Ann. Inst. Statist. Math.
Vol. 51, No. 4, 707–721 (1999)

GROWTH CURVE MODEL WITH HIERARCHICAL WITHIN-INDIVIDUALS DESIGN MATRICES

YASUNORI FUJIKOSHI[1], TAKASHI KANDA[2] AND MEGU OHTAKI[3]

[1] *Department of Mathematics, Hiroshima University, Kagamiyama, Higashi Hiroshima 739-8526, Japan*
[2] *Department of Environmental Design, Hiroshima Institute of Technology, Miyake, Saeki-ku, Hiroshima 731-5193, Japan*
[3] *Department of Environmentrics and Biometrics, Research Institute for Radiation Biology and Medicine, Hiroshima University, Kasumi, Minami-ku, Hiroshima 734-8553, Japan*

(Received August 21, 1997; revised May 20, 1998)

Abstract. This paper deals with some inferential problems under an extended growth curve model with several hierarchical within-individuals design matrices. The model includes the one whose mean structure consists of polynomial growth curves with different degrees. First we consider the case when the covariance matrix is unknown positive definite. We derive a LR test for examining the hierarchical structure for within-individuals design matrices and a model selection criterion. Next we consider the case when a random coefficients covariance structure is assumed, under certain assumption of between-individual design matrices. Similar inferential problems are also considered. The dental measurement data (see, e.g., Potthoff and Roy (1964, *Biometrika*, **51**, 313–326)) is reexamined, based on extended growth curve models.

Key words and phrases: Extended growth curve model, hierarchical within-individuals design matrices, inferential problems, random-coefficient model.

1. Introduction

In this paper we are concerned with an extended growth curve model with several hierarchical within-individuals design matrices. For an $N \times p$ random matrix Y, the model is defined by

$$Y = A_1\Xi_1X_{(1)} + \cdots + A_k\Xi_kX_{(k)} + \mathcal{E}, \qquad E(\mathcal{E}) = O, \tag{1.1}$$

where A_i are known $N \times r_i$ between-individuals design matrices of ranks r_i, $X_{(i)}$ are known $q_i \times p$ within-individuals design matrices with ranks q_i, Ξ_i are unknown $r_i \times q_i$ parameter matrices, and $\mathcal{E}$ is an error matrix. Here it is assumed that $X_{(i)}$ has a hierarchical structure

$$X_{(i)} = \begin{bmatrix} X_1 \\ \vdots \\ X_i \end{bmatrix}, \qquad i = 1, \ldots, k, \tag{1.2}$$

and hence $q_1 \leq \cdots \leq q_k$. By appropriately modifying between-individuals design matrices, without loss of generality we can assume $q_1 < \cdots < q_k$. First we consider the case when the rows of $\mathcal{E}$ are independently distributed as a p-variate normal distribution having unknown positive definite covariance matrix Σ, i.e., $\mathcal{E} \sim N_{N\times p}(O, I_N \otimes \Sigma)$. Then the model is expressed as

$$Y \sim N_{N\times p}(A_1\Xi_1 X_{(1)} + \cdots + A_k\Xi_k X_{(k)}, I_N \otimes \Sigma). \tag{1.3}$$

The model is an extension of the growth curve model by Potthoff and Roy (1964), but a special case of Verbyla and Venables (1988), etc. It may be noted that the model (1.1) includes an important one whose mean structure consists of polynomial growth curves with different degrees. In this case the between-individuals design matrices have a form of

$$A_1 = \begin{bmatrix} A_{11} \\ O \\ \vdots \\ O \end{bmatrix}, \ldots, A_k = \begin{bmatrix} O \\ \vdots \\ O \\ A_{kk} \end{bmatrix}, \tag{1.4}$$

where $A_{ii} : N_i \times r_i$ and $N_1 + \cdots + N_k = N$. Let $Y = [Y_1' \cdots Y_k']'$, $Y_i : N_i \times p$. Then, the special model is expressed as

$$Y_i \sim N_{N_i\times p}(A_{ii}\Xi_i X_{(i)}, I_{N_i} \otimes \Sigma), \qquad i = 1, \ldots, k, \tag{1.5}$$

where $Y_1, \ldots, Y_k$ are mutually independent.

Random-coefficient models for Potthoff and Roy (1964) were considered by Rao (1965) and Fearn (1975), etc. A natural random-coefficient model for (1.1) is defined as

$$Y_i = A_{ii}\Xi_i X_{(i)} + \mathcal{V}_i X_{(i)} + \mathcal{E}_i, \qquad i = 1, \ldots, k, \tag{1.6}$$

where $\mathcal{V}_1, \ldots, \mathcal{V}_k, \mathcal{E}_1, \ldots, \mathcal{E}_k$ are mutually independent, and

$$\mathcal{V}_i \sim N_{N_i\times q_i}(O, I_{N_i} \otimes \Delta^{(i)}), \qquad \mathcal{E}_i \sim N_{N_i\times q_i}(O, I_{N_i} \otimes \sigma^2 I_p). \tag{1.7}$$

Here $\Delta^{(i)}$ is the first $q_i \times q_i$ submatrix of $\Delta = \Delta^{(k)}$. It is assumed that $\Delta^{(i)}$ is unknown positive semi-definite, and σ^2 is unknown positive constant. Vonesh and Carter (1987) have considered a random-coefficient model for unbalanced longitudinal data. However, it may be noted that our model is not a special one of the model as in Vonesh and Carter (1987), in the sense that the random-coefficient components consist of k random matrices with different dimensions and a hierarchical structure.

The paper is organized in the following way. In Section 2 we see that the model (1.3) and (1.6) can be viewed from other points of views. Their canonical forms are also given. In Section 3 we consider the model (1.3). We derive a LR test for examining the hierarchical structure for within-individuals design matrices and a model selection criterion. In Section 4 we discuss similar inferential problems under the random-coefficient model (1.6). In Section 4 we illustrate our results in a special model and reexamine the dental measurements data (see, e.g., Potthoff and Roy (1964)) by considering the model (1.3).

