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 \( x_1, x_2, \ldots \) 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 \( \xi \ (p \times r), \Sigma \ (p \times p) \) 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 \( X_n \ (p \times n) \) 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 \( \hat{\xi}_n \) of the coefficient matrix \( \xi \) such that

\[
R(\omega, \hat{\xi}_n) = \mathbb{E}_\omega [n^{-1} \text{tr} Q(\hat{\xi}_n - \xi)A_n A_n' (\hat{\xi}_n - \xi)'] \leq \varepsilon
\]

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

Throughout the paper, let \( m_n \) be the smallest integer \( (\geq r) \) such that \( \text{rank}(A_{m_n})=r \). In the case where \( \Sigma \) is known, for integer \( n \ (\geq m_n) \), MLE of \( \xi \) is given by

\[
\hat{\xi}_n(n) = X_n A_n' (A_n A_n')^{-1}
\]

and from Muirhead (1982),

\[
R(\omega, \hat{\xi}_n(n)) = \mathbb{E}_\omega [n^{-1} \text{vec}(\hat{\xi}_n(n)-\xi)'(A_n A_n' \otimes Q)\text{vec}(\hat{\xi}_n(n)-\xi)]
\]

\[
= n^{-1} \text{tr} (A_n A_n' \otimes Q) \text{Cov} (\text{vec} \hat{\xi}_n(n))
\]

Received May 2, 1989.
where the notation \( \text{vec}\xi \) denotes \( pr \times 1 \) 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 \( R(\omega, \hat{\xi}_n(n)) \leq \varepsilon \) if and only if \( n \geq r \text{tr} Q \Sigma / \varepsilon \ (=n^*) \). Since \( \Sigma \) 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 \( a_i=(1, \ldots, 1) \) for each \( i \), Mukhopadhyay (1985) and Takada (1989) obtained three-stage and purely sequential procedures satisfying

\[
\lim_{n \to \infty} R(\omega, \hat{\xi}_n)/\varepsilon = 1. \quad \text{(asymptotic consistency)}
\]

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

\[
N = \text{Min} \left\{ n \geq m ; n \geq \frac{r}{\varepsilon(n-r)} \text{tr} Q S_n \right\},
\]

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

\[
\hat{\xi}_N=X_N A_N'(A_N A_N')^{-1},
\]

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

\[
\lim_{n \to \infty} E[N]/n^* = 1.
\]

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 \( \Psi \) and the common coefficient matrix \( \xi \). 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 \( p(m-r) \geq 3 \). 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 $n (\geq m \geq m_n)$, the $p \times p$ matrix $S_n = X_n(I_n - A_n'(A_n A_n')^{-1} A_n)X_n'$ is written as

\[
S_n = \sum_{i=m}^{n} T_i,
\]

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

(a) Each $T_i$ is a statistic based on only $x_i, \cdots, x_n$, 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 $X_n A_n'$.

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

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

Then we can express $S_n$ as

\[
S_n = S_{n-1} + X_n D_n X_n'.
\]

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

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

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

\[
S_n = \sum_{i=m}^{n} X_n D_i X_n'.
\]

It can be shown that $D_i = D_i$, $D_i D_j = 0$ ($i \neq j$) and $A_n D_i = 0$ for $i=m, \cdots, n$ that rank $(D_m) = m-r$, rank $(D_i) = 1$ for $i=m+1, \cdots, n$. Letting $T_i = X_n D_i X_n'$ establishes Lemma 2.1.

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

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

\[
E[(n^*/N)^2 I_{(n^*/N)\geq d}] \leq d^{-\delta} E[(n^*/N)^{\delta(1+\delta)}],
\]
where \( I_{\cdot,\cdot} \) designates the indicator function, so that it is sufficient to show that 
\[
\sup_{0 < \varepsilon < \varepsilon_0} E\left[ \left( \frac{n^*}{N} \right)^{\alpha(1+\delta)} \right] < \infty.
\]
Lemma 2.3 of Woodroofe (1977) gives that for 
\[
0 < \theta < 1,
\]
\[
P\left[ N \leq \theta n^* \right] = O\left( \varepsilon^{p(m-r)/2} \right) \quad \text{as } \varepsilon \to 0.
\]
[Woodroofe's notations \( c, \, t, \, m, \, \alpha, \, \beta, \, \mu, \, \tau, \, L_0, \, \lambda \) correspond to our \( \varepsilon/r, \, N-r, \, m-r, \, 2, \, 1, \, \text{tr} \, Q \Sigma, \, 2 \, \text{tr} \, (Q \Sigma)^2, \, r, \, n^* \), respectively.] Hence for \( 0 < \varepsilon < \varepsilon_0 \),

