on the set $\mathbb{N}_p$ for the outer product least squares model (6). A least squares solution $\widehat{\Sigma}_{ls}(\mathbf{Y})$ in the least squares problem (7) is said to be *a least squares estimator* to covariance $\Sigma$ if the $\widehat{\Sigma}_{ls}(\mathbf{Y})$ is unique.

### 2.2. Out product least squares problem and out product least squares solutions

Let

$$\mathscr{V} = \{M\mathrm{vec}(\mathbf{Y})\Box M\mathrm{vec}(\mathbf{Y})\colon \mathrm{vec}(\mathbf{Y}) \in \mathbb{R}^{np}\}, \qquad \mathscr{H}_{nnd} = \{M(I_n \otimes \Sigma)M\colon \Sigma \in \mathbb{N}_p\}$$

and

$$\mathscr{H} = \{M(I_n \otimes V)M\colon V \in \mathbb{M}_{p\times p}\}.$$

Obviously, the trace inter product space $(\mathbb{M}_{np\times np}, \langle\cdot,\cdot\rangle)$ is an Euclidean space. $\mathscr{H}$ is a subspace of $\mathbb{M}_{np\times np}$. And $\mathscr{H}_{nnd}$ is a convex cone (set) of $\mathscr{H}$, also a convex cone of $\mathbb{M}_{np\times np}$.

Since $\mathscr{H}_{nnd}$ is a convex cone of $\mathbb{M}_{np\times np}$, the optimization problem (7) is a convex optimization problem. Seeking a method for solving (7) from the convex optimization theory will not be a job of this paper.

Alternatively, to find least squares solutions in the least squares problem (7), we expand the domain $\mathbb{N}_p$ to the space $\mathbb{M}_{p\times p}$ for the least squares problem (7). This yields the following *out product least squares problem for covariance* (8): finding a matrix in $\mathbb{M}_{p\times p}$, written as $\widehat{\Sigma}_{copls}(\mathbf{Y})$, such that

$$\widehat{\Sigma}_{copls}(\mathbf{Y}) = \underset{V\in\mathbb{M}_{p\times p}}{\mathrm{argmin}}\, D(V, \mathbf{Y}). \tag{8}$$

Here, $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is said to *an out product least squares solution of covariance* $\Sigma$. Moreover, $\widehat{\Sigma}_{copls}(\mathbf{Y})$ in the problem (8) is said to be *an outer product least squares estimator for covariance* $\Sigma$ (written as COPLSE or COPLS estimator) if the $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is unique. It will be seen that the outer product least squares solution $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is unique under a very mild condition.

Note that $\mathbb{M}_{p\times p}$ or $\mathscr{H}$ is a subspace while $\mathbb{N}_p$ or $\mathscr{H}_{nnd}$ is a convex cone.

If an outer product least squares estimator for covariance $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is an element in the set $\mathbb{N}_p$, then $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is a least squares estimator $\widehat{\Sigma}_{ls}(\mathbf{Y})$ of the least squares problem (7). The problem (7) has been solved.

Based on the above discussion, a procedure for the framework doing least squares estimation to covariance is designed below.

(1) To find an outer product least squares solution $\widehat{\Sigma}_{copls}(\mathbf{Y})$ for the optimization problem (8), which is an outer product least squares estimator under a mild condition, see Theorem 2.1 in the next subsection.
(2) To find a least square estimator $\widehat{\Sigma}_{ls}(\mathbf{Y})$, namely, an outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ with nonnegative definiteness.

### 2.3. The normal equations for outer product least squares solutions

If $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is an outer product least squares solution to the outer product least squares problem (8), it follows from projection theory that, for any $V \in \mathbb{M}_{p\times p}$, $M\,\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})'M - M(I\otimes\widehat{V}(\mathbf{Y}))M$ and $M(I\otimes V)M$ are trace orthogonal, namely, $\langle M\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})'M - M(I\otimes\widehat{\Sigma}_{copls}(\mathbf{Y}))M,\ M(I\otimes V)M\rangle = 0$ for any $V \in \mathbb{M}_{p\times p}$. This yields the following equations

$$\mathrm{tr}\left(\left[M\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})'M - M(I\otimes\widehat{\Sigma}_{copls}(\mathbf{Y}))M\right](I\otimes V)\right) = 0 \quad \text{for any } V \in \mathbb{M}_{p\times p}. \tag{9}$$

Note that

$$\mathrm{tr}\left(\left[M\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})'M - M(I\otimes\widehat{\Sigma}_{copls}(\mathbf{Y}))M\right](I\otimes V)\right) = \mathrm{tr}\left(\sum_{i=1}^{n} M_i'(\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})' - I\otimes\widehat{\Sigma}_{copls}(\mathbf{Y}))M_iV\right),$$

where $M = (M_1,\ldots,M_n)$ with $np\times p$ matrix $M_i$, $i = 1,\ldots,n$. The arbitrariness of $V$ in the space $\mathbb{M}_{p\times p}$ implies

$$\sum_{i=1}^{n} M_i'\left(\mathrm{vec}(\mathbf{Y})\mathrm{vec}(\mathbf{Y})' - I\otimes\widehat{\Sigma}_{copls}(\mathbf{Y})\right)M_i = 0. \tag{10}$$

Further blocking matrices causes

$$\sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}'\widehat{\Sigma}_{copls}(\mathbf{Y})M_{ji} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} M_{ji}\mathbf{Y}_j\right)\left(\sum_{l=1}^{n} M_{li}\mathbf{Y}_l\right)' \tag{11}$$

where $\mathbf{Y} = (\mathbf{Y}_1,\ldots,\mathbf{Y}_n)'$ and $M = (M_{ij})$ with $p\times p$ matrix $M_{ij}$, $i,j = 1,\ldots,n$.

Any of the Eqs. (9)–(11) is said to be *the normal equations* for an outer product least squares problem (8). Let

$$H = \sum_{i,j=1}^{n} M_{ij}\otimes M_{ij} \quad \text{and} \quad C(\mathbf{Y}) = \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n} M_{ik}\otimes M_{jk}\right)\mathrm{vec}(\mathbf{Y}_i\mathbf{Y}_j'). \tag{12}$$

Then the normal equations (11) can be rewritten as

$$H\text{vec}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})\right) = C(\mathbf{Y}). \tag{13}$$

The original motivation for the normal equations was to solve the outer product least squares problem (8). Taking the following result, now we complete that discussion about relationship between the normal equations and the outer product least squares problem (8).

**Theorem 2.1.** *A matrix $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is an outer product least squares solution if and only if the matrix $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is a solution to the normal equations. Moreover, $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is unique under $r(X) < n$ for a given observation $\mathbf{Y}$.*

**Proof.** It suffices to show sufficiency. Assume that $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is a solution to the normal equations. Then, from (9), we have the following inequality

$$D(V, \mathbf{Y}) = D(\widehat{\Sigma}_{copls}(\mathbf{Y}), \mathbf{Y}) + \|M(I_n \otimes (\widehat{\Sigma}_{copls}(\mathbf{Y}) - V))M\|^2 \geq D(\widehat{\Sigma}_{copls}(\mathbf{Y}), \mathbf{Y}),$$

for any $V \in \mathbb{M}_{p\times p}$. So $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is an outer product least squares solution to the outer product least squares problem (8).

Assume that $\widehat{V}_1(\mathbf{Y})$ and $\widehat{V}_2(\mathbf{Y})$ both are solutions of the normal equations (9). Let $V(\mathbf{Y}) = \widehat{V}_1(\mathbf{Y}) - \widehat{V}_2(\mathbf{Y})$, then, with $M = M_X \otimes I + P_X \otimes M_Z$, we have

$$\text{tr}\left((M_X \otimes I + P_X \otimes M_Z)(I \otimes V(\mathbf{Y}))(M_X \otimes I + P_X \otimes M_Z)(I \otimes S)\right) = 0 \tag{14}$$

for any $S \in \mathbb{M}_{p\times p}$. After routine tensor product operations, (14) is equivalent to

$$(n - r(X))V(\mathbf{Y}) + r(X)M_Z V(\mathbf{Y})M_Z = \mathbf{0}. \tag{15}$$

Multiplying both sides of (15) by $M_Z$ yields $nM_Z V(\mathbf{Y})M_Z = \mathbf{0}$. Due to $r(X) < n$, it follows from (15) that $V(\mathbf{Y}) = \mathbf{0}$, namely, $\widehat{V}_1(\mathbf{Y}) = \widehat{V}_2(\mathbf{Y})$. Therefore, we complete the proof of theorem. □

*2.4. The out product least squares estimator of covariance*

We shall use the normal equations (13) to obtain the outer product least squares estimator $\widehat{\Sigma}_{opls}(\mathbf{Y})$ for covariance $\Sigma$ in the model (1). The result is represented in the following theorem.