2. Canonical forms for the models (1.3) and (1.6)

An extended growth curve model has been introduced by considering some prior restrictions in the growth curve model due to Potthoff and Roy (1964), see, e.g., Banken (1984), Kariya (1985), etc. Related to this, consider the ordinary growth curve model given by

$$
\begin{aligned}
(2.1) \qquad E(Y) &= A\Xi X \\
&= [A_1 \cdots A_k] \begin{bmatrix} \Xi_{11} & \cdots & \Xi_{1k} \\ \vdots & \ddots & \vdots \\ \Xi_{k1} & \cdots & \Xi_{kk} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix},
\end{aligned}
$$

where $A_i : N \times r_i$, $X_i : b_i \times p$, $b_1 = q_1$, $b_2 = q_2 - q_1, \ldots, b_k = q_k - q_{k-1}$. Then the mean structure in the model (1.1) can be expressed as (2.1) with the following restrictions

$$
(2.2) \qquad \forall i = 1, \ldots, k-1,\ \forall j = i+1, \ldots, k : C_i \Xi D_j = O,
$$

where $C_i = [O \cdots O\ I_{r_i}\ O \cdots O]$, $D_j = [O \cdots O\ I_{b_j}\ O \cdots O]'$, $i, j = 1, \ldots, k$.

We can also write the mean structure (1.1) as

$$
(2.3) \qquad E(Y) = A_{[1]}\Xi_{[1]}X_{(1)} + \cdots + A_{[k]}\Xi_{[k]}X_{(k)},
$$

where

$$
A_{[i]} = [A_i \cdots A_k], \qquad \Xi_{[i]} = \begin{bmatrix} \Xi_{ii} \\ \vdots \\ \Xi_{ki} \end{bmatrix}.
$$

Here we note that $A_{[i]}$'s and $X_{(i)}$'s satisfy

$$
(2.4) \qquad \mathcal{R}[A_{[1]}] \supset \cdots \supset \mathcal{R}[A_{[k]}]
$$

and

$$
(2.5) \qquad \mathcal{R}[X'_{(1)}] \subset \cdots \subset \mathcal{R}[X'_{(k)}],
$$

respectively, where $\mathcal{R}[A]$ denotes the space spanned by the column vectors of A. von Rosen (1989, 1990) has considered an extended growth curve model (2.3) with (2.4) and any $X_{(i)}$'s, and has studied the MLE's of $E(Y)$ and Σ.

Anderson *et al.* (1993) have studied a general totally ordered multivariate linear model including the above models, relating to an algebraic condition of invariance under a full block-triangular group. Banken (1984) has given a canonical form for the growth curve model (2.1) with a general set of restrictions as in (2.2). The set of our restrictions is a special case and hence we will get a canonical form for the model (1.3). However, in order to get a clear correspondence between the original expressions and the transformed expressions, we give a direct derivation, starting from the model (1.3). Applying the Gram-Schmidt orthogonalization method to $A = [A_1 \cdots A_k]$ and $X = X_{(k)}$, we can choose $H = [H_1 \cdots H_\ell] \in O(N)$ and $B = [B_1' \cdots B_\ell']' \in O(p)$ such that

$$A = [H_1 \cdots H_k] \begin{bmatrix} L_{11} & O & \cdots & O \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & O \\ L_{k1} & \cdots & \cdots & L_{kk} \end{bmatrix} = [H_1 \cdots H_k] L \tag{2.6}$$

and

$$X = \begin{bmatrix} G_{11} & O & \cdots & O \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & O \\ G_{k1} & \cdots & \cdots & G_{kk} \end{bmatrix} \begin{bmatrix} B_1 \\ \vdots \\ B_k \end{bmatrix} = G \begin{bmatrix} B_1 \\ \vdots \\ B_k \end{bmatrix}, \tag{2.7}$$

where $H_i : N \times r_i$, $B_i : b_i \times p$, $\ell = k+1$, $n = N - r_1 - \cdots - r_k$ and $O(m)$ denotes the set of all the orthogonal matrices of order m. Here the notation ℓ is used for denoting $k+1$ simply. Consider the transformation from Y to

$$\begin{aligned} Z &= H'YB' \\ &= \begin{bmatrix} Z_{11} & \cdots & Z_{1k} & Z_{1\ell} \\ \vdots & & \vdots & \vdots \\ Z_{k1} & \cdots & Z_{kk} & Z_{k\ell} \\ Z_{\ell 1} & \cdots & Z_{\ell k} & Z_{\ell\ell} \end{bmatrix}, \end{aligned} \tag{2.8}$$

where $Z_{ij} = H_i'YB_j : r_i \times b_j$, $r_\ell = n$ and $b_\ell = p - q_k$. Let

$$\Theta = L\Xi G = \begin{bmatrix} \Theta_{11} & O & \cdots & O \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & O \\ \Theta_{k1} & \cdots & \cdots & \Theta_{kk} \end{bmatrix}, \tag{2.9}$$

$$\Omega = B\Sigma B' = \begin{bmatrix} \Omega_{11} & \cdots & \Omega_{1\ell} \\ \vdots & & \vdots \\ \Omega_{\ell 1} & \cdots & \Omega_{\ell\ell} \end{bmatrix}. \tag{2.10}$$

Then, for the transformed matrix Z we have

$$Z \sim N_{N\times p}(E(Z), I_N \otimes \Omega), \tag{2.11}$$

where Ω is an unknown positive definite matrix and

$$E(Z) = \begin{bmatrix} \Theta_{11} & O & \cdots & O \\ \vdots & \ddots & \ddots & \vdots \\ \Theta_{k1} & \cdots & \Theta_{kk} & O \\ O & \cdots & O & O \end{bmatrix}. \tag{2.12}$$

As a significance test for the hierarchical structure (1.2) we consider to test the null hypothesis $H_0 : \Xi_{[2]} = O, \ldots, \Xi_{[k]} = O$, which is equivalent to

$$H_0 : \Theta_{[2]} = O, \ldots, \Theta_{[k]} = O, \tag{2.13}$$

where $\Theta_{[i]} = [\Theta'_{ii}, \ldots, \Theta'_{ki}]'$, $i = 1, \ldots, k$.

