HIROSHIMA MATH. J.
26 (1996), 635–647

Estimation and model selection in an extended growth curve model

Yasunori FUJIKOSHI and Kenichi SATOH

(Received September 5, 1995)

ABSTRACT. In this paper we consider the growth curve model with two different within-individuals design matrices. The MLE's of the model are obtained in closed forms, based on a canonical form. Some basic properties of the MLE's are given. The problem of selecting such a model is also considered. We derive a corrected version of *AIC* (Akaike Information Criterion, Akaike (1973)) which will be useful in small samples.

1. Introduction

An extension of the usual growth curve model introduced by Potthoff and Roy (1964) has been proposed by Verbyla and Venable (1988), who considered the model with several different within-individuals design matrices. In this paper we consider the growth curve model with two different within-individuals design matrices as follows: Let Y denote an $n \times p$ data matrix whose rows consist of observations on distinct experimental units. We assume that

$$(1.1) \qquad Y = A_1\Theta_1 X_1 + A_2\Theta_2 X_{(2)} + \mathscr{E},$$

where A_i are $n \times k_i$ between-individuals design matrices of ranks k_i, X_1 and $X_{(2)}$ are $q_1 \times p$ and $q \times p$ $(q = q_1 + q_2 \leq p)$ within-individuals design matrices of ranks q_1 and q, respectively, Θ_1 and Θ_2 are $k_1 \times q_1$ and $k_2 \times q$ matrices of unknown parameters and the rows of $\mathscr{E}$ are independently distributed, each with a p-variates normal distribution having mean zero and unknown covariance matrix Σ. Further, it is assumed that

$$(1.2) \qquad X_{(2)} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

It may be noted that our model can be applied for the case where the n individuals have been measured at the same p different times or occasions, and consist of two types of polynomial growth curves with different degrees. In a polynomial growth curve model the design matrices within individuals are

1991 *Mathematics Subject Classification.* 62H12, 62F7.
Key words and phrases. Extended growth curve model, maximum likelihood estimators, model selection, corrected version of *AIC*.

defined by

$$X_1 = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_p \\ \vdots & & \vdots \\ t_1^{q_1-1} & \cdots & t_p^{q_1-1} \end{bmatrix}, \quad X_{(2)} = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_p \\ \vdots & & \vdots \\ t_1^{q-1} & \cdots & t_p^{q-1} \end{bmatrix}. \tag{1.3}$$

Our model also includes as a special case, the usual growth curve model when $X_{(2)} = O$ or $A_2 = O$, and a mixture of MANOVA and GMANOVA models (Chinchilli and Elswick (1985)) when $X_{(2)} = I_p$. On the other side, the model (1.1) is similar to the usual growth curve model since

$$E(Y) = [A_1\ A_2]\begin{bmatrix} \Theta_1 & O \\ & \Theta_2 \end{bmatrix} X_{(2)}. \tag{1.4}$$

That is, the model (1.1) can be regarded as one of the growth curve models with certain linear restrictions. For a review on the latter models, see, e.g., Kariya (1985).

The purpose of this paper is to study the problems of estimating the unknown parameters Θ_1, Θ_2 and Σ, and selecting the degrees of polynomials in the model (1.1). In Section 2 a canonical form for the model (2.1) is given. In Section 3 the MLE's of Θ_1, Θ_2 and Σ are obtained in closed forms. Some basic properties of the MLE's are also given. The model (1.1) can be considered as a fitting model for the true model defined by setting $q_1 = q_1^*$ and $q_2 = q_2^*$. For such a situation, we can apply AIC (Akaike Information Criterion, Akaike (1973)) for the selection of good models. In Section 4 we present a corrected version of AIC which will be useful in small samples.

2. A canonical form

Since the model (1.1) is closely related to the usual growth curve model, we can obtain a canonical form for (1.1) by the same technique as in Gleser and Olkin (1970). However, in order to clarify the correspondence between the quantities in a canonical form and the original form, we give a concrete transformation. The result is also useful in defining the models for selecting the degrees of polynomials. Applying the Gram-Schmidt orthonormalization method to $A = [A_1\ A_2]$, we can choose an orthogonal matrix $H = [H_1\ H_2\ H_3]$ of order n such that

$$\begin{aligned} [A_1\ A_2] &= [H_1\ H_2]\begin{bmatrix} L_{11} & O \\ L_{21} & L_{22} \end{bmatrix} \\ &= H_{(2)}L, \end{aligned}$$

where $H_i : n \times k_i$, and $L_{ii} : k_i \times k_i$ are lower triangular matrices. Similarly, let $B = [B_1' \ B_2' \ B_3']'$ be an orthogonal matrix of order p such that

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} G_{11} & \mathrm{O} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$
$$= GB_{(2)},$$

where $B_i : q_i \times p$, and $G_{ii} : q_i \times q_i$ are lower triangular matrices. Then we can write

$$A_1\Theta_1 X_1 + A_2\Theta_2 X_{(2)} = H_1\Xi_1 B_1 + H_2\Xi_2 B_{(2)}, \tag{2.1}$$

where

$$\Xi_1 = L_{11}\Theta_1 G_{11},$$
$$\Xi_2 = L_{22}\Theta_2 G + L_{21}[\Theta_1 \quad \mathrm{O}]G.$$

