数理解析研究所講究録
第682巻 1989年 123-141

SEQUENTIAL ESTIMATION IN A GROWTH CURVE MODEL

Tatsuya Kubokawa (久保川達也)
University of Tsukuba

Abstract. For a coefficient matrix in the growth curve model, the paper gives a two-stage estimation procedure such that its risk function relative to arbitrary quadratic loss is bounded above by a preassigned constant, and the asymptotic efficiency is discussed. Also for powers of the generalized variance, a procedure with a bounded risk is developed and some domination results are shown.

Key words and phrases: Growth curve model; Two-stage sampling rule; Asymptotic efficiency; Point estimation.

1. INTRODUCTION

Let $x_1, x_2, \ldots$ be a sequence of mutually independent random vectors, x_i having p-variate normal distribution $N_p(B\xi a_i, \Sigma)$ where B $(p\times q)$ is a known matrix of rank q, a_i $(r\times 1)$ is a known vector and ξ $(q\times r)$, Σ $(p\times p)$ are unknown matrices. Denote $X_n=(x_1, x_2, \ldots, x_n)$, $A_n=(a_1, a_2, \ldots, a_n)$ and $\omega=(\xi, \Sigma)$. Then the random matrix X_n $(p\times n)$ has $N_{p,n}(B\xi A_n; \Sigma, I_n)$, which is called a growth curve model by Potthoff and Roy (1964) and practical meaning and applications can be seen in their paper. If we put $q=p$ and $B=I$, we get an ordinary multivariate regression model.

Here we consider two problems for estimation of the coefficient matrix ξ and powers of the generalized variance $|\Sigma|^{\alpha}$, $\alpha\neq 0$ which are described below.

Problem (1). Given preassigned number $\varepsilon>0$, we want to construct estimator $d_1=d_1(\varepsilon)$ of ξ such that

$R_1(\omega, d_1) = E_\omega[\mathrm{tr}\ Q(d_1-\xi)(d_1-\xi)'] \leq \varepsilon$ for all ω, (consistency)

where Q $(p\times p)$ is a positive definite matrix.

Problem (2). For given $\varepsilon>0$, we want to find estimator $d_2=d_2(\varepsilon)$ of powers of the generalized variance $|\Sigma|^\alpha$ such that

$$R_2(\omega,d_2) = E_\omega[(d_2-|\Sigma|^\alpha)^2] \leq \varepsilon \qquad \text{for all } \omega \quad \text{(consistency)}.$$

Throughout the paper, let m_0 be the smallest integer n ($\geq$r) such that the rank of A_n is r. In the case where Σ is known, we shall find a procedure satisfying the requirement of Problem (1). For integer $n\geq m_0$, MLE of ξ is given by

$$\hat{\xi}_0(n) = (B'\Sigma^{-1}B)^{-1}B'\Sigma^{-1}X_nA_n'(A_nA_n')^{-1},$$

and from Muirhead (1982) and Sugiura and Kubokawa (1988),

$$(1.1) \qquad R_1(\omega,\hat{\xi}_0(n))=E[trQ(\hat{\xi}_0(n)-\xi)(\hat{\xi}_0(n)-\xi)']$$
$$=E[\{vec(\hat{\xi}_0(n)-\xi)\}'(I\otimes Q)vec(\hat{\xi}_0(n)-\xi)]$$
$$=tr(I\otimes Q)Cov(vec\ \hat{\xi}_0(n))$$
$$=tr(I\otimes Q)\{(A_nA_n')^{-1}\otimes(B'\Sigma^{-1}B)^{-1}\}$$
$$=tr(A_nA_n')^{-1}trQ(B'\Sigma^{-1}B)^{-1},$$

where the notation vec ξ denotes $qr\times1$ vector $(\xi_1',\ldots,\xi_r')'$ for $\xi=(\xi_1,\ldots,\xi_r)$ and $A\otimes B$ stands for kronecker product defined by $(a_{ij}B)$ for $A=(a_{ij})$. Hence we get that $R_1(\omega,\hat{\xi}_0(n^*))\leq\varepsilon$, all ω for integer n^* defined by

n^* = smallest integer n ($\geq m_0$) such that

$$(1.2) \qquad [tr(A_nA_n')^{-1}]^{-1} \geq trQ(B'\Sigma^{-1}B)^{-1}/\varepsilon.$$

Since Σ is unknown, however, there does not exist any fixed sample size such that Problem (1) holds. So, it is desired to obtain an estimation procedure resolving Problem (1). Recently Kubokawa (1988e) treated a similar problem with respect to the covariance criterion. For our purpose, arguments used there are heavily exploited.

In Section 2, based on the idea of Stein (1945), we construct a two-stage estimation rule for Problem (1), which is a multivariate extension of Rao (1973). Also the efficiency is discussed. A two-stage procedure proposed by Takada (1988b) in the ordinary multivariate case [$p=q, r=1, B=I, a_i=1$] is consistent, but it is not asymptotically efficient. The procedure given in Section 2 just possesses both properties. Section 3 shows that the two-stage estimator in Section 2 is improved on by combined estimators when an additional sample is taken. Two-stage procedures resolving Problem (2) are developed in Section 4.

2. TWO-STAGE ESTIMATION PROCEDURES FOR ξ

2.1 A consistent two-stage procedure. For preassigned number $\varepsilon>0$, we propose the following two-stage sampling rule.

(i) Start with m observations $x_1, x_2, \ldots, x_m$ for $m \geq \max(m_0, p-q+r+3)$, each x_i having $N_p(B\xi a_i, \Sigma)$.

(ii) Define the stopping number by

N = smallest integer n ($\geq m$) such that

$$[\mathrm{tr}(A_nA_n')^{-1}]^{-1} \geq k_m \mathrm{tr}Q(B'S_m^{-1}B)^{-1}/\varepsilon, \tag{2.1}$$

where

$$k_m = (m-r-1)/\{(m-p+q-r-1)(m-p+q-r-2)\} \tag{2.2}$$

and $S_m = X_m(I-A_m'(A_mA_m')^{-1}A_m)X_m'$ with $X_m=(x_1,\ldots,x_m)$, $A_m=(a_1,\ldots,a_m)$.