Next we consider a canonical form for the model (1.6). Consider the transformation from Y_i to

$$[U_i\ V_i] = Y_i[X'_{(i)}R^{(i)}\ \tilde{X}'_{(i)}], \tag{2.14}$$

where $R^{(i)} = (X_{(i)}X'_{(i)})^{-1}$ and $\tilde{X}_{(i)}$ is a $(p-q_i) \times p$ matrix satisfying $X_{(i)}\tilde{X}'_{(i)} = O$ and $\tilde{X}_{(i)}\tilde{X}'_{(i)} = I_{p-q_i}$. Such matrices may be defined in terms of B and G in (2.7) as

$$X'_{(i)}R^{(i)} = B_{(i)}G_{(i)}, \qquad R^{(i)} = (G_{(i)}G'_{(i)})^{-1}, \tag{2.15}$$

where

$$B_{(i)} = \begin{bmatrix} B_1 \\ \vdots \\ B_i \end{bmatrix}, \qquad G_{(i)} = \begin{bmatrix} G_{11} & O & \cdots & O \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & O \\ G_{i1} & \cdots & \cdots & G_{ii} \end{bmatrix}.$$

Then, we have

$$[U_i\ V_i] \sim N_{N_i \times p}\left[(A_{ii}\Xi_i\ O), I_{N_i} \otimes \begin{bmatrix} \Psi^{(i)} & O \\ O & \sigma^2 I_{p-q_i} \end{bmatrix}\right], \tag{2.16}$$

where $\Psi^{(i)} = \Delta^{(i)} + \sigma^2 R^{(i)}$.

3. Inferential problems under the model (1.3)

3.1 *MLE's*

The MLE's under extended growth curve models have been studied by Banken (1984), von Rosen (1989, 1990), Fujikoshi and Satoh (1996), etc. Here we attempt to give simple expressions of the MLE's of Θ and Ω, especially certain submatrices of Θ and certain transformed matrices of Ω, starting the canonical form (2.11). From these expressions we shall obtain distributional results on the corresponding MLE's. We use the following notations:

$$V^{(0)} = (Z_{\ell 1}, \ldots, Z_{\ell\ell})'(Z_{\ell 1}, \ldots, Z_{\ell\ell}),$$

and for $i = 1, \ldots, k$,

$$\begin{aligned} U^{(i)} &= (Z_{i1}, \ldots, Z_{i\ell})'(Z_{i1}, \ldots, Z_{i\ell}), \\ V^{(i)} &= V^{(0)} + U^{(1)} + \cdots + U^{(i)} \\ &= \begin{bmatrix} V_{11}^{(i)} & \cdots & V_{1\ell}^{(i)} \\ \vdots & \ddots & \vdots \\ V_{\ell 1}^{(i)} & \cdots & V_{\ell\ell}^{(i)} \end{bmatrix}, \\ Z_{i(1\cdots i)} &= [Z_{i1}, \ldots, Z_{ii}], \\ Z_{i(i+1\cdots\ell)} &= [Z_{ii+1}, \ldots, Z_{i\ell}], \\ V_{ii\cdot i+1\cdots\ell}^{(i-1)} &= V_{ii}^{(i-1)} - V_{i(i+1\cdots\ell)}^{(i-1)}\{V_{(i+1\cdots\ell)(i+1\cdots\ell)}^{(i-1)}\}^{-1}V_{(i+1\cdots\ell)i}^{(i-1)}. \end{aligned}$$

Similar notations are used for matrices of Θ, Ω, etc., partitioned in the same way. Further let

$$\begin{aligned}\mathcal{B}_{(i+1\cdots\ell)(1\cdots i)} &= \Omega^{-1}_{(i+1\cdots\ell)(i+1\cdots\ell)}\Omega_{(i+1\cdots\ell)(1\cdots i)} \\ &= (\mathcal{B}_{(i+1\cdots\ell)1}\cdots\mathcal{B}_{(i+1\cdots\ell)i}), \qquad (i=1,\ldots,k).\end{aligned}$$

It is well known that there exists a one-to-one correspondence between Ω and the set $\{\Omega_{11\cdot2\cdots\ell},\ldots,\Omega_{kk\cdot\ell},\Omega_{\ell\ell},\mathcal{B}_{(2\cdots\ell)1},\ldots,\mathcal{B}_{(k\ell)k-1},\mathcal{B}_{\ell k}\}$.

First we maximize the joint density $f(Z;\Theta,\Omega)$ of Z with respect to Θ. Considering the conditional density of $Z_{i(1\cdots i)}$ given $Z_{i(i+1\cdots\ell)}$ we can see that the maximum occurs at

$$\begin{aligned}(3.1)\qquad \tilde{\Theta}_i &= [\tilde{\Theta}_{i1},\ldots,\tilde{\Theta}_{ii}] \\ &= Z_{i(1\cdots i)} - Z_{i(i+1\cdots\ell)}\mathcal{B}_{(i+1\cdots\ell)(1\cdots i)}, \qquad i=1,\ldots,k.\end{aligned}$$

Then

$$\begin{aligned}(3.2)\qquad g(\Omega) &= f(Z;\tilde{\Theta},\Omega) \\ &= (2\pi)^{-pN/2}|\Omega|^{-N/2} \\ &\quad\times \exp\Big[-\frac{1}{2}\{\operatorname{tr}\Omega^{-1}V^{(0)} + \operatorname{tr}\Omega^{-1}_{(2\cdots\ell)(2\cdots\ell)}U^{(1)}_{(2\cdots\ell)(2\cdots\ell)} + \cdots \\ &\qquad\qquad + \operatorname{tr}\Omega^{-1}_{\ell\ell}U^{(k)}_{\ell\ell}\}\Big].\end{aligned}$$

For the maximization of (3.2), Gleser and Olkin (1970) studied the case $k=2$. The maximization under the general case can be obtained by using their idea repeatedly. This reduction has been used by Banken (1984). Summarizing these results, we have the following results.