There exists a one-to-one correspondence between the original parameter matrices $\{\Theta_1, \Theta_2\}$ and the transformed parameter matrices $\{\Xi_1, \Xi_2\}$. In fact, Θ_1 and Θ_2 can be expressed in terms of Ξ_1 and Ξ_2 as

$$\begin{aligned} \Theta_1 &= L_{11}^{-1}\Xi_1 G_{11}^{-1}, \\ \Theta_2 &= L_{22}^{-1}\Xi_2 G^{-1} - L_{22}^{-1}L_{21}[L_{11}^{-1}\Xi_1 G_{11}^{-1} \quad \mathrm{O}], \end{aligned} \tag{2.2}$$

or equivalently

$$\begin{bmatrix} \Theta_1 & \mathrm{O} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} = L^{-1} \begin{bmatrix} \Xi_1 & \mathrm{O} \\ \Xi_{21} & \Xi_{22} \end{bmatrix} G^{-1},$$

where $\Theta_2 = [\Theta_{21} \ \Theta_{22}]$ and $\Xi_2 = [\Xi_{21} \ \Xi_{22}]$.

Now consider the transformation from Y to

$$\begin{aligned} Z &= H'YB' \\ &= \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix}, \end{aligned} \tag{2.3}$$

where $Z_{ij} = H_i'YB_j'$. Then it is easily seen that the rows of the random matrix $Z : n \times p$ are independently distributed, each with a p-variate normal distribution with unknown covariance matrix $\Omega = B\Sigma B'$ and means

$$E(Z) = \begin{bmatrix} \Xi_1 & \mathrm{O} & \mathrm{O} \\ \Xi_{21} & \Xi_{22} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix}. \tag{2.4}$$

The canonical form is related to the one (Gleser and Olkin (1970)) for testing a linear hypothesis in the usual growth curve model, which is to test "$E(Z_{12}) = O$". Here we note that the hypothesis

$$\Theta_1 = [\Theta_{10}^* \quad O] \quad \text{and} \quad \Theta_2 = [\Theta_{20}^* \quad O] \tag{2.5}$$

for some $q_1^* \le q_2^*$ $(q_i^* \le q_i)$ can be expessed as

$$\Xi_1 = [\Xi_{10}^* \quad O] \quad \text{and} \quad \Xi_2 = [\Xi_{20}^* \quad O] \tag{2.6}$$

in the canonical form, where $\Theta_{i0}^* : k_i \times q_i^*, \Xi_{i0}^* : k_i \times q_i^*$ satisfy the relation (2.2).

In the following we list some notations we are using in the subsequent sections. Let

$$\begin{aligned} U &= [Z_{11}\ Z_{12}\ Z_{13}]'[Z_{11}\ Z_{12}\ Z_{13}] \\ &= [U_{ij}], \quad U_{ij} = Z_{1i}'Z_{1j}, \\ V &= [Z_{21}\ Z_{22}\ Z_{23}]'[Z_{21}\ Z_{22}\ Z_{23}] \\ &= [V_{ij}], \quad V_{ij} = Z_{2i}'Z_{2j}, \\ W &= [Z_{31}\ Z_{32}\ Z_{33}]'[Z_{31}\ Z_{32}\ Z_{33}] \\ &= [W_{ij}], \quad W_{ij} = Z_{3i}'Z_{3j}, \\ T &= U + W \\ &= [T_{ij}], \quad T_{ij} = U_{ij} + W_{ij}. \end{aligned} \tag{2.7}$$

We decompose the covariance matrix Ω as

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix}, \qquad \Omega_{ij} = B_i \Sigma B_j'. \tag{2.8}$$

Further, we denote

$$Z_i = [Z_{i1}\ Z_{i2}\ Z_{i3}], \qquad Z_{i(12)} = [Z_{i1}\ Z_{i2}],$$

$$\Omega_{(23)1} = \begin{bmatrix} \Omega_{21} \\ \Omega_{31} \end{bmatrix}, \qquad \Omega_{(23)(23)} = \begin{bmatrix} \Omega_{22} & \Omega_{23} \\ \Omega_{32} & \Omega_{33} \end{bmatrix},$$

$$\Omega_{11\cdot 23} = \Omega_{11} - \Omega_{1(23)}\Omega_{(23)(23)}^{-1}\Omega_{(23)1}, \qquad \text{etc.}$$

Similar notations are used for matrices of U, V, W and T partitioned in the above manner.

3. Estimation of Θ_1, Θ_2 and Σ

Verbyla and Venables (1988) have proposed a procedure for estimation in a general extended growth curve model, by reducing the model to the seemingly unrelated form. However, the property of such estimation has not been demonstrated. Our model is a special case of the general extended model, and, furthermore, is closely related to the usual growth curve model. This speciality makes the maximum likelihood approach feasible in a convenient form, which will be demonstrated in this section.

