TSUKUBA J. MATH.
Vol. 14 No. 1 (1990), 79–89

SEQUENTIAL POINT ESTIMATION WITH BOUNDED RISK IN A MULTIVARIATE REGRESSION MODEL

By

Tatsuya KUBOKAWA

For the coefficient matrix of the multivariate regression model, consider the problem of finding an estimator with asymptotically bounded risk. The paper proposes a sequential procedure resolving the problem and investigates the asymptotic properties. Also it is shown that if additional observations with the same coefficient matrix are available, then the sequential estimator is improved on by a combined procedure.

1. Introduction

Let $x_1, x_2, \cdots$ be a sequence of mutually independent random vectors, x_i having p-variate normal distribution $N_p(\xi a_i, \Sigma)$ where a_i $(r\times 1)$ is a known vector and ξ $(p\times r)$, Σ $(p\times p)$ are unknown matrices. Denote $X_n=(x_1, x_2, \cdots, x_n)$, $A_n=(a_1, a_2, \cdots, a_n)$ and $\omega=(\xi, \Sigma)$. Then X_n $(p\times n)$ has $N_{p,n}(\xi A_n; \Sigma, I_n)$, being a multivariate regression model.

For a preassigned constant $\varepsilon>0$, we consider the problem of finding an estimator $\hat{\xi}_\varepsilon$ of the coefficient matrix ξ such that

$$(1.1)\qquad R(\omega, \hat{\xi}_\varepsilon)=E_\omega[n^{-1}\operatorname{tr} Q(\hat{\xi}_\varepsilon-\xi)A_nA_n'(\hat{\xi}_\varepsilon-\xi)']\leqq\varepsilon$$

for all ω, where Q $(p\times p)$ is a positive definite matrix.

Throughout the paper, let m_0 be the smallest integer $(\geqq r)$ such that $\operatorname{rank}(A_{m_0})=r$. In the case where Σ is known, for integer n $(\geqq m_0)$, MLE of ξ is given by

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

and from Muirhead (1982),

$$(1.2)\qquad R(\omega, \hat{\xi}_0(n))=E_\omega[n^{-1}\{\operatorname{vec}(\hat{\xi}_0(n)-\xi)\}'(A_nA_n'\otimes Q)\operatorname{vec}(\hat{\xi}_0(n)-\xi)]$$
$$=n^{-1}\operatorname{tr}(A_nA_n'\otimes Q)\operatorname{Cov}(\operatorname{vec}\hat{\xi}_0(n))$$

Received May 2, 1989.

$$=n^{-1}\,\mathrm{tr}(A_nA_n'\otimes Q)\{(A_nA_n')^{-1}\otimes\Sigma\}$$
$$=n^{-1}\,\mathrm{tr}\,Q\Sigma\,,$$

where the notation vec ξ denotes $pr\times 1$ vector $(\xi_1',\cdots,\xi_r')'$ for $\xi=(\xi_1,\cdots,\xi_r)$ and $A\otimes B$ stands for kronecker product defined by $(a_{ij}B)$ for $A=(a_{ij})$. Hence we get that $R(\omega,\hat{\xi}_0(n))\leqq\varepsilon$ if and only if $n\geqq r\,\mathrm{tr}\,Q\Sigma/\varepsilon$ $(=n^*)$. Since Σ is unknown, there is no fixed sample size rule to achieve the goal.

For the univariate case, Rao (1973, pp. 486-487) provided a two-stage rule solving the problem (1.1) and multivariate extensions were given by Takada (1988) and Kubokawa (1989, 90). When $r=1$ and $a_i=(1,\cdots,1)$ for each i, Mukhopadhyay (1985) and Takada (1989) obtained three-stage and purely sequential procedures satisfying

(1.3) $$\lim_{\varepsilon\to 0}R(\omega,\hat{\xi}_\varepsilon)/\varepsilon=1\,. \quad \textit{(asymptotic consistency)}$$

In the above multivariate regression model, we consider the purely sequential rule of the form

(1.4) $$N=\mathrm{Min}\left\{n\geqq m\,;\;n\geqq\frac{r}{\varepsilon(n-r)}\mathrm{tr}\,QS_n\right\},$$

where $S_n=X_n(I_n-A_n'(A_nA_n')^{-1}A_n)X_n'$ and m $(\geqq\max\{m_0,r+1\})$ is the first sample size. When ξ is estimated by

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

Section 2 demonstrates asymptotic consistency of $\hat{\xi}_N$ and *asymptotic efficiency* of N, that is,

(1.5) $$\lim_{\varepsilon\to 0}E[N]/n^*=1\,.$$

The asymptotic expansions of $E[N]$ and $R(\omega,\hat{\xi}_N)$ are also developed based on Woodroofe (1977). These are extensions of the results given by Takada (1989).