THEOREM 3.1. *The MLE's of* Θ_i, $\mathcal{B}_{(i+1\cdots\ell)i}$, $\Omega_{ii\cdot(i+1\cdots\ell)}$, $i=1,\ldots,k$, $\Omega_{\ell\ell}$ *are given as follows*:

$$(3.3)\qquad \begin{aligned}\hat{\Theta}_i &= Z_{i(1\cdots i)} - Z_{i(i+1\cdots\ell)}\hat{\mathcal{B}}_{(i+1\cdots\ell)(1\cdots i)}, \\ \hat{\mathcal{B}}_{(i+1\cdots\ell)i} &= \{V^{(i-1)}_{(i+1\cdots\ell)(i+1\cdots\ell)}\}^{-1}V^{(i-1)}_{(i+1\cdots\ell)i}, \qquad i=1,\ldots,k, \\ N\hat{\Omega}_{ii\cdot(i+1\cdots\ell)} &= V^{(i-1)}_{ii\cdot(i+1\cdots\ell)}, \\ N\hat{\Omega}_{\ell\ell} &= V^{(k)}_{\ell\ell}.\end{aligned}$$

Further the maximum of the likelihood is

$$\begin{aligned}(3.4)\qquad g(\hat{\Omega}) &= (2\pi Ne)^{-pN/2}\{|N\hat{\Omega}_{11\cdot(2\cdots\ell)}|\cdot|N\hat{\Omega}_{22\cdot(3\cdots\ell)}|\cdots|N\hat{\Omega}_{\ell\ell}|\}^{-pN/2} \\ &= (2\pi Ne)^{-pN/2}\{|V^{(0)}_{11\cdot(2\cdots\ell)}|\cdot|V^{(1)}_{22\cdot(3\cdots\ell)}|\cdots|V^{(k)}_{\ell\ell}|\}^{-pN/2}.\end{aligned}$$

Theorem 3.1 implied the following distributional results.

THEOREM 3.2.

(i) $$\hat{\Theta}_{ii} - \Theta_{ii} = [Z_{ii} - \Theta_{ii}\ Z_{i(i+1\cdots\ell)}]\begin{bmatrix} I_{b_i} \\ -\hat{\mathcal{B}}_{(i+1\cdots\ell)i}\end{bmatrix},$$

where $\hat{\mathcal{B}}_{(i+1\cdots\ell)i} = \{V^{(i-1)}_{(i+1\cdots\ell)(i+1\cdots\ell)}\}^{-1}V^{(i-1)}_{(i+1\cdots\ell)i}$. *Here* $[Z_{ii} - \Theta_{ii}\ Z_{i(i+1\cdots\ell)}]$ *and* $V^{(i-1)}_{(i\cdots\ell)(i\cdots\ell)}$ *(hence* $\hat{\mathcal{B}}_{(i+1\cdots\ell)i}$*) are mutually independent,*

$$[Z_{ii} - \Theta_{ii}\ Z_{i(i+1\cdots\ell)}] \sim N_{r_i\times(b_{i+1}+\cdots+b_\ell)}[O, \Omega_{(i\cdots\ell)(i\cdots\ell)} \otimes I_{r_i}], \qquad and$$
$$V^{(i-1)}_{(i\cdots\ell)(i\cdots\ell)} \sim W_{(b_i+\cdots+b_\ell)}(n + r_1 + \cdots + r_{i-1}, \Omega_{(i\cdots\ell)(i\cdots\ell)}).$$

(ii) $N\hat{\Omega}_{11\cdot(i\cdots\ell)(i\cdots\ell)}, \ldots, N\hat{\Omega}_{kk\cdot\ell}$, $N\hat{\Omega}_{\ell\ell}$ *are mutually independent and for* $i = 1, \ldots, k$,

$$N\hat{\Omega}_{ii\cdot(i+1\cdots\ell)} \sim W_{b_i}(n + r_1 + \cdots + r_{i-1} - b_{i+1} - \cdots - b_\ell, \Omega_{ii\cdot(i+1\cdots\ell)}),$$
$$N\hat{\Omega}_{\ell\ell} \sim W_{b_i}(n + r_1 + \cdots + r_k, \Omega_{\ell\ell}).$$

3.2 *LR test*

We consider to test the hypothesis (2.13). Let $f_0(Z; \Xi_{[1]}, \Omega)$ be the density function of Z under the hypothesis (2.13). The maximum of $f_0(Z; \Xi_{[1]}, \Omega)$ with respect to $\Xi_{[1]}$ is given by

$$g_0(\Omega) = (2\pi)^{-pN/2}|\Omega|^{-N/2} \cdot \exp\left[-\frac{1}{2}\{\mathrm{tr}\,\Omega^{-1}V^{(0)} + \mathrm{tr}\,\Omega^{-1}_{(2\cdots\ell)(2\cdots\ell)}(U^{(1)} + \cdots + U^{(k)})_{(2\cdots\ell)(2\cdots\ell)}\}\right]$$

By the same way as in the maximization of (3.2), we have

$$\sup_{\Omega>O} g_0(\Omega) = (2\pi Ne)^{-pN/2}\{|V^{(0)}_{11\cdot2\cdots\ell}| \cdot |V^{(k)}_{(2\cdots\ell)(2\cdots\ell)}|\}^{-N/2}.$$

Therefore, the LR criterion is an increasing function of

(3.5) $$\Lambda = \frac{|V^{(1)}_{22\cdot3\cdots\ell}|\cdots|V^{(k-1)}_{kk\cdot\ell}|\cdot|V^{(k)}_{\ell\ell}|}{|V^{(k)}_{(2\cdots\ell)(2\cdots\ell)}|} = \frac{|V^{(1)}_{22\cdot3\cdots\ell}|\cdots|V^{(k-1)}_{kk\cdot\ell}|}{|V^{(k)}_{22\cdot3\cdots\ell}|\cdots|V^{(k)}_{kk\cdot\ell}|} = \Lambda_{(1)}\cdots\Lambda_{(k-1)},$$

where

(3.6) $$\Lambda_{(i-1)} = |V^{(i-1)}_{ii\cdot(i+1)\cdots\ell}|/|V^{(k)}_{ii\cdot(i+1)\cdots\ell}|, \qquad i = 2, \ldots, k.$$

THEOREM 3.3. *The LR criterion for testing the hypothesis* (2.13) *is an increasing function of* $\Lambda = \Lambda_{(1)} \cdots \Lambda_{(k-1)}$, *where* $\Lambda_{(i)}$*'s are defined by* (3.6). *Under the hypothesis* $\Lambda_{(i)}$*'s are mutually independent, and* $\Lambda_{(i)}$ *is distributed as a Wilks* Λ*-distribution* $\Lambda_{b_{i+1}}(r_{i+1} + \cdots + r_k, n + r_1 + \cdots + r_i - b_{i+2} - \cdots - b_\ell)$.