Since there exists a one-to-one correspondence between $\{\Theta_1, \Theta_2, \Sigma\}$ and $\{\Xi_1, \Xi_2, \Omega\}$, maximum likelihood estimators of $\{\Xi_1, \Xi_2, \Omega\}$ yields maximum likelihood estimators of $\{\Theta_1, \Theta_2, \Sigma\}$. These estimators can be obtained by slightly extending Gleser and Olkin's argument. First we maximize the joint density function $f(Z; \Xi_1, \Xi_2, \Omega)$ with respect to (Ξ_1, Ξ_2). Considering the conditional distribution of Z_{11} and $(Z_{21}\ Z_{22})$ given $(Z_{12}\ Z_{13})$ and Z_{23}, we can see that the maximum occurs at

$$
\begin{aligned}
\bar{\bar{\Xi}}_1 &= Z_{11} - (Z_{12}\ Z_{13})\Omega_{(23)(23)}^{-1}\Omega_{(23)1}, \\
\bar{\bar{\Xi}}_2 &= (Z_{21}\ Z_{22}) - Z_{23}\Omega_{33}^{-1}\Omega_{3(12)}.
\end{aligned} \tag{3.1}
$$

Then

$$
\begin{aligned}
g(\Omega) &= f(Z; \bar{\bar{\Xi}}_1, \bar{\bar{\Xi}}_2, \Omega) \\
&= (2\pi)^{-np/2}|\Omega|^{-n/2} \exp[-\tfrac{1}{2}\{\mathrm{tr}\,\Omega^{-1}W \\
&\quad + \mathrm{tr}\,\Omega_{(23)(23)}^{-1}U_{(23)(23)} + \mathrm{tr}\,\Omega_{33}^{-1}V_{33}\}].
\end{aligned} \tag{3.2}
$$

The exponent part of $g(\Omega)$ consists of the sum of three hierarchical terms. For the case of the sum of such two terms, Gleser and Olkin (1970) obtained its maximization. Using their idea repeatedly it can be seen that the maximum of $g(\Omega)$ over $\Omega > O$ is achieved at

$$
\begin{aligned}
n\hat{\Omega}_{11\cdot 23} &= W_{11\cdot 23}, \\
\hat{\Omega}_{1(23)}\hat{\Omega}_{(23)(23)}^{-1} &= W_{1(23)}W_{(23)(23)}^{-1}, \\
n\hat{\Omega}_{22\cdot 3} &= T_{22\cdot 3}, \\
\hat{\Omega}_{23}\hat{\Omega}_{33}^{-1} &= T_{23}T_{33}^{-1}, \\
n\hat{\Omega}_{33} &= T_{33} + V_{33},
\end{aligned} \tag{3.3}
$$

or equivalently at

$$n\hat{\Omega} = W + \begin{bmatrix} W_{1(23)}W_{(23)(23)}^{-1} \\ I_{p-q_1} \end{bmatrix} Q[W_{(23)(23)}^{-1}W_{(23)1}\ I_{p-q_1}], \tag{3.4}$$

where

$$Q = U_{(23)(23)} + \begin{bmatrix} T_{23}T_{33}^{-1} \\ I_{q_3} \end{bmatrix} V_{33}[T_{33}^{-1}T_{32}\ I_{q_3}], \qquad \text{and} \qquad q_3 = p - q.$$

Noting that

$$\hat{\Omega}_{13}\hat{\Omega}_{33}^{-1} = W_{1(23)}W_{(23)(23)}^{-1} \begin{bmatrix} T_{23}T_{33}^{-1} \\ I_{q_3} \end{bmatrix},$$

we have

$$\hat{\Xi}_1 = Z_{11} - Z_{1(23)}W_{(23)(23)}^{-1}W_{(23)1}, \tag{3.5}$$

$$\hat{\Xi}_2 = Z_{2(12)} - Z_{23}T_{33}^{-1}[T_{3(23)}W_{(23)(23)}^{-1}W_{(23)1}\ T_{32}]. \tag{3.6}$$

For studying the distributions of $\hat{\Xi}_1$, $\hat{\Xi}_2$ and $\hat{\Omega}$, we use the following reductions. Let F be the $p \times p$ lower triangular matrix such that

$$F'\Omega F = I_p.$$

Consider the transformation

$$\tilde{Z} = \begin{bmatrix} Z_{11} - \Xi_1 & Z_{12} & Z_{13} \\ Z_{21} - \Xi_{21} & Z_{22} - \Xi_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} F,$$

and let $\tilde{U}$, $\tilde{V}$, $\tilde{W}$ and $\tilde{T}$ be the matrices defined by (2.7) for Z replaced by $\tilde{Z}$. Let F, $\tilde{Z}$, $\tilde{U}$, etc. be partitioned in the same manner as Ω. Then we can see that

$$(\hat{\Xi}_1 - \Xi_1)F_{11} = \tilde{Z}_{11} - \tilde{Z}_{1(23)}\tilde{W}_{(23)(23)}^{-1}\tilde{W}_{(23)1},$$

$$(\hat{\Xi}_2 - \Xi_2)F_{(12)(12)} = \tilde{Z}_{2(12)} - \tilde{Z}_{23}\tilde{T}_{33}^{-1}[\tilde{T}_{3(23)}\tilde{W}_{(23)(23)}^{-1}\tilde{W}_{(23)1}\tilde{T}_{32}],$$

$nF'\hat{\Omega}F =$ "*the quantity obtained from the right-hand side of* (3.4) *by replacing* Z *by* $\tilde{Z}$"

These reductions may be summarized as in the following Lemma.

