HIROSHIMA MATH. J.
27 (1997), 487–512

Efficient tests for mean structure in random effects models

Hironori FUJISAWA

(Received December 9, 1996)

ABSTRACT. The main purpose of the present paper is to study efficient tests for mean structure in random effects models. We are mainly concerned with a multivariate one-way classification model with random effects. A simplified test is naturally constructed. The test has uniformly higher power than the Wald-type test and the likelihood ratio test. The idea can be applied to other related models including a random coefficient growth curve model.

0. Introduction

A multivariate one-way classification model with random effects is given by

$$\boldsymbol{y}_{ij} = \boldsymbol{\mu} + \boldsymbol{b}_i + \boldsymbol{e}_{ij}, \qquad i = 1, \dots, n, \ j = 1, \dots, k, \tag{0.1}$$

where $\boldsymbol{y}_{ij}$ is a p-component vector of the j-th repeated observation of the i-th individual, $\boldsymbol{\mu}$ is a total mean parameter, $\boldsymbol{b}_i$ is a random effect of the i-th individual, $\boldsymbol{e}_{ij}$ is a noise, n is the number of individuals, and k is the number of repeated observations. Assume that $\boldsymbol{b}_i$'s and $\boldsymbol{e}_{ij}$'s are mutually independent, $\boldsymbol{b}_i$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix Γ, and $\boldsymbol{e}_{ij}$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix Σ. The covariance matrix Γ expresses a dispersion of the randomly chosen individuals. We also deal with the case where Σ has an appropriate structure.

Let us give an example. Suppose that there are p types of machines in a factory. The n workers are randomly chosen. An observation y_{ija} is the amount of the j-th product made by the i-th individual using the a-th machine. Letting $\boldsymbol{y}_{ij} = (y_{ij1}, \dots, y_{ijp})'$, we may use the multivariate one-way classification model with random effects to analyze the data.

In the above model, a considerable work has been done. The maximum likelihood estimators were obtained by Anderson et al. [3] and Anderson [5]. Their asymptotic properties were investigated by Remadi and Amemiya [21]. The tests for rank of random effects were discussed by Amemiya et al.

1991 *Mathematics Subject Classification*. 62H15, 62H10.

Keywords and phrases. Growth curve model, Likelihood ratio test, Multivariate one-way classification model, Power, Random effect, Test for mean structure, Wald-type test.

[2], Anderson [6], Kuriki [16], and Schott and Saw [25]. For the unbalanced case, i.e., when the numbers of repeated observations are not equal, the maximum likelihood estimators are not obtained as a closed form. Some estimation procedures were suggested by Amemiya [1]. The tests for rank of random effects were extended by Anderson and Amemiya [7]. In balanced multivariate variance components models, Calvin and Dykstra [9] proposed a computational algorithm of the estimate which is guaranteed to converge to the restricted maximum likelihood estimators. That method was applied to some models by Calvin [8] and Calvin and Dykstra [10].

In the present paper we consider testing the linear hypothesis

$$\mathrm{H} : C\boldsymbol{\mu} = \mathbf{0}, \tag{0.2}$$

against the alternative hypothesis $K : C\boldsymbol{\mu} \neq \mathbf{0}$, where C is a $q \times p$ known design matrix. If we set the $(p-1) \times p$ matrix

$$C = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{pmatrix},$$

the null hypothesis is the same as $\mathrm{H} : \mu_1 = \cdots = \mu_p$, where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$. On the above example, the equality of qualities of the machines is tested. Similar testing problems were discussed by Fujisawa [11], Fujisawa [12], Reinsel [20], Suzukawa [28], Vonesh and Carter [29], and Yokoyama [34] in random coefficient growth curve models, and by Yokoyama and Fujikoshi [33] in a parallel profile model with random effects.

We consider the Wald-type test and the likelihood ratio test. Type I errors of these tests generally depend on nuisance parameters (see Suzukawa [28]). Therefore we need to evaluate their supremums. If they are obtained, we can perform the Wald-type test and the likelihood ratio test. When $\Sigma = \sigma^2 I$, the supremums can be derived by using the same methods as in Fujisawa [11], [12]. The results are extended to a more general situation, which is mentioned in Sections 4 and 8.

In the present paper we suggest the simplified tests which are the Wald-type test and the likelihood ratio test based on the sample means $\bar{y}_{i\cdot}$'s only, where $\bar{y}_{i\cdot} = \sum_{j=1}^{k} y_{ij}/k$. The test statistics are expressed as a closed form and their Type I errors are simple. It is shown that the simplified tests have uniformly higher power than the Wald-type test and the likelihood ratio test. For details, see Sections 4 and 8. In Part I, some preliminary results are presented. The testing problem is reduced to a simple canonical form in Section 1. In the subsequent sections, we use the canonical form in place of the original testing problem. In Parts II and III, the Wald-type test and the likelihood ratio test are discussed. Each part consists of four sections: Main results, Examples, Test statistics and their properties, Proofs. In Part IV, the methods are applied to other related models including a random coefficient growth curve model.

Part I. Preliminaries

In this part, the original testing problem (0.2) under the model (0.1) is reduced to a simple canonical form and some preliminary results are presented.

1. Canonical form

Let $\boldsymbol{y}_i = (\boldsymbol{y}'_{i1}, \ldots, \boldsymbol{y}'_{ik})'$ and $\boldsymbol{e}_i = (\boldsymbol{e}'_{i1}, \ldots, \boldsymbol{e}'_{ik})'$. The model (0.1) is expressed as

$$\boldsymbol{y}_i = (\mathbf{1}_k \otimes I_p)(\boldsymbol{\mu} + \boldsymbol{b}_i) + \boldsymbol{e}_i,$$

where $\mathbf{1}_k = (1, \ldots, 1)'$, $\otimes$ denotes the Kronecker product, and $\boldsymbol{e}_i$'s are independently normally distributed as $N_{kp}(\mathbf{0}, I_k \otimes \Sigma)$. Let A_0 be a $k \times k$ orthogonal matrix whose first column is $\mathbf{1}_k/\sqrt{k}$ and let $A = A_0 \otimes I_p$. Then, the orthogonal transformation $\boldsymbol{x}_i = A'\boldsymbol{y}_i$ leads to

$$\boldsymbol{x}_i = \begin{pmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \\ \vdots \\ \boldsymbol{x}_{ik} \end{pmatrix} = \begin{pmatrix} \sqrt{k}(\boldsymbol{\mu} + \boldsymbol{b}_i) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} + A'\boldsymbol{e}_i.$$

Here $\boldsymbol{x}_{ij}$'s are $p \times 1$ vectors and in particular $\boldsymbol{x}_{i1} = \sqrt{k}\bar{\boldsymbol{y}}_{i\cdot\cdot}$. Decompose $C = P_C Q'_C$ where P_C is a $q \times q$ non-singular matrix and Q_C is a $p \times q$ matrix such that $Q'_C Q_C = I_q$. Let Q_0 be a $p \times p$ orthogonal matrix whose first q-column is Q_C. Then, the null hypothesis $\mathrm{H} : C\boldsymbol{\mu} = \mathbf{0}$ corresponds to $\mathrm{H} : \boldsymbol{\theta}^{(1)} = \mathbf{0}$, where $\boldsymbol{\theta}^{(1)}$ is the first q-component vector of $\boldsymbol{\theta} = \sqrt{k}Q'_0\boldsymbol{\mu}$, because

$$\mathbf{0} = C\boldsymbol{\mu} = P_C Q'_C \boldsymbol{\mu} = P_C Q'_C Q_0 Q'_0 \boldsymbol{\mu} = \frac{1}{\sqrt{k}} P_C (I_q O)\boldsymbol{\theta} = \frac{1}{\sqrt{k}} P_C \boldsymbol{\theta}^{(1)}.$$

Let $Q = I_k \otimes Q_0$. The orthogonal transformation $\boldsymbol{z}_i = Q'\boldsymbol{x}_i (= Q'A'\boldsymbol{y}_i)$ yields to

$$\boldsymbol{z}_i = \begin{pmatrix} \boldsymbol{z}_{i1} \\ \boldsymbol{z}_{i2} \\ \vdots \\ \boldsymbol{z}_{ik} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\theta} + \boldsymbol{b}^*_i \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} + \boldsymbol{e}^*_i,$$

where $z_{ij} = Q_0' x_{ij}$, $z_{i1} = \sqrt{k} Q_0' \bar{y}_{i\cdot}$, $b_i^* = \sqrt{k} Q_0' b_i$ is normally distributed as $N_p(\mathbf{0}, \Psi)$, $\Psi = kQ_0' \Gamma Q_0$, $e_i^* = Q' A' e_i$ is normally distributed as $N_{kp}(\mathbf{0}, I_k \otimes \Phi)$, and $\Phi = Q_0' \Sigma Q_0$. It is seen that z_{ij}'s are mutually independent, z_{i1} is distributed as $N_p(\theta, \Delta)$, where $\Delta = \Psi + \Phi$, and $z_{ij} (j \geq 2)$ is distributed as $N_p(\mathbf{0}, \Phi)$.