**Theorem 2.2.** *If the rank of the treatment design matrix $X$ is less than the number of observations, i.e. $r(X) < n$, the outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ for the growth curve model (1) without assumption of normality is given by*

$$\widehat{\Sigma}_{copls}(\mathbf{Y}) = \frac{1}{n-r}\mathbf{Y}'M_X\mathbf{Y} + \frac{1}{n}M_Z\mathbf{Y}'\mathbf{Y}M_Z - \frac{1}{n-r}M_Z\mathbf{Y}'M_X\mathbf{Y}M_Z, \tag{16}$$

*where $\mathbf{Y}$ is an $n \times p$ matrix of observations.*

**Proof.** Let $M_T = \left(M_{ij}\right)$, $1 \leq i, j \leq n$, $M_X = \left(m_{ij}^X\right)_{n\times n}$ and $P_X = \left(p_{ij}^X\right)_{n\times n}$, then, due to $M_T = M_X \otimes I + P_X \otimes M_Z$, $M_{ij}$ can be expressed as

$$M_{ij} = M_{ji} = m_{ij}^X I + p_{ij}^X M_Z. \tag{17}$$

Thus using (17) to replace $M_{ij}$ in (12) yields

$$\begin{aligned} H &= \sum_{i,j=1}^{n}\left(m_{ij}^X I + p_{ij}^X M_Z\right) \otimes \left(m_{ji}^X I + p_{ji}^X M_Z\right) \\ &= \text{tr}(M_X^2)I + \text{tr}(M_X P_X)(I \otimes M_Z + M_Z \otimes I) + \text{tr}(P_X^2)M_Z \otimes M_Z \\ &= (n-r)I + rM_Z \otimes M_Z. \end{aligned}$$

With $r < n$, $H$ is nonsingular and $H^{-1} = \frac{1}{n-r}\left(I - \frac{r}{n}M_Z \otimes M_Z\right)$. Also, due to $M = M_X \otimes P_Z + I \otimes M_Z$, we have

$$\begin{aligned} C(\mathbf{Y}) &= \text{vec}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left(m_{ij}^X P_Z + \delta_{ij}M_Z\right)\mathbf{Y}_j\sum_{l=1}^{n}\mathbf{Y}_l'\left(m_{il}^X P_Z + \delta_{il}M_Z\right)'\right) \\ &= \text{vec}(P_Z\mathbf{Y}'M_X\mathbf{Y}P_Z + P_Z\mathbf{Y}'M_X\mathbf{Y}M_Z + M_Z\mathbf{Y}'M_X\mathbf{Y}P_Z + M_Z\mathbf{Y}'\mathbf{Y}M_Z) \\ &= \text{vec}(\mathbf{Y}'M_X\mathbf{Y} - M_Z\mathbf{Y}'M_X\mathbf{Y}M_Z + M_Z\mathbf{Y}'\mathbf{Y}M_Z), \end{aligned}$$

where $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n]'$, $\delta_{ij} = 1$ for $i = j$ and 0 for any distinct $i, j$. The solution of the normal equations (13) is uniquely determined by

$$\text{vec}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})\right) = \frac{1}{n-r}\left(I - \frac{r}{n}M_Z \otimes M_Z\right)C(\mathbf{Y}).$$

A simple computation causes (16). So, the proof is complete. □

When $Z$ is a nonsingular square matrix, $P_Z = I$ and $M_Z = \mathbf{0}$. The expression (16) reduces to the expression

$$\widehat{\Sigma}_{copls}(\mathbf{Y}) = \frac{1}{n-r}\mathbf{Y}'M_X\mathbf{Y} \equiv \widehat{\Sigma}_{ols}^{multi}(\mathbf{Y}) \tag{18}$$

the least squares estimator of covariance for the multivariate linear model, see Chapter 19 of Arnold [1].

Regarding the outer product least squares estimator $\hat{\Sigma}(\mathbf{Y})$ given in (16), it was first derived by Xu and Yang [24] via an estimating class, in which it was said to be LSE in a sense. The literature focused on providing necessary and sufficient conditions for a parameter function $\mathrm{tr}(C\hat{\Sigma}_{copls}(\mathbf{Y}))$ to be UMVIQLUE of its expectation $\mathrm{tr}(C\Sigma)$, a result similar to multivariate Hsu's Theorem [7]. Along the topic with this objective, Yang and some scholars extended Xu and Yang's results, e.g. see Yang [25]. The recent research came from Wu et al. [23], who extended Yang's results from the growth curve model to a generalized growth curve model.

Let $\mathscr{G}$ be a group of transformations defined by

$$\mathscr{G} = \{g_{\boldsymbol{\mu}}(\mathbf{Y}): g_{\boldsymbol{\mu}}(\mathbf{Y}-\boldsymbol{\mu}) = g_{\mathbf{0}}(\mathbf{Y}),\ \text{where } \boldsymbol{\mu} = X\Theta Z'\}.$$

Then the invariance of $\widehat{\Sigma}_{copls}(\mathbf{Y})$ on $\mathscr{G}$ and its unbiasedness are summarized in the following proposition.

**Proposition 2.1.** *The outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ given in (16) is unbiased and invariant under the group $\mathscr{G}$ of transformations, in particular, $\widehat{\Sigma}_{copls}(\mathbf{Y}) = \widehat{\Sigma}_{copls}(\mathcal{E})$.*

**Proof.** Recall that $E(\mathbf{Y}'A\mathbf{Y}) = (X\Theta Z')'A(X\Theta Z') + \mathrm{tr}(A)\Sigma$ for any symmetric $A$. Taking expectation on both sides of (16) yields

$$E\left(\widehat{\Sigma}_{copls}(\mathbf{Y})\right) = \frac{1}{n-r}E(\mathbf{Y}'M_X\mathbf{Y}) + \frac{1}{n}M_Z E(\mathbf{Y}'\mathbf{Y})M_Z - \frac{1}{n-r}M_Z E(\mathbf{Y}'M_X\mathbf{Y})M_Z.$$

A simple computation shows that $E\left(\widehat{\Sigma}_{copls}(\mathbf{Y})\right) = \Sigma$ for all $\Sigma \in \mathbb{N}_p$.

The normal equation (9) is invariant under the group $\mathscr{G}$ of transformations. So the outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is invariant under the group $\mathscr{G}$. In particular, $\widehat{\Sigma}_{copls}(\mathbf{Y}) = \widehat{\Sigma}_{copls}(\mathcal{E})$. The proof is complete. □

The above proof is straightforward. Proposition 2.1 can also be derived from the unbiased and invariant of $\mathrm{tr}(C\hat{\Sigma}_{copls}(\mathbf{Y}))$ due to arbitrariness of the matrix $C$, see Yang [25].

## 3. Distribution of the outer product least squares estimator under assumption of normality

From the proposed framework, we have seen that the outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ for the model (1) has invariance and unbiasedness for the random errors $\mathcal{E}$ with a continuous distribution. If a normal distribution imposes on the random errors $\mathcal{E}$, can we get the exact distribution of the estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$? The answer is yes. The following theorem provides its exact distribution of $\widehat{\Sigma}_{copls}(\mathbf{Y})$.

**Theorem 3.1.** *Suppose that the error matrix $\mathcal{E} \sim N_{np}(\mathbf{0}, I_n \otimes \Sigma)$. The outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ of covariance $\Sigma$ in the model (1) has the same distribution as the following random matrix*

$$\frac{1}{n-r}\mathbf{R}_1 + \frac{1}{n}M_Z\mathbf{R}_0M_Z - \frac{r}{n(n-1)}M_Z\mathbf{R}_1M_Z \tag{19}$$

*with $\mathbf{R}_0 \sim W_p^0(r, \Sigma)$ and $\mathbf{R}_1 \sim W_p^1(n-r, \Sigma)$, where $W_p^0(r, \Sigma)$ and $W_p^1(n-r, \Sigma)$ are two independent Wishart distributions.*

**Proof.** Since $P_X + M_X = I$, there exists an orthogonal matrix $U$ such that $U'P_XU = \mathrm{diag}(I_r, \mathbf{0})$ and $U'M_XU = \mathrm{diag}(\mathbf{0}, I_{n-r})$. Let $\boldsymbol{\eta} = (\boldsymbol{\eta}_1', \ldots, \boldsymbol{\eta}_n')' = U'\mathcal{E}$. Obviously, $\boldsymbol{\eta} \sim N_{np}(\mathbf{0}, I \otimes \Sigma)$. Then

$$\mathbf{R}_0 = \mathcal{E}'P_X\mathcal{E} = \boldsymbol{\eta}'\mathrm{diag}(\mathbf{0}, I_r)\boldsymbol{\eta}, \qquad \mathbf{R}_1 = \mathcal{E}'M_X\mathcal{E} = \boldsymbol{\eta}'\mathrm{diag}(\mathbf{0}, I_{n-r})\boldsymbol{\eta}. \tag{20}$$

From the Eq. (16), we have the following matrix decompositions

$$\begin{aligned}\widehat{\Sigma}_{copls}(\mathbf{Y}) &= \widehat{\Sigma}_{copls}(\mathcal{E}) = (n-r)^{-1}\mathcal{E}'M_X\mathcal{E} + n^{-1}M_Z\mathcal{E}'\mathcal{E}M_Z - (n-r)^{-1}M_Z\mathcal{E}'M_X\mathcal{E}M_Z \\ &= (n-r)^{-1}\mathbf{R}_1 + n^{-1}M_Z(\mathbf{R}_0 + \mathbf{R}_1)M_Z - (n-1)^{-1}M_Z\mathbf{R}_1M_Z \\ &= (n-r)^{-1}\mathbf{R}_1 + n^{-1}M_Z\mathbf{R}_0M_Z + (n^{-1} - (n-1)^{-1})M_Z\mathbf{R}_1M_Z.\end{aligned}$$

By Theorem 3.2 of Hu [5], the random matrices $\mathbf{R}_0$ and $\mathbf{R}_1$ follow independent Wishart distributions $W_p^0(r, \Sigma)$ and $W_p^1(n-r, \Sigma)$, respectively. Hence, $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is equal to the random matrix $\frac{1}{n-r}\mathbf{R}_1 + \frac{1}{n}M_Z\mathbf{R}_0M_Z - \frac{r}{n(n-1)}M_Z\mathbf{R}_1M_Z$ in distribution, and the proof is complete. □

When $Z$ is nonsingular, the growth curve model reduces to the multivariate linear model. By (19) in Theorem 3.1, the outer product least squares estimator of covariance, which is nothing but the least squares estimator, follows a Wishart distribution $W_p(n-r, (n-r)^{-1}\Sigma)$, see Theorem 19.1 in [1]. For $p = 1$, $(n-r)\widehat{\sigma}^2(\mathbf{Y})$ follows a chi-squared distribution with degrees of freedom $n - r$ which is the famous result in the standard statistical inference textbooks.

When $m \geq p$, Wishart distribution $W_p(m, \Sigma)$ has a density function, e.g., see Chapters 3 and 10 of Muirhead [13], Chapter 8 of Eaton [2]. When $m < p$, Wishart distribution $W_p(m, \Sigma)$ is singular. Srivastava [21] investigates the probability density function of a singular Wishart distribution.

## 4. Asymptotic properties of the outer product least squares estimator

In this section, we shall investigate asymptotic properties of the outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$.

**Theorem 4.1.** *The outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ given by (16) is strong consistent to covariance $\Sigma$.*

**Proof.** There exist an orthogonal matrix $Q$ of order $p$ and an orthogonal matrix $U$ of order $n$ such that $Q'P_ZQ = \begin{pmatrix} I_{r_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, $Q'M_ZQ = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{p-r_1} \end{pmatrix}$, and $U'M_XU = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{n-r} \end{pmatrix}$ with $r_1 = r(Z)$ and $r = r(X)$.