(iii) Take a sample of size N-m, and estimate ξ by

$$\hat{\xi}_N = (B'S_m^{-1}B)^{-1}B'S_m^{-1}X_NA_N'(A_NA_N')^{-1}, \tag{2.3}$$

for $X_N=(X_m, x_{m+1}, \ldots, x_N)$ and $A_N=(A_m, a_{m+1}, \ldots, a_N)$.

Then we get

Theorem 2.1. (consistency) The estimator $\hat{\xi}_N$ given by (2.3) is a solution of Problem (1).

Proof. Since $X_N A_N' = X_m A_m' + x_{m+1} a_{m+1}' + \cdots + x_N a_N'$, the independence of $X_m A_m'$ and S_m gives that the conditional distribution of $X_N A_N'$ given S_m has $N_{p,r}(B\xi A_N A_N'; \Sigma, A_N A_N')$. Thereby,

$$\begin{aligned} \mathrm{Cov}(\mathrm{vec}\ \hat{\xi}_N) &= E[\mathrm{Cov}(\mathrm{vec}\ \hat{\xi}_N | S_m)] \\ &= E[(A_N A_N')^{-1} \otimes G(S_m, \Sigma)], \end{aligned}$$

where

$$G(S_m, \Sigma) = (B' S_m^{-1} B)^{-1} B' S_m^{-1} \Sigma S_m^{-1} B (B' S_m^{-1} B)^{-1}$$

and $\mathrm{Cov}(\cdot | S_m)$ designates the conditional covariance matrix given S_m. Note that $G(S_m, \Sigma)$ is independent of $B' S_m^{-1} B$ by lemma 2.1 of Sugiura and Kubokawa (1988). Then $G(S_m, \Sigma)$ is independent of N, which implies that

$$\begin{aligned} \mathrm{Cov}(\mathrm{vec}\ \hat{\xi}_N) &= E[(A_N A_N')^{-1}] \otimes E[G(S_m, \Sigma)] \\ &= \frac{m-r-1}{m-p+q-r-1} E[(A_N A_N')^{-1}] \otimes (B' \Sigma^{-1} B)^{-1}, \end{aligned}$$

since $E[G(S_m, \Sigma)] = \{(m-r-1)/(m-p+q-r-1)\}(B' \Sigma^{-1} B)^{-1}$ as shown by Rao (1967), Williams (1967), Gleser and Olkin (1972) and Sugiura and Kubokawa (1988). Hence from (1.1) and the definition of N given by (2.1), we obtain that

$$\begin{aligned} R_1(\omega, \hat{\xi}_N) &= \mathrm{tr}(I \otimes Q) \mathrm{Cov}(\mathrm{vec}\ \hat{\xi}_N) \\ &= \frac{m-r-1}{m-p+q-r-1} E[\mathrm{tr}(A_N A_N')^{-1}] \mathrm{tr} Q (B' \Sigma^{-1} B)^{-1} \qquad (2.4) \\ &\le \varepsilon (m-p+q-r-2) E[\mathrm{tr} Q (B' \Sigma^{-1} B)^{-1} / \mathrm{tr} Q (B' S_m^{-1} B)^{-1}]. \end{aligned}$$

Here, $\mathrm{tr} Q (B' S_m^{-1} B)^{-1} = \mathrm{tr} W (B' \Sigma^{-1} B)^{-1/2} Q (B' \Sigma^{-1} B)^{-1/2}$ for $W = (B' \Sigma^{-1} B)^{1/2} (B' S_m^{-1} B)^{-1} (B' \Sigma^{-1} B)^{1/2}$ having $W_q(m-p+q-r, I)$. Denote $\mathrm{diag}(\sigma_1, \ldots, \sigma_q) = H' (B' \Sigma^{-1} B)^{-1/2} Q (B' \Sigma^{-1} B)^{-1/2} H$ for some orthogonal matrix H. From the Bartlett's decomposition, we have

$$\mathrm{tr} Q (B' S_m^{-1} B)^{-1} = \Sigma_{i=1}^{q} \sigma_i w_i, \qquad (2.5)$$

where $W_1, \ldots, W_q$ are mutually independent random variables, each being distributed as $\chi^2_{m-p+q-r}$. Letting

(2.6) $\lambda_i = \sigma_i / \Sigma_{j=1}^{q} \sigma_j, \quad i=1,\ldots,q,$

we can see that

(2.7) $E[trQ(B'\Sigma^{-1}B)^{-1}/trQ(B'S_m^{-1}B)^{-1}] = E[(\Sigma_{i=1}^{q}\lambda_i W_i)^{-1}],$

(2.8) $E[W_i^{-1}] = (m-p+q-r-2)^{-1}, \quad i=1,\ldots,q.$

Therefore from (2.4), (2.7) and (2.8), we can get the required conclusion if the following inequality holds:

(2.9) $E[(\Sigma_{i=1}^{q}\lambda_i W_i)^{-1}] \leqq \max_{1\leqq i\leqq q}\{E[W_i^{-1}]\}.$

The case of q=2 follows from the convexity of $f(\lambda_1)=E[\{\lambda_1 W_1+(1-\lambda_1)W_2\}^{-1}]$. When $q\geqq 3$, by induction, we have

$$\begin{aligned} E[(\Sigma_{i=1}^{q}\lambda_i W_i)^{-1}] &= E[\{\lambda_q W_q + (1-\lambda_q)(\Sigma_{i=1}^{q-1}\lambda_i W_i/\Sigma_{j=1}^{q-1}\lambda_j)\}^{-1}] \\ &\leqq \max\{E[W_q^{-1}],\ E[\{\Sigma_{i=1}^{q-1}(\lambda_i/\Sigma_{j=1}^{q-1}\lambda_j)W_i\}^{-1}]\} \\ &\leqq \max_{1\leqq i\leqq q}\{E[W_i^{-1}]\}, \end{aligned}$$

which establishes (2.9), and the proof of Theorem 2.1 is complete.