\[
E\left[ \left( \frac{n^*}{N} \right)^{\alpha(1+\delta)} \right] \leq E\left[ \left( \frac{n^*}{m} \right)^{\alpha(1+\delta)} P\left[ N \leq \theta n^* \right] + \theta^{-\lambda(1+\delta)} \right]
\]
\[
\leq K \varepsilon^{p(m-r)/2-\lambda(1+\delta)} + \theta^{-\lambda(1+\delta)}
\]
\[
\leq K \varepsilon^{p(m+r)/2-\lambda(1+\delta)} + \theta^{-\lambda(1+\delta)},
\]
where \( K \) is a constant independent of \( \varepsilon \). The fourth inequality in (2.3) follows from the fact that there exists a positive \( \delta \) satisfying \( p(m-r)/2-\lambda(1-\delta) \geq 0 \), which always holds if \( p(m-r)/2 > \lambda \).

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

\[
\text{tr} \, QT_i = \sum_{j=1}^p \sigma_j W_{ij},
\]
where \( W_{i1}, \ldots, W_{ip} \) are mutually independent random variables, \( W_{ij} \sim \mathcal{N}_{m-r} \), and for \( i = m+1, \ldots, n \), \( W_{ij} \sim \mathcal{N}_r \). Then,

\[
\text{tr} \, QS_n = \sum_{i=m}^n \sum_{j=1}^p \sigma_j W_{ij} = \sum_{j=1}^p \sum_{i=m}^n \sigma_j Q_j^{(n-r)},
\]
where \( Q_j^{(n-r)} = \sum_{i=m}^n W_{ij} \) having \( \mathcal{N}_{n-r} \). Also from the definition of \( N \),

\[
N < \frac{1}{\varepsilon (N-r-1)} \text{tr} \, Q \Sigma N-I_{\{N \geq m+1\}} + 1 + m.
\]
Let \( \tau = -\lambda(1+\delta) \). In the rhs of the inequality (2.2), from (2.5) and (2.6), we can see that for \( 0 < \varepsilon < \varepsilon_0 \),

\[
E\left[ \left( \frac{N}{n^*} \right)^r \right] < E\left[ \left( \text{tr} \, Q \Sigma \right)^{-1} \sum_{j=1}^p \sigma_j Q_j^{(N-r-1)/(N-r-1)} \right] + \left( 1 + m \right) / n^* \]
\[
< \sum_{j=1}^p C_j E\left[ \left( Q_j^{(N-r-1)/(N-r-1)} \right)^r \right] + C_0 \varepsilon_0^r
\]
\[
< \frac{1}{\varepsilon_0} \sum_{j=1}^p C_j E\left[ \sup_{n \geq m} \left( Q_j^{(n-r)/(n-r)} \right)^r \right] + C_0 \varepsilon_0^r,
\]
where $C_n, C_1, \ldots, C_p$ are constants independent of $\varepsilon$. The Doob's maximal inequality for reversed martingale sequence $\{Q^{(n-r)}/(n-r)\}_{n \geq m}$ gives that $E[\sup_{n \geq m} \{Q^{(n-r)}/(n-r)\}] < \infty$. Therefore the uniform integrability of $(n^*/N)^2$ is completely proved.