Let $\mathbf{W} = U'\mathcal{E}Q$, then $\mathrm{Cov}(\mathbf{W}) = I \otimes \Sigma_1$ where $\Sigma_1 = Q\Sigma Q'$ is positive definite. Partitioning $\mathbf{W} = (\mathbf{W}_1, \mathbf{W}_2)$, where $\mathbf{W}_1$ is $n \times r_1$ and $\mathbf{W}_2$ is $n \times (p - r_1)$, and $\Sigma_1 = \begin{pmatrix} \Sigma_1^{11} & \Sigma_1^{12} \\ \Sigma_1^{21} & \Sigma_1^{22} \end{pmatrix}$, where $\Sigma_1^{11}$ is $r_1 \times r_1$ and $\Sigma_1^{22}$ is $(p - r_1) \times (p - r_1)$, we have $\mathrm{Cov}(\mathbf{W}_1) = I \otimes \Sigma_1^{11} > \mathbf{0}$ and $\mathrm{Cov}(\mathbf{W}_2) = I \otimes \Sigma_1^{22} > \mathbf{0}$.

Furthermore, partitioning $\mathbf{W}$ into $\begin{pmatrix} \mathbf{w}'_{11} & \mathbf{w}'_{12} \\ \vdots & \vdots \\ \mathbf{w}'_{n1} & \mathbf{w}'_{n2}, \end{pmatrix}$, where $\mathbf{w}_{11}, \ldots, \mathbf{w}_{n1}$ are $r_1$-dimensional independent and iid random vectors and $\mathbf{w}_{12}, \ldots, \mathbf{w}_{n2}$ are $n-r_1$-dimensional iid random vectors. From the Eq. (16), the $Q'\widehat{\Sigma}_{copls}(\mathbf{Y})Q$ can be decomposed as

$$Q'\widehat{\Sigma}_{copls}(\mathbf{Y})Q = Q'\widehat{\Sigma}_{copls}(\mathcal{E})Q = \frac{1}{n-r}\mathbf{W}'\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{n-r} \end{pmatrix}\mathbf{W} + \frac{1}{n}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{p-r_1} \end{pmatrix}\mathbf{W}'\mathbf{W}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{p-r_1} \end{pmatrix}$$
$$- \frac{1}{n-r}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{p-r_1} \end{pmatrix}\mathbf{W}'\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{n-r} \end{pmatrix}\mathbf{W}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{p-r_1} \end{pmatrix}.$$

Simple matrix operations yield

$$Q'\widehat{\Sigma}_{copls}(\mathbf{Y})Q = \begin{pmatrix} \dfrac{1}{n-r}\displaystyle\sum_{i=r+1}^{n}\mathbf{w}_{i1}\mathbf{w}'_{i1} & \dfrac{1}{n-r}\displaystyle\sum_{i=1}^{n-r}\mathbf{w}_{i1}\mathbf{w}'_{i2} \\ \dfrac{1}{n-r}\displaystyle\sum_{i=1}^{n-r}\mathbf{w}_{i1}\mathbf{w}'_{i2} & \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\mathbf{w}_{i2}\mathbf{w}'_{i2} \end{pmatrix}.$$

By the strong law of large number, $\frac{1}{n-r}\sum_{i=1}^{n-r}\mathbf{w}_{i1}\mathbf{w}'_{i1}$ converges to $\Sigma_1^{11}$ with probability 1. Similarly, with probability 1, $\frac{1}{n-r}\sum_{i=1}^{n-r}\mathbf{w}_{i1}\mathbf{w}'_{i2}$ converges to $\Sigma_1^{12}$, $\frac{1}{n-r}\sum_{i=1}^{n-r}\mathbf{w}_{i2}\mathbf{w}'_{i1}$ converges to $\Sigma_1^{21}$ and $\frac{1}{n}\sum_{i=1}^{n}\mathbf{w}_{i2}\mathbf{w}'_{i2}$ converges to $\Sigma_1^{22}$. Thus $Q'\widehat{\Sigma}_{opls}(\mathbf{Y})Q$ converges to $\Sigma_1$ with probability 1. It follows that $\widehat{\Sigma}_{copls}(\mathbf{Y})$ converges to covariance $\Sigma$ with probability 1. Hence, the proof is complete. □

Generally speaking, the $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is not nonnegative definite. However, Theorem 4.1 tells us that the $\widehat{\Sigma}_{copls}(\mathbf{Y})$ given by (16) is asymptotically positive definite. When the sample size is sufficiently large, the $\widehat{\Sigma}_{copls}(\mathbf{Y}) > \mathbf{0}$. For finite samples, our simulation studies show that $\widehat{\Sigma}_{copls}(\mathbf{Y})$ seems to be positive definite only if $n - (r(X) + p)$ keeps an appropriately small integer, which can be easily satisfied in the repeated measurement experiments over multiple time points.

The problem (7) has been solved or the least squares estimator of covariance has been obtained in a sense of asymptotically positive definite and the performance of finite sample simulation studies.

Next, we shall investigate asymptotic normality of the statistics $\widehat{\Sigma}_{opls}(\mathbf{Y})$. The fourth-order moment of the errors will be needed in the following discussion.

**Assumption 1.** $E(\boldsymbol{\varepsilon}_1) = \mathbf{0}$, $E(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}'_1) = \Sigma > \mathbf{0}$, $E(\boldsymbol{\varepsilon}_1 \otimes \boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}'_1) = \mathbf{0}_{p^2\times p}$ and $E\|\boldsymbol{\varepsilon}_1\|^4 < \infty$, where $\boldsymbol{\varepsilon}'_1$ is the first row vector of the error matrix $\mathcal{E}$.

**Theorem 4.2.** *Under Assumption 1, the statistic $\sqrt{n}\left(\widehat{\Sigma}_{copls}(\mathbf{Y}) - \Sigma\right)$ converges in distribution to the multivariate normal distribution $N_{p^2}(\mathbf{0}, \mathrm{Cov}(\boldsymbol{\varepsilon}_1 \otimes \boldsymbol{\varepsilon}_1))$.*

**Proof.** $\sqrt{n}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)$ can be decomposed into

$$\sqrt{n}\left(\frac{1}{n}\mathcal{E}'\mathcal{E}-\Sigma\right)+Q\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{0}\end{pmatrix}Q'$$

with $\mathbf{A}_{kl}=\sqrt{n}\left(\frac{1}{n-r}\sum_{i=1}^{n-r}\mathbf{w}_{ik}\mathbf{w}_{il}'-\frac{1}{n}\sum_{i=1}^{n}\mathbf{w}_{ik}\mathbf{w}_{il}'\right)$ for $k,l=1,2$ except $k=l=2$. Since

$$\mathbf{A}_{kl}=\frac{r\sqrt{n}}{n-r}\frac{1}{n}\sum_{i=1}^{n}\mathbf{w}_{ik}\mathbf{w}_{il}'-\frac{\sqrt{n}}{n-r}\sum_{i=n-r+1}^{n}\mathbf{w}_{ik}\mathbf{w}_{il}',$$

$\mathbf{A}_{kl}$ converges to $\mathbf{0}$ in probability. By assumption, the first item converges to $N_{p^2}(\mathbf{0},\Phi_2)$ in distribution, where $\Phi_2=\mathrm{Cov}(\boldsymbol{\varepsilon}_1\otimes\boldsymbol{\varepsilon}_1)$. Hence, it follows from Slutsky's Theorem, see Lehmann and Romano [12], that the $\sqrt{n}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)$ converges in distribution to $N_{p^2}(\mathbf{0},\mathrm{Cov}(\boldsymbol{\varepsilon}_1\otimes\boldsymbol{\varepsilon}_1))$, completing the proof. □

## 5. Two-stage GLS estimator for the regression coefficient matrix

Note that the regression coefficient matrix $\Theta$ in model (1) is defined before a design is planned and an observation value matrix $Y$ is obtained. And the rows of the treatment design matrix $X$ in model (1) are added one after another and the profile matrix $Z$ in model (1) does not depend on the sample size $n$. So, in the repeated measurement experiments over multiple time points, the design matrix $X$ and the profile matrix $Z$ usually are of full rank. It is reasonable for us to only consider the case of full-rank matrices $X$ and $Z$. Assume that $X$ and $Z$ are of full rank in the sequent discussions.

To seek the least squares estimators for regression coefficient matrix $\Theta$ in the model (1), we usual use the two-stage generalized least squares estimation. That is, first, based on data $Y$, find a first-stage estimator $\tilde{\Sigma}$ of $\Sigma$; and secondly, replace the unknown $\Sigma$ with the first-stage estimator $\tilde{\Sigma}$ and then find the two-stage generalized least squares estimator $\widehat{\Theta}(\mathbf{Y})$ through the normal equations to regression coefficient matrix.

Taking the outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ given in (16) as the first-stage estimator of covariance $\Sigma$ and according to the theory of least squares, we have the normal equation to regression coefficient matrix in the model (1)

$$X'X\Theta Z'\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z=X'Y\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z,$$

where $\left(Z'\widehat{\Sigma}_{copls}(\mathbf{Y})Z\right)^{-1}$ exists with probability 1.

Then the two-stage least squares estimator, written as $\widehat{\Theta}_{copls}(\mathbf{Y})$, of $\Theta$ is given by

$$\widehat{\Theta}_{copls}(\mathbf{Y})=(X'X)^{-1}X'Y\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z\left(Z'\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z\right)^{-1}. \tag{21}$$

The estimator $\widehat{\Theta}_{copls}(\mathbf{Y})$ is easily shown to have unbiasedness under the assumption of $\mathcal{E}$ being symmetric about the origin.