Therefore, the above canonical form is expressed as follows: $z_{ij} (i = 1, \dots, n, j = 1, \dots, k)$ are mutually independent,

$$z_{i1} \sim N_p(\theta, \Delta), \qquad z_{ij} \sim N_p(\mathbf{0}, \Phi) \qquad \text{for } j \geq 2,$$

where Δ and Φ are positive definite matrices such that $\Delta \geq \Phi$. We also deal with the case that Φ has an appropriate structure. The null hypothesis is

$$\mathrm{H} : \theta^{(1)} = \mathbf{0},$$

against the alternative hypothesis $\mathrm{K} : \theta^{(1)} \neq \mathbf{0}$.

Let $Z_1 = (z_{11} \cdots z_{n1})$, $Z_2 = (z_{12} \cdots z_{n2} z_{13} \cdots z_{nk})$, and $Z = (Z_1, Z_2)$. If the restriction $\Delta \geq \Phi$ is neglected, Z_1(i.e. $\bar{y}_{i\cdot}$'s) is useful but Z_2 is not useful to test the null hypothesis $\mathrm{H} : \theta^{(1)} = \mathbf{0}$ because Z_2 has no information about θ. Due to its property, the former observations may be characterized as the main information and the latter observations as the additional information.

We use the following notations: For any vector ϕ, $\phi^{(1)}$ denotes the first q-component vector of ϕ. For any matrix K, we consider a partition $K = (K_{ij})_{i,j=1,2}$ such that K_{11} is the first $q \times q$ submatrix of K. We denote by $F_{q,n-q}$ a random variate according to F-distribution with q and $n - q$ degrees of freedom.

2. Some properties of likelihood

The likelihood of Z is

$$l = (2\pi)^{-nkp/2} |\Delta|^{-n/2} |\Phi|^{-n(k-1)/2} \exp\Big[-\frac{n}{2} \operatorname{tr}\{(\bar{z}_{\cdot 1} - \theta)(\bar{z}_{\cdot 1} - \theta)' \Delta^{-1}\} - \frac{n}{2} \operatorname{tr}(U \Delta^{-1}) - \frac{n(k-1)}{2} \operatorname{tr}(V \Phi^{-1}) \Big], \tag{2.1}$$

where

$$\bar{z}_{\cdot 1} = \frac{1}{n} \sum_{i=1}^{n} z_{i1}, \quad U = \frac{1}{n} \sum_{i=1}^{n} (z_{i1} - \bar{z}_{\cdot 1})(z_{i1} - \bar{z}_{\cdot 1})', \quad V = \frac{1}{n(k-1)} \sum_{i=1}^{n} \sum_{j=2}^{k} z_{ij} z_{ij}'.$$

The maximum likelihood estimator of θ is $\hat{\theta} = \bar{z}_{\cdot 1}$. The likelihood maximized with respect to θ is given by

$$h(\Delta, \Phi) = (2\pi)^{-nkp/2} |\Delta|^{-n/2} |\Phi|^{-n(k-1)/2} \exp\Big[-\frac{n}{2} \operatorname{tr}(U \Delta^{-1}) - \frac{n(k-1)}{2} \operatorname{tr}(V \Phi^{-1}) \Big]. \tag{2.2}$$

When Φ is unrestricted or Φ is diagonal with equal variance, the maximum likelihood estimators of Δ and Φ were obtained by Anderson et al. [3]. In general, if Φ has a structure, it is difficult to obtain the maximum likelihood estimators as a closed form. If the restriction $\Delta \geq \Phi$ is neglected, the maximum likelihood estimator of Δ is U and the maximum likelihood estimator of Φ is based on V (i.e. Z_2). If Φ is known, we can derive the following lemma.

Lemma 2.1 *Suppose that Φ is known. Then under the restriction $\Delta \geq \Phi$ the function $h(\Delta, \Phi)$ is maximized at*

$$\bar{\Delta} = \Phi^{1/2} Q_G \dot{D}_G Q'_G \Phi^{1/2},$$

where $G = \Phi^{-1/2} U \Phi^{-1/2}$, $G = Q_G D_G Q'_G$, $D_G = \mathrm{diag}(g_1, \dots, g_p)$, $g_1 \geq \cdots \geq g_p > 0$ are the ordered eigenvalues of G, Q_G is an orthogonal matrix, $m_G = \max\{a : g_a > 1\}$, and $\dot{D}_G = \mathrm{diag}(g_1, \dots, g_{m_G}, 1, \dots, 1)$. Further, it holds that

$$\sup_{\Delta \geq \Phi} h(\Delta, \Phi) = |G|^{-n/2} h^*(G) h^{**}(\Phi), \tag{2.3}$$

where

$$h^*(G) = \left(\prod_{a=m_G+1}^{p} g_a \right)^{n/2} \exp\left[-\frac{n}{2} \left(m_G + \sum_{a=m_G+1}^{p} g_a \right) \right],$$

$$h^{**}(\Phi) = (2\pi)^{-nkp/2} |\Phi|^{-nk/2} \exp\left[-\frac{n(k-1)}{2} \mathrm{tr}(V \Phi^{-1}) \right].$$

3. Proof

Proof of Lemma 2.1. Let $\Omega = \Phi^{-1/2} \Delta \Phi^{-1/2}$. Then we have $\Omega \geq I$ and

$$h(\Delta, \Phi) = |\Omega|^{-n/2} \exp\left[-\frac{n}{2} \mathrm{tr}(G \Omega^{-1}) \right] h^{**}(\Phi).$$

Since $\Omega \geq I$, we have a decompose $\Omega = Q_\Omega D_\Omega Q'_\Omega$, where Q_Ω is an orthogonal matrix, $D_\Omega = \mathrm{diag}(\omega_1, \dots, \omega_p), \omega_1 \geq \cdots \geq \omega_p \geq 1$. Here we have

$$\log h(\Delta, \Phi) = -\frac{n}{2} \left\{ \sum_{a=1}^{p} \log \omega_a + \mathrm{tr}[D_G Q'_G Q_\Omega D_\Omega^{-1} Q'_\Omega Q_G] \right\} + \log h^{**}(\Phi).$$

First, we consider the case that g_a's are all different. Since $Q'_G Q_\Omega$ is the orthogonal matrix, the expression is maximized at $Q'_G Q_\Omega = I$, that is, $Q_\Omega = Q_G$

(see von Neumann's lemma in Anderson et al. [3]). Then, the resultant expression is

$$-\frac{n}{2}\sum_{a=1}^{p}\left\{\log\omega_a+\frac{g_a}{\omega_a}\right\}+\log h^{**}(\Phi).$$

Note that $\log x + a/x$ $(a > 0, x > 0)$ is minimized at $x = a$. Under the restriction $\omega_1 \geq \cdots \geq \omega_p \geq 1$, the above formula is maximized at $\overline{\omega}_a = g_a$ for $a = 1, \ldots, m_G$, and $\overline{\omega}_a = 1$ for $a = m_G + 1, \ldots, p$. Therefore, $\overline{\Omega} = Q_G \dot{D}_G Q_G'$ and $\overline{\Delta} = \Phi^{1/2}\overline{\Omega}\Phi^{1/2} = \Phi^{1/2} Q_G \dot{D}_G Q_G' \Phi^{1/2}$. Also, by simple calculations, we can show (2.3). For the case that some of g_a's are equal, the similar arguments can be performed. □

Part II. Wald-type test

In this part, the Wald-type test based on original information is compared with the Wald-type test based on the main information only, say the simplified test.

4. Main results

Consider the Wald-type test statistic denoted by W. Its Type I error, that is, $\mathrm{P}_H(W > c)$ generally depends on nuisance parameters. Then we need to evaluate its supremum, in other words, $\sup_H \mathrm{P}_H(W > c)$. If it is known, the Wald-type test with significance level α has the reject region $\{W > c_\alpha\}$ such that $\sup_H \mathrm{P}_H(W > c_\alpha) = \alpha$.

When $\Phi = \sigma^2 I$ (i.e. $\Sigma = \sigma^2 I$), the supremum of Type I error can be derived by using the same method as in Fujisawa [12]. Letting $F^*_{q,n-q} = [nq/(n-q)]F_{q,n-q}$, we have

Theorem 4.1. *Suppose that $\Phi = \sigma^2 I$. Then*

$$\sup_H \mathrm{P}_H(W > c) = \mathrm{P}(F^*_{q,n-q} > c).$$

The above result can be extended to a more general structure of Φ, satisfying the following conditions:

(*W1) the parameter space of Φ has the zero-matrix as a boundary,

(*W2) there exist the maximum likelihood estimators $\hat{\Delta}$ and $\hat{\Phi}$,

(*W3) $\hat{\Phi}$ converges in probability to the zero-matrix as Φ tends to the zero-matrix, more precisely,

$$\lim_{\Phi\to O} \mathrm{P}(\mathrm{tr}[\hat{\Phi}^2] > \varepsilon) = 0 \qquad \text{for any } \varepsilon > 0.$$

Such examples are given in Section 5. Theorem 4.1 is extended as follows.

THEOREM 4.2 *Suppose that Φ has an appropriate structure provided that* (*W1)–(*W3). *Then*

$$\sup_H \mathrm{P}_H(W > c) = \mathrm{P}(F^*_{q,n-q} > c).$$

Using the canonical form and noting that Z_1 is the main information, we construct the simplified test which is the Wald-type test based on Z_1 (i.e. $\bar{y}_{i\cdot}$'s). If the condition (*W1) is satisfied, the test statistic is expressed as a closed form (see Section 6.1), given by

$$w = n\bar{z}_{\cdot 1}^{(1)\prime} U_{11}^{-1} \bar{z}_{\cdot 1}^{(1)}. \tag{4.1}$$

Its Type I error, that is, $\mathrm{P}_H(w > c)$ does not depend on nuisance parameters. The simplified test with significance level α has the reject region $\{w > c_\alpha\}$ such that $\mathrm{P}_H(w > c_\alpha) = \alpha$.