In Section 3, we assume that additional observations Y $(p\times l)$ are taken where Y has $N_{p,l}(\xi C\,;\Psi,I_l)$ with known design matrix C $(r\times l)$, unknown positive definite matrix Ψ and the common coefficient matrix ξ. Using information of additional sample, we construct a combined estimator and prove, by the method of Ghosh, Nickerson and Sen (1987), that it exactly dominates $\hat{\xi}_N$. A second order asymptotic comparison of their risks is presented in Section 4.

2. Asymptotic properties

THEOREM 2.1. *The sequential procedure* $\hat{\xi}_N$ *is asymptotically consistent for* $p(m-r)\geqq 3$. *The stopping number* N *given by* (1.4) *is asymptotically efficient.*

To prove the theorem, we need the following lemmas.

LEMMA 2.1. *For integer* n $(\geqq m \geqq m_0)$, *the* $p\times p$ *matrix* $S_n = X_n(I_n - A_n'(A_nA_n')^{-1}A_n)X_n'$ *is written as*

$$S_n = \sum_{i=m}^{n} T_i\,, \tag{2.1}$$

where $T_m, \cdots, T_n$ *satisfy the following conditions:*

(a) *Each* T_i *is a statistic based on only* $x_1, \cdots, x_i$, *that is, independent of* $x_{i+1}, \cdots, x_n$.

(b) $T_m, \cdots, T_n$ *are independently distributed as* $T_m \sim W_p(\Sigma, m-r)$ *and* $T_i \sim W_p(\Sigma, 1)$ *for* $i = m+1, \cdots, n$.

(c) $(T_m, \cdots, T_n)$ *is independent of* X_nA_n'.

PROOF. Let $A_n = (A_{n-1}, a)$, $A_{n-1} = A$, $\alpha_n = a'(AA')^{-1}a$ and

$$D_n = \frac{1}{1+\alpha_n}\begin{pmatrix} A'(AA')^{-1}aa'(AA')^{-1}A & -A'(AA')^{-1}a \\ -a'(AA')^{-1}A & 1 \end{pmatrix}.$$

Then we can express S_n as

$$S_n = S_{n-1} + X_nD_nX_n'\,.$$

Further letting $A_{n-1} = (A_{n-2}, b)$, $A_{n-2} = B$, $\alpha_{n-1} = b'(BB')^{-1}b$ and

$$D_{n-1} = \frac{1}{1+\alpha_{n-1}}\begin{pmatrix} B'(BB')^{-1}bb'(BB')^{-1}B & -B'(BB')^{-1}b & 0 \\ -b'(BB')^{-1}B & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

we have $S_{n-1} = S_{n-2} + X_nD_{n-1}X_n'$. By the same consideration, consequently, we get

$$S_n = \sum_{i=m}^{n} X_nD_iX_n'\,.$$

It can be shown that $D_i^2 = D_i$, $D_iD_j = 0$ $(i \neq j)$ and $A_nD_i = 0$ for $i = m, \cdots, n$ that $\mathrm{rank}(D_m) = m-r$, $\mathrm{rank}(D_i) = 1$ for $i = m+1, \cdots, n$. Letting $T_i = X_nD_iX_n'$ establishes Lemma 2.1.

LEMMA 2.2. *Assume that* $p(m-r)/2 > \lambda > 0$ *or* $\lambda < 0$. *Then* $(n^*/N)^\lambda$ *is uniformly integrable for* $0 < \varepsilon < \varepsilon_0$ *(specified).*

PROOF. Consider the case of $\lambda > 0$. We first have that for $d, \delta > 0$,

$$E[(n^*/N)^\lambda I_{[(n^*/N)^\lambda > d]}] \leqq d^{-\delta}E[(n^*/N)^{\lambda(1+\delta)}]\,, \tag{2.2}$$

where $I_{[\cdot]}$ designates the indicator function, so that it is sufficient to show that $\sup_{0<\varepsilon<\varepsilon_0}\{E[(n^*/N)^{\lambda(1+\delta)}]\}<\infty$. Lemma 2.3 of Woodroofe (1977) gives that for $0<\theta<1$,

$$P[N\leqq\theta n^*]=O(\varepsilon^{p(m-r)/2}) \quad \text{as } \varepsilon\to 0 .$$