**Assumption 2.** Assume that $\lim_{n\to\infty}\frac{1}{n}X'X=R>\mathbf{0}$.

**Theorem 5.1.** *Under Assumption 2, the two-stage generalized least squares estimator $\widehat{\Theta}_{copls}(\mathbf{Y})$ is consistent to regression coefficients $\Theta$.*

**Proof.** $\widehat{\Theta}_{copls}(\mathbf{Y})$ can be decomposed into $\Theta+\left(\frac{X'X}{n}\right)^{-1}\frac{X'\mathcal{E}}{n}\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z\left(Z'\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z\right)^{-1}$. Note that $\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})$ and $\left(Z'\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z\right)^{-1}$ both are bounded with probability 1. With Assumption 2, $\frac{1}{n}X'\mathcal{E}$ converges to $\mathbf{0}$ in probability due to

$$P\left(\left\|\frac{1}{n}X'\mathcal{E}\right\|\geq\varepsilon\right)\leq\frac{1}{n^2\varepsilon^2}E(\mathrm{tr}(X'\mathcal{E}\mathcal{E}'X))=\frac{1}{n\varepsilon^2}\mathrm{tr}\left(\frac{1}{n}X'X\right)\mathrm{tr}(\Sigma)$$

for any $\varepsilon>0$. Thus, the statistic $\widehat{\Theta}(\mathbf{Y})$ converges to $\Theta$ in probability, completing the proof. □

**Theorem 5.2.** *Under Assumptions 1 and 2, then*

(a) $\sqrt{n}\left(\widehat{\Theta}_{copls}(\mathbf{Y})-\Theta\right)$ *converges in distribution to the multivariate normal distribution* $N_{mq}\left(0,R^{-1}\otimes(Z'\Sigma^{-1}Z)^{-1}\right)$, *and*
(b) $\sqrt{n}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)$ *and* $\sqrt{n}\left(\widehat{\Theta}_{copls}(\mathbf{Y})-\Theta\right)$ *are asymptotically independent.*

**Proof.** (a) $\sqrt{n}\left(\widehat{\Theta}_{copls}(\mathbf{Y})-\Theta\right)$ can be decomposed into

$$\sqrt{n}\left((X'X)^{-1}X'\mathcal{E}\right)\left(\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z(Z'\widehat{\Sigma}_{copls}^{-1}(\mathbf{Y})Z)^{-1}\right). \tag{22}$$

Let $L_n=(X'X)^{-1}X'\mathcal{E}$. By Theorem 4.2 of Hu and Yan [6], $\sqrt{n}L_n$ converges in distribution to the normal distribution $N_{mp}(0,R^{-1}\otimes\Sigma)$. Therefore, $\sqrt{n}\left(\widehat{\Theta}_{copls}(\mathbf{Y})-\Theta\right)$ converges in distribution to $N_{mq}(0,R^{-1}\otimes(Z'\Sigma^{-1}Z)^{-1})$.

(b) By the Eq. (22), it suffices to prove the asymptotically independence between $\frac{1}{\sqrt{n}}\mathrm{vec}(X'\mathcal{E})$ and $\sqrt{n}\mathrm{vec}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)$. Let $Q_n = X'\mathcal{E} = (x_1,\ldots,x_n)(\boldsymbol{\varepsilon}_1,\ldots,\boldsymbol{\varepsilon}_n)'$. Then

$$\begin{aligned}\mathrm{Cov}\left(Q_n\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)\right) &= \mathrm{Cov}\left(\left(\sum_{i=1}^n x_i\boldsymbol{\varepsilon}_i'\right)\left(\frac{1}{n}\sum_{i=1}^n \boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'-\Sigma\right)\right)+o_p(\mathbf{1})\\ &= E\left(\left(\sum_{i=1}^n x_i\otimes\boldsymbol{\varepsilon}_i'\right)\left(\sum_{i=1}^n \boldsymbol{\varepsilon}_i\otimes\boldsymbol{\varepsilon}_i'-\Sigma\right)\right)+o_p(\mathbf{1}).\end{aligned}$$

According to Assumption 2, $\mathrm{Cov}\left(\left(\frac{1}{\sqrt{n}}X'\mathcal{E}\right)\sqrt{n}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)\right)$ converges to $\mathbf{0}$ in probability, implying that $\frac{1}{\sqrt{n}}\mathrm{vec}\left(X'\mathcal{E}\right)$ and $\sqrt{n}\mathrm{vec}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)$ are asymptotically independent. Therefore, $\sqrt{n}\left(\widehat{\Sigma}_{copls}(\mathbf{Y})-\Sigma\right)$ and $\sqrt{n}\left(\widehat{\Theta}_{copls}(\mathbf{Y})-\Theta\right)$ are also asymptotically independent. So, the proof is complete. □

## 6. Simulation studies and a numerical example

### 6.1. Simulation studies

In this section, we conduct some simulation studies to show the finite sample performance of the proposed procedure in previous sections. The data are generated from the following growth curve model

$$\mathbf{Y}_{n\times p} = X_{n\times s}\Theta_{s\times q}Z'_{p\times q}+\mathcal{E}_{n\times p},\qquad \mathrm{Cov}(\mathcal{E}) = I_n\otimes\Sigma$$

where $p = 4, s = 2, q = 3$, the same size $n = 20, 30, 50, 100$ with the design matrices $X = \mathrm{diag}\left(\mathbf{1}_{\frac{1}{n}},\mathbf{1}_{\frac{1}{n}}\right)$, respective, and the true $\Theta = \begin{pmatrix}-1 & 1 & 2\\ 1 & 3 & 5\end{pmatrix}$. For covariance $\Sigma$, we take two cases. The first one is an arbitrary positive definite structure

$$\Sigma_1 = \begin{pmatrix}1 & 0.8 & 0.5 & 0.4\\ 0.8 & 1 & 0.6 & 0.2\\ 0.5 & 0.6 & 1 & 0.7\\ 0.4 & 0.2 & 0.7 & 1\end{pmatrix}.$$

The second one is the autoregressive structure, namely

$$\Sigma_2 = \begin{pmatrix}1 & \rho & \rho^2 & \rho^3\\ \rho & 1 & \rho & \rho^2\\ \rho^2 & \rho & 1 & \rho\\ \rho^3 & \rho^2 & \rho & 1\end{pmatrix}\Big/(1-\rho^2)\quad \text{with } \rho = 0.6.$$

Under assumption of normality, the maximum likelihood estimators, written as $\widehat{\Theta}_{ml}$ and $\widehat{\Sigma}_{ml}$, are reexpressed, respectively, by

$$\begin{aligned}\widehat{\Theta}_{ml} &= (X'X)^{-1}X'\mathbf{Y}\left(\mathbf{Y}'M_X\mathbf{Y}\right)^{-1}Z\left(Z'\left(\mathbf{Y}'M_X\mathbf{Y}\right)^{-1}Z\right)^{-1}\\ \widehat{\Sigma}_{ml} &= \frac{1}{n}(Y-X\widehat{\Theta}^{ml}Z')'(Y-X\widehat{\Theta}^{ml}Z'),\end{aligned}\tag{23}$$

see Khatri [8] or von Rosen [22] for details.

In each case, the number of simulated realizations is 1000. Regarding the two-stage GLS estimator $\widehat{\Theta}_{copls}$, given by (21), and the MLE $\widehat{\Theta}_{ml}$, given by (23), of the parametric matrix $\Theta$, for a fixed sample size, the sample means (*sms*), bias (the difference between estimated values and the corresponding true values), the standard deviations (*stds*), the mean squared error (*mses*), and coverage of the 95% nominal confidence intervals (*cps*) are obtained. The results are summarized in Tables 1 and 2.

From the Tables, we make the following observations:

(1) The standard deviations and the mean squared error of the proposed GLS estimator $\widehat{\Theta}_{copls}$ decrease as $n$ increases. The proposed GLS estimator $\widehat{\Theta}_{copls}$ performs a little bit more efficient than MLE $\widehat{\Theta}_{ml}$ does in sense MSE.
(2) The biases both decrease as $n$ increases. The bias of the proposed GLS estimator $\widehat{\Theta}_{copls}$ is almost smaller than that of MLE $\widehat{\Theta}_{ml}$.
(3) Each coverage of the 95% nominal confidence interval of the proposed GLS estimator $\widehat{\Theta}_{copls}$ is more satisfactory than that of MLE $\widehat{\Theta}_{ml}$ does.