THEOREM 4.3 *Suppose that Φ has an appropriate structure provided that* $(*W1)$. *Then*

$$P_H(W > c) = P(F^*_{q,n-q} > c).$$

The above two test statistics w and W have a relation (see Section 6.3)

$$w \geq W. \tag{4.2}$$

This relation and the above theorems imply the following main theorem which insists that the simplified test is more efficient than the usual Wald-type test.

THEOREM 4.4 *Suppose that Φ has an appropriate structure provided that* (*W1)–(*W3). *Letting c^*_α be the upper α point of the variate $F^*_{q,n-q}$, we have*

$$\mathrm{P}_H(w > c^*_\alpha) = a, \qquad \sup_H \mathrm{P}_H(W > c^*_\alpha) = \alpha,$$

and

$$\mathrm{P}_K(w > c^*_\alpha) \geq \mathrm{P}_K(W > c^*_\alpha).$$

5. Examples

The various structures will be satisfied with $(*W1) \sim (*W3)$. In this section, two examples are given.

EXAMPLE 5.1 *The case that $\Phi = \sigma^2 I$.* The condition (*W1) is satisfied. The maximum likelihood estimators were obtained by Anderson et al. [3] as follows: Decompose

$$U = Q_1 D_{u1} Q_1',$$

where Q_1 is an orthogonal matrix, $D_{u_1} = \mathrm{diag}(u_{11}, \dots, u_{1p})$, and $u_{11} > \cdots > u_{1p} > 0$ (distinct with probability 1) are the ordered eigenvalues of U. Define

$$s_a = \{(k-1)\mathrm{tr}[V] + u_{1,a+1} + \cdots + u_{1p}\}/(kp - a),$$

$$m_1 = \max\{a : u_{1a} > s_a\},$$

$$\dot{D}_{u_1} = \mathrm{diag}(u_{11}, \dots, u_{1m_1}, s_{m_1}, \dots, s_{m_1}).$$

Then, the maximum likelihood estimators of Δ and σ^2 are

$$\hat{\Delta} = Q_1 \dot{D}_{u_1} Q_1', \qquad \hat{\sigma}^2 = s_{m_1}.$$

The condition (*W2) is satisfied. It can be proved that $\hat{\sigma}^2 \le s_p = \mathrm{tr}[V]/p$ as shown later. Since $n(k-1)\mathrm{tr}[V]/\sigma^2$ has a chi-squared distribution with $n(k-1)p$ degrees of freedom, $\hat{\Phi} = \hat{\sigma}^2 I$ converges in probability to the zero-matrix as Φ tends to the zero-matrix (in other words, σ^2 tends to zero). The condition (*W3) is satisfied.

First we prove the above inequality $\hat{\sigma}^2 \le s_p$ for the case $m_1 = p$. Since $\hat{\sigma}^2 = s_p$, the inequality is true. Next we consider the case $m_1 < p$. The definition of m_1 follows that $u_{1p} \le s_p$. Also,

$$u_{1,p-1} \le s_{p-1} = \frac{(k-1)\mathrm{tr}[V] + u_{1p}}{kp - p + 1} \le \frac{(k-1)ps_p + s_p}{kp - p + 1} = s_p.$$

By induction, it is shown that $u_{1a} \le s_p$ for $a = m_1 + 1, \dots, p$. Hence

$$\hat{\sigma}^2 = s_{m_1} = \frac{(k-1)\mathrm{tr}[V] + u_{1,m_1+1} + \cdots + u_{1p}}{kp - m_1} \le \frac{(k-1)ps_p + (p - m_1)s_p}{kp - m_1} = s_p.$$

EXAMPLE 5.2. *The case that Φ is unrestricted.* The condition (*W1) is satisfied. The maximum likelihood estimators were obtained by Anderson et al. [3] as follows: Decompose

$$U = Q_2 D_{u_2} Q_2', \qquad V = Q_2 Q_2',$$

where Q_2 is a non-singular matrix, $D_{u_2} = \mathrm{diag}(u_{21}, \dots, u_{2p})$, and $u_{21} > \cdots > u_{2p} > 0$ (distinct with probability 1) are the ordered eigenvalues of UV^{-1}. Define

$$m_2 = \max\{a : u_{2a} > 1\},$$

$$\dot{D}_{u_2} = \mathrm{diag}(u_{21}, \dots, u_{2m_2}, 1, \dots, 1),$$

$$\ddot{D}_{u_2} = \mathrm{diag}(1, \dots, 1, u_{2,m_2+1}, \dots, u_{2p}).$$

Then, the maximum likelihood estimators of Δ and Φ are

$$\hat{\Delta} = Q_2 \dot{D}_{u_2} Q_2', \qquad \hat{\Phi} = V - \frac{1}{k} Q_2 (I_p - \ddot{D}_{u_2}) Q_2'.$$

The condition (*W2) is satisfied. Note that $I \geq \ddot{D}_{u_2}$. Therefore, it is seen that $\hat{\Phi} \leq V$. Since $n(k-1)V$ has a Wishart distribution $W_p(n(k-1), \Phi)$, $\hat{\Phi}$ converges in probability to the zero-matrix as Φ tends to the zero-matrix. The condition (*W3) is satisfied.

6. Test statistics and their properties

6.1. Wald-type test statistic based on main information

We derive the Wald-type test statistic w based on Z_1. Remember that z_{i1}'s are independently distributed as $N_p(\boldsymbol{\theta}, \Delta)$. Let $\boldsymbol{\tau}$ be the parameter vector of Δ and let $\xi = (\boldsymbol{\theta}'\boldsymbol{\tau}')'$. Then the Fisher information matrix is

$$-E\left[\frac{\partial^2}{\partial\xi\partial\xi'}(\text{likelihood of } Z_1)\right] = \begin{pmatrix} n\Delta^{-1} & \mathrm{O} \\ \mathrm{O} & * \end{pmatrix}.$$

The condition (*W1) implies that the parameter space of Δ is positive definite. So, the maximum likelihood estimators of $\boldsymbol{\theta}$ and Δ are $\tilde{\boldsymbol{\theta}} = \bar{z}_{\cdot 1}$ and $\tilde{\Delta} = U$. By using a $q \times p$ matrix $C_* = (I_q\mathrm{O})$, the null hypothesis $\mathrm{H} : \boldsymbol{\theta}^{(1)} = 0$ can be written as $\mathrm{H} : C_*\boldsymbol{\theta} = \mathrm{O}$. Therefore, the Wald-type test statistic is given by

$$w = (C_*\tilde{\boldsymbol{\theta}})'[C_*(n\tilde{\Delta}^{-1})^{-1}C_*']^{-1}(C_*\tilde{\boldsymbol{\theta}}) = n\bar{z}_{\cdot 1}^{(1)'}U_{11}^{-1}\bar{z}_{\cdot 1}^{(1)}. \tag{6.1}$$

6.2. Wald-type test statistic based on original information

We derive the Wald-type test statistic W based on Z. Remember the canonical form. Let $\boldsymbol{\tau}$ be the vector of the covariance parameters and let $\xi = (\boldsymbol{\theta}'\boldsymbol{\tau}')'$. Then the Fisher information matrix is

$$-E\left[\frac{\partial^2}{\partial\xi\partial\xi'}(\text{likelihood of } Z)\right] = \begin{pmatrix} n\Delta^{-1} & \mathrm{O} \\ \mathrm{O} & * \end{pmatrix}.$$

The maximum likelihood estimator of $\boldsymbol{\theta}$ is $\hat{\boldsymbol{\theta}} = \bar{z}_{\cdot 1}$. Let $\hat{\Delta}$ be the maximum likelihood estimator of Δ. Then the Wald-type test statistic is

$$W = (C_*\hat{\boldsymbol{\theta}})'[C_*(n\hat{\Delta}^{-1})^{-1}C_*']^{-1}(C_*\hat{\boldsymbol{\theta}}) = n\bar{z}_{\cdot 1}^{(1)'}\hat{\Delta}_{11}^{-1}\bar{z}_{\cdot 1}^{(1)}. \tag{6.2}$$

6.3. Their properties

First we demonstrate some properties of $\hat{\Delta}$. The condition (*W2) admits that $\hat{\Phi}$ is given. Remember the function $h(\Delta, \Phi)$ defined by (2.2). The maximum likelihood estimator $\hat{\Delta}$ is characterized by

$$\sup_{\Delta \geq \hat{\Phi}} h(\Delta, \hat{\Phi}) = h(\hat{\Delta}, \hat{\Phi}).$$