[Woodroofe's notations c, t_c, m, α, β, μ, τ^2, L_0, λ correspond to our ε/r, $N-r$, $m-r$, 2, 1, $\mathrm{tr}\, Q\Sigma$, $2\,\mathrm{tr}(Q\Sigma)^2$, r, n^*, respectively.] Hence for $0<\varepsilon<\varepsilon_0$,

$$
\begin{aligned}
(2.3)\qquad E[(n^*/N)^{\lambda(1+\delta)}] &\leqq E[(n^*/N)^{\lambda(1+\delta)} I_{[N\leqq\theta n^*]}]+\theta^{-\lambda(1+\delta)}\\
&\leqq (n^*/m)^{\lambda(1+\delta)} P[N\leqq\theta n^*]+\theta^{-\lambda(1+\delta)}\\
&\leqq K\varepsilon^{p(m-r)/2-\lambda(1+\delta)}+\theta^{-\lambda(1+\delta)}\\
&\leqq K\varepsilon_0^{p(m+r)/2-\lambda(1+\delta)}+\theta^{-\lambda(1+\delta)} ,
\end{aligned}
$$

where K is a constant independent of ε. The fourth inequality in (2.3) follows from the fact that there exists a positive δ satisfying $p(m-r)/2-\lambda(1-\delta)\geqq 0$, which always holds if $p(m-r)/2>\lambda$.

When $\lambda<0$, from Lemma 2.1, note that $\mathrm{tr}\, QS_n=\sum_{i=m}^{n}\mathrm{tr}\, QT_i$. Here $\mathrm{tr}\, QT_i =\mathrm{tr}\, W_i\Sigma^{1/2}Q\Sigma^{1/2}$ for $W_i=\Sigma^{-1/2}T_i\Sigma^{-1/2}$, $W_m\sim W_p(I, m-r)$ and for $i=m+1,\cdots,n$, $W_i\sim W_p(I,1)$. Denote $\mathrm{diag}(\sigma_1,\cdots,\sigma_p)=H'\Sigma^{1/2}Q\Sigma^{1/2}H$ for some orthogonal matrix H. From the Bartlett's decomposition, we have

$$(2.4)\qquad \mathrm{tr}\, QT_i=\sum_{j=1}^{p}\sigma_j W_{ij} ,$$

where $W_{i1},\cdots,W_{ip}$ are mutually independent random variables, $W_{mj}\sim\chi^2_{m-r}$ and for $i=m+1,\cdots,n$, $W_{ij}\sim\chi^2_1$. Then,

$$(2.5)\qquad \mathrm{tr}\, QS_n=\sum_{i=m}^{n}\sum_{j=1}^{p}\sigma_j W_{ij}=\sum_{j=1}^{p}\sigma_j Q_j^{(n-r)} ,$$

where $Q_j^{(n-r)}=\sum_{i=m}^{n}W_{ij}$ having χ^2_{n-r}. Also from the definition of N,

$$(2.6)\qquad N<\frac{r}{\varepsilon(N-r-1)}\mathrm{tr}\, QS_{N-1}I_{[N\geqq m+1]}+1+m .$$

Let $\tau=-\lambda(1+\delta)$. In the *rhs* of the inequality (2.2), from (2.5) and (2.6), we can see that for $0<\varepsilon<\varepsilon_0$,

$$
\begin{aligned}
E[(N/n^*)^\tau] &< E[\{(\mathrm{tr}\, Q\Sigma)^{-1}\sum_{j=1}^{p}\sigma_j(Q_j^{(N-r-1)}/(N-r-1))I_{[N\geqq m+1]}+(1+m)/n^*\}^\tau]\\
&<\sum_{j=1}^{p}C_jE[\{Q_j^{(N-r-1)}/(N-r-1)\}^\tau I_{[N\geqq m+1]}]+C_0\varepsilon_0^\tau\\
&<\sum_{j=1}^{p}C_jE[\sup_{n\geqq m}\{Q_j^{(n-r)}/(n-r)\}^\tau]+C_0\varepsilon_0^\tau ,
\end{aligned}
$$

where $C_0, C_1, \cdots, C_p$ are constants independent of ε. The Doob's maximal inequality for reversed martingale sequence $\{Q_j^{(n-r)}/(n-r)\}_{n\geq m}$ gives that $E[\sup_{n\geq m}\{Q_j^{(n-r)}/(n-r)\}^{\tau}]<\infty$. Therefore the uniform integrability of $(n^*/N)^{\lambda}$ is completely proved.

PROOF OF THEOREM 2.1 Denote by $\mathcal{F}_n$ the σ-algebra generated by $T_m, \cdots, T_n$ given in Lemma 2.1. Similar to (1.2), the risk function of $\hat{\xi}_N$ is represented as

$$R(\omega, \hat{\xi}_N)=\sum_{n=m}^{\infty} n^{-1}\operatorname{tr}(A_nA_n'\otimes Q)E[(\operatorname{vec}\hat{\xi}_n-\operatorname{vec}\xi)(\operatorname{vec}\hat{\xi}_n-\operatorname{vec}\xi)'I_{[N=n]}] .$$