PROOF. The distributional result is proved by using the following result repeatedly. Let W and B be independently distributed as $W_p(n, \Sigma)$ and $W_p(q, \Sigma)$, respectively. Let $T = W + B$, and decompose W and T as

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, \qquad T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix},$$

where $W_{12} : p_1 \times p_2$ and $T_{12} : p_1 \times p_2$. Then it is known (see, e.g., Siotani *et al.* (1985, p. 577)) that $|W_{11\cdot 2}|/|T_{11\cdot 2}| \sim \Lambda_{p_1}(q, n - p_2)$, and $|W_{11\cdot 2}|/|T_{11\cdot 2}|$ is independent of W_{22} and T_{22}. First we use the result to see that $\Lambda_{(1)}$ is distributed as $\Lambda_{b_2}(r_2 + \cdots + r_k, n + r_1 - b_3 - \cdots - b_\ell)$ and is independent of $\{V^{(1)}_{(3\cdots\ell)(3\cdots\ell)}, V^{(k)}_{(3\cdots\ell)(3\cdots\ell)}\}$, and hence of $\{V^{(2)}_{(3\cdots\ell)(3\cdots\ell)}, V^{(k)}_{(3\cdots\ell)(3\cdots\ell)}\}$. Next we use the distributional result to $\Lambda_{(2)}$, and so on.

We note that the hypothesis (2.13) can be also expressed as the set of additional restrictions

$$\forall i = 1, \ldots, k-1,\ \forall j = 2, \ldots, i : C_i \Xi D_j = O. \tag{3.7}$$

Banken (1984) has studied LR test for testing a general hypothesis as in (3.5). The above result can be confirmed from his general result.

The limiting distribution of $-n \log \Lambda_{(i)}$ when n is large is a χ^2 distribution with $f_i = b_{i+1}(r_{i+1} + \cdots + r_k)$ degrees of freedom. A better χ^2 approximation is obtained by the statistic with Bartlett adjustment $-n\rho_i \log \Lambda_{(i)}$, where

$$n\rho_i = n + r_1 + \cdots + r_i - b_{i+2} - \cdots - b_{k+1} + \frac{1}{2}(r_{i+1} + \cdots + r_k - b_{i+1} - 1). \tag{3.8}$$

Since $\Lambda_{(i)}$'s are independent, we have that the limiting distribution of $-n \log \Lambda$ is a χ^2 distribution with $f = f_1 + \cdots + f_{k-1}$ degree of freedom. Further, the statistic with Bartlett adjustment is given by $-n\rho \log \Lambda$, where

$$\rho = \frac{1}{f}(f_1\rho_1 + \cdots + f_{k-1}\rho_{k-1}). \tag{3.9}$$

In fact, it holds that

$$P(-n\rho \log \Lambda \le x) = P(\chi^2_f \le x) + O(n^{-2}). \tag{3.10}$$

In the special case $k = 2$ we can write the LR statistic as

$$\Lambda = |V^{(1)}_{22\cdot 3}|/|V^{(2)}_{22\cdot 3}| \tag{3.11}$$

whose null distribution is $\Lambda_{b_2}(r_2, n+r_1-b_3)$. The exact distributions of $\Lambda_{b_2}(r_2, n+r_1-b_3)$ can be used (see, e.g. Anderson (1984, p. 304) and Siotani *et al.* (1985, p. 248)) for $b_2 = 1, 2$ and/or $r_2 = 1, 2$, as follows:

$$(3.12) \qquad \frac{n}{m} \cdot \frac{1-\Lambda_1(m,n)}{\Lambda_1(m,n)} \sim F_{m,n}, \qquad \frac{n-p+1}{p} \cdot \frac{1-\Lambda_p(1,n)}{\Lambda_p(1,n)} \sim F_{p,n-p+1}.$$

For tables of the distribution of Λ, see, e.g., Table 47 of Biometrika Tables for Statisticians, Volume 2, edited by Pearson and Hartly (1972).

3.3 *AIC for the model* (1.3)

We derive Akaike Information Criterion (Akaike (1973)) for the model (1.3). Let the model denote by $M_k(q_1, \ldots, q_k)$. Since the likelihood function of Y is the same as the one of Z in (2.11) with (2.12), we can write

$$\begin{aligned} &\mathrm{AIC}(M_k(q_1, \ldots, q_k)) \\ &\quad = -2 \log f(Z; \hat{\Theta}, \hat{\Omega}) \\ &\qquad + 2 \times \text{the number of independent parameters under } M_k(q_1, \ldots, q_k). \end{aligned}$$

Using (3.4), we obtain

$$(3.13) \quad \begin{aligned} \mathrm{AIC}(M_k(q_1, \ldots, q_k)) = {} & N \log \left\{ \left(\frac{1}{N} \right)^p |V_{11\cdot(2\cdots\ell)}^{(0)}| \cdot |V_{22\cdot(3\cdots\ell)}^{(1)}| \cdots |V_{\ell\ell}^{(k)}| \right\} \\ & + pN(\log(2\pi) + 1) \\ & + 2 \left\{ \sum_{i=1}^{k} r_i(b_1 + \cdots + b_i) + \frac{1}{2} p(p+1) \right\}. \end{aligned}$$

The criterion can be used in selecting an appropriate model from a set of candidate models with different values of k and $q_1 < \cdots < q_k$. For a refinement of AIC, see Fujikoshi and Satoh (1996), who obtained a finite correction for $\mathrm{AIC}(M_2(q_1, q_2))$.

