Estimation for parameters of interest in random effects growth curve models

Wai-Cheung Ip\textsuperscript{a,\ast}, Mi-Xia Wu\textsuperscript{b}, Song-Gui Wang\textsuperscript{b}, Heung Wong\textsuperscript{a}

\textsuperscript{a}Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong
\textsuperscript{b}College of Applied Sciences, Beijing University of Technology, Beijing 100022, China

Received 1 March 2005
Available online 8 September 2006

Abstract

In this paper, we consider the general growth curve model with multivariate random effects covariance structure and provide a new simple estimator for the parameters of interest. This estimator is not only convenient for testing the hypothesis on the corresponding parameters, but also has higher efficiency than the least-square estimator and the improved two-stage estimator obtained by Rao under certain conditions. Moreover, we obtain the necessary and sufficient condition for the new estimator to be identical to the best linear unbiased estimator. Examples of its application are given.

© 2006 Elsevier Inc. All rights reserved.

AMS 1991 subject classification: 62H12; 62H15; 62F10

Keywords: Growth curve models; Random-effects; Exact test; Best linear unbiased estimator; Maximum likelihood estimator

1. Introduction

The linear mixed model is a popular choice for the treatment of longitudinal data with random effects, see Laird and Ware [6], Diggle et al. [3], Srivastava and VonRosen [14] and Verbeke and Molenberghs [15]. In this paper, we consider the multivariate model in which \( m \) distinct characteristics on each of \( N \) individuals taken from \( r \) different groups are measured on each of \( p \) different occasions. The \( j \)th characteristic on the \( i \)th individual can be assumed to follow the mixed model

\[
y_{ij} = X_j \beta_j^{(k)} + Z_j u_{ij} + \varepsilon_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, m, \quad i \in k\text{th group},
\]

\( \ast \) Corresponding author. Fax: +852 23629045.
E-mail address: mathipwc@polyu.edu.hk (W.C. Ip).

0047-259X/S - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2006.04.001
where $y_{ijl} = (y_{ijl1}, \ldots, y_{ijlp})'$, $y_{ijl}$ is the measurement of the $j$th characteristic at occasion $l$ on the $i$th individual, $X_j$ and $Z_j$ are, respectively, known $p \times q$ and $p \times c$ design matrices of full column rank, $\beta_{jl}^{(k)}$ is the $q \times 1$ vector of regression parameters on the $j$th characteristic in the treatment group $k$ ($k = 1, \ldots, r$), $u_{ij}$ is $c \times 1$ random effect vector, and $\varepsilon_{ijl}$ is $p \times 1$ random error vector.

If each of the $m$ characteristics that we consider follows a response curve of the same general type over the $p$ occasions, that is, $X_j = X$ and $Z_j = Z$, then for $i$th individual,

$$y_i = (X \otimes I_m)\beta_j^{(k)} + (Z \otimes I_m)u_i + \varepsilon_i,$$

where

$$y_i = \text{Vec}((y_{i1}, \ldots, y_{im})'), \quad \beta_j^{(k)} = \text{Vec}((\beta_{i1}^{(k)}, \ldots, \beta_{im}^{(k)})'),$$

$$\varepsilon_i = \text{Vec}((\varepsilon_{i1}, \ldots, \varepsilon_{im})'), \quad u_i = \text{Vec}((u_{i1}, \ldots, u_{im})').$$

$\otimes$ denotes the Kronecker product, and $\text{Vec}(-)$ operator stacks the columns of a matrix below one another to form a column vector. The individual random effects $u_i$ are assumed to be distributed independently as $N(0, \Sigma_u)$, and independent of the error $\varepsilon_i$, with distribution $N(0, I_p \otimes \Sigma_e)$, where $\Sigma_e$ is a $m \times m$ positive definite matrix and $\Sigma_u$ is a $cm \times cm$ nonnegative definite matrix, that is, $\Sigma_e > 0$ and $\Sigma_u \geq 0$. Thus the covariance matrix of $y_i$ is

$$\text{Cov}(y_i) = I_p \otimes \Sigma_e + (Z \otimes I_m)\Sigma_u (Z' \otimes I_m) = \Omega \quad \text{(say)}.$$  \hspace{1cm} (2)

Let $Y = (y_1, \ldots, y_N)$, $U = (u_1, \ldots, u_N)$, and $E = (\varepsilon_1, \ldots, \varepsilon_N)$, the full multivariate model (1) can be expressed as

$$Y = (X \otimes I_m)\zeta A + (Z \otimes I_m)U + E,$$

$$\text{Cov}(\text{Vec}(Y)) = I_N \otimes \Omega,$$  \hspace{1cm} (3)

which is a general growth curve model with multivariate random effects covariance structure, where $\zeta = (\beta^{(1)}, \ldots, \beta^{(r)})$ is the $qm \times r$ matrix of the growth curve coefficients, and $A = (a_1, \ldots, a_N)$ is the $r \times N$ matrix with full row rank. In particular, if its elements are either 1 or 0 indicating the group from which an observation comes, $A$ is called the 'group indicator' matrix in literature. When $m = 1$, model (3) becomes $Y = X\zeta A + ZU + E$, and $\Omega = \sigma_e^2 I_p + Z\Sigma_u Z'$, reducing to the single-variable case.