Since $\hat{\xi}_n$ and $(T_m, \cdots, T_n)$ are independent by Lemma 2.1, we have

$$\begin{aligned} &E[(\operatorname{vec}\hat{\xi}_n-\operatorname{vec}\xi)(\operatorname{vec}\hat{\xi}_n-\operatorname{vec}\xi)'I_{[N=n]}] \\ &\quad =E[I_{[N=n]}E[(\operatorname{vec}\hat{\xi}_n-\operatorname{vec}\xi)(\operatorname{vec}\hat{\xi}_n-\operatorname{vec}\xi)'\mid\mathcal{F}_n]] \\ &\quad =E[I_{[N=n]}\operatorname{Cov}(\operatorname{vec}\hat{\xi}_n)] \\ &\quad =\{(A_nA_n')^{-1}\otimes\Sigma\}E[I_{[N=n]}] , \end{aligned}$$

which yields that

$$R(\omega, \hat{\xi}_N)=\sum_{n=m}^{\infty} rn^{-1}\operatorname{tr} Q\Sigma E[I_{[N=n]}]=\varepsilon E[n^*/N] . \tag{2.7}$$

Since $n^*/N\to 1$ a.s. as $\varepsilon\to 0$, applying Lemma 2.2 with $\lambda=1$ proves that $R(\omega, \hat{\xi}_N)/\varepsilon\to 1$ as $\varepsilon\to 0$. The asymptotic efficiency of N is trivial from Lemma 2.2, and the proof is complete.

Theorem 2.1 shows the first order asymptotic efficiency and consistency. More detailed, the second order asymptotic expansions for $E[N]$ and $R(\omega, \hat{\xi}_N)$ are presented based on Woodroofe (1977).

THEOREM 2.2. *For* $p(m-r)\geqq 5$,

$$E[N]=n^*+\frac{\nu}{\operatorname{tr} Q\Sigma}-2\frac{\operatorname{tr}(Q\Sigma)^2}{(\operatorname{tr} Q\Sigma)^2}+o(1) , \tag{2.8}$$

$$R(\omega, \hat{\xi}_N)=\varepsilon+\frac{\varepsilon^2}{r\operatorname{tr} Q\Sigma}\left\{4\frac{\operatorname{tr}(Q\Sigma)^2}{(\operatorname{tr} Q\Sigma)^2}-\frac{\nu}{\operatorname{tr} Q\Sigma}\right\}+o(\varepsilon^2) , \tag{2.9}$$

where ν *is defined by* (2.4) *in Woodroofe* (1977).

The expansion (2.8) is from Woodroofe (1977). Note that $n^*/N=(N-n^*)^2/(n^*N)-N/n^*+2$ and that $(N-n^*)N^{-1/2}\to N(0, 2\operatorname{tr}(Q\Sigma)^2/(\operatorname{tr} Q\Sigma)^2)$ as $\varepsilon\to 0$. Hence (2.9) can be derived by combining (2.7), (2.8) and the following lemma.

LEMMA 2.3. *Assume that* $0<\lambda<p(m-r)/2$. *Then* $(|N-n^*|/N^{1/2})^{\lambda}$ *is uniformly integrable.*

PROOF. First, observe that for $d, \delta>0$,

$$\begin{aligned}E[(|N-n^*|/N^{1/2})^{\lambda}I_{[(|N-n^*|/N^{1/2})^{\lambda}>d]}]\\ &\leqq d^{-\delta}E[(|N-n^*|/N^{1/2})^{\lambda(1+\delta)}]\\ &\leqq d^{-\delta}\{E[(n^*/N)^{\lambda(1+\delta)}]E[\{|N-n^*|/(n^*)^{1/2}\}^{2\lambda(1+\delta)}]\}^{1/2}.\end{aligned}$$

From Lemma 2.2, $(n^*/N)^{\lambda(1+\delta)}$ is uniformly integrable for some $\delta>0$ under the condition $0<\lambda<p(m-r)/2$. Also, Theorem 2.3 of Woodroofe (1977) demonstrates that $\{(N-n^*)^2/n^*\}^{\lambda(1+\delta)}$ is uniformly integrable under the same condition. Hence there exists some constant M independent of ε such that

$$E[(n^*/N)^{\lambda(1+\delta)}]E[\{|N-n^*|/(n^*)^{1/2}\}^{2\lambda(1+\delta)}]<M,$$

for $0<\varepsilon<\varepsilon_0$, which establishes Lemma 2.3.

3. Improving on the sequential procedure when an additional sample is available

In this section, we discuss two-sample problem. Assume that for the principal estimation of ξ, sample $x_1, \cdots, x_N$ is obtained based on the sequential sampling rule in Section 2, each x_i having $N_p(\xi a_i, \Sigma)$. We further assume that supplementary observations $Y(p\times l)$ are taken where Y has $N_{p,l}(\xi C; \Psi, I_l)$ with known matrix C $(r\times l)$, 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 $\hat{\xi}_N$.