**Table 1**
Finite sample performances of COPLS estimators and MLEs under Case 1.

| $n$ | | COPLS estimators | | | | | ML estimators | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | sm | bias | std | mse | cp | sm | bias | std | mse | cp |
| 20 | $\hat{\theta}_{11}$ | −0.9994 | 0.0006 | 0.5778 | 0.3335 | 0.9410 | −1.0007 | −0.0007 | 0.5807 | 0.3369 | 0.9380 |
| | $\hat{\theta}_{12}$ | 1.0043 | 0.0043 | 0.5022 | 0.2520 | 0.9430 | 1.0052 | 0.0052 | 0.5047 | 0.2545 | 0.9420 |
| | $\hat{\theta}_{13}$ | 1.9987 | −0.0013 | 0.1044 | 0.0109 | 0.9380 | 1.9986 | −0.0014 | 0.1049 | 0.0110 | 0.9350 |
| | $\hat{\theta}_{21}$ | 1.0256 | 0.0256 | 0.5624 | 0.3167 | 0.9470 | 1.0248 | 0.0248 | 0.5656 | 0.3202 | 0.9450 |
| | $\hat{\theta}_{22}$ | 2.9633 | −0.0367 | 0.4909 | 0.2421 | 0.9540 | 2.9641 | −0.0359 | 0.4937 | 0.2448 | 0.9510 |
| | $\hat{\theta}_{23}$ | 5.0076 | 0.0076 | 0.1027 | 0.0106 | 0.9520 | 5.0074 | 0.0074 | 0.1033 | 0.0107 | 0.9520 |
| | $\hat{\sigma}_{11}$ | 1.0247 | 0.0247 | 0.3451 | 0.1196 | – | 0.9315 | −0.0685 | 0.3154 | 0.1041 | – |
| | $\hat{\sigma}_{12}$ | 0.8169 | 0.0169 | 0.3095 | 0.0960 | – | 0.7456 | −0.0544 | 0.2851 | 0.0842 | – |
| | $\hat{\sigma}_{13}$ | 0.4998 | −0.0002 | 0.2706 | 0.0732 | – | 0.4445 | −0.0555 | 0.2494 | 0.0652 | – |
| | $\hat{\sigma}_{14}$ | 0.4050 | 0.0050 | 0.2632 | 0.0692 | – | 0.3598 | −0.0402 | 0.2430 | 0.0606 | – |
| | $\hat{\sigma}_{22}$ | 1.0065 | 0.0065 | 0.3397 | 0.1153 | – | 0.9217 | −0.0783 | 0.3149 | 0.1052 | – |
| | $\hat{\sigma}_{23}$ | 0.5971 | −0.0029 | 0.2767 | 0.0765 | – | 0.5253 | −0.0747 | 0.2534 | 0.0697 | – |
| | $\hat{\sigma}_{24}$ | 0.2064 | 0.0064 | 0.2470 | 0.0610 | – | 0.1724 | −0.0276 | 0.2296 | 0.0534 | – |
| | $\hat{\sigma}_{33}$ | 0.9821 | −0.0179 | 0.3230 | 0.1045 | – | 0.9167 | −0.0833 | 0.3074 | 0.1013 | – |
| | $\hat{\sigma}_{34}$ | 0.6863 | −0.0137 | 0.2730 | 0.0746 | – | 0.6462 | −0.0538 | 0.2586 | 0.0697 | – |
| | $\hat{\sigma}_{44}$ | 0.9842 | −0.0158 | 0.3126 | 0.0979 | – | 0.9136 | −0.0864 | 0.2963 | 0.0952 | – |
| 30 | $\hat{\theta}_{11}$ | −1.0173 | −0.0173 | 0.4414 | 0.1950 | 0.9540 | −1.0179 | −0.0179 | 0.4428 | 0.1962 | 0.9540 |
| | $\hat{\theta}_{12}$ | 1.0107 | 0.0107 | 0.3930 | 0.1544 | 0.9500 | 1.0110 | 0.0110 | 0.3941 | 0.1553 | 0.9470 |
| | $\hat{\theta}_{13}$ | 1.9980 | −0.0020 | 0.0825 | 0.0068 | 0.9470 | 1.9980 | −0.0020 | 0.0826 | 0.0068 | 0.9470 |
| | $\hat{\theta}_{21}$ | 1.0057 | 0.0057 | 0.4499 | 0.2022 | 0.9530 | 1.0052 | 0.0052 | 0.4502 | 0.2025 | 0.9510 |
| | $\hat{\theta}_{22}$ | 2.9951 | −0.0049 | 0.3968 | 0.1573 | 0.9500 | 2.9954 | −0.0046 | 0.3974 | 0.1578 | 0.9500 |
| | $\hat{\theta}_{23}$ | 5.0004 | 0.0004 | 0.0828 | 0.0069 | 0.9480 | 5.0004 | 0.0004 | 0.0830 | 0.0069 | 0.9470 |
| | $\hat{\sigma}_{11}$ | 1.0076 | 0.0076 | 0.2707 | 0.0733 | – | 0.9444 | −0.0556 | 0.2548 | 0.0679 | – |
| | $\hat{\sigma}_{12}$ | 0.8043 | 0.0043 | 0.2443 | 0.0596 | – | 0.7557 | −0.0443 | 0.2310 | 0.0553 | – |
| | $\hat{\sigma}_{13}$ | 0.5000 | 0.0000 | 0.2084 | 0.0434 | – | 0.4625 | −0.0375 | 0.1970 | 0.0402 | – |
| | $\hat{\sigma}_{14}$ | 0.3990 | −0.0010 | 0.2003 | 0.0401 | – | 0.3686 | −0.0314 | 0.1890 | 0.0367 | – |
| | $\hat{\sigma}_{22}$ | 1.0040 | 0.0040 | 0.2668 | 0.0711 | – | 0.9458 | −0.0542 | 0.2537 | 0.0672 | – |
| | $\hat{\sigma}_{23}$ | 0.6076 | 0.0076 | 0.2133 | 0.0455 | – | 0.5580 | −0.0420 | 0.2010 | 0.0421 | – |
| | $\hat{\sigma}_{24}$ | 0.2069 | 0.0069 | 0.1919 | 0.0368 | – | 0.1841 | −0.0159 | 0.1823 | 0.0334 | – |
| | $\hat{\sigma}_{33}$ | 1.0107 | 0.0107 | 0.2537 | 0.0644 | – | 0.9638 | −0.0362 | 0.2455 | 0.0615 | – |
| | $\hat{\sigma}_{34}$ | 0.7010 | 0.0010 | 0.2199 | 0.0483 | – | 0.6719 | −0.0281 | 0.2116 | 0.0455 | – |
| | $\hat{\sigma}_{44}$ | 0.9913 | −0.0087 | 0.2517 | 0.0634 | – | 0.9418 | −0.0582 | 0.2411 | 0.0614 | – |
| 50 | $\hat{\theta}_{11}$ | −1.0073 | −0.0073 | 0.3529 | 0.1245 | 0.9460 | −1.0073 | −0.0073 | 0.3531 | 0.1246 | 0.9450 |
| | $\hat{\theta}_{12}$ | 1.0021 | 0.0021 | 0.3096 | 0.0958 | 0.9480 | 1.0021 | 0.0021 | 0.3098 | 0.0959 | 0.9480 |
| | $\hat{\theta}_{13}$ | 1.9998 | −0.0002 | 0.0643 | 0.0041 | 0.9440 | 1.9998 | −0.0002 | 0.0643 | 0.0041 | 0.9440 |
| | $\hat{\theta}_{21}$ | 0.9944 | −0.0056 | 0.3406 | 0.1159 | 0.9490 | 0.9944 | −0.0056 | 0.3408 | 0.1160 | 0.9480 |
| | $\hat{\theta}_{22}$ | 2.9983 | −0.0017 | 0.3088 | 0.0953 | 0.9510 | 2.9984 | −0.0016 | 0.3089 | 0.0954 | 0.9500 |
| | $\hat{\theta}_{23}$ | 5.0008 | 0.0008 | 0.0648 | 0.0042 | 0.9480 | 5.0007 | 0.0007 | 0.0648 | 0.0042 | 0.9480 |
| | $\hat{\sigma}_{11}$ | 0.9976 | −0.0024 | 0.1991 | 0.0396 | – | 0.9594 | −0.0406 | 0.1916 | 0.0383 | – |
| | $\hat{\sigma}_{12}$ | 0.8027 | 0.0027 | 0.1740 | 0.0303 | – | 0.7729 | −0.0271 | 0.1678 | 0.0289 | – |
| | $\hat{\sigma}_{13}$ | 0.5068 | 0.0068 | 0.1547 | 0.0239 | – | 0.4840 | −0.0160 | 0.1498 | 0.0227 | – |
| | $\hat{\sigma}_{14}$ | 0.3990 | −0.0010 | 0.1527 | 0.0233 | – | 0.3807 | −0.0193 | 0.1477 | 0.0222 | – |
| | $\hat{\sigma}_{22}$ | 1.0072 | 0.0072 | 0.1901 | 0.0362 | – | 0.9715 | −0.0285 | 0.1840 | 0.0346 | – |
| | $\hat{\sigma}_{23}$ | 0.6083 | 0.0083 | 0.1633 | 0.0267 | – | 0.5782 | −0.0218 | 0.1576 | 0.0253 | – |
| | $\hat{\sigma}_{24}$ | 0.1993 | −0.0007 | 0.1434 | 0.0206 | – | 0.1860 | −0.0140 | 0.1387 | 0.0194 | – |
| | $\hat{\sigma}_{33}$ | 1.0073 | 0.0073 | 0.2067 | 0.0427 | – | 0.9788 | −0.0212 | 0.2026 | 0.0414 | – |
| | $\hat{\sigma}_{34}$ | 0.7015 | 0.0015 | 0.1812 | 0.0328 | – | 0.6837 | −0.0163 | 0.1777 | 0.0318 | – |
| | $\hat{\sigma}_{44}$ | 1.0024 | 0.0024 | 0.2060 | 0.0424 | – | 0.9718 | −0.0282 | 0.2016 | 0.0414 | – |
| 100 | $\hat{\theta}_{11}$ | −0.9888 | 0.0112 | 0.2407 | 0.0580 | 0.9680 | −0.9888 | 0.0112 | 0.2408 | 0.0580 | 0.9680 |
| | $\hat{\theta}_{12}$ | 0.9894 | −0.0106 | 0.2088 | 0.0437 | 0.9610 | 0.9894 | −0.0106 | 0.2089 | 0.0437 | 0.9610 |
| | $\hat{\theta}_{13}$ | 2.0021 | 0.0021 | 0.0436 | 0.0019 | 0.9500 | 2.0021 | 0.0021 | 0.0436 | 0.0019 | 0.9500 |
| | $\hat{\theta}_{21}$ | 1.0058 | 0.0058 | 0.2534 | 0.0642 | 0.9460 | 1.0058 | 0.0058 | 0.2535 | 0.0642 | 0.9460 |
| | $\hat{\theta}_{22}$ | 2.9927 | −0.0073 | 0.2203 | 0.0485 | 0.9480 | 2.9927 | −0.0073 | 0.2204 | 0.0486 | 0.9470 |
| | $\hat{\theta}_{23}$ | 5.0015 | 0.0015 | 0.0457 | 0.0021 | 0.9430 | 5.0015 | 0.0015 | 0.0457 | 0.0021 | 0.9420 |
| | $\hat{\sigma}_{11}$ | 1.0006 | 0.0006 | 0.1397 | 0.0195 | – | 0.9813 | −0.0187 | 0.1371 | 0.0191 | – |
| | $\hat{\sigma}_{12}$ | 0.7997 | −0.0003 | 0.1272 | 0.0162 | – | 0.7850 | −0.0150 | 0.1250 | 0.0158 | – |
| | $\hat{\sigma}_{13}$ | 0.4990 | −0.0010 | 0.1088 | 0.0118 | – | 0.4873 | −0.0127 | 0.1071 | 0.0116 | – |
| | $\hat{\sigma}_{14}$ | 0.3995 | −0.0005 | 0.1053 | 0.0111 | – | 0.3900 | −0.0100 | 0.1036 | 0.0108 | – |
| | $\hat{\sigma}_{22}$ | 1.0006 | 0.0006 | 0.1448 | 0.0209 | – | 0.9828 | −0.0172 | 0.1426 | 0.0206 | – |
| | $\hat{\sigma}_{23}$ | 0.6008 | 0.0008 | 0.1139 | 0.0129 | – | 0.5855 | −0.0145 | 0.1118 | 0.0127 | – |