Now we show the asymptotic efficiency of the stopping number in the sense of Chow and Robbins (1965), that is, $\lim_{\varepsilon\to 0}E[N]/n^*=1$. The method is due to that of Mukhopadyay (1980) for the univariate case.

Theorem 2.2. (asymptotic efficiency) Assume that $n^{-1}(A_n A_n') \to \Omega > 0$ as $n\to\infty$, and that $m=O(\varepsilon^{-d})$ for $0<d<1$. Then $\lim_{\varepsilon\to 0}E[N]/n^*=1$.

Proof. For simplicity, denote $g(n)=n\,tr(A_nA_n')^{-1}$, which converges to $tr\Omega^{-1}$ as n tends to infinity. From the definitions of n^* and N given by (1.2) and (2.1) respectively,

$$k_m g(N)trQ(B'S_m^{-1}B)^{-1}[\frac{n^*}{n^*-1}g(n^*-1)trQ(B'\Sigma^{-1}B)^{-1}+m_0\varepsilon]^{-1} < N/n^*$$

$$< [k_m\cdot\frac{N}{N-1}g(N-1)trQ(B'S_m^{-1}B)^{-1}+\varepsilon m][g(n^*)trQ(B'\Sigma^{-1}B)^{-1}]^{-1}.$$

Note that if ε tends to zero, then $m\to\infty$, $\varepsilon m\to 0$, $n^*\to\infty$, $N\to\infty$ a.s., $trQ(B'S_m^{-1}B)^{-1}/(m-p+q-r)\to trQ(B'\Sigma^{-1}B)^{-1}$ a.s.. Hence we can get the required conclusion if the followings hold:

(2.10) $$\lim_{\varepsilon\to 0}E[g(N)trQ(B'S_m^{-1}B)^{-1}/(m-p+q-r)]=tr\Omega^{-1}trQ(B'\Sigma^{-1}B)^{-1},$$

(2.11) $$\lim_{\varepsilon\to 0}E[\frac{N}{N-1}g(N-1)trQ(B'S_m^{-1}B)^{-1}/(m-p+q-r)]=tr\Omega^{-1}trQ(B'\Sigma^{-1}B)^{-1}.$$

For (2.10), it must be shown that $g(N)trQ(B'S_m^{-1}B)^{-1}/(m-p+r-r)$ is bounded above by an integrable function which is independent of ε. Since $\lim_{\varepsilon\to 0}g(n)=tr\Omega^{-1}$, there is some n_0 such that for any $n>n_0$, $g(n)<tr\Omega^{-1}+1$. Hence for $N>n_0$, $g(N)<tr\Omega^{-1}+1$, so that for any $N\geq 1$,

$$g(N) < g(1)+\cdots+g(n_0)+\{tr\Omega^{-1}+1\} = c_0 \quad \text{(say)}.$$

Then from (2.5),

$$g(N)trQ(B'S_m^{-1}B)^{-1}/(m-p+q-r) \leq c_0\Sigma_{i=1}^q\sigma_iW_i/(m-p+q-r)$$
$$\leq qc_0\max_{1\leq i\leq q}(\sigma_i)\sup_{m\geq m_0}\{T_m\},$$

where $T_m=\Sigma_{i=1}^qW_i/\{q(m-p+q-r)\}$. Since $\Sigma_{i=1}^qW_i \sim \chi^2_{q(m-p+q-r)}$, lemma 3 of Simons (1968) implies that T_m is a reversed martingale, so that by Doob's inequality,

$$E[\sup_{m\geq m_0}(T_m)] \leq \frac{e}{e-1}\{1+E[T_{m_0}\log^+T_{m_0}] < \infty,$$

where $\log^+u=\max(\log u, 1)$. This inequality means that $g(N)trQ(B'S_m^{-1}B)^{-1}/(m-p+q-r)$ is bounded above by the integrable function independent of ε. Finally, applying the dominated convergence theorem, we get (2.10). Similarly, we can verify (2.11) and Theorem 2.2 is proved.

Remark. In the ordinary multivariate case $[p=q, r=1, B=I, a_i=1]$, Takada (1988b) proposed the stopping number

$$N = \max\{[\frac{p}{\varepsilon(m-p-2)} ch_{max}(S_m)]+1,\ m\},$$

where $[u]$ designates the integer part of u and $ch_{max}(S_m)$ is the largest characteristic root of S_m. As remarked in the paper, his two-stage procedure is consistent but is not asymptotically efficient when $p \geqq 2$. For any p, our procedure satisfies the requirements of both consistency and asymptotic efficiency.

As a measure of stronger efficiency, Simons (1968) and Star and Woodroofe (1968) considered an asymptotically bounded regret, that is, $\lim_{\varepsilon\to 0} E[N-n^*]<\infty$, and Ghosh and Mukhopadyay (1981) called it second order efficiency. Takada (1988a) demonstrated that Rao's two-stage method is not second order asymptotically efficient. It is shown that the same result holds in our general model.

Theorem 2.3. Assume that $ntr(A_nA_n')^{-1}=tr\Omega^{-1}+O(n^{-1})$, and that $m=O(\varepsilon^{-d})$, $\lim_{\varepsilon\to 0}\varepsilon^d m>0$ for $1/2\leqq d<1$. Then $\lim_{\varepsilon\to 0}E[N-n^*]=\infty$.