Let $T = \hat{\Phi}^{-1/2} U \hat{\Phi}^{-1/2}$ and decompose $T = Q_T D_T Q_T'$, where Q_T is an orthogonal matrix, $D_T = \mathrm{diag}(t_1, \ldots, t_q)$, and $t_1 \geq \cdots \geq t_p > 0$ are the ordered eigenvalues of T. As a result of Lemma 2.1, the maximum likelihood estimator $\hat{\Delta}$ is expressed as

$$\hat{\Delta} = \hat{\Phi}^{1/2} Q_T \dot{D}_T Q_T' \hat{\Phi}^{1/2},$$

where $\dot{D}_T = \mathrm{diag}(t_1, \ldots, t_{m_T}, 1, \ldots, 1)$ and $m_T = \max\{a : t_a > 1\}$. Noting that $\dot{D}_T \geq D_T$, we know the relation between $\hat{\Delta}$ and U, given by

$$\hat{\Delta} = \hat{\Phi}^{1/2} Q_T \dot{D}_T Q_T' \hat{\Phi}^{1/2} \geq \hat{\Phi}^{1/2} Q_T D_T Q_T' \hat{\Phi}^{1/2} = \hat{\Phi}^{1/2} T \hat{\Phi}^{1/2} = U. \tag{6.3}$$

From (6.1), (6.2), and (6.3), it follows that

$$w \geq W. \tag{6.4}$$

Next we describe a property of m_T. Remember that $T = \hat{\Phi}^{-1/2} U \hat{\Phi}^{-1/2}$ and nU has a Wishart distribution as $W_p(n-1, \Delta)$, where $\Delta = \Psi + \Phi$. The condition (*W3) implies that $\lim_{\mathscr{A}} \mathrm{P}(t_p > 1) = 1$, where $\mathscr{A} = \{\Psi > \mathrm{O}, \Phi \to \mathrm{O}\}$. Since the case $t_p > 1$ is the same as the case $m_T = p$,

$$\lim_{\mathscr{A}} \mathrm{P}(m_T = p) = 1. \tag{6.5}$$

Also, when $m_T = p$, we see that $\hat{\Delta} = U$ and $w = W$. So

$$\lim_{\mathscr{A}} \mathrm{P}(w = W) = 1. \tag{6.6}$$

7. Proofs

PROOF OF THEOREM 4.2. By the virtue of Theorem 4.3 and the inequality $w \geq W$, it holds that

$$\mathrm{P}(F_{q,n-q}^* > c) = \mathrm{P}_H(w > c) \geq \mathrm{P}_H(W > c).$$

If the supremum of the right-hand term attains the upper bound, the theorem is proved. Using the properties illustrated in the previous section, we have

$$\begin{aligned} \mathrm{P}(F_{q,n-q}^* > c) &\geq \sup_H \mathrm{P}_H(W > c) \\ &\geq \lim_{\mathscr{A}} \mathrm{P}_H(W > c) \end{aligned}$$

$$= \lim_{\mathscr{A}} \mathrm{P}_H(W > c, m_T = p)$$

$$= \lim_{\mathscr{A}} \mathrm{P}_H(w > c, m_T = p)$$

$$= \lim_{\mathscr{A}} \mathrm{P}_H(w > c)$$

$$= \mathrm{P}(F^*_{q,n-q} > c).$$

The proof is complete. □

PROOF OF THEOREM 4.3. We know that z_{i1}'s are independently distributed as $\mathrm{N}_p(\boldsymbol{\theta}, \Delta)$. The Wald-type test statistic is the same as Hotelling's T^2-statistic. Its distribution is well-known (see, e.g., Anderson [4]). □

PROOF OF THEOREM 4.4. The result follows from the above theorems and the inequality $w \geq W$. □

Part III. Likelihood ratio test

In this part, the likelihood ratio test based on original information is compared with the likelihood ratio test based on the main information only, say the simplified test.

8. Main results

Consider the likelihood ratio test statistic denoted by Λ. Its Type I error, that is, $\mathrm{P}_H(\Lambda < c)$ generally depends on nuisance parameters. Then we need to evaluate its supremum, in other words, $\sup_H \mathrm{P}_H(\Lambda < c)$. If it is known, the likelihood ratio test with signeficance level α has the reject region $\{\Lambda < c_\alpha\}$ such that $\sup_H \mathrm{P}_H(\Lambda < c_\alpha) = \alpha$.

When $\Phi = \sigma^2 I$(i.e. $\Sigma = \sigma^2 I$), the supremum of Type I error can be derived by using the same method as in Fujisawa [11]. Letting $F^{**}_{q,n-q} = \{1 + F^{**}_{q,n-q}/n\}^{-n/2}$, we have

THEOREM 8.1 *Suppose that* $\Phi = \sigma^2 I$. *Then*

$$\sup_H \mathrm{P}_H(\Lambda < c) = \mathrm{P}(F^{**}_{q,n-q} < c).$$

The above result can be extended to a more general structure of Φ, satisfying the following conditions:

(∗L1) the parameter space of Φ has the zero-matrix as a boundary,

(∗L2) there exist the maximum likelihood estimators $\hat{\Delta}$ and $\hat{\Phi}$ based on Z and there exists the maximum likelihood estimator $\tilde{\Phi}$ based on Z_2, (*L3) $\tilde{\Phi}$ converges in probability to the zero-matrix as Φ tends to the zero-matrix, more precisely,

$$\lim_{\Phi \to O} \mathrm{P}(\mathrm{tr}[\tilde{\Phi}^2] > \varepsilon) = 0 \qquad \text{for any } \varepsilon > 0.$$

The conditions (*L1)–(*L3) are slightly different from the conditions (*W1)–(*W3). Such examples are given in Section 9. Theorem 8.1 is extended as follows.

Theorem 8.2 *Suppose that Φ has an appropriate structure provided that* (*L1)–(*L3). *Then*

$$\sup_H \mathrm{P}_H(\Lambda < c) = \mathrm{P}(F^{**}_{q,n-q} < c).$$

Using the canonical form and noting that Z_1 is the main information, we construct the simplified test which is the likelihood ratio test based on Z_1 (i.e. $\bar{y}_{i\cdot}$'s). If the condition (*L1) is satisfied, the test statistic is expressed as a closed form (see Section 10.1), given by

$$\lambda = \{1 + w/n\}^{-n/2}, \tag{8.1}$$

where w is presented by (4.1). Its Type I error, that is, $\mathrm{P}_H(\lambda < c)$ does not depend on nuisance parameters. The simplified test with significance level α has the reject region $\{\lambda < c_\alpha\}$ such that $\mathrm{P}_H(\lambda < c_\alpha) = \alpha$.

Theorem 8.3 *Suppose that Φ has an appropriate structure provided that* (*L1). *Then*

$$\mathrm{P}_H(\lambda < c) = \mathrm{P}(F^{**}_{q,n-q} < c).$$

The above two test statistics λ and Λ have a relation (see Section 10.3)

$$\lambda \le \Lambda. \tag{8.2}$$

This relation and the above theorems imply the following main theorem which insists that the simplified test is more efficient than the usual likelihood ratio test.

Theorem 8.4 *Suppose that Φ has an appropriate structure provided that* (*L1)–(*L3). *Letting c^*_α be the upper α point of the variate $F^*_{q,n-q}$ and $c^{**}_\alpha = \{1 + c^*_\alpha/n\}^{-n/2}$, we have*

$$\mathrm{P}_H(\lambda < c^{**}_\alpha) = \alpha, \qquad \sup_H \mathrm{P}_H(\Lambda < c^{**}_\alpha) = \alpha.$$

and

$$\mathrm{P}_K(\lambda < c^{**}_\alpha) \ge \mathrm{P}_K(\Lambda < c^{**}_\alpha).$$

From expression (8.1) of λ, it is seen that λ is a monotone decreasing function of w. So, the reject region $\{\lambda < c_\alpha^{**}\}$ is the same as the reject region $\{w > c_\alpha^*\}$. That is, based on Z_1, the likelihood ratio test is the same as the Wald-type test. Therefore, two simplified tests are identical. Also, it is well known that the simplified test is uniformly most powerful among some invariant tests based on Z_1.

9. Examples

The various structures will be satisfied with (*L1)–(*L3). In fact, it is shown that the same two examples as in Section 5 satisfy the conditions.

EXAMPLE 9.1. *The case that* $\Phi = \sigma^2 I$. The condition (*L1) is satisfied. The maximum likelihood estimators $\hat{\Delta}$ and $\hat{\sigma}^2$ based on Z are obtained (see Example 5.1). Also, the maximum likelihood estimator of σ^2 based on Z_2 is well-known, given by

$$\tilde{\sigma}^2 = \mathrm{tr}[V]/p.$$

The condition (*L2) is satisfied. Since $n(k-1)\mathrm{tr}[V]/\sigma^2$ has a chi-squared distribution with $n(k-1)p$ degrees of freedom, $\tilde{\Phi} = \tilde{\sigma}^2 I$ converges in probability to the zero-matrix as Φ tends to the zero-matrix (in other words, σ^2 tends to zero). The condition (*L3) is satisfied.

EXAMPLE 9.2. *The case that* Φ *is unrestricted.* The condition (*L1) is satisfied. The maximum likelihood estimators $\hat{\Delta}$ and $\hat{\Phi}$ based on Z are obtained (see Example 5.2). Also, the maximum likelihood estimator of Φ based on Z_2 is well-known, given by

$$\tilde{\Phi} = V.$$

The condition (*L2) is satisfied. Since $n(k-1)V$ has a Wishart distribution $W_p(n(k-1), \Phi), \tilde{\Phi}$ converges in probability to the zero-matrix as Φ tends to the zero-matrix. The condition (*L3) is satisfied.