Reinsel [11,12], Azzalini [1], Lange and Laird [7], and Nummi [8,9] considered two special cases: $Z = X$ and $Z = X_e$, where $X = (X_e : X_c)$, and showed that the maximum likelihood estimator (MLE) of $\zeta$ is identical to its least-squares estimator (LSE)

$$\hat{\zeta} = ((X'X)^{-1}X' \otimes I_m)Y A' (A A')^{-1}.$$  \hspace{1cm} (4)

However, this is not the case for a general design matrix $Z$, where the explicit MLE of $\zeta$ usually does not exist. Moreover, both the two-stage estimator provided by Khatri [5] and the LSE $\hat{\zeta}$ ignore the structure information on the covariance matrix $\Omega$. One alternative is to adopt the improved two-stage estimator suggested by Rao [10] using the structure covariance matrix of the mixed effect model.

In many practical situations, only a part of the parameters of model (3) are meaningful to the researcher. For example, in the rat data of Verbeke and Molenberghs [15], of primary interest is the estimation of changes over time and testing whether these changes are treatment dependent. For the mixed linear model with one random effect, Wu and Wang [17] gave a simple estimator
of the part parameter by the reduced model which is often used to remove nuisance parameters. In this paper, we mainly consider this problem under the general growth curve model (3) with random effects.

In Section 2, we introduce a new simple estimator for the part parameter of \( \xi \) by a reduced model, which can be superior to the LSE under certain conditions. In Section 3, we consider the optimality of the new estimator, and obtain the necessary and sufficient condition the new estimator to be identical to the best linear unbiased estimator (BLUE). Furthermore, we show that the new estimator is superior to the corresponding two-stage estimator under certain conditions. Examples are presented in Section 4.

2. New estimator

Without loss of generality, we partition \( X = (X_1 : X_2) \) in model (3) such that

\[
\mathcal{M}(X_1) \subseteq \mathcal{M}(Z), \quad \mathcal{M}(X_2) \cap \mathcal{M}(Z) = \{0\},
\]

where \( X_1 \) and \( X_2 \) are \( p \times l \) and \( p \times (q - l) \) matrices, respectively, \( (0 \leq l \leq q) \), and \( \mathcal{M}(C) \) is the range space of any matrix \( C \). Partition \( \tilde{\xi} = (\tilde{\xi}_1 : \tilde{\xi}_2)' \) conformably. Thus model (3) can be rewritten as

\[
Y = (X_1 \otimes I_m)\tilde{\xi}_1 A + (X_2 \otimes I_m)\tilde{\xi}_2 A + (Z \otimes I_m)U + E,
\]

where \( \tilde{\xi}_2 \) is the parameter matrix of primary interest. In what follows, we mainly consider the estimation of \( \tilde{\xi}_2 \).

Denote \( Q_1 \) the \( p \times (p - c) \) matrix such that \( Q_1'Q_1 = I_{p-c} \) and \( Q_1'Z = 0 \). Thus \( Q_1'Q_1' = I_p - P_Z \), where \( P_Z = Z(Z'Z)^{-1}Z' \) is the orthogonal projector on the space \( \mathcal{M}(Z) \). Let \( M_Z = I_p - P_Z = Q_1'Q_1' \), premultiplying model (6) by \( Q_1' \otimes I_m \), we can obtain the reduced model

\[
(Q_1' \otimes I_m)Y = (Q_1'X_2 \otimes I_m)\tilde{\xi}_2 A + \varepsilon, \quad \text{Cov} (\text{Vec} (\varepsilon)) = I_{N(p-c)} \otimes \Sigma_\varepsilon, (7)
\]

which is relevant to the matrix parameters \( \tilde{\xi}_2 \) and \( \Sigma_\varepsilon \). It is readily to see that, under model (7), the MLE of \( \tilde{\xi}_2 \) equals to its LSE

\[
\tilde{\xi}_2 = \left[(X_2'M_ZX_2)^{-1}X_2'M_Z \otimes I_m\right] Y A'(AA')^{-1}.
\]

Let \( \tilde{\xi}_2 \) be partitioned as \( \tilde{\xi}_2 = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)})' \), where \( \tilde{\xi}_{2i} (i = 1, \ldots, q-l) \) are \( m \times r \) matrices. Assume that the elements of \( X \) are functions of the time variable \( t \), such as \( X = (f_i(t_j)) \), then \( \tilde{\xi}_{2i} \) represent the regression coefficients attached to \( f_{i+1}(t), \ldots, f_{q-l}(t) \), for the \( m \) characteristics and \( r \) groups. Denote \( \Psi = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)}) \). Applying definition of the restricted maximum likelihood (REML) estimator, we can obtain the REML estimator of \( \Sigma_\varepsilon \) under reduced model (7), that is,

\[
\tilde{\Sigma}_\varepsilon = \frac{\sum_{i=1}^{N} (Y_iQ_1 - \tilde{\Psi}(X'_2Q_1 \otimes a_i))(Y_iQ_1 - \tilde{\Psi}(X'_2Q_1 \otimes a_i))'}{k},
\]

where \( k = N(p-c) - (q-l)r \), \( Y_i = (y_{i1}, \ldots, y_{im})' \), and \( \tilde{\Psi} = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)}) \).