Proof. From the definitions of n^* and N,

$$E[N-n^*] \geqq \frac{1}{\varepsilon}\{k_m E[g(N)trQ(B'S_m^{-1}B)^{-1}]-\frac{n^*}{n^*-1}g(n^*-1)trQ(B'\Sigma^{-1}B)^{-1}\}-m_0$$

$$(2.12) \qquad = \frac{1}{\varepsilon}\{k_m E[trQ(B'S_m^{-1}B)^{-1}]-\frac{n^*}{n^*-1}trQ(B'\Sigma^{-1}B)^{-1}\}tr\Omega^{-1}-m_0$$

$$+\frac{1}{\varepsilon}k_m E[D_N trQ(B'S_m^{-1}B)^{-1}]-\frac{n^*}{n^*-1}\cdot\frac{1}{\varepsilon}D_{n^*-1}trQ(B'\Sigma^{-1}B)^{-1},$$

where $D_n = g(n) - \mathrm{tr}\Omega^{-1}$. Also the definition of N gives that

$$\frac{1}{\varepsilon}k_m \mathrm{tr}Q(B'S_m^{-1}B)^{-1} > [\mathrm{tr}(A_{N-1}A'_{N-1})^{-1}]^{-1} = \frac{N-1}{g(N-1)},$$

so that

$$\frac{1}{\varepsilon}k_m E[D_N \mathrm{tr}Q(B'S_m^{-1}B)^{-1}] > E[\frac{1}{g(N-1)}\cdot\frac{N-1}{N}ND_N]. \tag{2.13}$$

Since $g(n-1)\to\mathrm{tr}\Omega^{-1}$ and $nD_n = O(1)$ as $n\to\infty$, it can be seen that the r.h.s. in (2.13) is bounded below. Clearly, $n^*\{\varepsilon(n^*-1)\}^{-1}D_{n^*-1}$ $\mathrm{tr}Q(B'\Sigma^{-1}B)^{-1}$ is bounded above, so that for some constant M_0,

$$\begin{aligned} E[N-n^*] &\geq \frac{1}{\varepsilon}\{k_m(m-p+q-r)-1\}\mathrm{tr}\Omega^{-1}\mathrm{tr}Q(B'\Sigma^{-1}B)^{-1} + M_0 \\ &= \frac{p-q+2}{\varepsilon m}\cdot\frac{m^2}{(m-p+q-r-1)(m-p+q-r-2)}\mathrm{tr}\Omega^{-1}\mathrm{tr}Q(B'\Sigma^{-1}B)^{-1} \\ &\quad +(\varepsilon m^2)^{-1}O(1) + M_0. \end{aligned} \tag{2.14}$$

Noting that $\lim_{\varepsilon\to 0}\varepsilon m = 0$ and $\lim_{\varepsilon\to 0}\varepsilon m^2 > 0$ by the assumptions, we can see that the r.h.s. in the equality of (2.14) approaches infinity as ε tends to zero, which establishes Theorem 2.3.

2.2. An asymptotically consistent procedure. Instead of Problem 1, Takada (1988a) considered the relaxed problem requiring asymptotic consistency and found a procedure with both asymptotic consistency and an asymptotically bounded risk in the univariate nonparametric model.

Problem (1'). (asymptotic consistency) We want to estimator $d = d(\varepsilon)$ of ξ such that

$$\lim_{\varepsilon\to 0} R_1(\omega, d)/\varepsilon \leq 1.$$

Of course, the two-stage procedure in Section 2.1 is asymptotically consistent. To obtain a procedure with an asymptotically bounded risk, we modify the stopping number given by (2.1) as follows:

N = smallest integer n ($\geq$m) such that

(2.15) $$[tr(A_n A_n')^{-1}]^{-1} \geq (m-p+q)^{-1} trQ(B'S_m^{-1}B)^{-1}/\varepsilon.$$

Then from the proof of Theorem 2.1 we can see that $R_1(\omega, \hat{\xi}_N)/\varepsilon \leq (m-r-1)(m-p+q-r)/\{(m-p+q-r-1)(m-p+q-r-2)\}$, which implies that $\lim_{\varepsilon\to 0}\{R_1(\omega,\hat{\xi}_N)/\varepsilon\} \leq 1$ if $m\to\infty$ as $\varepsilon\to 0$.

Theorem 2.4. (asymptotic consistency) Assume that $m=O(\varepsilon^{-d})$, $0<d<1$, and that the stopping number N is given by (2.15). Then the estimator $\hat{\xi}_N$ is asymptotically consistent.

By the same arguments as used in the proofs of Theorems 2.2 and 2.3, we can show the following theorems.

Theorem 2.5. (asymptotic efficiency) Under the conditions of Theorem 2.2, $\lim_{\varepsilon\to 0} E[N]/n^* = 1$.

Theorem 2.6. Under the conditions of Theorem 2.3, $\lim_{\varepsilon\to 0} E[N-n^*] < \infty$.

3. IMPROVING ON THE TWO-STAGE ESTIMATION PROCEDURES WHEN AN ADDITIONAL SAMPLE IS AVAILABLE

In this section, we discuss two-sample problem. Assume that for the principal estimation of ξ, sample $x_1,\ldots,x_N$ is obtained based on the two-stage sampling rule in Section 2, each x_i having $N_p(B\xi a_i, \Sigma)$. We further assume that supplementary observations Y ($p\times\ell$) are taken where Y has $N_{p,\ell}(D\xi C; \Psi, I_\ell)$ with known matrices D ($p\times q$, kank q), C ($r\times\ell$, rank r), unknown positive definite matrix Ψ and the common coefficient matrix ξ. Using information of the additional sample, we want to construct an estimator superior to $\hat{\xi}_N$ defined by (2.3).

This problem of estimating the common parameters has been studied by Graybill and Deal (1959), Brown and Cohen (1974), Khatri and Shah (1974), Bhattacharya (1980), Sugiura and Kubokawa (1988) and Kubokawa (1987a,b,c;1988a,b). Recently Kubokawa (1988c) presented an improved combined procedure in the above situation with p=q=r=1. This multivariate extension is given here. Since MLE based on only Y is

$$\hat{\xi}_Y = (D'TD)^{-1}D'T^{-1}YC'(CC')^{-1} \quad \text{for } T=Y(I-C'(CC')^{-1}C)Y',$$

we consider a combined estimator of $\hat{\xi}_N$ and $\hat{\xi}_Y$ of the form

$$(3.1) \qquad \tilde{\xi}_N(a,b_N) = \hat{\xi}_N + a(1+R_N)^{-1}(\hat{\xi}_Y-\hat{\xi}_N),$$

where

$$R_N = \frac{(m-p+q-r-1)(\ell-r-1)tr(CC')^{-1}trQ(D'T^{-1}D)^{-1}}{(m-r-1)(\ell-p+q-r-1)tr(A_NA_N')^{-1}trQ(B'S_m^{-1}B)^{-1}}b_N,$$

where a and b_n are positive constants.