10. Test statistics and their properties

10.1. Likelihood ratio test statistic based on main information

We derive the likelihood ratio test statistic λ based on Z_1. Remember that z_{i1}'s are independently distributed as $N_p(\boldsymbol{\theta}, \Delta)$, the null hypothesis is $\mathrm{H} : \boldsymbol{\theta}^{(1)} = \mathbf{0}$, and the alternative hypothesis is $\mathrm{K} : \boldsymbol{\theta}^{(1)} \neq \mathbf{0}$. The condition (*L1) implies that the parameter space of Δ is positive definite. So, the maximum likelihood estimators of $\boldsymbol{\theta}$ and Δ are $\tilde{\boldsymbol{\theta}} = \bar{z}_{\cdot 1}$ and $\tilde{\Delta} = U$. Hence, the likelihood ratio test statistic is given by

$$\lambda = \{1 + w/n\}^{-n/2}. \tag{10.1}$$

10.2. Likelihood ratio test statistic based on original information

We investigate the likelihood ratio test statistic Λ based on Z. Using the formula (2.1) of the likelihood l and considering its maximization with respect to $\boldsymbol{\theta}$, the likelihood ratio test statistic is expressed as

$$\Lambda = \frac{\sup_H l}{\sup_{H \cup K} l} = \frac{\sup_{\Delta \geq \Phi} h_0}{\sup_{\Delta \geq \Phi} h},$$

where the function $h(\Delta, \Phi)$ is defined by (2.2), $H_{11} = \bar{z}_{\cdot 1}^{(1)} \bar{z}_{\cdot 1}^{(1)\prime}$,

$$h_0(\Delta, \Phi) = (2\pi)^{-nkp/2} |\Delta|^{-n/2} |\Phi|^{-n(k-1)/2} \exp\left[-\frac{n}{2} \operatorname{tr}(H_{11} \Delta_{11}^{-1}) - \frac{n}{2} \operatorname{tr}(U \Delta^{-1}) - \frac{n(k-1)}{2} \operatorname{tr}(V \Phi^{-1}) \right]. \tag{10.2}$$

Let U partition as $U = (U_{ij})_{i,j=1,2}$ and let

$$\begin{aligned} S &= \begin{pmatrix} I & \mathrm{O} \\ U_{21} U_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} U_{11} + H_{11} & \mathrm{O} \\ \mathrm{O} & U_{22.1} \end{pmatrix} \begin{pmatrix} I & U_{11}^{-1} U_{12} \\ \mathrm{O} & I \end{pmatrix} \\ &= U + \begin{pmatrix} I \\ U_{21} U_{11}^{-1} \end{pmatrix} H_{11} (I \quad U_{11}^{-1} U_{12}), \end{aligned}$$

where $U_{22.1} = U_{22} - U_{21} U_{11}^{-1} U_{12}$. If the restriction $\Delta \geq \Phi$ is neglected, the function $h_0(\Delta, \Phi)$ is maximized at $\Delta = S$ and $\Phi = \tilde{\Phi}$.

It is difficult to express the numerator of Λ as a closed form. However, we can construct an upper bound and a lower bound of Λ. The upper bound of Λ is suggested as follows:

$$\Lambda = \frac{\sup_{\Delta \geq \Phi} h_0}{\sup_{\Delta \geq \Phi} h} \leq \frac{\sup h_0}{\sup_{\Delta \geq \Phi} h} = \frac{\sup_{\Phi = \tilde{\Phi}} h_0}{\sup_{\Delta \geq \Phi} h} \leq \frac{\sup_{\Phi = \tilde{\Phi}} h_0}{\sup_{\Delta \geq \Phi = \tilde{\Phi}} h} = \overline{\Lambda}.$$

The numerator of $\overline{\Lambda}$ is given by

$$\sup_{\Phi = \tilde{\Phi}} h_0(\Delta, \Phi) = h_0(S, \tilde{\Phi}) = |S|^{-n/2} |\tilde{\Phi}|^{n/2} \exp\left[-\frac{np}{2} \right] h^{**}(\tilde{\Phi}).$$

Let $R = \tilde{\Phi}^{-1/2} U \tilde{\Phi}^{-1/2}, r_1 \geq \cdots \geq r_p > 0$ be the ordered eigenvalues of R, and $m_R = \max\{a : r_a > 1\}$. By the virtue of Lemma 2.1, it is seen that the denominator of $\overline{\Lambda}$ is

$$\sup_{\Delta \geq \Phi = \tilde{\Phi}} h(\Delta, \Phi) = \sup_{\Delta \geq \tilde{\Phi}} h(\Delta, \tilde{\Phi}) = |R|^{-n/2} h^*(R) h^{**}(\tilde{\Phi}).$$

Therefore

$$\overline{\Lambda} = \frac{|S|^{-n/2} |\tilde{\Phi}|^{n/2} \exp[-np/2] h^{**}(\tilde{\Phi})}{|R|^{-n/2} h^*(R) h^{**}(\tilde{\Phi})} = \left\{ \frac{|R||\tilde{\Phi}|}{|S|} \right\}^{n/2} \times \frac{\exp[-np/2]}{h^*(R)} = \lambda \times \overline{v},$$

where $\overline{v} = \exp[-np/2]/h^*(R)$. The result can be dedcuced by using the following relations:

$$|R||\tilde{\Phi}| = |U| = |U_{11}||U_{22.1}|, \qquad |S| = |U_{11} + H_{11}||U_{22.1}|, \qquad H_{11} = \overline{z}_{.1}^{(1)} \overline{z}_{.1}^{(1)'},$$

and

$$\left\{ \frac{|R||\tilde{\Phi}|}{|S|} \right\}^{n/2} = \left\{ \frac{|U_{11}|}{|U_{11} + H_{11}|} \right\}^{n/2} = \left\{ \frac{1}{1 + \overline{z}_{.1}^{(1)'} U_{11}^{-1} \overline{z}_{.1}^{(1)}} \right\}^{n/2} = \{1 + w/n\}^{-n/2} = \lambda.$$

Note that

$$\operatorname{tr}(S\Delta^{-1}) - \operatorname{tr}(U\Delta^{-1}) - \operatorname{tr}(H_{11}\Delta_{11}^{-1}) = \operatorname{tr}[H_{11}(U_{11}^{-1}U_{12} - \Delta_{11}^{-1}\Delta_{12})\Delta_{22.1}^{-1}(U_{11}^{-1}U_{12} - \Delta_{11}^{-1}\Delta_{12})'] \geq 0.$$

Therefore, from the expression (10.2) of the function h_0, we have

$$h_0 \geq \underline{h}_0(\Delta, \Phi) = (2\pi)^{-nkp/2} |\Delta|^{-n/2} |\Phi|^{-n(k-1)/2} \exp\left[-\frac{n}{2} \operatorname{tr}(S\Delta^{-1}) - \frac{n(k-1)}{2} \operatorname{tr}(V\Phi^{-1}) \right].$$

The lower bound of Λ is suggested as follows:

$$\Lambda = \frac{\sup_{\Delta \geq \Phi} h_0}{\sup_{\Delta \geq \Phi} h} \geq \frac{\sup_{\Delta \geq \Phi} \underline{h}_0}{\sup_{\Delta \geq \Phi} h} = \frac{\sup_{\Delta \geq \Phi} \underline{h}_0}{\sup_{\Delta \geq \Phi = \hat{\Phi}} h} \geq \frac{\sup_{\Delta \geq \Phi = \hat{\Phi}} \underline{h}_0}{\sup_{\Delta \geq \Phi = \hat{\Phi}} h} = \underline{\Lambda}.$$

Let $F = \hat{\Phi}^{-1/2} S \hat{\Phi}^{-1/2}$, $f_1 \geq \cdots \geq f_p > 0$ be the ordered eigenvalues of F, and $m_F = \max\{a : f_a > 1\}$. Using Lemma 2.1, we can see that the numerator of

$\underline{\Lambda}$ is

$$\sup_{\Delta \geq \Phi = \hat{\Phi}} \underline{h}_0(\Delta, \Phi) = \sup_{\Delta \geq \hat{\Phi}} \underline{h}_0(\Delta, \hat{\Phi}) = |F|^{-n/2} h^*(F) h^{**}(\hat{\Phi}),$$

and the denominator of $\underline{\Lambda}$ is

$$\sup_{\Delta \geq \Phi = \hat{\Phi}} h(\Delta, \Phi) = \sup_{\Delta \geq \hat{\Phi}} h(\Delta, \hat{\Phi}) = |T|^{-n/2} h^*(T) h^{**}(\hat{\Phi}),$$

where $T = \hat{\Phi}^{-1/2} U \hat{\Phi}^{-1/2}$ (defined in Section 6.3). Therefore

$$\begin{aligned} \Lambda &= \frac{|F|^{-n/2} h^*(F) h^{**}(\hat{\Phi})}{|T|^{-n/2} h^*(T) h^{**}(\hat{\Phi})} \\ &= \left\{ \frac{|T|}{|F|} \right\}^{n/2} \times \frac{h^*(F)}{h^*(T)} \\ &= \lambda \times \underline{\nu}, \end{aligned}$$

where $\underline{\nu} = h^*(F)/h^*(T)$. The result can be deduced by using the following relations:

$$|T||\hat{\Phi}| = |U| = |U_{11}||U_{22.1}|, \quad |F||\hat{\Phi}| = |S| = |U_{11} + H_{11}||U_{22.1}|, \quad H_{11} = \bar{z}_{.1}^{(1)} \bar{z}_{.1}^{(1)\prime},$$

and

$$\left\{ \frac{|T|}{|F|} \right\}^{n/2} = \left\{ \frac{|U_{11}|}{|U_{11} + H_{11}|} \right\}^{n/2} = \lambda.$$