Table 1 (*continued*)

| $n$ | COPLS estimators | | | | | ML estimators | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | sm | bias | std | mse | cp | sm | bias | std | mse | cp |
| $\hat{\sigma}_{24}$ | 0.2013 | 0.0013 | 0.0980 | 0.0096 | – | 0.1941 | −0.0059 | 0.0965 | 0.0093 | – |
| $\hat{\sigma}_{33}$ | 1.0035 | 0.0035 | 0.1369 | 0.0187 | – | 0.9894 | −0.0106 | 0.1356 | 0.0185 | – |
| $\hat{\sigma}_{34}$ | 0.7013 | 0.0013 | 0.1192 | 0.0142 | – | 0.6927 | −0.0073 | 0.1182 | 0.0140 | – |
| $\hat{\sigma}_{44}$ | 0.9986 | −0.0014 | 0.1412 | 0.0199 | – | 0.9836 | −0.0164 | 0.1399 | 0.0198 | – |

Regarding the $\widehat{\Sigma}_{copls}$, given by (16), and the MLE $\widehat{\Sigma}_{ml}$, see (23), for a fixed sample size, the sample means (*sms*), bias (the difference between estimated values and the corresponding true values), standard deviations (*stds*) and the mean squared error (*mses*) are obtained. The results are summarized in Tables 1 and 2. From the Tables, we make the following observations:

(i) The sample means (*sms*) of $\widehat{\Sigma}_{copls}$ are closer to the corresponding true values than that of $\widehat{\Sigma}_{ml}$. They both are closer and closer to the corresponding true values as $n$ increases.

(ii) The biases and standard deviations both decrease as $n$ increases for two estimators. The bias of $\widehat{\Sigma}_{copls}$ is much smaller than that of $\widehat{\Sigma}_{ml}$ while the standard deviation and the mean squared error of $\widehat{\Sigma}_{ml}$ are smaller than those of the $\widehat{\Sigma}_{copls}$. Also see the following geometrical presentation.

The above observations imply that $\widehat{\Theta}_{copls}$ has more efficient (more precision and more accuracy) than that of $\widehat{\Theta}_{ml}$ from the small sample performance, and $\widehat{\Sigma}_{copls}$ is more efficient than $\widehat{\Sigma}_{ml}$ in the sense of accuracy. These small sample simulation studies also show that the proposed method or the COPLS approach is an alternate competitor for parameter estimation in the growth curve model.

### 6.2. A geometrical presentation

The conclusion from above simulations is natural. A geometrical interpretation is presented. Let us re-look at the maximum likelihood estimator $\widehat{\Theta}_{ml}(\mathbf{Y})$. If taking $\widehat{\Sigma}^{multi}_{ols}$ given by (18) as the first-stage estimator of covariance, we have

$$\begin{aligned}\widehat{\Theta}_{2sls}(\mathbf{Y}) &= (X'X)^{-1}X'\mathbf{Y}\left(\widehat{\Sigma}^{multi}_{ols}(\mathbf{Y})\right)^{-1}Z\left(Z'\left(\widehat{\Sigma}^{multi}_{ols}(\mathbf{Y})^{-1}\right)Z\right)^{-1}\\ &= (X'X)^{-1}X'\mathbf{Y}\,(\mathbf{Y}M_X\mathbf{Y})^{-1}Z\left(Z'\,(\mathbf{Y}M_X\mathbf{Y})^{-1}Z\right)^{-1}\\ &= \widehat{\Theta}_{ml}(\mathbf{Y}).\end{aligned} \tag{24}$$

That is to say, $\widehat{\Theta}_{ml}(\mathbf{Y})$ is a two-stage GLS estimator of $\Theta$ in the growth curve model. From the discussion in Section 2.4, $\widehat{\Sigma}^{multi}_{ols}$ in (18) is the COPLS estimator of covariance $\Sigma$ for the multivariate linear model. Compared to the COPLS estimator given in (16), the $\widehat{\Sigma}^{multi}_{ols}$ is obtained under completely ignoring $Z$ when the growth curve model is considered. We believe that the price will be paid for ignoring the profile matrix $Z$ due to $r(Z) < r(I) = p$.

For more clarity, we rewrite the multivariate linear model and the growth curve model with the same covariance structure $I \otimes \Sigma$ as follows:

$$\mathrm{vec}(\mathbf{Y}) = (X \otimes I)\mathrm{vec}(\Theta) + \mathcal{E} \tag{25}$$

and

$$\mathrm{vec}(\mathbf{Y}) = (X \otimes Z)\mathrm{vec}(\Theta) + \mathcal{E}. \tag{26}$$

Owing to $r(Z) < r(I) = p$, $\mathscr{C}(X \otimes Z) \subset \mathscr{C}(X \otimes I)$. It implies that the model (26) is a reduced model to the full model (25). Assume that the reduced model (26) is true. It means that the full model (25) is also true. The error space of the reduced model is bigger than that of the full model. The error space of the full model is a subspace of the error space of the reduced model. The $\widehat{\Sigma}_{copls}$ is obtained by using the bigger error space while $\widehat{\Sigma}^{multi}_{ols}$ is obtained by using the smaller error space. The bigger one contains more information about covariance. There is no reason for us to give up the COPLS estimator in (16) and choose the estimator in (18). And any estimator of the regression coefficient matrix strongly relies on the preestimated covariance matrix $\Sigma$. So, it is reasonable for us to believe that, for the growth curve model, $\widehat{\Theta}_{copls}(\mathbf{Y})$ is more competitive than the $\widehat{\Theta}_{2sls}(\mathbf{Y})$ or $\widehat{\Theta}_{ml}(\mathbf{Y})$, see (24). In one word, there are practical and statistical advantages to take the COPLS estimator given by (16) as the first-stage estimate of covariance when doing the two-stage GLS for the regression coefficient matrix, see Hu and Yan [6].

### 6.3. A numerical example

The numerical example, stated in [17], about measurements on 11 girls and 16 boys at 4 different ages is employed here to illustrate the calculation of using $\widehat{\Theta}_{opls}(\mathbf{Y})$ to estimate the regression coefficients of the growth curve for 11 girls and 16 boys (see Table 3).