The problem of estimating the common parameters in the fixed sample size case has been studied by several authors. [For the brief bibliography, see Kubokawa (1988).] Since MLE based on only Y is $\hat{\xi}_Y=YC'(CC')^{-1}$, we consider a combined estimator of $\hat{\xi}_N$ and $\hat{\xi}_Y$ of the form

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

where

$$R_N=b\operatorname{tr}A_NA_N'(CC')^{-1}\operatorname{tr}QT/(rv_N),$$

$$v_N=\operatorname{tr}QS_N/\{3(N-1)\},$$

$$T=Y(I_l-C'(CC')^{-1}C)Y',$$

and a, b are positive constants.

THEOREM 3.1. *Assume that* $a \leqq \min\{1, 2(l-r-4)b\}$. *Then* $R(\omega, \tilde{\xi}_N(a, b)) \leqq R(\omega, \hat{\xi}_N)$ *for all* ω.

PROOF. By using Lemma 2.1, the risk difference is written as

$$\begin{aligned}(3.2)\qquad \Delta &= R(\omega, \hat{\xi}_N) - R(\omega, \tilde{\xi}_N(a, b)) \\ &= E\Big[\frac{2ar}{(1+R_N)N}\operatorname{tr} Q\Sigma - \frac{a^2}{(1+R_N)^2}\Big\{\frac{r}{N}\operatorname{tr} Q\Sigma + \frac{1}{N}\operatorname{tr}(A_N A_N')(CC')^{-1}\operatorname{tr} Q\Psi\Big\}\Big] \\ &= ar \operatorname{tr} Q\Sigma E[N^{-1}\{2(1+\theta_N U_N)^{-1} - a(1+\theta_N)(1+\theta_N U_N)^{-2}\}],\end{aligned}$$

where

$$(3.3)\qquad \theta_N = \operatorname{tr} A_N A_N'(CC')^{-1} \operatorname{tr} Q\Psi/(r \operatorname{tr} Q\Sigma), \quad U_N = b(\operatorname{tr} Q\Sigma)(\operatorname{tr} QT)/(v_N \operatorname{tr} Q\Psi).$$

Here by the inequality (2.5) of Kubokawa (1988),

$$(3.4)\qquad \frac{2}{1+\theta_N U_N} - \frac{(1+\theta_N)a}{(1+\theta_N U_N)^2} \geqq \frac{a}{1+\theta_N a}(2U_N^{-1} - aU_N^{-2}),$$

which yields that $\Delta \geqq 0$ for all ω if

$$(3.5)\qquad E\Big[g(N)N^{-1}v_N\Big\{2b\frac{\operatorname{tr} Q\Psi}{\operatorname{tr} QT} - a\Big(\frac{\operatorname{tr} Q\Psi}{\operatorname{tr} QT}\Big)^2 v_N/\sigma\Big\}\Big] \geqq 0 \qquad \text{for all } \omega,$$

where $g(n) = (1+\theta_n a)^{-1}$ and $\sigma = \operatorname{tr} Q\Sigma$. Similar to (2.5),

$$(3.6)\qquad \frac{E[\operatorname{tr} Q\Psi/\operatorname{tr} QT]}{E[(\operatorname{tr} Q\Psi/\operatorname{tr} QT)^2]} = \frac{E[(\sum_{i=1}^p \eta_i w_i)^{-1}]}{E[(\sum_{i=1}^p \eta_i w_i)^{-2}]} \geqq \min_{1\leq i\leq p}\Big\{\frac{E[w_i^{-1}]}{E[w_i^{-2}]}\Big\},$$

where $w_1, \cdots, w_p$ are mutually independent random variables, each w_i having χ^2_{l-r} and $\eta_1, \cdots, \eta_p$ are parameters satisfying $\sum_{i=1}^d \eta_i = 1$ and $\eta_i > 0$, $i = 1, \cdots, p$. Here the inequality in (3.6) follows from theorem 2.2 of Bhattacharya (1984). From the condition $a \leqq 2(l-r-4)b$ and the fact that $E[w_i^{-1}]/E[w_i^{-2}] = l-r-4$, the inequality (3.5) holds if $E[g(N)N^{-1}v_N(1-v_N/\sigma)] \geqq 0$ for all ω, which is rewritten as

$$(3.7)\qquad \sum_{n=m}^{\infty} g(n)E[n^{-1}v_n(1-v_n/\sigma)I_{[N=n]}] \geqq 0 \qquad \text{for all } \omega.$$