10.3. Their properties

In this subsection we concern with basic properties of $\underline{\nu} = h^*(F)/h^*(T)$ and $\bar{\nu} = \exp[-np/2]/h^*(R)$. The function $h^*(\Omega)$ is monotone increasing for the eigenvalues of Ω (see Section 11). Since $S \geq U$ and $F \geq T$, we have $\underline{\nu} = h^*(F)/h^*(T) \geq 1$. Therefore

$$\lambda \leq \lambda \times \underline{\nu} = \underline{\Lambda} \leq \Lambda \leq \bar{\Lambda} = \lambda \times \bar{\nu}. \tag{10.3}$$

By the similar ways as in Section 6.3, it is shown that

$$\lim_{\mathscr{A}} \mathrm{P}(m_R = p) = 1.$$

Also, when $m_R = p$, it is seen that $\bar{\nu} = 1$. So,

$$\lim_{\mathscr{A}} \mathrm{P}(\bar{\nu} = 1) = 1.$$

11. Proofs

PROOF OF "*the function $h^*(\Omega)$ is monotone increasing for the eigenvalues of Ω*". The logarithm of $2/n$ power of $h^*(\Omega)$ is given by

$$\rho(\omega) = \sum_{a=m_\Omega+1}^{p} \{\log \omega_a - \omega_a + 1\} - p,$$

where $\omega = (\omega_1, \ldots, \omega_p)$, $\omega_1 \geq \cdots \geq \omega_p > 0$ are the ordered eigenvalues of Ω, and $m_\Omega = \max\{a : \omega_a > 1\}$. So, it is seen that $\omega_a \leq 1$ for $a = m_\Omega + 1, \ldots, p$. The function $\kappa(x) = \log x - x + 1$ is monotone increasing and non-positive on $0 < x \leq 1$. Let $\omega^* = (\omega_1^*, \ldots, \omega_p^*), \omega_a^* \geq \omega_a$ for $a = 1, \ldots, p$, and $m_{\Omega^*} = \max\{a : \omega_a^* > 1\}(\geq m_\Omega)$. Then

$$\rho(\omega^*) - \rho(\omega) = \sum_{a=m_{\Omega^*}+1}^{p} \{\kappa(\omega_a^*) - \kappa(\omega_a)\} - \sum_{a=m_\Omega+1}^{m_{\Omega^*}} \kappa(\omega_a) \geq 0. \quad \square$$

PROOF OF THEOREM 8.2. By the virtue of Theorem 8.3 and the inequality (10.3), it holds that

$$\mathrm{P}(F_{q,n-q}^{**} < c) = \mathrm{P}_H(\lambda < c) \geq \mathrm{P}_H(\Lambda < c) \geq \mathrm{P}_H(\lambda \times \bar{\nu} < c).$$

If the supremum of the right-hand term attains the upper bound, the theorem is proved. Using the properties illustrated in the previous section, we have

$$\begin{aligned} \mathrm{P}(F_{q,n-q}^{**} < c) &\geq \sup_H \mathrm{P}_H(\Lambda < c) \\ &\geq \sup_H(\lambda \times \bar{\nu} < c) \\ &\geq \lim_{\mathscr{A}} \mathrm{P}_H(\lambda \times \bar{\nu} < c) \\ &= \lim_{\mathscr{A}} \mathrm{P}_H(\lambda \times \bar{\nu} < c, m_R = p) \\ &= \lim_{\mathscr{A}} \mathrm{P}_H(\lambda < c, m_R = p) \\ &= \lim_{\mathscr{A}} \mathrm{P}_H(\lambda < c) \\ &= \mathrm{P}(F_{q,n-q}^{**} < c). \end{aligned}$$

This completes the proof. $\square$

PROOF OF THEOREM 8.3. This is a direct consequence from expression (8.1) of λ and Theorem 4.3. $\square$

PROOF OF THEOREM 8.4. This theorem follows from the above theorems and the inequality $\lambda \leq \Lambda$. □

Part IV. Other related models

The idea established in previous sections can be applied to other related models. In this part, some results are outlined. The detailed proofs are omitted since the results are similarly proved.

12. Multivariate nested classification model with random effects

A multivariate nested classification model with random effects is given by

$$y_{i_1 i_2 \cdots i_k} = \boldsymbol{\mu} + \boldsymbol{b}_{i_1}^{(1)} + \boldsymbol{b}_{i_1 i_2}^{(2)} + \cdots + \boldsymbol{b}_{i_1 i_2 \cdots i_{k-1}}^{(k-1)} + \boldsymbol{e}_{i_1 i_2 \cdots i_k},$$

$$i_j = 1, \ldots, n_j, \quad j = 1, \ldots, k,$$

where $\boldsymbol{y}_{i_1 i_2 \cdots i_k}$ is a p-component observation vector, $\boldsymbol{\mu}$ is a total mean parameter, $\boldsymbol{b}_{i_1 i_2 \cdots i_j}^{(j)}$ is a ramdom effect, and $\boldsymbol{e}_{i_1 i_2 \cdots i_k}$ is a noise. Assume that $\boldsymbol{b}_{i_1 i_2 \cdots i_j}^{(j)}$'s and $\boldsymbol{e}_{i_1 i_2 \cdots i_k}$'s are mutually independent, $\boldsymbol{b}_{i_1 i_2 \cdots i_j}^{(j)}$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix Σ_j, and $\boldsymbol{e}_{i_1 i_2 \cdots i_k}$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix Σ_k. We also deal with the case that $\Sigma_j (j \geq 2)$ has an appropriate structure. We consider testing the linear hypothesis

$$\mathrm{H} : C\boldsymbol{\mu} = \mathbf{0},$$

against the alternative hypothesis $\mathrm{K} : C\boldsymbol{\mu} \neq \mathbf{0}$, where C is a $q \times p$ known design matrix.

The testing problem is reduced to a canonical form: $z_{ij} (i = 1, \ldots, N_j, j = 1, \ldots, k)$ are mutually independent,

$$z_{i1} \sim \mathrm{N}_p(\boldsymbol{\theta}, \Delta), \qquad z_{ij} \sim \mathrm{N}_p(\mathbf{0}, \Phi_j) \qquad \text{for } j \geq 2,$$

where $\Delta = \Phi_1 \geq \Phi_2 \geq \cdots \geq \Phi_k > \mathrm{O}$, and let $n = n_1 = N_1$. We also deal with the case that $\Phi_j (j \geq 2)$ has an appropriate structure. The null hypothesis becomes

$$\mathrm{H} : \boldsymbol{\theta}^{(1)} = \mathbf{0},$$

against the alternative hypothesis $\mathrm{K} : \boldsymbol{\theta}^{(1)} \neq \mathbf{0}$.

Write $Z_1 = (z_{11} \cdots z_{n1}), Z_2 = (z_{12} \cdots z_{N_2 2} z_{13} \cdots z_{N_k k})$, and $Z = (Z_1 Z_2)$. If the restriction $\Delta = \Phi_1 \geq \Phi_2 \geq \cdots \geq \Phi_k > O$ is neglected, Z_1 is useful but Z_2 is not useful to test the null hypothesis $\mathrm{H} : \boldsymbol{\theta}^{(1)} = \mathbf{0}$ because Z_2 has no information about $\boldsymbol{\theta}$. Due to this property, the former observations may be characterized as the main information and the latter observations as the additional information.

We are interested in comparing two Wald-type test statistics. One is based on Z. The other is based on Z_1, say the simplified test. Let W and w be the corresponding test statistics. If the condition ($*$W1), stated later, is satisfied, it is shown that w is expressed as a closed form (W is not always), given by

$$w = n\bar{\boldsymbol{z}}_{\cdot 1}^{(1)\prime} U_{11}^{-1} \bar{\boldsymbol{z}}_{\cdot 1}^{(1)},$$

where U is the sample covariance matrix of z_{i1}'s. The following theorem insists that the simplified test is more efficient that the usual Wald-type test.

THEOREM 12.1 *Suppose that* ($*$W1) *the parameter space of* Φ_2 *has the zero-matrix as a boundary,* ($*$W2) *there exist the maximum likelihood estimators* $\hat{\Phi}_j$*'s based on* Z, ($*$W3) $\hat{\Phi}_2$ *converges in probability to the zero-matrix as* Φ_2 *tends to the zero-matrix. Letting* c_α^* *be the upper* α *point of the variate* $F_{q,n-q}^*$, *we have*

$$\mathrm{P}_H(w > c_\alpha^*) = \alpha, \qquad \sup_H \mathrm{P}_H(W > c_\alpha^*) = \alpha,$$

and

$$\mathrm{P}_K(w > c_\alpha^*) \geq \mathrm{P}_K(W > c_\alpha^*).$$

Similar results hold for comparing two likelihood ratio test statistics. One is based on Z. The other is based on Z_1, say the simplified test. Let Λ and λ be the corresponding test statistics. If the condition ($*$L1), stated later, is satisfied, it is shown that λ is expressed as a closed form (Λ is not always), given by

$$\lambda = \{1 + w/n\}^{-n/2}.$$

The following theorem insists that the simplified test is more efficient that the usual likelihood ratio test.