4. Inferential problems under random-coefficient model (1.6)

4.1 *MLE's and unbiased estimators*

In this section we assume the random-coefficient model (1.6). It is easily seen that the MLE of Ξ is given by

$$(4.1) \qquad \hat{\Xi}_i = (A_{ii}' A_{ii})^{-1} A_{ii}' U_i$$

and hence

$$(4.2) \qquad \hat{\Xi}_i \sim N_{r_i \times q_i}[\Xi_i, \Psi^{(i)} \otimes (A_{ii}' A_{ii})^{-1}].$$

Let

$$(4.3) \quad W^{(i)} = U_i'(I_{N_i} - P_{A_{ii}})U_i, \quad (i = 1, \ldots, k), \qquad s^2 = \frac{1}{f} \sum_{i=1}^{k} \mathrm{tr}\, V_i' V_i = \frac{1}{f} v^2,$$

where $P_{A_{ii}} = A_{ii}(A'_{ii}A_{ii})^{-1}A'_{ii}$ and $f = \sum_{i=1}^{k} N_i(p - q_i)$. Then, the negative of twice the log likelihood function after maximization with respect to Ξ_i, $i = 1, \ldots, k$, is given by

$$(4.4) \qquad d_k(\sigma^2, \Delta) = \sum_{i=1}^{k} N_i \log|\Psi^{(i)}| + \sum_{i=1}^{k} \mathrm{tr}(\Psi^{(i)})^{-1}W^{(i)} + f\left(\log\sigma^2 + \frac{s^2}{\sigma^2}\right).$$

The MLE's of σ^2 and Δ can be obtained by minimizing (4.4) subject to $\sigma^2 > 0$ and $\Delta \geq O$. A difficulty in this minimization problem comes from the fact that

(i) $\Psi^{(i)}$ must satisfy a restriction $\Psi^{(i)} - \sigma^2 R^{(i)} \equiv \Delta^{(i)} \geq O$,

(ii) $\Psi^{(i)}$'s have a hierarchical structure, i.e., $\Psi^{(i)}$ is the first $q_i \times q_i$ matrix of $\Psi^{(k)} = \Delta + \sigma^2 R^{(k)}$.

The explicit expression is available for the case $k = 1$ (see, Khatri and Rao (1988)). For simplicity, we consider the case when $k = 2$ and $X_1X_2' = O$. Partition $\Psi^{(2)}$ and $W^{(2)}$ as

$$\Psi^{(2)} = \Psi = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}, \qquad W^{(2)} = \begin{bmatrix} W_{11}^{(2)} & W_{12}^{(2)} \\ W_{21}^{(2)} & W_{22}^{(2)} \end{bmatrix},$$

where $\Psi_{12} : q_1 \times q_2$, $W_{12}^{(2)} : q_1 \times q_2$. Then we can write

$$(4.5) \qquad d_2(\sigma^2, \Delta) = N\log|\Psi_{11}| + \mathrm{tr}\,\Psi_{11}^{-1}(W^{(1)} + W_{11}^{(2)}) + N_2\log|\Psi_{22\cdot1}| + \mathrm{tr}\,\Psi_{22\cdot1}^{-1}(-\Gamma\ I_{p-q_2})W^{(2)}(-\Gamma\ I_{p-q_2})' + f\left(\log\sigma^2 + \frac{s^2}{\sigma^2}\right),$$

where $\Psi_{22\cdot1} = \Psi_{22} - \Psi_{21}\Psi_{11}^{-1}\Psi_{12}$ and $\Gamma = \Psi_{21}\Psi_{11}^{-1}$. This implies that

$$(4.6) \qquad \begin{aligned} N\hat{\Psi}_{11} &= W^{(1)} + W_{11}^{(2)}, \qquad \hat{\Gamma} = W_{21}^{(2)}(W_{11}^{(2)})^{-1}, \\ N\hat{\Psi}_{22\cdot1} &= W_{22\cdot1}^{(2)}, \qquad \hat{\sigma}^2 = s^2. \end{aligned}$$

Therefore

$$(4.7) \qquad \begin{aligned} \hat{\Delta}_{11} &= \frac{1}{N}(W^{(1)} + W_{11}^{(2)}) - \hat{\sigma}^2 R^{(1)}, \\ \hat{\Delta}_{21} &= \frac{1}{N}W_{21}^{(2)}(W_{11}^{(2)})^{-1}(W^{(1)} + W_{11}^{(2)}), \\ \hat{\Delta}_{22} &= \frac{1}{N}W_{22}^{(2)} + \frac{1}{N}W_{21}^{(2)}(W_{11}^{(2)})^{-1}W^{(1)}(W_{11}^{(2)})^{-1}W_{12}^{(2)} - \hat{\sigma}^2 R_{22}^{(2)}. \end{aligned}$$

We note that the solutions in (4.6) are the MLE's if $\hat{\Psi} - \hat{\sigma}^2 R^{(2)} = \hat{\Delta}$ is positive semi-definite.

It is difficult to give an explicit expression of the MLE in the general case $k > 1$. However, we can propose an unbiased estimator for σ^2 and Δ, based on $W^{(1)}, \dots, W^{(k)}$ and s^2. Note that

$$(4.8) \qquad \begin{aligned} &W^{(i)} \sim W_{b_1+\cdots+b_i}(n_i, \Delta^{(i)} + \sigma^2 R^{(i)}), \\ &fs^2 \sim \sigma^2 \chi_f^2, \end{aligned}$$

and they are independent, where $n_i = N_i - r_i$. Therefore, we have

$$(4.9) \qquad E\left(\frac{1}{n_i} W^{(i)}\right) = \Delta^{(i)} + \sigma^2 R^{(i)}, \qquad E(s^2) = \sigma^2.$$

Partition $W^{(i)}$ and $R^{(i)}$ in $b_1, \dots, b_i$ rows and columns as

$$(4.10) \qquad W^{(i)} = \begin{bmatrix} W_{11}^{(i)} & \cdots & W_{1i}^{(i)} \\ \vdots & \ddots & \vdots \\ W_{i1}^{(i)} & \cdots & W_{ii}^{(i)} \end{bmatrix}, \qquad R^{(i)} = \begin{bmatrix} R_{11}^{(i)} & \cdots & R_{1i}^{(i)} \\ \vdots & \ddots & \vdots \\ R_{i1}^{(i)} & \cdots & R_{ii}^{(i)} \end{bmatrix}.$$