Under the original model (6), estimators \( \tilde{\xi}_2 \) and \( \tilde{\Sigma}_\varepsilon \) are also unbiased for \( \tilde{\xi}_2 \) and \( \Sigma_\varepsilon \), respectively.
Theorem 2.1.
(a) \( \tilde{\xi}_2 \sim N(\tilde{\xi}_2, (AA')^{-1} \otimes (X'_2MZX_2)^{-1} \otimes \Sigma_e) \),
(b) \( k \cdot \tilde{\Sigma}_e \sim W_m(k, \Sigma_e) \), and independent of \( \tilde{\xi}_2 \),

where \( W_m(k, \Sigma_e) \) is the m-dimensional Wishart distribution with k degrees of freedom and parameter matrix \( \Sigma_e \).

Proof. (a) is obvious so we only need to prove (b). Noting \( \text{Vec}(AXC) = (C' \otimes A) \text{Vec}(X) \), thus model (7) can be rewritten as

\[
Y_i Q_1 = \Psi(X'_2Q_1 \otimes a_i) + E_i, \quad \text{Cov}(\text{Vec}(E_i)) = I_{p-c} \otimes \Sigma_e, \quad i = 1, \ldots, N,
\]

where \( Y_i \) is the same as that in (9), \( \Psi = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)}) \), and \( E_i = (\varepsilon_{i1}, \ldots, \varepsilon_{im})' Q_1 \).

Denote

\[
Y_0 = (Y_1Q_1, \ldots, Y_NQ_1), \quad X_0 = (X'_2Q_1 \otimes a_1, \ldots, X'_2Q_1 \otimes a_N), \quad E_0 = (E_1, \ldots, E_N),
\]

then model (7) can also be rewritten as

\[
Y_0 = \Psi X_0 + E_0, \quad \text{Cov}(\text{Vec}(E_0)) = I_{N(p-c)} \otimes \Sigma_e.
\]

Thus we get the equivalent forms of \( \tilde{\Psi} = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)}) \) and \( \tilde{\Sigma}_e \), respectively,

\[
\tilde{\Psi} = Y_0X'_0(X_0X'_0)^{-1} = \sum_{i=1}^{N} Y_i Q_1 \left( Q'_1 X_2(X'_2MZX_2)^{-1} \otimes a'_i(AA')^{-1} \right),
\]

\[
\tilde{\Sigma}_e = Y_0(I_{N(p-c)} - X'_0(X_0X'_0)^{-1}X_0)Y'_0/k,
\]

\[
= \left( \sum_{i=1}^{N} Y_i MZX'_i - \tilde{\Psi}(X'_2MZX_2 \otimes AA')\tilde{\Psi}' \right) / k.
\]

It follows readily from (11) that \( \tilde{\Sigma}_e \) is independent of \( \tilde{\Psi} \) and \( k \cdot \tilde{\Sigma}_e \sim W_m(k, \Sigma_e) \). The proof of Theorem 2.1 is completed. \( \square \)

Remark 2.1. \( \text{Cov}(\text{Vec}(\tilde{\xi}_2)) \) does not depend on the matrix \( \Sigma_u \).

According to Theorem 2.1, the new estimator \( \tilde{\xi}_2 \) can be used to construct an exact test on \( \Psi \), i.e. on \( \xi_2 \) for the general linear hypothesis \( H_0 : L\Psi G = 0 \). In fact, the Wilks’s \( \Lambda \) is given by

\[
\Lambda = \frac{|W|}{|W + H|},
\]

where \( W = k(L\tilde{\Sigma}_e L') \) and \( H = (L\tilde{\Psi}G) \left[ G' (X'_2MZX_2 \otimes AA')^{-1} G \right]^{-1} (G'\tilde{\Psi}'L') \).