LEMMA 3.1. *Let $\hat{\Xi}_1$, $\hat{\Xi}_2$ and $\hat{\Omega}$ be the maximum likelihood estimators of Ξ_1, Ξ_2 and Ω, respectively. Then,* $(\hat{\Xi}_1 - \Xi_1)F_{11}$, $(\hat{\Xi}_2 - \Xi_2)F_{(12)(12)}$ *and* $nF'\hat{\Omega}F$

are distributed like the quantities of the right-hand side of (3.5), (3.6) *and* (3.4), *respectively, with the assumptions of* $\Xi_1 = O$, $\Xi_2 = O$ *and* $= I_p$.

Using Lemma 3.1 we obtain the following fundamental properties of these estimators.

THEOREM 3.1. *Let* $\hat{\Xi}_1$, $\hat{\Xi}_2$ *and* $\hat{\Omega}$ *be the maximum likelihood estimators of* Ξ_1, Ξ_2 *and* Ω, *respectively. Then*

(i) $E[\hat{\Xi}_i] = \Xi_i$, $i = 1, 2$,

(ii) $\mathrm{Var}[\mathrm{vec}(\hat{\Xi}_1)] = \{1 + (p - q_1)/n_1\}\Omega_{11\cdot 23} \otimes I_{k_1}$,

(iii) $\mathrm{Var}[\mathrm{vec}(\hat{\Xi}_2)] = \left[\{n_3 q_3/(n_1 n_2)\}\Omega_{(12)(12)\cdot 3} - \{k_1 q_3/(n_1 n_2)\}\begin{bmatrix} O & O \\ O & \Omega_{22\cdot 3} \end{bmatrix}\right] \otimes I_{k_2}$,

(iv) $\mathrm{Cov}[\mathrm{vec}(\hat{\Xi}_1), \mathrm{vec}(\hat{\Xi}_2)] = O$,

(v) $E[\hat{\Omega}] = \Omega + \dfrac{1}{n} F'^{-1} \begin{bmatrix} c_1 I_{q_1} & O & O \\ O & c_2 I_{q_2} & O \\ O & O & c_3 I_{q_3} \end{bmatrix} F^{-1}$,

where $\mathrm{vec}(A)$ *denotes the vector obtained by stacking the columns of* A *under each other,*

$$
\begin{aligned}
& n_1 = n - k - p + q_1 - 1, \qquad n_2 = n - k_2 - p + q - 1, \\
& n_3 = n - k_2 - p + q_1 - 1, \\
& c_1 = \{k(p - q) + k_1 q_2\}/n_1 - \{k_2 q_2 (p - q)\}/(n_1 n_2), \\
& c_2 = k_1 + \{k_2(p - q)\}/n_2, \qquad c_3 = k.
\end{aligned} \tag{3.7}
$$

PROOF. These results are obtained by using Lemma 3.1. (i) is easily seen. The distribution of $\hat{\Xi}_1$ is essentially the same as in the usual growth curve model, and hence (ii) has been shown in Grizzle and Allen (1969). For (iii), first we consider the expectation with respect to Z_2, getting

$$\mathrm{Var}[\mathrm{vec}(\hat{\Xi}_2 - \Xi_2)] = [F'^{-1}_{(12)(12)}\{I_q + E[M]\}F^{-1}_{(12)(12)}] \otimes I_{k_2},$$

where

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

$$M_{11} = W_{1(23)} W^{-1}_{(23)(23)} T_{(23)3} T_{33}^{-2} T_{3(23)} W^{-1}_{(23)(23)} W_{(23)1},$$

$$M_{12} = M'_{21} = W_{1(23)} W^{-1}_{(23)(23)} T_{(23)3} T_{33}^{-2} T_{32}, \qquad \text{and } M_{22} = T_{23} T_{33}^{-2} T_{32}.$$

Note that $W \sim W_p(I_p, n - k)$ and $T_{(23)(23)} \sim W_{p-q_1}(I_{p-q_1}, n - k_2)$. Now we use the fact (see, e.g., Siotani, Hayakawa and Fujikoshi (1985)) that $W_{11\cdot 23} \sim$