THEOREM 12.2 *Suppose that* ($*$L1) *the parameter space of* Φ_2 *has the zero-matrix as a boundary,* ($*$L2) *there exist the maximum likelihood estimators* $\hat{\Phi}_j$*'s based on* Z *and there exist the maximum likelihood estimators* $\tilde{\Phi}_j$*'s based on* Z_2, ($*$L3) $\tilde{\Phi}_2$ *converges in probability to the zero-matrix as* Φ_2 *tends to the zero-matrix. Letting* c_α^* *be the upper* α *point of the variate* $F_{q,n-q}^*$ *and* $c_\alpha^{**} = \{1 + c_\alpha^*/n\}^{-n/2}$, *we have*

$$\mathrm{P}_H(\lambda < c_\alpha^{**}) = \alpha, \qquad \sup_H \mathrm{P}_H(\Lambda < c_\alpha^{**}) = \alpha,$$

and

$$\mathrm{P}_K(\lambda < c_\alpha^{**}) \geq \mathrm{P}_K(\Lambda < c_\alpha^{**}).$$

Note that the likelihood of Z is

$$l = (2\pi)^{-Np/2}|\Delta|^{-N_1/2}\prod_{j=2}^{p}|\Phi_j|^{-N_j/2}\exp\left[\frac{N_1}{2}\mathrm{tr}\{(\bar{z}_{\cdot 1} - \boldsymbol{\theta})(\bar{z}_{\cdot 1} - \boldsymbol{\theta})'\Delta^{-1}\}\right.$$
$$\left. -\frac{N_1}{2}\mathrm{tr}(U\Delta^{-1}) - \sum_{j=2}^{k}\frac{N_j}{2}\mathrm{tr}(V_j\Phi_j^{-1})\right],$$

where $N = \sum_{j=1}^{k} N_j$,

$$V_j = \frac{1}{N_j}\sum_{i=1}^{N_j} z_{ij}z_{ij}'.$$

This expression is similar to (2.1). Therefore, the above theorems can be proved by the same ways as in Parts II and III. It seems that the various structures will be satisfied with the above conditions. However, such an example has not been considered yet.

13. Multivariate one-way classification model with random effects and some groups

A multivariate one-way classification model with random effects and q groups is given by

$$\boldsymbol{y}_{gij} = \boldsymbol{\mu}_g + \boldsymbol{b}_{gi} + \boldsymbol{e}_{gij}, \qquad g = 1,\dots,q, \quad i = 1,\dots,n_g, \quad j = 1,\dots,k,$$

where $\boldsymbol{y}_{gij}$ is a p-component vector of the j-th repeated observation of the i-th individual of the g-th group, $\boldsymbol{\mu}_g$ is a total mean parameter of the g-th group, $\boldsymbol{b}_{gi}$ is a random effect, $\boldsymbol{e}_{gij}$ is a noise, q is the number of groups, n_g is the number of individuals of the g-th group, and k is the number of repeated observations. Assume that $\boldsymbol{b}_{gi}$'s and $\boldsymbol{e}_{gij}$'s are mutually independent, $\boldsymbol{b}_{gi}$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix Γ, and $\boldsymbol{e}_{gij}$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix Σ. We also deal with the case that Σ has an appropriate structure. We consider testing the general linear hypothesis

$$\mathrm{H} : C_1\Theta C_2 = \mathrm{O},$$

against the alternative hypothesis $\mathrm{K} : C_1\Theta C_2 \neq \mathrm{O}$, where $\Theta = (\boldsymbol{\mu}_1 \cdots \boldsymbol{\mu}_q)$, C_1 and C_2 are $c_1 \times p$ and $q \times c_2$ known design matrices, respectively.

The testing problem is reduced to a canonical form: Let Z be the $n \times r$ observation matrix, where $n = \sum_{g=1}^{q} n_g$ and $r = pk$. Partition

$$Z = \begin{pmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{pmatrix},$$

and let $Z_{1(12)} = (Z_{11}Z_{12}), Z_{1(123)} = (Z_{11}Z_{12}Z_{13})$,

$$Z_{(12)(12)} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix},$$

and so on, where Z_{11} is the $c_2 \times c_1$ observation matrix, $Z_{(12)(12)}$ is the $q \times p$ observation matrix. The rows of Z are independent normally distributed with covariance matrix

$$\Omega = \begin{pmatrix} \Delta & \mathrm{O} \\ \mathrm{O} & I_{k-1} \otimes \Phi \end{pmatrix},$$

where Δ and Φ are $p \times p$ positive definite matrices such that $\Delta \geq \Phi$, and

$$E(Z) = \begin{pmatrix} \Xi_{11} & \Xi_{12} & \mathrm{O} \\ \Xi_{21} & \Xi_{22} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} \end{pmatrix},$$

where Ξ_{11} is a $c_2 \times c_1$ matrix of mean parameters, $\Xi = \Xi_{(12)(12)}$ is a $q \times p$ matrix of mean parameters. The null hypothesis becomes

$$\mathrm{H} : \Xi_{11} = \mathrm{O},$$

against the alternative hypothesis $\mathrm{K} : \Xi_{11} \neq \mathrm{O}$.

Let

$$U = \frac{1}{n} Z'_{3(12)} Z_{3(12)}, \qquad V = \frac{1}{n(k-1)} \sum_{j=1}^{k-1} Z_3^{(j)\prime} Z_3^{(j)},$$

where $Z_{(123)3} = (Z_3^{(1)} \cdots Z_3^{(k-1)})$, $Z_3^{(j)}$ is the $n \times p$ matrix. The likelihood of Z is

$$l = (2\pi)^{-nkp/2} |\Delta|^{-n/2} |\Phi|^{-n(k-1)/2} \exp\left[-\frac{1}{2} \mathrm{tr}\{(Z_{(12)(12)} - \Xi)'(Z_{(12)(12)} - \Xi)\Delta^{-1}\} \right.$$
$$\left. - \frac{n}{2} \mathrm{tr}(U\Delta^{-1}) - \frac{n(k-1)}{2} \mathrm{tr}(V\Phi^{-1}) \right].$$

When the restriction $\Delta \geq \Phi$ is neglected, it is seen that $Z_{(12)(12)}$ and U (i.e. $Z_{(123)(12)}$) are useful but V (i.e. $Z_{(123)3}$) is not useful to test the null hyothesis

$H : \Xi_{11} = O$ because $Z_{(123)3}$ has no information about Ξ. Therefore, the former observations may be characterized as the main information and the latter observations as the additional information.

We are interested in comparing two Wald-type test statistics. One is based on Z. The other is based on $Z_{(123)(12)}$, say the simplified test. Let W and w be the corresponding test statistics. If the condition (*W1), stated later, is satisfied, it is shown that w is expressed as a closed form (W is not always), given by

$$w = n\,\mathrm{tr}(H_{11} U_{11}^{-1}),$$

where $H_{11} = Z_{11}' Z_{11}/n, U_{11}$ is the first $c_1 \times c_1$ matrix of U. That is called Lawley-Hotelling-type test statistic. The following theorem insists that the simplified test is more efficient that the usual Wald-type test.

THEOREM 13.1 *Suppose that* (*W1) *the parameter space of* Φ *has the zero-matrix as a boundary,* (*W2) *there exist the maximum likelihood estimators* $\hat{\Lambda}$ *and* $\hat{\Phi}$ *based on* Z, (*W3) $\hat{\Phi}$ *converges in probability to the zero-matrix as* Φ *tends to the zero-matrix. Let* c_α *be the upper* α *point of the variate* $n\ \mathrm{tr}(S_1 S_2^{-1})$ *where* S_1 *and* S_2 *are independently distributed as* $W_{c_1}(c_2, I)$ *and* $W_{c_1}(n-q, I)$, *respectively. Then we have*

$$\mathrm{P}_H(w > c_\alpha) = \alpha, \qquad \sup_H \mathrm{P}_H(W > c_\alpha) = \alpha,$$

and

$$\mathrm{P}_K(w > c_\alpha) \geq \mathrm{P}_K(W > c_\alpha).$$

Similar results hold for comparing two likelihood ratio test statistics. One is based on Z. The other is based on $Z_{(123)(12)}$, say the simplified test. Let Λ and λ be the corresponding test statistics. If the condition (*L1), stated later, is satisfied, it is shown that λ is expressed as a closed form (Λ is not always), given by

$$\lambda = (|U_{11}|/|H_{11} + U_{11}|)^{n/2}.$$

That is called Λ-type test statistic. The following theorem insists that the simplified test is more efficient that the usual likelihood ratio test.

THEOREM 13.2 *Suppose that* (*L1) *the parameter space of* Φ *has the zero-matrix as a boundary,* (*L2) *there exist the maximum likelihood estimators* $\hat{\Lambda}$ *and* $\hat{\Phi}$ *based on* Z *and there exists the maximum likelihood estimator* $\tilde{\Phi}$ *based on* $Z_{(123)3}$, (*L3) $\tilde{\Phi}$ *converges in probability to the zero-matrix as* Φ *tends to the zero-matrix. Let* c_α *be the lower* α *point of the variate* $(|S_2|/|S_1 + S_2|)^{n/2}$ *where*

S_1 and S_2 are independently distributed as $W_{c_1}(c_2, I)$ and $W_{c_1}(n-q, I)$, respectively. Then we have

$$\mathrm{P}_H(\lambda < c_\alpha) = \alpha, \qquad \sup_H \mathrm{P}_H(\Lambda < c_\alpha) = \alpha,$$

and

$$\mathrm{P}_K(\lambda < c_\alpha) \geq \mathrm{P}_K(\Lambda < c_\alpha).$$

The expression of the likelihood is also similar to (2.1). The above theorems can be proved by the same ways as in Parts II and III. The above conditions are satisfied for two cases: Φ is unrestricted, $\Phi = \sigma^2 I$. It may be noted that in this model w and λ are not identical.