To prove (3.7), the arguments used in Ghosh, Nickerson and Sen (1987) are available. Let n_0 denote the smallest integer ($\geqq m$) such that $\varepsilon n(n-r)/\{3r(n-1)\} \geqq \sigma$. It should be noted that n_0 is uniquely determined. Then we write

the lhs of (3.7)

$$(3.8)\qquad = \sum_{n=m}^{n_0-1} g(n)E\Big[\frac{1}{n}v_n(1-v_n/\sigma)I_{[N=n]}\Big] + g(n_0)E\Big[\frac{1}{n_0}v_{n_0}(1-v_{n_0}/\sigma)I_{[N\geqq n_0]}\Big]$$

$$+\sum_{n=n_0}^{\infty}\Big\{g(n+1)E\Big[\frac{1}{n+1}v_{n+1}(1-v_{n+1}/\sigma)I_{[N\geq n+1]}\Big]$$
$$-g(n)E\Big[\frac{1}{n}v_n(1-v_n/\sigma)I_{[N\geq n+1]}\Big]\Big\},$$

where the first term in the *rhs* of (3.8) should be interpreted as zero if $n_0=m$. Note that for $n\geqq n_0$, on the set $\{N\geqq n+1\}$, $v_n>\varepsilon n(n-r)/\{3r(n-1)\}\geqq\sigma$. Since $g(n)$ is nonincreasing,

third term in the rhs of (3.8)

$$\geqq\sum_{n=n_0}^{\infty}g(n+1)E\Big[\Big\{\frac{1}{n+1}v_{n+1}(1-v_{n+1}/\sigma)-\frac{1}{n}v_n(1-v_n/\sigma)\Big\}I_{[N\geq n+1]}\Big].$$

Note that $I_{[N\geq n+1]}$ is a $\mathcal{F}_n$-measurable function. Also from Lemma 2.1, $v_{n+1}=(n-1)n^{-1}v_n+(3n)^{-1}u$ for $u=\mathrm{tr}\,QT_{n+1}$. Then,

$$\text{(3.9)}\quad E\Big[\Big\{\frac{1}{n+1}v_{n+1}(1-v_{n+1}/\sigma)-\frac{1}{n}v_n(1-v_n/\sigma)\Big\}I_{[N\geq n+1]}\,|\,\mathcal{F}_n\Big]$$
$$=I_{[N\geq n+1]}\Big[\frac{1}{n+1}\Big\{\frac{n-1}{n}v_n+\frac{1}{3n}E[u]\Big\}$$
$$-\frac{1}{n+1}\Big\{\Big(\frac{n-1}{n}\Big)^2v_n^2+\frac{2(n-1)}{3n^2}v_nE[u]+\frac{1}{9n^2}E[u^2]\Big\}\frac{1}{\sigma}-\frac{1}{n}v_n+\frac{1}{n\sigma}v_n^2\Big]$$
$$\geqq I_{[N\geq n+1]}\Big[\frac{3n-1}{n^2(n+1)}v_n^2/\sigma-\frac{8n-2}{3n^2(n+1)}v_n+\frac{n-1}{3n^2(n+1)}\sigma\Big]$$

since $E[u]=\sigma$ and $E[u^2]\leqq 3\sigma^2$. Note that the multiple of $I_{[N\geq n+1]}$ in the extreme *rhs* of (3.9) is a convex function of v_n, where the minimum occurs at $v_n=(4n-1)\sigma/\{3(3n-1)\}$ $(<\sigma)$. Here, recalling that on the set $\{N\geqq n+1\}$, $v_n>\sigma$, it follows that

extreme rhs of (3.9)

$$\geqq I_{[N\geq n+1]}\Big[\frac{3n-1}{n^2(n+1)}-\frac{8n-2}{3n^2(n+1)}+\frac{n-1}{3n^2(n+1)}\Big]\sigma$$
$$\geqq 0\,.$$

Next, note that $I_{[N\geq m]}=1$ with probability 1 and that

$$\text{(3.10)}\quad E[v_m(1-v_m/\sigma)]=E[v_m]-E[v_m^2]/\sigma$$
$$\geqq\frac{m-r}{3(m-1)}\sigma-\frac{(m-r)(m-r+2)}{\{3(m-1)\}^2}\sigma$$
$$=(m-r)(2m+r-5)\sigma/\{9(m-1)^2\}$$
$$\geqq 0\,.$$

Thus, if $n_0=m$ then the first two terms in the *rhs* of (3.8)$\geqq 0$. For $n_0>m$, first note that for $n\leqq n_0-1$, on the set $\{N=n\}$, $v_n\leqq \varepsilon n(n-r)/\{3r(n-1)\}<\sigma$. Then

the first two terms in the rhs of (3.8)

$$\geqq g(n_0)\left\{\sum_{n=m}^{n_0-1} E[n^{-1}v_n(1-v_n/\sigma)I_{[N=n]}]+E[n_0^{-1}v_{n_0}(1-v_{n_0}/\sigma)I_{[N\geq n_0]}]\right\}. \tag{3.11}$$