**Table 2**
Finite sample performances of COPLS estimators and MLEs under Case 2.

| $n$ | | COPLS estimators | | | | | ML estimators | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | sm | bias | std | mse | cp | sm | bias | std | mse | cp |
| 20 | $\hat{\theta}_{11}$ | −0.9935 | 0.0065 | 0.8065 | 0.6498 | 0.9440 | −0.9926 | 0.0074 | 0.8087 | 0.6534 | 0.9410 |
| | $\hat{\theta}_{12}$ | 1.0114 | 0.0114 | 0.6933 | 0.4804 | 0.9490 | 1.0102 | 0.0102 | 0.6969 | 0.4852 | 0.9480 |
| | $\hat{\theta}_{13}$ | 1.9962 | −0.0038 | 0.1350 | 0.0182 | 0.9400 | 1.9965 | −0.0035 | 0.1356 | 0.0184 | 0.9400 |
| | $\hat{\theta}_{21}$ | 0.9578 | −0.0422 | 0.8205 | 0.6744 | 0.9440 | 0.9581 | −0.0419 | 0.8232 | 0.6787 | 0.9430 |
| | $\hat{\theta}_{22}$ | 3.0227 | 0.0227 | 0.7189 | 0.5168 | 0.9360 | 3.0224 | 0.0224 | 0.7214 | 0.5204 | 0.9350 |
| | $\hat{\theta}_{23}$ | 4.9971 | −0.0029 | 0.1395 | 0.0195 | 0.9370 | 4.9972 | −0.0028 | 0.1401 | 0.0196 | 0.9370 |
| | $\hat{\sigma}_{11}$ | 1.5435 | −0.0190 | 0.5216 | 0.2721 | – | 1.3983 | −0.1642 | 0.4731 | 0.2505 | – |
| | $\hat{\sigma}_{12}$ | 0.9122 | −0.0253 | 0.4224 | 0.1789 | – | 0.8243 | −0.1132 | 0.3857 | 0.1614 | – |
| | $\hat{\sigma}_{13}$ | 0.5342 | −0.0283 | 0.3812 | 0.1460 | – | 0.4868 | −0.0757 | 0.3522 | 0.1297 | – |
| | $\hat{\sigma}_{14}$ | 0.3155 | −0.0220 | 0.3716 | 0.1385 | – | 0.2863 | −0.0512 | 0.3374 | 0.1164 | – |
| | $\hat{\sigma}_{22}$ | 1.5338 | −0.0287 | 0.5329 | 0.2845 | – | 1.4120 | −0.1505 | 0.4987 | 0.2712 | – |
| | $\hat{\sigma}_{23}$ | 0.9093 | −0.0282 | 0.4144 | 0.1723 | – | 0.8017 | −0.1358 | 0.3787 | 0.1617 | – |
| | $\hat{\sigma}_{24}$ | 0.5404 | −0.0221 | 0.3881 | 0.1510 | – | 0.4921 | −0.0704 | 0.3579 | 0.1329 | – |
| | $\hat{\sigma}_{33}$ | 1.5593 | −0.0032 | 0.5228 | 0.2731 | – | 1.4381 | −0.1244 | 0.4950 | 0.2602 | – |
| | $\hat{\sigma}_{34}$ | 0.9339 | −0.0036 | 0.4148 | 0.1719 | – | 0.8457 | −0.0918 | 0.3841 | 0.1558 | – |
| | $\hat{\sigma}_{44}$ | 1.5505 | −0.0120 | 0.5195 | 0.2698 | – | 1.4061 | −0.1564 | 0.4721 | 0.2471 | – |
| 30 | $\hat{\theta}_{11}$ | −1.0090 | −0.0090 | 0.6574 | 0.4318 | 0.9500 | −1.0093 | −0.0093 | 0.6597 | 0.4348 | 0.9480 |
| | $\hat{\theta}_{12}$ | 1.0084 | 0.0084 | 0.5708 | 0.3256 | 0.9480 | 1.0082 | 0.0082 | 0.5724 | 0.3274 | 0.9450 |
| | $\hat{\theta}_{13}$ | 1.9988 | −0.0012 | 0.1095 | 0.0120 | 0.9470 | 1.9989 | −0.0011 | 0.1098 | 0.0120 | 0.9460 |
| | $\hat{\theta}_{21}$ | 1.0010 | 0.0010 | 0.6476 | 0.4190 | 0.9550 | 1.0015 | 0.0015 | 0.6485 | 0.4202 | 0.9530 |
| | $\hat{\theta}_{22}$ | 2.9949 | −0.0051 | 0.5643 | 0.3182 | 0.9430 | 2.9945 | −0.0055 | 0.5655 | 0.3195 | 0.9430 |
| | $\hat{\theta}_{23}$ | 5.0013 | 0.0013 | 0.1103 | 0.0122 | 0.9400 | 5.0014 | 0.0014 | 0.1106 | 0.0122 | 0.9400 |
| | $\hat{\sigma}_{11}$ | 1.5252 | −0.0373 | 0.4113 | 0.1704 | – | 1.4278 | −0.1347 | 0.3853 | 0.1665 | – |
| | $\hat{\sigma}_{12}$ | 0.9053 | −0.0322 | 0.3304 | 0.1101 | – | 0.8469 | −0.0906 | 0.3122 | 0.1056 | – |
| | $\hat{\sigma}_{13}$ | 0.5312 | −0.0313 | 0.3034 | 0.0930 | – | 0.4976 | −0.0649 | 0.2876 | 0.0868 | – |
| | $\hat{\sigma}_{14}$ | 0.3142 | −0.0233 | 0.2777 | 0.0776 | – | 0.2939 | −0.0436 | 0.2600 | 0.0694 | – |
| | $\hat{\sigma}_{22}$ | 1.5404 | −0.0221 | 0.4182 | 0.1752 | – | 1.4582 | −0.1043 | 0.3995 | 0.1703 | – |
| | $\hat{\sigma}_{23}$ | 0.9110 | −0.0265 | 0.3414 | 0.1171 | – | 0.8363 | −0.1012 | 0.3208 | 0.1130 | – |
| | $\hat{\sigma}_{24}$ | 0.5444 | −0.0181 | 0.3006 | 0.0906 | – | 0.5104 | −0.0521 | 0.2840 | 0.0833 | – |
| | $\hat{\sigma}_{33}$ | 1.5511 | −0.0114 | 0.4190 | 0.1755 | – | 1.4668 | −0.0957 | 0.3992 | 0.1684 | – |
| | $\hat{\sigma}_{34}$ | 0.9298 | −0.0077 | 0.3339 | 0.1114 | – | 0.8686 | −0.0689 | 0.3145 | 0.1036 | – |
| | $\hat{\sigma}_{44}$ | 1.5645 | 0.0020 | 0.4096 | 0.1676 | – | 1.4636 | −0.0989 | 0.3836 | 0.1568 | – |
| 50 | $\hat{\theta}_{11}$ | −1.0228 | −0.0228 | 0.5004 | 0.2507 | 0.9600 | −1.0229 | −0.0229 | 0.5007 | 0.2510 | 0.9590 |
| | $\hat{\theta}_{12}$ | 1.0178 | 0.0178 | 0.4374 | 0.1914 | 0.9420 | 1.0179 | 0.0179 | 0.4377 | 0.1917 | 0.9420 |
| | $\hat{\theta}_{13}$ | 1.9968 | −0.0032 | 0.0848 | 0.0072 | 0.9470 | 1.9968 | −0.0032 | 0.0849 | 0.0072 | 0.9470 |
| | $\hat{\theta}_{21}$ | 0.9876 | −0.0124 | 0.5125 | 0.2625 | 0.9490 | 0.9877 | −0.0123 | 0.5129 | 0.2630 | 0.9470 |
| | $\hat{\theta}_{22}$ | 3.0105 | 0.0105 | 0.4394 | 0.1930 | 0.9540 | 3.0104 | 0.0104 | 0.4398 | 0.1934 | 0.9540 |
| | $\hat{\theta}_{23}$ | 4.9991 | −0.0009 | 0.0854 | 0.0073 | 0.9550 | 4.9991 | −0.0009 | 0.0854 | 0.0073 | 0.9540 |
| | $\hat{\sigma}_{11}$ | 1.5497 | −0.0128 | 0.3117 | 0.0972 | – | 1.4891 | −0.0734 | 0.2995 | 0.0950 | – |
| | $\hat{\sigma}_{12}$ | 0.9332 | −0.0043 | 0.2583 | 0.0667 | – | 0.8958 | −0.0417 | 0.2494 | 0.0639 | – |
| | $\hat{\sigma}_{13}$ | 0.5701 | 0.0076 | 0.2406 | 0.0579 | – | 0.5486 | −0.0139 | 0.2323 | 0.0541 | – |
| | $\hat{\sigma}_{14}$ | 0.3424 | 0.0049 | 0.2322 | 0.0539 | – | 0.3289 | −0.0086 | 0.2232 | 0.0498 | – |
| | $\hat{\sigma}_{22}$ | 1.5589 | −0.0036 | 0.3169 | 0.1004 | – | 1.5072 | −0.0553 | 0.3078 | 0.0977 | – |
| | $\hat{\sigma}_{23}$ | 0.9392 | 0.0017 | 0.2693 | 0.0725 | – | 0.8928 | −0.0447 | 0.2592 | 0.0691 | – |
| | $\hat{\sigma}_{24}$ | 0.5671 | 0.0046 | 0.2476 | 0.0613 | – | 0.5456 | −0.0169 | 0.2390 | 0.0573 | – |
| | $\hat{\sigma}_{33}$ | 1.5618 | −0.0007 | 0.3291 | 0.1082 | – | 1.5102 | −0.0523 | 0.3201 | 0.1051 | – |
| | $\hat{\sigma}_{34}$ | 0.9356 | −0.0019 | 0.2660 | 0.0707 | – | 0.8982 | −0.0393 | 0.2569 | 0.0675 | – |
| | $\hat{\sigma}_{44}$ | 1.5526 | −0.0099 | 0.3128 | 0.0979 | – | 1.4918 | −0.0707 | 0.3006 | 0.0953 | – |
| 100 | $\hat{\theta}_{11}$ | −0.9854 | 0.0146 | 0.3532 | 0.1248 | 0.9570 | −0.9853 | 0.0147 | 0.3533 | 0.1249 | 0.9560 |
| | $\hat{\theta}_{12}$ | 0.9905 | −0.0095 | 0.3124 | 0.0976 | 0.9450 | 0.9906 | −0.0094 | 0.3125 | 0.0976 | 0.9430 |
| | $\hat{\theta}_{13}$ | 2.0014 | 0.0014 | 0.0602 | 0.0036 | 0.9510 | 2.0014 | 0.0014 | 0.0602 | 0.0036 | 0.9510 |
| | $\hat{\theta}_{21}$ | 0.9968 | −0.0032 | 0.3599 | 0.1294 | 0.9470 | 0.9969 | −0.0031 | 0.3601 | 0.1295 | 0.9470 |
| | $\hat{\theta}_{22}$ | 2.9953 | −0.0047 | 0.3073 | 0.0943 | 0.9520 | 2.9953 | −0.0047 | 0.3074 | 0.0944 | 0.9510 |
| | $\hat{\theta}_{23}$ | 5.0010 | 0.0010 | 0.0587 | 0.0034 | 0.9560 | 5.0010 | 0.0010 | 0.0587 | 0.0034 | 0.9560 |
| | $\hat{\sigma}_{11}$ | 1.5500 | −0.0125 | 0.2147 | 0.0462 | – | 1.5194 | −0.0431 | 0.2104 | 0.0461 | – |
| | $\hat{\sigma}_{12}$ | 0.9292 | −0.0083 | 0.1733 | 0.0301 | – | 0.9104 | −0.0271 | 0.1703 | 0.0297 | – |
| | $\hat{\sigma}_{13}$ | 0.5688 | 0.0063 | 0.1598 | 0.0256 | – | 0.5579 | −0.0046 | 0.1571 | 0.0247 | – |
| | $\hat{\sigma}_{14}$ | 0.3414 | 0.0039 | 0.1594 | 0.0254 | – | 0.3346 | −0.0029 | 0.1562 | 0.0244 | – |
| | $\hat{\sigma}_{22}$ | 1.5561 | −0.0064 | 0.2140 | 0.0458 | – | 1.5299 | −0.0326 | 0.2111 | 0.0456 | – |
| | $\hat{\sigma}_{23}$ | 0.9413 | 0.0038 | 0.1815 | 0.0329 | – | 0.9179 | −0.0196 | 0.1782 | 0.0321 | – |

Table 2 (*continued*)

| $n$ | | COPLS estimators | | | | | ML estimators | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | sm | bias | std | mse | cp | sm | bias | std | mse | cp |
| | $\hat{\sigma}_{24}$ | 0.5632 | 0.0007 | 0.1699 | 0.0288 | – | 0.5523 | −0.0102 | 0.1671 | 0.0280 | – |
| | $\hat{\sigma}_{33}$ | 1.5702 | 0.0077 | 0.2233 | 0.0499 | – | 1.5441 | −0.0184 | 0.2202 | 0.0488 | – |
| | $\hat{\sigma}_{34}$ | 0.9409 | 0.0034 | 0.1889 | 0.0357 | – | 0.9220 | −0.0155 | 0.1856 | 0.0347 | – |
| | $\hat{\sigma}_{44}$ | 1.5686 | 0.0061 | 0.2237 | 0.0500 | – | 1.5376 | −0.0249 | 0.2192 | 0.0486 | – |