Using (4.9) and making pooled unbiased estimators for submatrices of Δ, we can get a natural unbiased estimator of σ^2 and Δ given by

$$(4.11) \qquad \begin{aligned} &\tilde{\sigma}^2 = s^2, \\ &\tilde{\Delta}_{(1\cdots i)i} = \frac{1}{n_i + \cdots + n_k}\{W_{(1\cdots i)i}^{(i)} + \cdots + W_{(1\cdots i)i}^{(k)} \\ &\qquad\qquad - n_i \tilde{\sigma}^2 R_{(1\cdots i)i}^{(i)} - \cdots - n_k \tilde{\sigma}^2 R_{(1\cdots i)i}^{(k)}\}. \end{aligned}$$

4.2 *Test and AIC*

We consider testing the null hypothesis (2.13) under (1.6). From (4.2) it follows that

$$(4.12) \qquad \mathrm{vec}(\hat{\Xi}'_{i(2\cdots i)}) \sim N(\mathrm{vec}(\Xi'_{i(2\cdots i)}),\ (A'_{ii}A_{ii})^{-1} \otimes Q'_i \Psi^{(i)} Q_i),$$

where

$$(4.13) \qquad Q_i = \begin{bmatrix} O & \cdots & \cdots & O \\ I_{b_2} & \ddots & & \vdots \\ O & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & O \\ O & \cdots & O & I_{b_i} \end{bmatrix}.$$

This suggests that

$$(4.14) \qquad T = \sum_{i=2}^{k} \mathrm{vec}(\hat{\Theta}'_{i(2\cdots i)})'[(A'_{ii}A_{ii})^{-1} \otimes Q'_i \tilde{\Psi}^{(i)} Q_i]^{-1} \mathrm{vec}(\hat{\Theta}_{i(2\cdots i)})$$

should be used as a Wald type test, where $\tilde{\Psi}^{(i)}$ is the unbiased estimator of $\Psi^{(i)}$ obtained from (4.11). The limiting null distribution of T is a χ^2 distribution with $f = \sum_{i=2}^{k} r_i(b_2 + \cdots + b_i)$ degrees of freedom as n_i's tend to infinity.

Let the model (1.6) for Y be denoted by $\tilde{M}_k(q_1, \dots, q_k)$. Then the likelihood function of Y can be written as

$$L(\Xi, \sigma^2, \Delta) = \prod_{i=1}^{k} |X_{(i)}X'_{(i)}|^{-N_i/2} f(U_i, V_i; \Xi_i, \sigma^2, \Delta), \tag{4.15}$$

where $f(U_i, V_i; \Xi_i, \sigma^2, \Delta)$ is the joint density function of $[U_i \ V_i]$ in (2.14). Therefore, we can write AIC for the model $\tilde{M}_k(q_1, \dots, q_k)$ as

$$\begin{aligned} &\mathrm{AIC}(\tilde{M}_k(q_1, \dots, q_k)) \\ &= \sum_{i=1}^{k} N_i \log |\hat{\Psi}^{(i)}| + \sum_{i=1}^{k} \mathrm{tr}(\hat{\Psi}^{(i)})^{-1} W^{(i)} \\ &\quad + f\left(\log \hat{\sigma}^2 + \frac{s^2}{\hat{\sigma}^2}\right) + \sum_{i=1}^{k} N_i \log |X_{(i)}X'_{(i)}| + Np \log(2\pi) \\ &\quad + 2\left\{\sum_{i=1}^{k} r_i(b_1 + \cdots + b_i) + \frac{1}{2} q_k(q_k + 1) + 1\right\}. \end{aligned} \tag{4.16}$$

5. A special case and numerical example

To illustrate our results, we consider the case where the data consist of two groups, and observations in two groups are measured at the same p time-points $t_1, \dots, t_p$ or occasion. Further, suppose that the growth curve is linear for the first group, and is quadratic for the second group. Then, we have an extended growth curve model

$$E(Y) = \begin{bmatrix} \mathbf{1}_{N_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{N_2} \end{bmatrix} \begin{bmatrix} \xi_{11} & \xi_{12} & 0 \\ \xi_{21} & \xi_{22} & \xi_{23} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_p \\ t_1^2 & \cdots & t_p^2 \end{bmatrix}, \tag{5.1}$$

where $\mathbf{1}_N$ is the N dimensional column vector whose elements are all one. It is assumed that each rows of Y are independently distributed as $N_p(\cdot, \Sigma)$, where Σ is unknown positive definite. Note that we assume that the data of two groups have an identical covariance matrix. Our interest is to see whether the growth curve for the second group is also linear or not, i.e., ξ_{23} is zero or not.

Let $\bar{\boldsymbol{y}}_1$ and $\bar{\boldsymbol{y}}_2$ be the sample mean vectors, and S the pooled matrix of sums of squares and cross-products for the observation matrix Y. The LR statistic (3.6) is based on

$$V^{(i)} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} S^{(1)} [B'_1 \ B'_2 \ B'_3], \tag{5.2}$$

where $S^{(1)} = S + N_1\bar{y}_1\bar{y}_1'$ and $S^{(2)} = S^{(1)} + N_2\bar{y}_2\bar{y}_2'$. Here $B = [B_1'\ B_2'\ B_3']'$ is an orthogonal matrix of order p such that

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} G_{11} & O \\ G_{21} & g_{22} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_p \\ t_1^2 & \cdots & t_p^2 \end{bmatrix}, \tag{5.3}$$

where $B_1 : 2 \times p$, $B_2 : 1 \times p$, $G_{11} : 2 \times 2$ and $G_{21} : 1 \times 2$. We can write $V_{22\cdot 3}^{(i)}$ as

$$\begin{aligned} V_{22\cdot 3}^{(i)} &= B_2\{S^{(i)} - S^{(i)}B_3'(B_3S^{(i)}B_3')^{-1}B_3S^{(i)}\}B_2' \\ &= [0\ 0\ 1]\left\{\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}(S^{(i)})^{-1}[B_1'\ B_2']\right\}^{-1}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \text{the } (3,3) \text{ elements of } \left\{\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}(S^{(i)})^{-1}[B_1'\ B_2']\right\}^{-1}. \end{aligned} \tag{5.4}$$