On the other hand, from (4), it is easy to obtain that the LSE of \( \xi_2 \) under model (6) is

\[
\hat{\xi}_2 = \left( (X'_2 M_{X_1} X_2)^{-1} X'_2 M_{X_1} \otimes I_m \right) YA'(AA')^{-1},
\]

and

\[
\text{Cov}(\text{Vec}(\hat{\xi}_2)) = (AA')^{-1} \otimes \left[ (X'_2 M_{X_1} X_2)^{-1} \otimes \Sigma_e + \left( (X'_2 M_{X_1} X_2)^{-1} X'_2 M_{X_1} Z \otimes I_m \right) \Sigma_u \right.
\]

\[ \times \left. \left( Z'M_{X_1} X_2 (X'_2 M_{X_1} X_2)^{-1} \otimes I_m \right) \right]. \]
Combining with Remark 2.1, it is reasonable to expect that \( \tilde{\xi}_2 \) may be superior to \( \hat{\xi}_2 \) under some conditions. Suppose that \( X_1, X_2 \) and \( Z \) in (5) satisfy the following conditions:

\[
X_1'X_2 = 0, \quad Z = (X_1, Z_0), \quad \mathcal{M}(Z_0) \cap \mathcal{M}(X_i) = [0], \quad i = 1, 2,
\]

then \( M_ZX_1 = 0, M_{X_1}X_2 = X_2 \) and

\[
M_Z = M_{X_1} - M_{X_1}Z_0(Z_0'M_{X_1}Z_0)^{-1}Z_0'M_{X_1}.
\]

For the proof of the last equality see [13]. By the use of (14) and the matrix identity

\[
(A + BCB')^{-1} = A^{-1} - A^{-1}B(B'\hat{A}^{-1}B + C^{-1})^{-1}B'\hat{A}^{-1},
\]

we can obtain that

\[
(X_2'M_ZX_2)^{-1} = (X_2'M_ZX_2)^{-1} + (X_2'M_ZX_2)^{-1}X_2'Z_0(Z_0'M_{X}Z_0)^{-1}Z_0'X_2(X_2'M_ZX_2)^{-1}.
\]

Thus \( \text{Cov}(\text{Vec}(\tilde{\xi}_2)) \) can be rewritten as

\[
\text{Cov}(\text{Vec}(\tilde{\xi}_2)) = (AA')^{-1} \otimes \left[ (X_2'M_ZX_2)^{-1} \otimes \Sigma_e + \left( (X_2'M_ZX_2)^{-1}X_2'Z_0 \otimes I_m \right) \right.
\]

\[
\times \left. \left( (Z_0'M_{X}Z_0)^{-1} \otimes \Sigma_e \right) \left( Z_0'X_2(X_2'M_ZX_2)^{-1} \otimes I_m \right) \right].
\]

Furthermore, under the assumption (14), (13) can be simplified as

\[
\text{Cov}(\text{Vec}(\tilde{\xi}_2)) = (AA')^{-1} \otimes \left[ (X_2'M_ZX_2)^{-1} \otimes \Sigma_e \right.
\]

\[
\left. + \left( (X_2'M_ZX_2)^{-1}X_2'Z_0 \otimes I_m \right) (\Sigma_u)_{22} \left( Z_0'X_2(X_2'M_ZX_2)^{-1} \otimes I_m \right) \right],
\]

where \( (\Sigma_u)_{22} = (0, I_m(c-l)) \Sigma_u (0, I_m(c-l))' \) is \( m(c-l) \times m(c-l) \) nonnegative definite matrix.

Comparing (16) with (17), we obtain the condition for which the new estimator \( \tilde{\xi}_2 \) is superior to the LSE \( \hat{\xi}_2 \) in the following theorem.

**Theorem 2.2.** If conditions (14) hold, then \( \text{Cov}(\text{Vec}(\tilde{\xi}_2)) < \text{Cov}(\text{Vec}(\hat{\xi}_2)) \) if and only if

\[
X_2'Z_0 \neq 0 \quad \text{and} \quad ((Z_0'M_{X}Z_0)^{-1} \otimes \Sigma_e) < (\Sigma_u)_{22}.
\]

For example, \( X_1 = (1, 1, 1)' \), \( X_2 = (-1, 0, 1)' \), \( Z_0 = (1, 4, 9)' \), and \( (\Sigma_u)_{22} = a \cdot \Sigma_e \), then \( \tilde{\xi}_2 \) is superior to \( \hat{\xi}_2 \) only if \( a > \frac{3}{2} \).

**Remark 2.2.** The condition (18) is irrelevant to the other sub-matrices of \( \Sigma_u \) besides \( (\Sigma_u)_{22} \).

In the following section, we will show that the new estimator \( \tilde{\xi}_2 \) can also be superior to the improved two-stage estimator of \( \xi_2 \) under some conditions.

3. The optimality of the new estimator

In this section, we consider the optimality of the new estimator \( \tilde{\xi}_2 \) of \( \xi_2 \), the parameter matrix of primary interest, and give the necessary and sufficient condition under which \( \tilde{\xi}_2 \) is equal to the BLUE of \( \xi_2 \) for the original model (6).
Theorem 3.1. \( \tilde{\xi}_2 \) is the BLUE of \( \hat{\xi}_2 \) under model (6), if and only if

\[ Z'M_{X_1}X_2 = 0, \]  
where \( M_{X_1} = I_p - X_1(X_1'X_1)^{-1}X_1'. \)