Theorem 3.1. Assume that

(a) $tr(A_nA_n')^{-1}b_n$ is decreasing with $\lim_{n\to\infty}tr(A_nA_n')^{-1}b_n=0$,

(b) b_n is nondecreasing,

(c) $a\leq 2(\ell-p+q-r-4)b_m/(m-p+q-r+2)$ for $\ell>p-q+r+4$.

Then $R_1(\omega,\tilde{\xi}_N(a,b_N)) \leq R_1(\omega,\hat{\xi}_N)$ for all ω.

Proof. Without the loss of generality, suppose that the stopping number N is given by (2.1). For $\Phi_N=a(1+R_N)^{-1}$, the risk difference is represented as

$$R(\omega,\tilde{\xi}_N(a,b_N))-R(\omega,\hat{\xi}_N)$$
$$= E[\Phi_N^2\{trQ(\hat{\xi}_Y-\xi)(\hat{\xi}_Y-\xi)'+trQ(\hat{\xi}_N-\xi)(\hat{\xi}_N-\xi)'\}-2\Phi_N trQ(\hat{\xi}_N-\xi)(\hat{\xi}_N-\xi)']$$

$$= E[\phi_N^2\{\frac{\ell-r-1}{\ell-p+q-r-1}tr(CC')^{-1}trQ(D'\Psi^{-1}D)^{-1}$$

(3.2) $$+ \frac{m-r-1}{m-p+q-r-1}tr(A_NA_N')^{-1}trQ(B'\Sigma^{-1}B)^{-1}\}$$

$$-2\phi_N\cdot\frac{m-r-1}{m-p+q-r-1}tr(A_NA_N')^{-1}trQ(B'\Sigma^{-1}B)^{-1}]$$

$$= a\frac{m-r-1}{m-p+q-r-1}trQ(B'\Sigma^{-1}B)^{-1}E[tr(A_NA_N')^{-1}\{\frac{(1+\theta_N)a}{(1+\theta_NQ_N)^2} - \frac{2}{1+\theta_NQ_N}\}],$$

where

$$\theta_N = \frac{(m-p+q-r-1)(\ell-r-1)tr(CC')^{-1}trQ(D'\Psi^{-1}D)^{-1}}{(m-r-1)(\ell-p+q-r-1)tr(A_NA_N')^{-1}trQ(B'\Sigma^{-1}B)^{-1}},$$

and $Q_N=R_N/\theta_N$. Here by the inequality (2.5) of Kubokawa (1988c),

$$(1+\theta_N)a(1+\theta_NQ_N)^{-2}-2(1+\theta_NQ_N)^{-1}\leq a(1+\theta_Na)^{-1}(aQ_N^{-2}-2Q_N^{-1}),$$

which yields that $R_1(\omega,\tilde{\xi}_N(a,b_N))\leq R_1(\omega,\hat{\xi}_N)$ for all ω if

(3.3) $$E[\frac{tr(A_NA_N')^{-1}}{1+\theta_Na}\{a(\frac{trQ(D'\Psi^{-1}D)^{-1}}{trQ(D'T^{-1}D)^{-1}}\cdot\frac{trQ(B'S_m^{-1}B)^{-1}}{trQ(B'\Sigma^{-1}B)^{-1}}b_N^{-1})^2 -2\frac{trQ(D'\Psi^{-1}D)^{-1}}{trQ(D'T^{-1}D)^{-1}}\cdot\frac{trQ(B'S_m^{-1}B)^{-1}}{trQ(B'\Sigma^{-1}B)^{-1}}b_N^{-1}\}] \leq 0$$

for all ω. Similar to (2.7),

$$\frac{E[trQ(D'\Psi^{-1}D)^{-1}/trQ(D'T^{-1}D)^{-1}]}{E[\{trQ(D'\Psi^{-1}D)^{-1}/trQ(D'T^{-1}D)^{-1}\}^2]} = \frac{E[(\Sigma_{i=1}^q\eta_iV_i)^{-1}]}{E[(\Sigma_{i=1}^q\eta_iV_i)^{-2}]}$$

(3.4) $$\geq \min_{1\leq i\leq q}\{\frac{E[V_i^{-1}]}{E[V_i^{-2}]}\},$$

where $V_1,\ldots,V_q$ are mutually independent random variables, each $V_i \sim \chi^2_{\ell-p+q-r}$ and $\eta_1,\ldots,\eta_q$ are parameters satisfying $\Sigma_{i=1}^q\eta_i=1$ and $\eta_i>0$, $i=1,\ldots,q$. Here the inequality in (3.4) follows from theorem 2.2 of Bhattacharya (1984). Since $E[V_i^{-1}]/E[V_i^{-2}]=\ell-p+q-r-4$, the inequality (3.3) holds if $E[\alpha_NU_m(U_m-\beta_N\sigma)]\leq 0$ for all ω, which is rewritten as

(3.5) $h(\omega) \overset{def}{=} \Sigma_{n=m}^{\infty}\alpha_n E[U_m(U_m-\beta_n\sigma)I_{[N=n]}] \leq 0$ for all ω,

where $U_m=\mathrm{tr}Q(B'S_m^{-1}B)^{-1}$, $\sigma=\mathrm{tr}Q(B'\Sigma^{-1}B)^{-1}$, $\alpha_N=\mathrm{tr}(A_NA_N')^{-1}(1+\theta_Na)^{-1}b_N^{-2}$, $\beta_N=(2/a)(\ell-p+q-r-4)b_N$, and $I_{[N=n]}=1$ if $[\mathrm{tr}(A_{n-1}A_{n-1}')^{-1}]^{-1}<k_mU_m/\varepsilon \leq[\mathrm{tr}(A_nA_n')^{-1}]^{-1}$; $=0$ otherwise.