Since $v_{n_0}=(n_0-2)(n_0-1)^{-1}v_{n_0-1}+\{3(n_0-1)\}^{-1}\operatorname{tr} QT_{n_0}$, it can be seen that

$$\begin{aligned} E[n_0^{-1}v_{n_0}(1-v_{n_0}/\sigma)I_{[N\geq n_0]}\,|\,\mathcal{F}_{n_0-1}] &\geqq \{a_{n_0}v_{n_0-1}-b_{n_0}v_{n_0-1}^2/\sigma+\sigma c_{n_0}\}I_{[N\geq n_0]} \\ &\geqq b_{n_0}v_{n_0-1}(1-v_{n_0-1}/\sigma)I_{[N\geq n_0]}, \end{aligned} \tag{3.12}$$

where $a_{n_0}=(3n_0-5)(n_0-2)/\{3(n_0-1)^2n_0\}$, $b_{n_0}=(n_0-2)^2/\{(n_0-1)^2n_0\}$ and $c_{n_0}=$ $=(n_0-2)/\{3n_0(n_0-1)^2\}$. If $E[v_{n_0-1}(1-v_{n_0-1}/\sigma)I_{[N\geq n_0]}]\geqq 0$, noting again that $v_n<\sigma$ on the set $\{N=n\}$ for all $n\leqq n_0-1$, we prove that the *rhs* of (3.11)$\geqq 0$. Otherwise using the fact that $b_{n_0}<1/(n_0-1)$, we get from (3.12) that

$$\textit{the rhs of } (3.11)\quad \geqq g(n_0)\left\{\sum_{n=m}^{n_0-2} E[n^{-1}v_n(1-v_n/\sigma)I_{[N=n]}]\right.$$
$$\left. +E[(n_0-1)^{-1}v_{n_0-1}(1-v_{n_0-1}/\sigma)I_{[N\geq n_0-1]}]\right\}.$$

Proceed inductively to get

the rhs of (3.11) $\geqq g(n_0)E[m^{-1}v_m(1-v_m/\sigma)]\geqq 0$

as shown earlier, and the proof of Theorem 3.1 is complete.

4. Asymptotic risk expansion

Now we reveal the asymptotic risk expansion of $\tilde{\xi}_N(a, b)$ and asymptotically compare the risks of $\hat{\xi}_N$ and $\tilde{\xi}_N(a, b)$.

From (3.2), the risk difference is written as

$$\Delta=-\varepsilon^2 a(r\operatorname{tr} Q\Sigma)^{-1}E[(n^*/N)^2P_N], \tag{4.1}$$

where $P_N=N\{a(1+\theta_N)(1+\theta_N U_N)^{-2}-2(1+\theta_N U_N)^{-1}\}$. Then the following lemma is essential for our purpose.

LEMMA 4.1. *Assume that* $n^{-1}A_nA_n'\to\Omega>0$ *as* $n\to\infty$. *If* $p(m-r)>8$ *and* $l-r>8$, *then* $(n^*/N)^2N(1+\theta_N)(1+\theta_N U_N)^{-2}$ *and* $(n^*/N)^2N(1+\theta_N U_N)^{-1}$ *are uniformly integrable for* $0<\varepsilon<\varepsilon_0$.

PROOF. Put $Z_N=N(1+\theta_N)(1+\theta_N U_N)^{-2}$. Observe that for $d, \delta>0$,

$$(4.2)\qquad E[(n^*/N)^2 Z_N I_{[(n^*/N)^2 Z_N > d]}] \leqq d^{-\delta} E[(n^*/N)^{2(1+\delta)} Z_N^{1+\delta}]$$
$$\leqq d^{-\delta}\{E[(n^*/N)^{4(1+\delta)}]E[Z_N^{2(1+\delta)}]\}^{1/2}.$$

Since $(n^*/N)^{4(1+\delta)}$ is uniformly integrable under the condition $p(m-r)>8$ by Lemma 2.2, there exists some constant M_1 independent of ε such that $E[(n^*/N)^{4(1+\delta)}]<M_1$ for $0<\varepsilon<\varepsilon_0$. Also,

$$(4.3)\qquad Z_N = \frac{N}{1+\theta_N}\left(\frac{1+\theta_N}{1+\theta_N U_N}\right)^2 \leqq \frac{N}{1+\theta_N}\left(1+\frac{1}{\theta_N}\right)^2 U_N^{-2}.$$

Noting that $\theta_n\to\infty$ and $\theta_n/n\to \mathrm{tr}\, \Omega(CC')^{-1}\, \mathrm{tr}\,(Q\Psi)/(r\, \mathrm{tr}\, Q\Sigma)$ as $n\to\infty$, we can take a constant M_2 such that

$$(4.4)\qquad n(1+\theta_n)^{-1}(1+\theta_n^{-1})^2 \leqq M_2 \qquad \text{for all } n\geqq m.$$