**Table 3**
Measurements on 11 girls and 16 boys, at 4 different ages 8, 10, 12, 14.

| Girls | 8 | 10 | 12 | 14 | Boys | 8 | 10 | 12 | 14 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 21 | 20 | 21.5 | 23 | 1 | 26 | 25 | 29 | 31 |
| 2 | 21 | 21.5 | 24 | 25.5 | 2 | 21.5 | 22.5 | 23 | 26.5 |
| 3 | 20.5 | 24 | 24.5 | 26 | 3 | 23 | 22.5 | 24 | 27.5 |
| 4 | 23.5 | 24.5 | 25 | 26.5 | 4 | 25.5 | 27.5 | 26.5 | 27 |
| 5 | 21.5 | 23 | 22.5 | 23.5 | 5 | 20 | 23.5 | 22.5 | 26 |
| 6 | 20 | 21 | 21 | 22.5 | 6 | 24.5 | 25.5 | 27 | 28.5 |
| 7 | 21.5 | 22.5 | 23 | 25 | 7 | 22 | 22 | 24.5 | 26.5 |
| 8 | 23 | 23 | 23.5 | 24 | 8 | 24 | 21.5 | 24.5 | 25.5 |
| 9 | 20 | 21 | 22 | 21.5 | 9 | 23 | 20.5 | 31 | 26 |
| 10 | 16.5 | 19 | 19 | 19.5 | 10 | 27.5 | 28 | 31 | 31.5 |
| 11 | 24.5 | 25 | 28 | 28 | 11 | 23 | 23 | 23.5 | 25 |
| | | | | | 12 | 21.5 | 23.5 | 24 | 28 |
| | | | | | 13 | 17 | 24.5 | 26 | 29.5 |
| | | | | | 14 | 22.5 | 25.5 | 25.5 | 26 |
| | | | | | 15 | 23 | 24.5 | 26 | 30 |
| | | | | | 16 | 22 | 21.5 | 23.5 | 25 |
| Mean | 21.18 | 22.23 | 23.09 | 24.09 | Mean | 22.87 | 23.81 | 25.72 | 27.47 |

(a) First, we assume quadratic equations in time $t$ for the growth curves of 16 boys and 11 girls. Here $m = 2$, $p = 3$, $t_1 = -3$, $t_2 = -2$, $t_3 = 1$, $t_4 = 3$, design matrix $X = \begin{pmatrix} \mathbf{1}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{16} \end{pmatrix}$, profile matrix $Z' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \\ 9 & 1 & 1 & 9 \end{pmatrix}$ and regression coefficient $\Theta = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{pmatrix}$.

By (16) and (21), we obtain the estimate of the regression coefficient matrix

$$\widehat{\Theta}_{copls} = \begin{pmatrix} 22.6819 & 0.4783 & -0.0026 \\ 24.6444 & 0.7887 & 0.0501 \end{pmatrix}$$

and the least squares estimator of covariance

$$\widehat{\Sigma}_{copls} = \begin{pmatrix} 5.4081 & 2.7388 & 3.8882 & 2.7176 \\ 2.7388 & 4.1187 & 2.9932 & 3.2951 \\ 3.8882 & 2.9932 & 6.3896 & 4.1528 \end{pmatrix}.$$

In the above $\widehat{\Theta}_{copls}$, $\hat{\theta}_{13}$ and $\hat{\theta}_{23}$ are so close to zero that it motivates us to consider linear growth curve for 11 girls and 16 boys.

(b) We assume linear equations in time $t$ for the growth curves of 11 girls and 16 boys. Then $p = 2$, profile matrix $Z' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -3 & -1 & 1 & 3 \end{pmatrix}$ and regression coefficient $\Theta = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix}$. Other are same as that stated in (a).

Also by Eqs. (16) and (21), we obtain the estimate of regression coefficient

$$\widehat{\Theta}_{copls} = \begin{pmatrix} 22.6665 & 0.4765 \\ 24.9382 & 0.8255 \end{pmatrix}$$

and the least squares estimator of covariance

$$\widehat{\Sigma}_{copls} = \begin{pmatrix} 5.4262 & 2.708 & 3.8958 & 2.7228 \\ 2.708 & 4.1624 & 2.9985 & 3.2771 \\ 3.8958 & 2.9985 & 6.3563 & 4.1732 \\ 2.7228 & 3.2771 & 4.1732 & 4.9708 \end{pmatrix}.$$

## 7. Concluding remarks

When covariance in a linear model is known, the ordinary least squares method can give us a BLUE of the regression coefficient matrix. However, inference on the regression coefficient matrix strongly relies on the pre-estimated covariance matrix for the cases where covariance is unknown. So, before doing two-stage generalized least squares, we need a good first-stage estimator of covariance.

To provide a method to seek a good first-stage estimate of covariance, we develop a framework for directly doing least squares estimation to covariance in the growth curve model (1) without assumption of normality. Based on the idea of analogy, our consideration starts with using the outer product of the residual vector of data to estimate unknown covariances of random errors. An outer product least squares approach is formulated and an outer product least squares estimator $\widehat{\Sigma}_{copls}(\mathbf{Y})$ is obtained by the proposed framework. The COPLS estimator of covariance has an explicit expression in matrix quadratic forms and has been shown to have the properties: (1) following a linear transformation of two independent Wishart distribution for a normal error matrix; (2) having asymptotic normality for a nonnormal error matrix; and (3) having unbiasedness and invariance under a linear transformation group. These support the COPLS estimator as an excellent competitor to the maximum likelihood estimator of covariance. Taking the $\widehat{\Sigma}_{copls}(\mathbf{Y})$ as the first-stage estimator, we obtain the corresponding two-stage GLS estimator for the regression coefficient matrix and show its asymptotic normality.

Simulation studies with sample size 20, 30, 50, 100 to the growth curve model with a normal error demonstrate that the two-stage GLS estimators $\widehat{\Theta}_{copls}(\mathbf{Y})$ obtained by our framework are alternative competitors with more efficiency in the sense of MSE to the maximum likelihood estimators for the regression coefficients in finite samples.

The outer product least squares approach is suitable to estimate unknown parameters in covariance for a class of linear models with independent and identically distributed errors. The further development and applications of this approach are on progress.

## Acknowledgments

The authors are deeply grateful to two anonymous referees and the Associate Editor for their valuable comments which led to the improved version of this paper.

Hu's research was supported by National Natural Science Foundation of China (NSFC) Grant 10971126 and by Shanghai University of Finance and Economics for partial funding through Project 211 Phase III and Shanghai Leading Academic Discipline Project B803. Liu's research was supported by SUFE's Postgraduate Innovation Fund: CXJJ-2011-350 and NSFC Grant 11001162. Ahmed's research has been supported by grants from Natural Sciences and Engineering Research Council of Canadian Institute of Health Research.

## References

[1] S.F. Arnold, The Theory of Linear Models and Multivariate Analysis, Wiley, New York, 1981.
[2] M.L. Eaton, Multivariate Statistics: A Vector Space Approach, Wiley, New York, 1983.
[3] J.E. Grizzle, D.M. Allen, Analysis of growth and dose response curves, Biometrics 25 (1969) 357–381.
[4] D.A. Harville, Bayesian inference for variance components using only error contract, Biometrika 61 (1974) 383–385.
[5] J. Hu, Wishartness and independence of matrix quadratic forms in a normal random matrix, J. Multivariate Anal. 99 (2008) 555–571.
[6] J. Hu, G. Yan, Asymptotic normality and consistency of a two-stage generalized least squares estimator in the growth curve model, Bernoulli 14 (2008) 623–636.
[7] P.L. Hsu, On the best unbiased estimator of variance, Statist. Res. Mem. 2 (1938) 91–104.
[8] C.G. Khatri, A note on a MANOVA model applied to problems in growth curve, Ann. Inst. Statist. Math. 18 (1966) 75–86.
[9] T. Kollo, D. von Rosen, Advanced Multivariate Statsitics with Matrices, Mathematics and its Applications, vol. 579, Springer, New York, Dordrecht, 2005.
[10] N. Lange, N.M. Laird, The effect of covariance structure on variance estimation in balanced growth curve models with random parameters, J. Amer. Statist. Assoc. 84 (1989) 241–247.
[11] J.C. Lee, Prediction and estimation of growth curves with special covariance structures, J. Amer. Statist. Assoc. 83 (1988) 432–440.
[12] E.L. Lehmann, J.P. Romano, Testing Statistical Hypotheses, Springer, New York, 2005.
[13] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[14] M. Ohlson, D. von Rosen, Explicit estimators of parameters in the growth curve model with linearly structureed covariance matrices, J. Multivariate Anal. 101 (2010) 1284–1295.
[15] J.X. Pan, K.T. Fang, Growth Curve Models and Statistical Diagnostics, Science Press, Beijing, 2007.
[16] H.D. Patterson, R. Thompson, Recovery of interblock information when block sizes are unequal, Biometrika 58 (1971) 545–554.
[17] R.F. Potthoff, S.N. Roy, A generalized multivariate analysis of variance model useful especially for growth curve problems, Biometrika 51 (1964) 313–326.
[18] C.R. Rao, The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, Biometrika 52 (1965) 447–458.
[19] C.R. Rao, Prediction of future observations in growth curve models with comments, Statist. Sci. 4 (1987) 434–471.
[20] G. Reinsel, Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance struture, J. Amer. Statist. Assoc. 77 (1982) 190–195.
[21] M.S. Srivastava, Singular Wishart and multivariate beta distributions, Ann. Statist. 31 (2003) 1537–1560.
[22] D. von Rosen, Maximum likelihood estimators in multivariate linear normal models, J. Multivariate Anal. 31 (1989) 187–200.
[23] X. Wu, H. Liang, G. Zou, Unbiased invariant least squares estimation in a generalized growth curve model, Sankhyā A 71 (2009) 73–93.
[24] C.Y. Xu, W.L. Yang, The optiimality of a quadratic estimation in multivariate linear model (I), Chinese Ann. Math. Ser. A 4(5) (1983) 607–620.
[25] W.L. Yang, A result similar to multivariate Hsu's theorem, J. Beijing Normal Univ. (Natur. Sci.) (1) (1987) 1–7.
[26] W.L. Yang, The simultaneous estimation of $\mathrm{tr}(D'B) + \mathrm{tr}(C\Sigma)$ in the growth curve model, Sci. China Ser. A 37 (1994) 661–672.
[27] W.L. Yang, MINQE(U,I) and UMVIQUE of the covariance matrix in the growth curve model, Statistics 26 (1995) 49–59.
[28] W.L. Yang, W.J. Jiang, The best nonnegative estimator of covariance matrix in growth curve model, Acta Math. Appl. Sin. 15 (1992) 83–98.
[29] W.L. Yang, C.Y. Xu, The optimality of a quadratic estimation in multivariate linear model (II), Chinese Ann. Math. Ser. A 6(4) (1985) 505–513.