Proof. By Lemma 5.4.3 [16], \( \tilde{\xi}_2 \) is the BLUE of \( \hat{\xi}_2 \) under model (6) if and only if

\[ \text{Cov}(\text{Vec}(\tilde{\xi}_2)), (I_n - P_{A'} \otimes P_X \otimes I_m) \text{Vec}(Y)) = 0, \]  
where \( n = Npm, \) that is,

\[ (AA')^{-1}A \otimes ((X_2'M_ZX_2)^{-1}X_2'M_Z \otimes I_m)\Omega(M_X \otimes I_m) = 0, \]

which is equivalent to

\[ X_2'M_ZM_X = 0. \]  
From (5), we have \( X_1'M_Z = 0, \) thus

\[ (21) \iff X'P_ZM_X = 0 \iff P_X P_Z M_X = 0 \iff P_X P_Z = P_{X_1}. \]

Combining Theorem 1 (A10, A11) of [2] and the fact that

\[ P_X = P_{X_1} + M_{X_1}X_2(X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}, \]  
Eq. (21) can been simplified to \( Z'M_{X_1}X_2 = 0. \) The proof of the Theorem 3.1 is completed.

The BLUE of \( \hat{\xi}_2 \) under model (6) is

\[ \hat{\xi}_2^* = C_{22,1}^{-1}C_2M_{C_1}^{-1}\Omega^{-1/2}Y'A'(AA')^{-1}, \]

where \( C_i = (X_i' \otimes I_m))\Omega^{-1/2}, i = 1, 2, C_{22,1} = C_2M_{C_1}C_2, M_{C_1} = I - C_1'(C_1'C_1)^{-1}C_1. \)

Obviously, \( \hat{\xi}_2^* \) is an usually unfeasible estimator since \( \Omega \) in (23) includes two unknown covariance matrices \( \Sigma_u \) and \( \Sigma_e. \) However, if condition (19) holds, then by Theorem 3.1, we have \( \hat{\xi}_2^* = \tilde{\xi}_2, \) that means the existence of the explicit ML estimator of the part parameter \( \hat{\xi}_2. \)

Theorem 3.2. Under model (6), the following statements are equivalent:

(a) \( Z'M_{X_1}X_2 = 0, \)
(b) \( \tilde{\xi}_2 = \hat{\xi}_2, \)
(c) \( \hat{\xi}_2 = \hat{\xi}_2^*. \)

Proof. (i) Proof of (a) \( \iff \) (c). Similar to the proof of Theorem 3.1, we have

\[ \tilde{\xi}_2 = \hat{\xi}_2^* \iff \text{Cov}(\text{Vec}(\tilde{\xi}_2)), (I_n - P_{A'} \otimes P_X \otimes I_m) \text{Vec}(Y)) = 0, \]

which can be simplified as

\[ (X_2'M_{X_1}Z \otimes I_m)\Sigma_u(Z'M_X \otimes I_m) = 0. \]  
Since \( \Sigma_u \geq 0 \) is arbitrary, (24) is equivalent to \( Z'M_{X_1}X_2 = 0 \) or \( Z'M_X = 0. \) From (5), we have

\[ Z'M_X = 0 \iff \mathcal{M}(Z) = \mathcal{M}(X_1) \Rightarrow Z'M_{X_1}X_2 = 0, \]

thus (24) is equivalent to \( Z'M_{X_1}X_2 = 0, \) that is, (a) \( \iff \) (c).
Theorem 3.1 and (26), we can obtain (27).

To be superior to the improved two-stage estimator of

where

Postmultiplying (25) by \( Z \), we have \((X'_2 M_{X_1} X_2)^{-1} X'_2 M_{X_1} Z = 0\), which is equivalent to \( Z'M_{X_1} X_2 = 0 \). The proof of Theorem 3.2 is completed. \( \square \)

Theorem 3.2 shows that the new estimator \( \hat{\xi}_2 \) and the LSE \( \hat{\xi}_2 \) can achieve optimality simultaneously, and both necessary and sufficient conditions are \( Z'M_{X_1} X_2 = 0 \).

**Remark 3.1.** The condition \( Z'M_{X_1} X_2 = 0 \) is weaker than the necessary and sufficient condition for \( \hat{\xi} = \xi^* \) under model (6).

In fact, by Lemma 5.4.3 of [16], we can testify that \( \hat{\xi} = \xi^* \) if and only if \( Z'X = 0 \) or \( \mathcal{M}(Z) = \mathcal{M}(X_1) \), which are two special cases for \( Z'M_{X_1} X_2 = 0 \).

Remark 3.1 sharpens the intuition that while the explicit MLE of the whole parameter matrix does not exist, it may do so for the part parameter matrix.