To prove (3.5), the arguments used in Ghosh, Nickerson and Sen (1987) are available. Let n_0 denote the smallest integer ($\geq m$) such that $\varepsilon/\{k_m\mathrm{tr}(A_nA_n')^{-1}\}\geq\beta_n\sigma$. By the condition (a), the n_0 is uniquely determined. Then we write

$$h(\omega) = \Sigma_{n=m}^{n_0-1}\alpha_n E[U_m(U_m-\beta_n\sigma)I_{[N=n]}]+\alpha_{n_0}E[U_m(U_m-\beta_{n_0}\sigma)I_{[N\geq n_0]}]$$

$$(3.6)\quad +\Sigma_{n=n_0}^{\infty}\{\alpha_{n+1}E[U_m(U_m-\beta_{n+1}\sigma)I_{[N\geq n+1]}]-\alpha_n E[U_m(U_m-\beta_n\sigma)I_{[N\geq n+1]}]\},$$

where the first term in the r.h.s. of (3.6) should be interpreted as zero if $n_0=m$. From the definition of N, on the set $\{N\geq n+1\}$, $U_m>\varepsilon[k_m\mathrm{tr}(A_nA_n')^{-1}]^{-1}\geq\beta_n\sigma$ when $n\geq n_0$. Since α_n and β_n are nonincreasing and nondecreasing in n, respectively, by the condition (b), we get

<u>third term in the r.h.s. of (3.6)</u>

$$(3.7)\quad \leq \Sigma_{n=n_0}^{\infty}\alpha_{n+1}\{E[U_m(U_m-\beta_{n+1}\sigma)I_{[N\geq n+1]}]-E[U_m(U_m-\beta_n\sigma)I_{[N\geq n+1]}]\}$$

$$= \Sigma_{n=n_0}^{\infty}\alpha_{n+1}(\beta_n-\beta_{n+1})\sigma E[U_mI_{[N\geq n+1]}]$$

$$\leq 0.$$

Next, on the set $\{N=n\}$, $U_m<\beta_n\sigma$ when $n\leq n_0-1$. Thus from the monotonicity of β_n,

<u>first two terms in the r.h.s. of (3.6)</u>

$$\leq \alpha_{n_0}\Sigma_{n=m}^{n_0-1}E[U_m(U_m-\beta_n\sigma)I_{[N=n]}]+\alpha_{n_0}E[U_m(U_m-\beta_{n_0}\sigma)I_{[N\geq n_0]}]$$

$$\leq \alpha_{n_0}\Sigma_{n=m}^{n_0-1}E[U_m(U_m-\beta_m\sigma)I_{[N=n]}]+\alpha_{n_0}E[U_m(U_m-\beta_m\sigma)I_{[N\geq n_0]}]$$

(3.8) $\quad = \alpha_{n_0} E[U_m(U_m - \beta_m \sigma)]$

$$= \alpha_{n_0} \sigma E[U_m]\{E[U_m^2/\sigma^2]/E[U_m/\sigma] - \beta_m\}.$$

Finally from (2.7),

$$E[U_m^2/\sigma^2]/E[U_m/\sigma] = E[(\Sigma_{i=1}^q \lambda_i W_i)^2]/E[\Sigma_{i=1}^q \lambda_i W_i]$$

(3.9) $$\le \max_{1 \le i \le q}\{E[W_i^2]/E[W_i]\}.$$

The inequality in (3.9) can be verified by the same arguments as in the proof of theorem 2.2 of Bhattacharya (1984). Since $E[W_i^2]/E[W_i] = m-p+q-r+2$, it can be seen that the r.h.s. in the last equality of (3.8) is not positive by the condition (c). Hence, together with (3.6) and (3.7), the required inequality (3.5) is established and the proof of Theorem 3.1 is complete.

4. TWO-STAGE ESTIMATION PROCEDURES FOR POWERS OF THE GENERALIZED VARIANCE $|\Sigma|^\alpha$

4.1. A consistent two-stage procedure. Now we look for a solution of Problem 2. Define the statistic $f_n(x_1,\ldots,x_n)$ based on data $(x_1,\ldots,x_n)$ by

$$f_n(x_1,\ldots,x_n) = X_n(I - A_n'(A_nA_n')^{-1}A_n)X_n'$$

for $X_n = (x_1,\ldots,x_n)$ and $A_n = (a_1,\ldots,a_n)$, and denote

$$g_n = 1 - \{\Gamma_p(\tfrac{n-r}{2}+\alpha)\}^2/\{\Gamma_p(\tfrac{n-r}{2}+2\alpha)\Gamma_p(\tfrac{n-r}{2})\},$$

where $\Gamma_p(x) = \pi^{p(p-1)/4}\Pi_{i=1}^p \Gamma(x-(i-1)/2)$. Then we consider the following two-stage sampling rule.

(i) Start with m observations $x_1,\ldots,x_m$ for $m \ge \max(m_0, r+1, r+4\alpha)$.

(ii) Define the stopping number by

N = smallest integer $n \ge n_1 = \max(m_0, r+1, r-4\alpha)$ such that

(4.1) $$g_n^{-1} \ge h_m |S_m|^{2\alpha},$$

where $S_m = f_m(x_1, \ldots, x_m)$ and $h_m = \Gamma_p(\frac{m-r}{2}-2\alpha)/\{\varepsilon 2^{2\alpha p}\Gamma_p(\frac{m-r}{2})\}$.

(iii) Take a sample $x_{m+1}, \ldots, x_{m+N}$, and estimate $|\Sigma|^\alpha$ by

(4.2) $$\delta_N = c_N |T_N|^\alpha,$$

where $T_N = f_N(x_{m+1}, \ldots, x_{m+N})$ and

(4.3) $$c_N = \Gamma_p(\tfrac{N-r}{2}+\alpha)/\{2^{\alpha p}\Gamma_p(\tfrac{N-r}{2}+2\alpha)\}.$$

Theorem 4.1. (consistency) The estimator δ_N given by (4.2) is a solution of Problem 2, that is, $R_2(\omega, \delta_N) \leqq \varepsilon$ for all ω.