$W_{q_1}(I_{q_1}, n-k-p+q_1)$, $W_{(23)(23)} \sim W_{p-q_1}(I_{p-q_1}, n-k)$, the elements of $W_{1(23)}W_{(23)(23)}^{-1/2}$ are independently distributed as $N(0, 1)$, and these three random matrices are independent. Similar results hold for submatrices of $T_{(23)(23)}$. Using these properties we can see that $E[M_{11}] = m_{11}I_{q_1}$, $E[M_{12}] = \mathrm{O}$, $E[M_{22}] = E[\mathrm{tr}\ T_{33}^{-1}]I_{q_2} = \{q_3/n_2\}I_{q_2}$, where

$$m_{11} = E[\mathrm{tr}\ W_{(23)(23)}^{-1}T_{(23)3}T_{33}^{-2}T_{3(23)}].$$

Note that

$$W_{(23)(23)}^{-1} = \begin{bmatrix} \mathrm{O} & \mathrm{O} \\ \mathrm{O} & W_{33}^{-1} \end{bmatrix} + \begin{bmatrix} I_{q_2} \\ -W_{33}^{-1}W_{32} \end{bmatrix} W_{22\cdot 3}^{-1}[I_{q_2} - W_{23}W_{33}^{-1}].$$

Considering the expectation with respect to $W_{22\cdot 3}$, we have

$$\begin{aligned} m_{11} &= E[\mathrm{tr}\ W_{33}^{-1}] + \{E[\mathrm{tr}\ W_{23}W_{33}^{-2}W_{32}] \\ &\quad + E[\mathrm{tr}\ T_{23}T_{33}^{-2}T_{32}] - 2\ \mathrm{tr}\ E[\mathrm{tr}\ W_{23}W_{33}^{-1}T_{33}^{-1}T_{32}]\}/n_1 \\ &= q_3 n_3/(n_1 n_2). \end{aligned}$$

This proves (iii). The result (iv) and (v) are shown by similar reductions. The details are omitted.

The properties of the MLE's $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are obtained from the ones of $\hat{\Xi}_1$ and $\hat{\Xi}_2$ through the relation (2.2). In fact, these estimators are unbiased. Using $\mathrm{vec}(ABC) = [C' \otimes A]\,\mathrm{vec}(B)$, we have

$$\begin{aligned} \mathrm{Var}[\mathrm{vec}(\hat{\Theta}_1)] &= \mathrm{Var}[\{G_{11}'^{-1} \otimes L_{11}^{-1}\}\mathrm{vec}(\hat{\Xi}_1 - \Xi_1)] \\ &= \{1 + (p-q)/n_1\}[G_{11}'^{-1}\Omega_{11\cdot 23}G_{11}^{-1}] \otimes [L_{11}^{-1}L_{11}'^{-1}]. \end{aligned}$$

From the definition of L it follows that

$$\tilde{A}_1'\tilde{A}_1 = L_{11}'L_{11}, \qquad A_2'A_2 = L_{22}'L_{22}, \qquad (A_2'A_2)^{-1}A_2'A_1 = L_{22}^{-1}L_{21},$$

where $\tilde{A}_1 = (I_n - P_{A_2})A_1$ and $P_{A_2} = A_2(A_2'A_2)^{-1}A_2'$ is the projection matrix onto the space spanned by the culumn vectors of A_2. Futher

$$\begin{aligned} (3.8)\qquad \Omega_{11\cdot 23} &= B_1\Sigma B_1' - B_1\Sigma B_{(23)}'\{B_{(23)}\Sigma B_{(23)}'\}^{-1}B_{(23)}\Sigma B_1' \\ &= B_1\Sigma B_1' - B_1\Sigma\{\Sigma^{-1} - \Sigma^{-1}B_1'(B_1\Sigma^{-1}B_1')^{-1}B_1\Sigma^{-1}\}\Sigma B_1' \\ &= (B_1\Sigma^{-1}B_1')^{-1}. \end{aligned}$$

For the reduction of the second equality, see, e.g., Siotani, Hayakawa and

Fujikoshi [1985, p. 311]. Therefore, we can write

$$\mathrm{Var}[\mathrm{vec}(\hat{\Theta}_1)] = \{1 + (p-q)/n_1\}(X_1\Sigma^{-1}X_1')^{-1} \otimes (\tilde{A}_1'\tilde{A}_1)^{-1}.$$

Similarly it holds that

$$\begin{aligned}\mathrm{Var}[\mathrm{vec}(\hat{\Theta}_2)] &= \mathrm{Var}[\{G'^{-1} \otimes L_{22}^{-1}\}\mathrm{vec}(\hat{\Xi}_2)] \\ &\quad + \mathrm{Var}\left[\begin{bmatrix} \{I_{q_1} \otimes L_{22}^{-1}L_{21}\}\mathrm{vec}(\hat{\Theta}_1) \\ \mathrm{O} \end{bmatrix}\right] \\ &= [\{n_3q_3/(n_1n_2)\}D - \{k_1q_3/(n_1n_2)\}[-X_2X_1'(X_1X_1')^{-1}I_{q_2}]' \\ &\quad \times D_{22}[-X_2X_1'(X_1X_1')^{-1}I_{q_2}]] \otimes (A_2'A_2)^{-1} \\ &\quad + \{1+(p-q)/n_1\}\begin{bmatrix} M & \mathrm{O} \\ \mathrm{O} & \mathrm{O} \end{bmatrix},\end{aligned}$$

where $M = (X_1\Sigma^{-1}X_1')^{-1} \otimes (A_2'A_2)^{-1}A_2A_1(\tilde{A}_1'\tilde{A}_1)^{-1}A_1'A_2(A_2'A_2)^{-1}$, and

$$D = (X_{(2)}\Sigma^{-1}X_{(2)}')^{-1} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}, \qquad D_{11} : q_1 \times q_1.$$