Since $v_N=(N-r)\{3(N-1)\}^{-1}\sum_{j=1}^p \sigma_j\{Q_j^{(N-r)}/(N-r)\}$ by (2.5), we obtain from (4.3) and (4.4) that

$$(4.5)\qquad E[Z_N^{2(1+\delta)}] \leqq \sum_{j=1}^p C_j E[(\mathrm{tr}\, QT)^{-4(1+\delta)}]E[\sup_{n\geqq m}\{Q_j^{(n-r)}/(n-r)\}^{4(1+\delta)}]$$

for constants C_j independent of ε. From the proof of Lemma 2.2, it is seen that the *rhs* of (4.5) is finite for $0<\varepsilon<\varepsilon_0$ under the condition $l-r>8$. Hence the uniform integrability of $(n^*/N)^2 Z_N$ holds. Similarly we can show the uniform integrability of $(n^*/N)^2 N(1+\theta_N U_N)^{-1}$.

Note that $(n^*/N)^2\to 1$ a.s. and

$$P_N \longrightarrow \frac{r\, \mathrm{tr}\, Q\Sigma}{(3b)^2\, \mathrm{tr}\, \Omega(CC')^{-1}\, \mathrm{tr}\, Q\Psi}\left\{a\left(\frac{\mathrm{tr}\, Q\Psi}{\mathrm{tr}\, QT}\right)^2 - 6b\frac{\mathrm{tr}\, Q\Psi}{\mathrm{tr}\, QT}\right\} \qquad \text{a.s.}$$

as $\varepsilon\to 0$. Then from Lemma 4.1, we get

THEOREM 4.1. *Assume that* $n^{-1}A_nA_n'\to\Omega>0$ *as* $n\to\infty$. *If* $p(m-r)>8$ *and* $l-r>8$, *then*

$$R(\omega, \tilde{\xi}_N(a, b)) = R(\omega, \hat{\xi}_N)$$
$$+\varepsilon^2\frac{a}{(3b)^2\, \mathrm{tr}\, \Omega(CC')^{-1}\, \mathrm{tr}\, Q\Psi}\left\{aE\left[\left(\frac{\mathrm{tr}\, Q\Psi}{\mathrm{tr}\, QT}\right)^2\right] - 6bE\left[\frac{\mathrm{tr}\, Q\Psi}{\mathrm{tr}\, QT}\right]\right\} + o(\varepsilon^2).$$

From Theorem 4.1 and the inequality (3.6), we can see that $\tilde{\xi}_N(a, b)$ asymptotically dominates $\hat{\xi}_N$ if $a\leqq 6(l-r-4)b$. In the univariate case $[p=r=1,\ a_i=1,\ Q=1,\ \Sigma=\sigma^2,\ C=(1, \cdots, 1),\ \Psi=\phi^2]$, Theorems 2.2 and 4.1 give that

$$R(\omega, \bar{X}_N) = E_\omega[(\bar{X}_N-\xi)^2] = \varepsilon + \varepsilon^2\sigma^{-2}(4-\nu\sigma^{-2}) + o(\varepsilon^2).$$

$$R(\omega, \tilde{\xi}_N(a, b)) = R(\omega, \bar{X}_N) + \varepsilon^2\frac{al\{a-6(l-5)b\}}{9b^2(l-3)(l-5)\phi^2} + o(\varepsilon^2).$$

Acknowledgement

The author is grateful to the referee for his helpful comments.

References

Bhattacharya, C. G. (1984), Two inequalities with an application, Ann. Inst. Statist. Math. **36**, 129-134.
Ghosh, M., Nickerson, D. and Sen, P.K. (1987), Sequential shrinkage estimation, Ann. Statist. **15**, 817-829.
Kubokawa, T. (1988), Inadmissibility of the uncombined two-stage estimator when additional samples are available, Ann. Inst. Statist. Math. **40**, 555-563.
Kubokawa, T. (1989), Two-stage procedures for parameters in a growth curve model, J. Statist. Plan. Infer. **22**, 105-115.
Kubokawa, T. (1990), Two-stage estimators with bounded risk in a growth curve model. J. Japan Statist. Soc. **20**, 77-87.
Muirhead, R.J. (1982), Aspects of Multivariate Statistical Theory. Wiley, New York.
Mukhopadhyay, N. (1985), A note on three-stage and sequential point estimation procedures for a normal mean, Sequential Analysis **4**, 311-319.
Rao, C.R. (1973), Linear Statistical Inference and Its Applications, 2nd ed. Wiley, New York.
Takada, Y. (1988), Two-stage procedures for a multivariate normal distribution, Kumamoto J. Math. **1**, 1-8.
Takada, Y. (1990), Sequential point estimation rule with bounded risk for the mean of a multivariate normal distribution, Unpublished.
Woodroofe, M. (1977), Second order approximations for sequential point and interval estimation, Ann. Statist. **5**, 984-995.

Department of Mathematical Engineering and Information Physics
Faculty of Engineering
University of Tokyo
Bunkyo-ku, Tokyo 113, Japan