Proof. We first note that the conditional distribution of T_N given S_m has $W_p(N-r, \Sigma)$. From the definition (4.1) and the fact that $E[|T_N|^\alpha | S_m] = |\Sigma|^\alpha 2^{\alpha p}\Gamma_p((N-r)/2+\alpha)/\Gamma_p((N-r)/2)$, it follows that

$$R_2(\omega, \delta_N) = E[g_N]|\Sigma|^{2\alpha} \leqq h_m^{-1}E[|S_m|^{-2\alpha}]|\Sigma|^{2\alpha} = \varepsilon,$$

getting the desired result.

4.2. Improving on the two-stage procedure by using information in the first stage sample. By Theorem 4.1, the estimator δ_N is certainly a solution of Problem 2. However, the estimator seems to be unsatisfactory because only the data of the second stage are used for estimation of $|\Sigma|^\alpha$. So we ask the question whether it is possible to improve on the estimator by using the information gathered in the first stage. For this question, based on Kubokawa (1988d), we consider the combined estimator

(4.4) $$\delta_N^*(d_N) = \delta_N + d_N|S_m|^\alpha,$$

where d_n is a suitable constant depending on n.

Theorem 4.2. Assume that

(a) $(d_n/g_n)^2/g_n$ is increasing in n with $\lim_{n\to\infty}(d_n/g_n)^2/g_n = \infty$,

(b) d_n and d_n/g_n are nonincreasing in n,

(c) $d_{n_1}/g_{n_1} \le 2c_m$ where c_m is given by (4.3) with N=m.

Then $R_2(\omega, \delta_N^*(d_N)) \le R_2(\omega, \delta_N)$ for all ω.

Proof. The risk function of $\delta_N^*(d_N)$ is written by

$$R_2(\omega, \delta_N^*(d_N)) = R_2(\omega, \delta_N) + E[d_N^2 |S_m|^{2\alpha} - 2d_N|S_m|^{\alpha} E[(|\Sigma|^{\alpha} - \delta_N)|S_m]]$$

and we observe that $E[(|\Sigma|^{\alpha} - \delta_N)|S_m] = g_N$, so that it suffices to show that for all ω,

$$E[d_N^2 |S_m|^{\alpha}\{|S_m|^{\alpha} - (2g_N/d_N)|\Sigma|^{\alpha}\}] \le 0,$$

which can be proved by the same arguments as in the proof of Theorem 3.1 under the conditions (a),(b) and (c) of Theorem 4.2.

Letting $d_n = ag_n$ for positive constant a, we get a corollary by the following lemma.

Lemma 4.3. g_n is decreasing in n with $\lim_{n\to\infty} g_n = 0$.

Proof. From the definition of $\Gamma_p(\cdot)$,

$$\frac{\{\Gamma_p(n/2+\alpha)\}^2}{\Gamma_p(n/2+2\alpha)\Gamma_p(n/2)} = \prod_{i=1}^{p} \frac{\{\Gamma([n-i+1]/2+\alpha)\}^2}{\Gamma((n-i+1)/2+2\alpha)\Gamma((n-i+1)/2)},$$

so that it suffices to show that $\{\Gamma(t+\alpha)\}^2/\{\Gamma(t+2\alpha)\Gamma(t)\}$ is increasing in t. Since $\Gamma(t) = (te^{\gamma t})^{-1}\prod_{j=1}^{\infty}(1+t/j)^{-1}e^{t/j}$ by the Euler's infinite product [see Abramowitz and Stegun (1964,p.76)],

$$\text{(4.5)} \qquad \frac{\{\Gamma(t+\alpha)\}^2}{\Gamma(t+2\alpha)\Gamma(t)} = \prod_{j=1}^{\infty} \frac{(t+2\alpha)t}{(t+\alpha)^2} \cdot \frac{(j+t+2\alpha)(j+t)}{(j+t+\alpha)^2}.$$

Clearly, the each component in the r.h.s. of (4.5) is increasing in t. Also we can observe that

$$\lim_{t\to\infty}\{\Gamma(t+\alpha)\}^2/\{\Gamma(t+2\alpha)\Gamma(t)\} = 1,$$

establishing Lemma 4.3.

Corollary 4.4. The estimator $\delta_N^*(ag_N)=\delta_N+ag_N|S_m|^\alpha$ is better than δ_N if $a\leqq 2c_m$.

4.3. Further domination by Stein's method. On the basis of the result of Stein (1964), inadmissibility of the best affine equivariant estimator for $|\Sigma|^\alpha$ has been indicated by Shorrock and Zidek (1976), Sinha (1976), Sinha and Ghosh (1987), Sugiura (1988a, 1988b) and Sugiura and Konno (1987, 1988). The purpose of this section is to show that the estimator $\delta_N^*(d_N)$ given by (4.4) is further dominated by Stein's method when $\alpha>0$.

Denote $Z_N=(x_{m+1},\dots,x_{m+N})G_N'(G_NG_N')^{-1/2}$ and $\eta_N=B\xi(G_NG_N')^{-1/2}$ for $G_N=(a_{m+1},\dots,a_{m+N})$. The Stein type estimator we look at is of the form

$$\delta_N^{**}=\min\{\delta_N,\ \frac{\Gamma_p(N/2+\alpha)}{\Gamma_p(N/2+2\alpha)}2^{-\alpha p}|T_N+Z_NZ_N'|^\alpha\}+d_N|S_m|^\alpha,\ \alpha>0.$$

Note that Z_N and T_N are conditionally independent given S_m, and that given S_m, Z_N has $N_{p,r}(\eta_N;\Sigma,I_r)$. Then we get

Theorem 4.5. The estimator $\delta_N^*(d_N)$ is further dominated by δ_N^{**} when $\alpha>0$.