For the covariance matrix with random effects such as (2), Rao [10] presented an improved two-stage estimator of \( \xi \)

\[
\hat{\xi}_{T_1} = \hat{\xi} - ((X'X)^{-1} X' \otimes I_m) S Q (Q'SQ)^{-1} Q'Y A'(AA')^{-1},
\]

where \( Q = M_{XZ} \otimes I_m, T_1 = Q'Y, \) and \( S = Y(I_n - P_A)'Y' \). \( \hat{\xi}_{T_1} \) is derived by covariance adjustment in the LSE \( \hat{\xi} \) using the concomitant variable \( T_1 \), Rao [10] and Grizzle and Allen [4] proved

\[
\text{Cov}(\text{Vec}(\hat{\xi}_{T_1})) = \frac{N - r - 1}{N - r - mk_0 - 1} (AA')^{-1} \otimes \{(X' \otimes I_m) \Omega^{-1}(X \otimes I_m)\}^{-1}
\]

\[
= \frac{N - r - 1}{N - r - mk_0 - 1} \text{Cov}(\text{Vec}(\xi^*)) \geq \text{Cov}(\text{Vec}(\xi^*)),
\]

(26)

where \( k_0 = \text{rk}(M_{XZ}) \). Clearly, (26) takes the equality if and only if \( k_0 = 0 \), which requires \( \mathcal{M}(Z) \subseteq \mathcal{M}(X) \). In this case the improved two-stage estimator \( \hat{\xi}_{T_1} \) is equal to the LSE \( \hat{\xi} \), and \( \hat{\xi}_2 = \hat{\xi}_2 = \hat{\xi}_2 \), where \( \hat{\xi}_2T_1 \) is the corresponding improved two-stage estimator of \( \xi_2 \). \( \hat{\xi}_1T_1 = (\hat{\xi}'_{1T_1}, \hat{\xi}'_{2T_1})' \). In the following corollary, we will give a set of conditions for the new estimator \( \hat{\xi}_2 \) to be superior to the improved two-stage estimator \( \hat{\xi}_2T_1 \) under the case \( k_0 \neq 0 \).

**Corollary 3.1.** If \( \mathcal{M}(X) \cap \mathcal{M}(Z) = \mathcal{M}(X_1), \mathcal{M}(Z) \neq \mathcal{M}(X_1) \) and \( Z'M_{X_1} X_2 = 0 \), then

\[
\text{Cov}(\text{Vec}(\hat{\xi}_{2T_1})) = \text{Cov}(\text{Vec}(\xi^*)) = \text{Cov}(\text{Vec}(\hat{\xi}_2)) = \text{Cov}(\text{Vec}(\hat{\xi}_2)).
\]

(27)

The conditions \( \mathcal{M}(X) \cap \mathcal{M}(Z) = \mathcal{M}(X_1) \) and \( \mathcal{M}(Z) \neq \mathcal{M}(X_1) \) ensure \( k_0 \neq 0 \). Upon use of Theorem 3.1 and (26), we can obtain (27).

In order to understand the set of conditions in Corollary 3.1, without loss of generality, we assume \( Z = (X_1 : Z_0) \). By (5), we have

\[
\mathcal{M}(Z_0) \cap \mathcal{M}(X) = \{0\},
\]

and

\[
Z'M_{X_1} X_2 = 0 \iff Z_0'M_{X_1} X_2 = 0.
\]
According to Theorem 1 (A10, A11) of [2], $Z_0'M_{X_1}X_2 = 0$ is equivalent to $P_{Z_0}P_X = 0$, thus

$$Z'M_{X_1}X_2 = 0 \iff Z_0'X = 0,$$

and hence the conditions in Corollary 3.1 are equivalent to $Z_0'X = 0$, and $Z_0 \neq 0$. For example, $X_1 = (1, 1, 1, 1)'$, $X_2 = (1, 2, 3, 4)'$, $Z = (X_1, Z_0)$, where $Z_0 = (1, -1, -1, 1)'$, clearly, $Z_0 \neq 0$ and $Z_0'X = 0$.

4. Examples

In this section, we shall give two simple examples to illustrate the foregoing results.

Example 1. Consider the model for the rat data (see Verbeke and Molenberghs [14]):

$$y_{ij} = \begin{cases} 
\beta_{01} + u_{0j} + \beta_1 t_i + u_{1j}z_i + \epsilon_{ij} & \text{if low dose}, \\
\beta_{02} + u_{0j} + \beta_2 t_i + u_{1j}z_i + \epsilon_{ij} & \text{if high dose}, \\
\beta_{03} + u_{0j} + \beta_3 t_i + u_{1j}z_i + \epsilon_{ij} & \text{if control},
\end{cases}$$

(29)

where $y_{ij}$ is the observation for the $j$th individual at $t_i$ time point, $i = 1, \ldots, p$, $u_j = (u_{0j}, u_{1j})'$ is random effect vector with normal distribution $N(0, \Sigma_u)$, and random error $\epsilon_{ij}$ follows the distribution $N(0, \sigma^2_e)$. We assume that $u_1, \ldots, u_N, \epsilon_{11}, \ldots, \epsilon_{1N}, \ldots, \epsilon_{p1}, \ldots, \epsilon_{pN}$ are independent each other. Of primary interest is the estimation of the slopes $\beta_1, \beta_2$ and $\beta_3$, and testing whether these slopes are equal to each other.