14. Random coefficient growth curve model

A random coefficient growth curve model is given by

$$\boldsymbol{y}_i = X\boldsymbol{\beta}_i + \boldsymbol{e}_i, \qquad \boldsymbol{\beta}_i = \Theta\boldsymbol{a}_i + \boldsymbol{\eta}_i, \; i = 1, \ldots, n,$$

where $\boldsymbol{y}_i$ is a $r \times 1$ observation vector of the i-th individual, X is a $r \times p$ design matrix within individuals, $\boldsymbol{\beta}_i$ is a $p \times 1$ random coefficient vector, $\boldsymbol{e}_i$ is a $r \times 1$ noise vector, Θ is a $p \times q$ matrix of mean parameters, $A = (\boldsymbol{a}_1 \cdots \boldsymbol{a}_n)'$ is a $n \times q$ design matrix between individuals, $\boldsymbol{\eta}_i$ is a $p \times 1$ random effect vector, and n is the number of individuals. Assume that $\boldsymbol{e}_i$'s and $\boldsymbol{\eta}_i$'s are mutually independent, $\boldsymbol{e}_i$ is distributed as $\mathrm{N}(\boldsymbol{0}, \sigma^2 I_r)$, and $\boldsymbol{\eta}_i$ is distributed as $\mathrm{N}(\boldsymbol{0}, \Gamma)$. In this model the variations of the individuals are taken as random effect. We consider testing the general linear hypothesis

$$\mathrm{H}: C_1\Theta C_2 = \mathrm{O},$$

against the alternative hypothesis $\mathrm{K}: C_1\Theta C_2 \neq \mathrm{O}$, where C_1 and C_2 are $c_1 \times p$ and $q \times c_2$ known design matrices, respectively.

The testing problem is reduced to a canonical form, which is the almost same as in Section 13 except that the covariance matrix of each row of Z is

$$\Omega = \begin{pmatrix} \Delta & \mathrm{O} \\ \mathrm{O} & \sigma^2 I_{r-p} \end{pmatrix},$$

and the restriction is $\Delta \geq \sigma^2 I$. After the similar discussions as in Section 13, we obtain the following results.

We are interested in comparing two Wald-type test statistics. One is based on Z. The other is based on $Z_{(123)(12)}$, say the simplified test. Let W and w be the corresponding test statistics. It is shown that w is expressed as a closed form, given by

$$w = n\,\mathrm{tr}(H_{11}U_{11}^{-1}).$$

The following theorem insists that the simplified test is more efficient that the usual Wald-type test.

THEOREM 14.1 *Let c_α be the upper α point of the variate n $\mathrm{tr}(S_1S_2^{-1})$ where S_1 and S_2 are independently distributed as $W_{c_1}(c_2, I)$ and $W_{c_1}(n-q, I)$, respectively. Then we have*

$$\mathrm{P}_H(w > c_\alpha) = \alpha, \qquad \sup_H \mathrm{P}_H(W > c_\alpha) = \alpha,$$

and

$$\mathrm{P}_K(w > c_\alpha) \geq \mathrm{P}_K(W > c_\alpha).$$

Similar results are obtained for comparing two likelihood ratio test statistics. One is based on Z. The other is based on $Z_{(123)(12)}$, say the simplified test. Let Λ and λ be the corresponding test statistics. It is shown that λ is expressed as a closed form (Λ is not always), given by

$$\lambda = (|U_{11}|/|H_{11} + U_{11}|)^{n/2}.$$

The following theorem insists that the simplified test is more efficient that the usual likelihood ratio test.

THEOREM 14.2 *Let c_α be the lower α point of the variate $(|S_2|/|S_1 + S_2|)^{n/2}$ where S_1 and S_2 are independently distributed as $W_{c_1}(c_2, I)$ and $W_{c_1}(n-q, I)$, respectively. Then we have*

$$\mathrm{P}_H(\lambda < c_\alpha) = \alpha, \qquad \sup_H \mathrm{P}_H(\Lambda < c_\alpha) = \alpha,$$

and

$$\mathrm{P}_K(\lambda < c_\alpha) \geq \mathrm{P}_K(\Lambda < c_\alpha).$$

Acknowledgements

The author would like to express the deep gratitude to Professor Yasunori Fujikoshi of Hiroshima university for his encouragements. The author would like to thank Mr. Kanta Naito and Mr. Kenichi Satoh for their comments and suggestions.

References

[1] Y. Amemiya, What should be done when an estimated between-group covariance matrix is not nonnegative definite?, Amer. Statist., **39** (1985), 112–117.

[2] Y. Amemiya, T. W. Anderson, and P. A. W. Lewis, Percentage points for a test of rank in multivariate components of variance, Biometrika, **77** (1990), 637–641.

[3] B. M. Anderson, T. W. Anderson, and I. Olkin, Maximum likelihood estimators and likelihood ratio criteria in multivariate components of variance, Ann. Statist., **14** (1986), 405–417.

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

[5] T. W. Anderson, Estimating linear statistical relationships, Ann. Statist., **12** (1984), 1–45.

[6] T. W. Anderson, The asymptotic distribution of the likelihood ratio criterion for testing rank in multivariate components of variance, J. Multivariate Anal., **30** (1989), 72–79.

[7] T. W. Anderson and Y. Amemiya, Testing dimensionality in the multivariate analysis of variance, Statist. Probab. Lett., **12** (1991), 445–463.

[8] J. A. Calvin, REML estimation in unbalanced multivariate variance components models using an EM algorithm, Biometrics, **49** (1993), 691–701.

[9] J. A. Calvin and R. L. Dykstra, Maximum likelihood estimation of a set of covariance matrices under Löwner order restrictions with applications to balanced multivariate variance components models, Ann. Statist., **19** (1991), 850–869.

[10] J. A. Calvin and R. L. Dykstra, REML estimation of covariance matrices with restricted parameter spaces, J. Amer. Statist. Assoc., **90** (1995), 321–329.

[11] H. Fujisawa, Likelihood ratio criterion for mean structure in the growth curve model with random effects, J. Multivariate Anal., (1996) (to appear).

[12] H. Fujisawa, Wald-type criterion for mean structure in the growth curve model with random effects, Statistical research group, Hiroshima Univ., (1996) TR. 96–3.

[13] H. Fujisawa, Testing for mean structure in multivariate one-way classification model with random effects, Statistical research group, Hiroshima Univ., (1996) TR. 96–13.

[14] T. Isogai and Y. Fujikoshi, Lower bounds for the distributions of certain multivariate test statistics, J. Multivariate Anal., **6** (1976), 250–255.

[15] C. G. Khatri and C. R. Rao, Multivariate linear model with latent variables: problems of estimation, Center for Multivariate Analysis, Penn State Univ., (1988) TR. No.88–48.

[16] S. Kuriki, One-sided test for the equality of two covariance matrices, Ann. Statist., **21** (1993), 1379–1384.

[17] S. Kuriki, Likelihood ratio tests for covariance structure in random effects model, J. Multivariate Anal., **46** (1993), 175–197.

[18] J. R. Magnus and H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, Wiley, New York, (1988).

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

[20] G. Reinsel, Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure, J. Amer. Statist. Assoc., **77** (1982), 190–195.

[21] S. Remadi and Y. Amemiya, Asymptotic properties of the estimators for multivariate components of variance, J. Multivariate Anal., **49** (1994), 110–131.

[22] D. von Rosen, The growth curve model: a review, Commun. Statist. Theory Meth., **20** (1991), 2791–2822.

[23] G. Saw, A lower bound for the distribution of a partial product of latent roots, Commun. Statist. Theory Meth., **3** (1974), 665–669.

[24] J. R. Schott, Optimal bounds for the distributions of some test criteria for tests of dimensionality, Biometrika, **71** (1984), 561–567.

[25] J. R. Schott and J. G. Saw, A multivariate one-way classification model with random effects, J. Multivariate Anal., **15** (1984), 1–12.

[26] S. R. Searle, G. Casella, and C. E. McCulloch, Variance components, Wiley, New York, (1992).

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

[28] A. Suzukawa, Linear hypothesis testing in a random effects growth curve model, J. Japan Statist. Soc., (1996) (to appear).

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

[30] J. H. Ware, Linear models for the analysis of longitudinal studies, Amer. Statist., **39** (1985), 95–101.

[31] T. Yokoyama, LR test for random-effects covariance structure in a parallel profile model, Ann. Inst. Statist. Math., **47** (1995), 309–320.

[32] T. Yokoyama and Y. Fujikoshi, Tests for random-effects covariance structures in the growth curve model with covariates, Hiroshima Math. J., **22** (1992), 195–202.

[33] T. Yokoyama and Y. Fujikoshi, A parallel profile model with random-effects covariance structure, J. Japan Statist. Soc., **23** (1993), 83–89.

[34] T. Yokoyama, Extended growth curve models with random-effects covariance structures, Commun. Statist. Theory Meth., **25** (1996), 571–584.

Department of Mathematics
Faculty of Science
Hiroshima University
Higashi-Hiroshima 739, Japan