Proof. Letting

$$\delta_S=\min\{\delta_N,\ \Gamma_p(N/2+\alpha)\{\Gamma_p(N/2+2\alpha)\}^{-1}2^{-\alpha p}|T_N+Z_NZ_N'|^\alpha\},$$

we have

$$\begin{aligned}
&R_2(\omega,\delta_N^{**})-R_2(\omega,\delta_N^*(d_N))\\
&=R_2(\omega,\delta_S)-R_2(\omega,\delta_N)+2E[(\delta_S-\delta_N)d_N|S_m|^\alpha]\\
&\leqq R_2(\omega,\delta_S)-R_2(\omega,\delta_N)
\end{aligned}$$

$$= E[E[(\delta_S-|\Sigma|^{\alpha})^2|S_m]-E[(\delta_N-|\Sigma|^{\alpha})^2|S_m]].$$

Here from the result of Sugiura (1988a),

$$E[(\delta_S-|\Sigma|^{\alpha})^2|S_m]-E[(\delta_N-|\Sigma|^{\alpha})^2|S_m] \quad \text{for all } \omega,$$

which proves Theorem 4.5.

Remark. It is interesting if we could find an estimator dominating $\delta_N^*(d_N)$ for $\alpha<0$.

REFERENCES

Abramowitz, M. and Stegun, I.A. (1964). Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. National Bureau of Standards Appl. Math. Ser. 55.

Bhattacharya, C.G. (1980). Estimation of a common mean and recovery of interblock information. Ann. Statist. **8**, 205-211.

Bhattacharya, C.G. (1984). Two inequalities with an application. Ann. Inst. Statist. Math. **36**, 129-134.

Brown, L.D. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information. Ann. Statist. **2**, 963-976.

Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean. Ann. Math. Statist. **36**, 457-462.

Ghosh, M. and Mukhopadyay, N. (1976). On two fundamental problems of sequential estimation. Sankhya A**38**, 203-218.

Ghosh, M. and Mukhopadyay, N. (1981). Consistency and asymptotic efficiency of two stage and sequential estimation procedures. Sankhya A **43**, 220-227.

Ghosh, M., Nickerson, D. and Sen, P.K. (1987). Sequential shrinkage estimation. Ann. Statist. **15**, 817-829.

Gleser, L.J. and Olkin, I. (1972). Estimation for a regression model with an unknown covariance matrix. Proc. Sixth Berkeley Symp. Math. Statist. Prob. **1**, 541-568.

Graybill, F.A. and Deal, R.B. (1959). Combining unbiased estimators. Biometrics **15**, 543-550.

Khatri, C.G. and Shah, K.R. (1974). Estimation of location parameters from two linear models under normality. Comm. Statist. A-Theory Methods **3**, 647-663.

Kubokawa, T. (1987a). Estimation of a common mean with symmetrical loss. J. Japan Statist. Soc. **17**, 75-79.

Kubokawa, T. (1987b). Estimation of a common mean of two normal distributions. Tsukuba J. Math. **11**, 157-175.

Kubokawa, T. (1987c). Admissible minimax estimation of a common mean of two normal populations. Ann. Statist. **15**, 1245-1256.

Kubokawa, T. (1988a). The recovery of interblock information in balanced incomplete block designs. Sankhya B**50**, 78-89.

Kubokawa, T. (1988b). Inadmissibility of truncated estimators in balanced incomplete block designs. J. Japan Statist. Soc. **18**, 71-75.

Kubokawa, T. (1988c). Inadmissibility of the uncombined two-stage estimator when additional samples are available. Ann. Inst. Statist. Math. **40**, 555-563.

Kubokawa, T. (1988d). Improving on two-stage estimators for scale families. To appear in Metrika.

Kubokawa, T. (1988e). Two-stage procedures for parameters in a growth curve model. To appear in J. Statist. Plan. Infer..

Muirhead, R.J. (1982). Aspects of Multivariate Statistical Theory. Wiley, New York.

Mukhopadyay, N. (1980). A consistent and asymptotically efficient two-stage procedure to construct fixed width confidence intervals for the mean. Metrika **27**, 281-284.

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. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. Proc. Fifth Berkely Symp. Math. Statist. Prob. **1**, 355-372.

Rao, C.R. (1973). Linear Statistical Inference and Its Application. 2nd ed., Wiley, New York.

Shorrock, R.B. and Zidek, J.V. (1976). An improved estimator of the generalized variance. Ann. Statist. **4**, 629-638.

Simons, G. (1968). On the cost of not knowing the variance when making a fixed-width confidence interval for the mean. Ann. Math. Statist. **39**, 1946-1952.

Sinha, B.K. (1976). On improved estimators of the generalized variance. J. Mult. Analysis **6**, 617-625.

Sinha, B.K. and Ghosh, M. (1987). Inadmissibility of the best equivariant estimator of the variance-covariance matrix and the generalized variance. Statistics & Decisions **5**, 201-227.

Starr, N. and Woodroofe, M. (1968). Remarks on a stopping time. Proc. Nat. Acad. Sci. USA **61**, 1215-1218.

Stein, C. (1945). A two sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Statist. **16**, 243-258.

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Inst. Statist. Math. **16**, 155-160.

Sugiura, N. (1988a). A class of improved estimators of powers of the generalized variance and precision under squared loss. Statistical Theory and Data Analysis II (Ed. K. Matsusita), North Holland, Amsterdam, 421-428.

Sugiura, N. (1988b). Entropy loss and a class of improved estimators for powers of the generalized variance. To appear in Sankhya.

Sugiura, N. and Konno, Y. (1987). Risk of improved estimators for generalized variance and precision. Advances in Multivariate Statistical Analysis. D. Reidel Pub. Co. Holland, 353-371.

Sugiura, N. and Konno, Y. (1988). Entropy loss and risk of improved estimators for the generalized variance and precision. Ann. Inst. Statist. Math. **40**, 329-341.

Sugiura, N. and Kubokawa, T. (1988). Estimating common parameters of growth curve models. Ann. Inst. Statist. Math. **40**, 119-135.

Takada, Y. (1986). Non-existence of fixed sample size procedures for scale families. Sequential Analysis **5**, 93-101.

Takada, Y. (1988a). Asymptotically bounded regret sequential estimation of the mean. Sequential Analysis **7**, 253-262.

Takada, Y. (1988b). Two-stage procedures for a multivariate normal distribution. To appear in Kumamoto J. Math.

Williams, J.S. (1967). The variance of weighted regression estimators. J. Amer. Statist. Asoc. **62**, 1290-1301.