In order to express $\hat{\Theta}_1$, $\hat{\Theta}_2$ and $\hat{\Sigma}$ in terms of the original notations, let

$$\begin{aligned} S_e &= \{1/(n-k)\}Y'H_3H_3'Y = \{1/(n-k)\}Y'(I_n - P_A)Y, \\ S_t &= \{1/(n-k_2)\}Y'(H_1H_1' + H_3H_3')Y = \{1/(n-k_2)\}Y'(I_n - P_{A_2})Y, \\ \tilde{X}_2 &= X_2(I_p - P_{X_1'}).\end{aligned}$$

Then, using reductions similar to (3.8), we obtain the following expressions:

$$\begin{aligned}\hat{\Theta}_1 &= (\tilde{A}_1'\tilde{A}_1)^{-1}YS_e^{-1}X_1'(X_1S_e^{-1}X_1')^{-1}, \\ \hat{\Theta}_{21} &= \tilde{\Theta}_{21} - \hat{\Theta}_{22}[\tilde{X}_2S^{-1}X_1'(X_1S_e^{-1}X_1')^{-1} + X_2X_1'(X_1X_1')^{-1}] \\ &\quad - (A_2'A_2)^{-1}A_2'A_1\hat{\Theta}_1, \\ \hat{\Theta}_{22} &= (A_2'A_2)^{-1}A_2'YS_t^{-1}X_{(2)}'(X_{(2)}S_t^{-1}X_{(2)}')^{-1}\begin{bmatrix} \mathrm{O} \\ I_{q_2} \end{bmatrix},\end{aligned}$$

where

$$\tilde{\Theta}_{21} = (A_2'A_2)^{-1}A_2'YS_e^{-1}X_1'(X_1S_e^{-1}X_1')^{-1}.$$

Our derivation of the MLE's yields

$$n\hat{\Sigma} = (Y - A_1\hat{\Theta}_1X_1 - A_2\hat{\Theta}_2X_{(2)})'(Y - A_1\hat{\Theta}_1X_1 - A_2\hat{\Theta}_2X_{(2)}).$$

Letting $K = I_p - S_e^{-1}X_1'(X_1S_e^{-1}X_1')^{-1}X_1$, we obtain

$$A_1\hat{\Theta}_1X_1 + A_2\hat{\Theta}_2X_{(2)} = \tilde{A}_1\hat{\Theta}_1X_1 + A_2\tilde{\Theta}_{21}X_1 + A_2\hat{\Theta}_{22}\tilde{X}_2K.$$

This implies

$$n\hat{\Sigma} = (n-k)S_e + K'(P_AY - A_2\hat{\Theta}_{22}\tilde{X}_2)'(P_AY - A_2\hat{\Theta}_{22}\tilde{X}_2)K$$
$$= (n-k)S_e + K'(Y'P_{\tilde{A}_1}Y + J'Y'P_{A_2}YJ)K,$$

where $J = I_p - S_t^{-1}X_{(2)}'(X_{(2)}S_t^{-1}X_{(2)}')^{-1}(\mathrm{O} \quad I_q)'\tilde{X}_2$.

4. Model selection and a corrected *AIC*

The model (1.1) involves a polynomial growth curve model with two types of polynomials with different degrees. In this case the numbers $q_1 - 1$ and $q_1 + q_2 - 1$ means the degrees of two types of polynomials. Now we assume that the true model is the model (1.1) with $q_i = q_i^*$ $(q_1^* \le q_2^*)$, $\Theta_i = [\Theta_{i0}^* \quad \mathrm{O}]$ (or $\Xi_i = [\Xi_{i0}^* \quad \mathrm{O}]$) in (2.5) and $\Sigma = \Sigma^*$ (or $\Omega = \Omega^*$). Then, the model (1.1) can be regarded as a fitted model. In a practical situation, the degrees of the polynomials are unknown, and we need to select appropriate degrees, i.e., good models. Here we consider the Akaike information criterion (Akaike (1973)) as one of such methods. The criterion has been proposed as an approximately unbiased estimator of the risk defined by the transformed predicted probability density or the expected Kullback-Leibler information of a fitted model. In our problem we can write the risk as

$$R(q_1, q_2) = E_Y^*E_{Y_F}^*[-2\log f(Y_F; \hat{\Xi}_1, \hat{\Xi}_2, \hat{\Omega})], \tag{4.1}$$

where $f(Y_F; \hat{\Xi}_1, \hat{\Xi}_2, \hat{\Omega})$ is the predicted probability density for the future observation matrix Y_F; $n \times p$. Here it is assumed that Y_F is independent of Y. The expectations E_Y^* and $E_{Y_F}^*$ means the ones under the true model. A formal application of *AIC* to our model yields

$$AIC = -2\log f(Y; \hat{\Xi}_1, \hat{\Xi}_2, \hat{\Sigma}) + 2\{k_1q_1 + k_2q + \tfrac{1}{2}p(p+1)\}. \tag{4.2}$$