This reduction is obtained by using Lemma 7.5.1 in Siotani *et al.* (1985). For calculation of the LR test statistic Λ, we need not obtain B_3 and it is sufficient to obtain B_1 and B_2 by applying Gram-Schmidt orthogonalization to the three column vectors

$$\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \begin{bmatrix} t_1 \\ \vdots \\ t_p \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} t_1^2 \\ \vdots \\ t_p^2 \end{bmatrix}.$$

The LR statistic for testing the hypothesis "$\xi_{23} = 0$" is given by

$$\Lambda = V_{22\cdot 3}^{(1)}/V_{22\cdot 3}^{(2)} \tag{5.5}$$

or equivalently

$$F = \frac{1-\Lambda}{\Lambda} \cdot \frac{N-p+2}{1} \sim F_{1,N-p+2}. \tag{5.6}$$

The AIC in (3.13) can be expressed as

$$\begin{aligned} \text{AIC}(M_2(q_1,q_2)) &= N\log\left\{\left(\frac{1}{N}\right)^p |V_{11\cdot 23}^{(0)}| \cdot |V_{22\cdot 3}^{(1)}| \cdot |V_{33}^{(2)}|\right\} \\ &\quad + Np(\log(2\pi)+1) + 2\left\{q_1 + q_2 + \frac{1}{2}p(p+1)\right\}. \end{aligned} \tag{5.7}$$

For $k = 2$, Fujikoshi and Satoh (1996) obtained a corrected AIC as

$$\begin{aligned} &\text{CAIC}(M_2(q_1,q_2)) \\ &= N\log\left\{\left(\frac{1}{N}\right)^p |V_{11\cdot 3}^{(0)}| \cdot |V_{22\cdot 3}^{(1)}| \cdot |V_{22\cdot 3}^{(2)}|\right\} + Np\log(2\pi) \\ &\quad + \frac{(q_2-q_1)N}{N-p+q_1-2} + \frac{(p-q_2)N^2}{N-p+q_2-1} + \frac{q_1N}{N-p-3} \\ &\quad + \frac{q_1N}{(N-p-3)(N-p+q_1-3)}\{(N-3)(N+1)+p-q_1\} \\ &\quad + \frac{q_2-q_1}{(N-p-3)(N-p+q_2-3)}\left\{\frac{N-2}{N-p+q_1-3} + (p-q_2)N\right\}. \end{aligned} \tag{5.8}$$

We applied the above results to the dental measurement data (see Potthoff and Roy (1964)), which are made on each of 11 girls and 16 boys at ages 8, 10, 12, 14 years. We analyzed by deleting a boy's measurement, since it is regarded (Lee and Geisser (1975)) that the measurement is outlier. So, $N_1 = 11$, $N_2 = 15$ and $p = 4$. We assume the model (5.1) for the dental measurement data, which is denoted by $M_2(q_1, q_2)$ with $q_1 = 2$ and $q_2 = 3$. Then we have

$$\Lambda = V_{22\cdot 3}^{(1)}/V_{22\cdot 3}^{(2)} = 16.743/22.041 = 0.75960$$

and $F = 7.59567$. The p-value is 0.01099. This suggests that boy's data is not linear, but quadratic. As the other possible models, we consider the models $M_2(2,2)$; both girl's and boy's data are linear, $M_2(3,3)$; both girl's and boy's data are quadratic. The AIC and CAIC are given as follows:

	AIC	CAIC
$M_2(2,2)$	520.58	528.82
$M_2(3,3)$	521.40	532.11
$M_2(2,3)$	519.45	493.67

This suggests also that the model $M_2(2,3)$ is more appropriate, in the comparison with $M_2(2,2)$ and $M_2(3,3)$.

Acknowledgements

The authors would like to express their gratitude to Professor von Rosen for bringing our attention to Banken's thesis (1984).

References

Akaike, H. (1973). Information theory and extension of the maximum likelihood principle, *Second International Symposium on Information Theory* (eds. B. N. Petrov and F. Czáki), 267–281, Academiai Kiadó, Budapest.

Anderson, S. A., Marden J. I. and Perlman M. D. (1993). Totally ordered multivariate linear models, *Sankhyā Ser. A*, **55**, 370–394.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, New York.

Banken, L. (1984). Eine Verallgemeinerung des GMANOVA-Modells, Dissertation, University of Trier, Germany.

Fearn, T. (1975). A Bayesian approach to growth curves, *Biometrika*, **62**, 89–100.

Fujikoshi, Y. and Satoh, K. (1996). Estimation and model selection in an extended growth curve model, *Hiroshima Math. J.*, **26**(3), 635–647.

Gleser, L. J. and Olkin, I. (1970). Linear model in multivariate analysis, *Essays in Probability and Statistics* (ed. R. C. Bose), 267–292, University of North Carolina Press, Chapel Hill.

Kariya, T. (1985). *Testing in the Multivariate General Linear Model*, Kinokuniya, Tokyo.

Khatri, C. G. and Rao, C. R. (1988). Multivariate linear model with latent variables: problems of estimation, Tech. Report, 88–48, Center for Multivariate Analysis, Penn State University.

Lee, J. C. and Geisser, S. (1975). Applications of growth curve prediction, *Sankhyā Ser. A*, **37**, 239–256.

Pearson, E. S. and Hartley, H. O. (1972). *Biometrika Tables for Statisticians*, Vol. II, Cambridge University Press, Cambridge.

Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems, *Biometrika*, **51**, 313–326.

Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, *Biometrika*, **52**, 447–458.

Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*, American Sciences Press, Columbus, Ohio.

Verbyla, A. P. and Venables, W. N. (1988). An extension of the growth curve model, *Biometrika*, **75**, 129–138.

von Rosen, D. (1989). Maximum likelihood estimators in multivariate linear normal models, *J. Multivariate Anal.*, **31**, 187–200.

von Rosen, D. (1990). Moments for a multivariate linear model with an application to the growth curve model, *J. Multivariate Anal.*, **35**, 243–259.

Vonesh, E. F. and Carter, R. L. (1987). Efficient inference for random-coefficient growth curve models with unbalanced data, *Biometrics*, **43**, 617–628.