Without loss of generality, we assume that $Y_1 = (y_{11}, \ldots, y_{n1})$, $Y_2 = (y_{n1+1}, \ldots, y_{n2})$ and $Y_3 = (y_{n2+1}, \ldots, y_N)$ are the matrices of the observations for the low-dose group, high-dose group and control group, respectively, where $y_j = (y_{ij}, \ldots, y_{pj})'$. Denote

$$Y = (Y_1 : Y_2 : Y_3), \quad U = (u_1, \ldots, u_N), \quad T = (t_1, \ldots, t_p)',$$

$$Z_0 = (z_1, \ldots, z_p)', \quad X = (1_p : T), \quad Z = (1_p : Z_0)$$

and

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \beta_{01} & \beta_{02} & \beta_{03} \\ \beta_1 & \beta_2 & \beta_3, \end{pmatrix}, \quad A = \begin{pmatrix} 1'_{n_1} & 0_{1 \times (n_2 - n_1)} & 0_{1 \times (N - n_2)} \\
0_{1 \times n_1} & 1'_{(n_2 - n_1)} & 0_{1 \times (N - n_2)} \\
0_{1 \times n_1} & 0_{1 \times (n_2 - n_1)} & 1'_{(N - n_2)} \end{pmatrix},$$

then model (4.2) can be rewritten as

$$Y = X\xi A + ZU + E = 1_p\xi_1 A + T\xi_2 A + ZU + E.$$

(30)

The covariance matrix of $Y$ is

$$\text{Cov}(\text{Vec}(Y)) = I_N \otimes (Z\Sigma_u Z' + \sigma^2_e I_p).$$

Denote $Q_1$ the $p \times (p - 2)$ matrix, such that $Q_1Q_1' = M_Z = I_p - P_Z$, then the reduced model of (30) may be represented by

$$Q_1'Y = Q_1'T\xi_2 A + Q_1'E, \quad \text{Cov}(\text{Vec}(Q_1'Y)) = I_N \otimes (\sigma^2_e I_{p - 2}).$$

(31)

By (8) and (9), we obtain the new estimators of $\xi_2 = (\beta_1, \beta_2, \beta_3)$ and $\sigma^2$

$$\bar{\xi}_2 = \frac{1}{T'M_ZT}T'M_Z(\bar{Y}_1 : \bar{Y}_2 : \bar{Y}_3),$$

$$\hat{\sigma}^2 = \frac{\text{Res}_v}{N - p - 2}$$
or

\[ \hat{\beta}_r = \frac{1}{T'M_ZT}T'M_Z\bar{Y}_r, \quad r = 1, 2, 3, \quad (32) \]

and

\[ \tilde{\sigma}_e^2 = \frac{3}{r} \sum SS_r/k, \quad (33) \]

where \( k = N(p - 2) - 3 \), \( \bar{Y}_r = (\bar{Y}_{1r}, \ldots, \bar{Y}_{pr})' \),

\[ \bar{Y}_{i1} = \sum_{j=1}^{n_1} y_{ij}/n_1, \quad \bar{Y}_{i2} = \sum_{j=n_1+1}^{n_2} y_{ij}/(n_2 - n_1), \quad \bar{Y}_{i3} = \sum_{j=n_2+1}^{N} y_{ij}/(N - n_2), \]

\[ SS_1 = \sum_{j=1}^{n_1} (y_j - \tilde{\beta}_1 T)'M_Z(y_j - \tilde{\beta}_1 T), \]

\[ SS_2 = \sum_{j=n_1+1}^{n_2} (y_j - \tilde{\beta}_2 T)'M_Z(y_j - \tilde{\beta}_2 T), \quad \text{and} \]

\[ SS_3 = \sum_{j=n_2+1}^{N} (y_j - \tilde{\beta}_3 T)'M_Z(y_j - \tilde{\beta}_3 T). \]

By Theorem 2.1, \( \tilde{\zeta}_2 \) and \( \tilde{\sigma}_e^2 \) are independent, based on which we can construct a test statistic for testing \( H_0: \beta_1 = \beta_2 = \beta_3 \). Let

\[ H' = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \]

then \( H_0 \) is equivalent to testing \( \tilde{\zeta}_2 H = 0 \). The exact test statistic is

\[ F = \frac{\tilde{\zeta}_2 H(H'(AA')^{-1}H)^{-1}H'\tilde{\zeta}_2'/2}{\tilde{\sigma}_e^2(T'M_ZT)}, \]

which is an \( F \)-statistic and hence \( F \sim F_{2,k} \) if \( \tilde{\zeta}_2 H = 0 \) holds.