The criterion (4.2) is an approximately unbiased estimator for (4.1) when n is large. Some refinement of this criterion has been discussed in certain models (see, e.g., Sugiura (1978), Hurvich and Tsai (1989), Bedrick and Tsai (1994)). Sugiura reported a correction of *AIC* in the usual growth curve model in a research meeting in 1981, but the result is unpublished. We can write the risk $R(q_1, q_2)$ as

$$R(q_1, q_2) = E_Y^*[n\log|\hat{\Sigma}|] + R_1 + R_2 + R_3 + pn\log(2\pi) \tag{4.3}$$

where

$$(4.4)\qquad \begin{aligned} R_1 &= nE^*_Y[\text{tr}\,\Sigma^*\hat{\Sigma}^{-1}],\\ R_2 &= E^*_Y[\text{tr}(\hat{\Xi}_1 - \Xi^*_1)'(\hat{\Xi}_1 - \Xi^*_1)B_1\hat{\Sigma}^{-1}B'_1],\\ R_3 &= E^*_Y[\text{tr}(\hat{\Xi}_2 - \Xi^*_2)'(\hat{\Xi}_2 - \Xi^*_2)B_{(2)}\hat{\Sigma}^{-1}B'_{(2)}]. \end{aligned}$$

LEMMA 4.1. *Let* R_i, $i = 1, 2, 3$ *be the quantities defined by* (4.3) *or* (4.4). *Then*

(i) $R_1 = n\{nq_3/n_4 + nn_0q_1/(n_1n_5) + q_2(n - k_2 - 1)/(n_2n_3)\}$,

(ii) $R_2 = nk_1q_1n_0/(n_1n_5)$,

(iii) $R_3 = n\{k_2q_2/n_3 + k_2q_1/n_5 + k_2q_1(p - q_1)/(n_1n_5) + k_2q_2q_3/(n_2n_3)\}$,

where n_i, $i = 1, 2, 3$ *are given by* (3.7), *and* $n_0 = n - k - 1$, $n_4 = n - p + q - 1$, $n_5 = n - k - p - 1$.

PROOF. Let F^* be the $p \times p$ lower triangular matrix such that $F^{*\prime}\Omega F^* = I_p$. Then, using Lemma 3.1 with F^* instead of F we can write

$$R_1 = E[\text{tr}\,\hat{\Omega}^{-1}].$$

Here $\hat{\Omega}$ is given by (3.4) and has the distributional reduction in Lemma 3.1. Note that

$$\begin{aligned} \text{tr}\,\hat{\Omega}^{-1} &= \text{tr}\,\hat{\Omega}^{-1}_{11\cdot 23} + \text{tr}\,\hat{\Omega}^{-1}_{(23)(23)}\hat{\Omega}_{(23)1}\hat{\Omega}^{-1}_{11\cdot 23}\hat{\Omega}_{1(23)}\hat{\Omega}^{-1}_{(23)(23)}\\ &\quad + \text{tr}\,\hat{\Omega}^{-1}_{22\cdot 3} + \text{tr}\,\hat{\Omega}^{-1}_{33}\hat{\Omega}_{32}\hat{\Omega}^{-1}_{22\cdot 3}\hat{\Omega}_{23}\hat{\Omega}^{-1}_{33} + \text{tr}\,\hat{\Omega}^{-1}_{33}\\ &= n[\text{tr}\,W^{-1}_{11\cdot 23} + \text{tr}\,W^{-1}_{(23)(23)}W_{(23)1}W^{-1}_{11\cdot 23}W_{1(23)}W^{-1}_{(23)(23)}\\ &\quad + \text{tr}\,T^{-1}_{22\cdot 3} + \text{tr}\,T^{-1}_{33}T_{32}T^{-1}_{22\cdot 3}T_{23}T^{-1}_{33} + \text{tr}(T_{33} + V_{33})]. \end{aligned}$$

Using the above expression we obtain (i). We can rewrite

$$R_2 = E^*_Y[\text{tr}\{(\hat{\Xi}_1 - \Xi^*_1)F^*_{11}\}'\{(\hat{\Xi}_1 - \Xi^*_1)F^*_{11}\}[I_{q_1}\quad \text{O}\quad \text{O}](F^{*\prime}\hat{\Omega}F^*)^{-1}[I_{q_1}\quad \text{O}\quad \text{O}]'].$$

Using Lemma 3.1 with F^* instead of F yields

$$R_2 = nE[\text{tr}[I_{q_1} - W_{1(23)}W^{-1}_{(23)(23)}]Z'_1Z_1[I_{q_1} - W_{1(23)}W^{-1}_{(23)(23)}]'W^{-1}_{11\cdot 23}],$$

which gives (ii), after doing some computations similar to the ones in Theorem 3.1. A similar argument shows that

$$R_3 = nE[\text{tr}[I_{q_3} - \hat{\Omega}_{(23)3}\hat{\Omega}^{-1}_{33}]Z'_2Z_2[I_{q_3} - \hat{\Omega}_{(23)3}\hat{\Omega}^{-1}_{33}]'\hat{\Psi}_{(12)(12)}].$$