By Theorem 2.2, under the case \( T'Z_0 \neq 0 \), if the variance of the error \( \sigma^2 \) satisfies

\[ \sigma_e^2 < (Z_0'M_XZ_0)(\Sigma_u)_{22} = (Z_0'M_XZ_0)\text{Var}(u_{1,j}), \quad (34) \]

then \( \tilde{\zeta}_2 \) is superior to the LSE of \( \zeta_2 \)

\[ \hat{\zeta}_2 = \frac{1}{(T - \tilde{\gamma} \cdot 1_p)'(T - \tilde{\gamma} \cdot 1_p)}(T - \tilde{\gamma} \cdot 1_p)'(\bar{Y}_1 : \bar{Y}_2 : \bar{Y}_3). \quad (35) \]

By Theorem 3.2, if \( T'Z_0 = 0 \) and \( 1_pZ_0 = 0 \), then \( \tilde{\zeta}_2 = \hat{\zeta}_2 \), and \( \tilde{\zeta}_2 \) is the BLUE of \( \zeta_2 \) under model (29). Furthermore, if \( Z_0 \neq 0 \), by Corollary 3.1, then \( \tilde{\zeta}_2 \) is superior to the improved two-stage estimator of \( \tilde{\zeta}_2 \).
Example 2. Consider the model for systolic and diastolic blood pressure data:

\[
y_{ijl} = \begin{cases} 
\beta_{0j}^{(1)} + u_{ij} + \beta_{j}^{(1)} x_l + v_{ij} t_l + \varepsilon_{ijl} & \text{if treatment 1 group,} \\
\beta_{0j}^{(2)} + u_{ij} + \beta_{j}^{(2)} x_l + v_{ij} t_l + \varepsilon_{ijl} & \text{if treatment 2 group,}
\end{cases}
\]  

(36)

where \(y_{i1l}, y_{i2l}\) are the systolic and diastolic blood pressure for the \(i\)th patients (with moderate essential hypertension) at \(t_l\) time point, respectively, \(\beta_{0j}^{(1)}\) and \(\beta_{j}^{(1)}\) are fixed intercepts, \(\beta_{j}^{(1)}\) are the fixed effects of treatment 1 and 2 on systolic and diastolic blood pressure. According to Theorem 2.1, we can obtain the actual test statistic on time point, respectively, \(u_i = (u_{i1}, u_{i2}, v_{i1}, v_{i2})\) follows distribution \(N(0, \Sigma_u)\), and independent of the error \(\varepsilon_i = (\varepsilon_{i11}, \varepsilon_{i21}, \ldots, \varepsilon_{i1p}, \varepsilon_{i2p})\)' with distribution \(N(0, I_p \otimes \Sigma_e)\).

Clearly, (36) is the case of model (3) with \(m = 2\),

\[
\zeta = \begin{pmatrix} \varepsilon_{1l} \\ \varepsilon_{2l} \end{pmatrix} = \begin{pmatrix} \beta_{0j}^{(1)} & \beta_{0j}^{(2)} \\ \beta_{1j}^{(1)} & \beta_{1j}^{(2)} \\ \beta_{2j}^{(1)} & \beta_{2j}^{(2)} \end{pmatrix}, \quad A = \begin{pmatrix} 1_{n_1} & 0_{1 \times (N-n_1)} \\ 0_{1 \times n_1} & 1_{(N-n_1)} \end{pmatrix},
\]  

(37)

\[
X = (1_p, x), \quad Z = (1_p, t), \quad U = (u_1, \ldots, u_N).
\]

Here \(x = (x_1, \ldots, x_p)'\) and \(t = (t_1, \ldots, t_p)\) are designs vectors on dose and time point.

Of primary interest is the testing on \(\xi_2\), the effects of treatment 1 and 2 on systolic and diastolic blood pressure. According to Theorem 2.1, we can obtain the actual test statistic on \(\xi_2\), Wilks’ \(\Lambda\) statistic (12), which is constructed by the new estimator \(\tilde{\xi}_2\) given by (8).

Otherwise, the new estimator of \(\xi_2\) is superior to the LSE and two-stage estimator under the conditions of Theorem 2.2 and Corollary 3.1. See the following typical designs on dose \(x\) and time point \(t\).

Take \(x = (1, 1, 0, -1, -1)\), \(t = (1, 2, 3, 4, 5)\), then conditions (18) can be equivalent to

\[
\Sigma_e = \text{Cov}((\varepsilon_{i1}, \varepsilon_{i2})') < \text{Cov}((u_{i1}, u_{i2})').
\]  

(38)

That is, if the covariance matrix of error is lower than the covariance matrix of time effect under Löwner partial ordering, then the new estimator of \(\xi_2\) is superior to the LSE.

Take \(x = (0, 2, 2, 2, 0)\), \(t = (-2, -1, 0, 1, 2)\), it is easy to verify that

\[
Z'M_{X_1}x = (0, t'x)' = (0, 0)',
\]

so the new estimator of \(\xi_2\) is superior to the two-stage estimator provided by Rao.

The above two examples show that Theorem 2.2 and Corollary 3.1 are helpful for the study on optimal design of experiments too.

Acknowledgments

This research was supported by a grant from The Hong Kong Polytechnic University Research Committee. Wu’s and Wang’s research was also partially supported by National Natural Science...
Foundation of China. We are grateful to the two referees for their detailed suggestions which considerably improved the quality of the paper.

References