Here $\hat{\Omega}^{-1} = \hat{\Psi}$ and each of the terms in the above expression of R_3 has the distributional reduction as in Lemma 3.1. Note that $\hat{\Psi}_{11} = nW^{-1}_{22\cdot 3}$, $\hat{\Psi}_{12} =$

$-nW_{11\cdot 23}^{-1}W_{1(23)}W_{(23)(23)}^{-1}[I_{q_2}\quad O]'$, $\hat{\Psi}_{22} = \hat{\Omega}_{22\cdot 3}^{-1} - \hat{\Psi}_{21}\hat{\Omega}_{12\cdot 3}\hat{\Omega}_{22\cdot 3}^{-1} = nT_{22\cdot 3}^{-1} + n[I_{q_2}\quad O]W_{(23)(23)}^{-1}W_{(23)1}W_{11\cdot 23}^{-1}W_{1(23)}W_{(23)(23)}^{1}[I_{q_2}\quad O]'$, $\hat{\Omega}_{33}^{-1}\hat{\Omega}_{32} = T_{33}^{-1}T_{32}$, and $\hat{\Omega}_{33}^{-1}\hat{\Omega}_{32} = T_{33}^{-1}T_{3(23)}W_{(23)(23)}^{-1}W_{(23)1}$. Using these expressions we can obtain (iii).

Based on (4.3) and Lemma 4.1, we can propose a corrected version of AIC given by

$$\begin{aligned} C\text{-}AIC(q_1, q_2) = n \log|\hat{\Sigma}| + np\log(2\pi) \\ + nk_2q_2/n_3 + n^2q_3/n_4 + nk_2q_1/n_5 \\ + n\{nn_0q_1 + k_1n_0q_1 + k_2q_1(p - q_1)\}/(n_1n_5) \\ + \{q_2(n - k_2 - 1)/n_1 + nk_2q_2q_3\}/(n_2n_3), \end{aligned}$$

which has an unbiased property in the sense of the following Theorem 4.1.

THEOREM 4.1. *Suppose that the model* (1.1) *with* $(q_1, q_2) = (q_1^*, q_2^*)$ *and the parameters* (2.5) *or* (2.6) *is true. Let* $R(q_1, q_2)$ *be the risk defined by* (4.1) *for selecting a fitted model* (1.1). *Then, for any* (q_1, q_2) *such that* $q_i \geq q_i^*$, $i = 1, 2$,

$$E[C\text{-}AIC(q_1, q_2)] = R(q_1, q_2).$$

Thus $C\text{-}AIC$ is an exact unbiased estimator for $R(q_1, q_2)$ under certain conditions. So, we can expect that $C\text{-}AIC$ has a better behaviour than AIC in small samples. For large n, it holds that

$$C\text{-}AIC - AIC = O(1/n),$$

since

$$-2\log f(Y; \hat{\Theta}_1, \hat{\Theta}_2, \hat{\Sigma}) = n\log|\hat{\Sigma}| + pn\{1 + \log(2\pi)\}.$$

In a special case of the usual growth curve model, i.e., $A_2 = O$, we can write

$$C\text{-}AIC = n\log|\hat{\Sigma}| + pn\log(2\pi) + n^2(p - q)/n_4 + n(n + k)qn_0/(n_1n_5).$$

References

[1] H. Akaike, Information theory and an extension of the maximum likelihood principle, In: 2nd International Symposium on Information Theory, Eds. B. N. Petrov and F. Csáki, pp. 267–281, Akadémia Kiado, Budapest, 1973.

[2] E. J. Bedrick and C. L. Tsai, Model selection for multivariate regression in small samples, Biometrics, **50** (1994), 226–231.

[3] V. M. Chinchilli and R. K. Elswick, A mixture of the MANOVA and GMANOVA models, Comm. Staist.-Theor. Meth., **14**(12) (1985), 3075–3089.

[4] L. J. Gleser and I. Olkin, Linear models in multivariate analysis, Essays in Prob. Statist., (R. C. Bose and Others, eds.), 267–292, Univ. North Carolina Press, Chapel Hill, N.C., 1973.

[5] J. E. Grizzle and D. M. Allen, Analysis of growth and dose response curves, Biometrics, **25** (1969), 357–381.

[6] C. M. Hurvich and C. L. Tsai, Regression and times series model selection in small samples, Biometrika, **76** (1989), 297–307.

[7] T. Kariya, Testing in the multivariate general linear model, Kinokuniya, Tokyo, 1985.

[8] R. F. Potthoff and S. N. Roy, A generalized multivariate analysis of variance model useful especially for growth curve problems, Biometrika, **51** (1964), 313–327.

[9] M. Siotani, T. Hayakawa and Y. Fujikoshi, Modern Multivariate Statistical Analysis: A Graduate Course and Handbook, American Sciences, Ohio, 1985.

[10] N. Sugiura, Futher analysis of the data by Akaike's information criterion and the finite corrections, Commun. Statist.-Theor. Meth., **7**(1) (1978), 13–26.

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

Department of Mathmatics
Faculty of Science
Hiroshima University
Higashi-Hiroshima 739, *Japan*