**Proof of Theorem 2.1** Denote by $\mathcal{F}_n$ the $\sigma$-algebra generated by $T_m, \ldots, T_n$ given in Lemma 2.1. Similar to (1.2), the risk function of $\hat{\xi}_N$ is represented as

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

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

$$E[\{\text{vec } \hat{\xi}_n - \text{vec } \xi\} \text{vec } \hat{\xi}_n - \text{vec } \xi] I_{N=n=1}] = E[I_{N=n=1}] E[\{\text{vec } \hat{\xi}_n - \text{vec } \xi\}] I_{N=n} = \{E[I_{N=n}] \text{Cov } (\text{vec } \hat{\xi}_n)] = \{(A_n A_n')^{-1} \otimes \Sigma\} E[I_{N=n}].$$

which yields that

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

Since $n^*/N \rightarrow 1$ a.s. as $\varepsilon \rightarrow 0$, applying Lemma 2.2 with $\lambda = 1$ proves that $R(\omega, \hat{\xi}_N) /\varepsilon \rightarrow 1$ as $\varepsilon \rightarrow 0$. 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 $p(m-r) \geq 5$,

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

$$R(\omega, \hat{\xi}_N) = \varepsilon^2 + \frac{2 \varepsilon \text{tr} (Q \Sigma)^{\nu}}{\text{tr} \Sigma (Q \Sigma)_{\nu}^\nu} - \frac{\nu}{\text{tr} \Sigma} + o(\varepsilon^2),$$

where $\nu$ is defined by (2.4) in Woodroofe (1977).

The expansion (2.8) is from Woodroofe (1977). Note that $n^*/N = (N-n^*)/(n^* N) - N/n^* + 2$ and that $(N-n^*)N^{-1/2} \rightarrow N(0, 2 \text{tr} (Q \Sigma)^{\nu}/\text{tr} \Sigma (Q \Sigma)_{\nu}^\nu)$ as $\varepsilon \rightarrow 0$. 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 $(|N-n^*|/N^{1/2})^2$ is uniformly integrable.

Proof. First, observe that for $d, \delta>0$,
\[
E[(|N-n^*|/N^{1/2})^2I_{\{N-n^*|/N^{1/2}d\}}] \\
\leq d^{-\delta}E[(|N-n^*|/N^{1/2})^{2(1+\delta)}] \\
\leq d^{-\delta}E[(n^*/N)^{2(1+\delta)}]E[|N-n^*|/(n^*)^{1/2}]^{2(1+\delta)}]^{1/2}.
\]
From Lemma 2.2, $(n^*/N)^{2(1+\delta)}$ 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 $(|N-n^*|/n^*)^{2(1+\delta)}$ is uniformly integrable under the same condition. Hence there exists some constant $M$ independent of $\varepsilon$ such that
\[
E[(n^*/N)^{2(1+\delta)}]E[|N-n^*|/(n^*)^{1/2}]^{2(1+\delta)}]^{1/2} < M,
\]
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 $\xi$, 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 $\Psi$ and the common coefficient matrix $\xi$. 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
\[
\hat{\xi}_N(a, b)=\hat{\xi}_N + a(1+R_N)^{-1}(\hat{\xi}_Y-\hat{\xi}_N),
\]
where
\[
R_N=b \text{ tr } A_N A_N'(CC')^{-1} \text{ tr } QT/(r v_N),
\]
\[
v_N=\text{ tr } Q S_N/[3(N-1)],
\]
\[
T=Y(I_l-C'(CC')^{-1}C)'Y',
\]
and $a$, $b$ are positive constants.
Sequential point estimation

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

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

\begin{equation}
\Delta = R(\omega, \hat{\xi}_N) - R(\omega, \hat{\xi}_N(a, b))
= E\left[\frac{2ar}{(1+R_N)^N} \text{tr} Q \Sigma - \frac{a^z}{(1+R_N)^2} \text{tr} Q \Sigma + \frac{1}{N} \text{tr} (A_N^1 A_N^2) (CC')^{-1} \text{tr} Q \Psi \right]
= ar \text{tr} Q \Sigma E^N \left[ N^{-1} \{2(1+\theta_N U_N)^{-1} - a(1+\theta_N)(1+\theta_N U_N)^{-z} \} \right],
\end{equation}

where

\begin{equation}
\theta_N = \text{tr} A_N^1 A_N^2 (CC')^{-1} \text{tr} Q \Psi / (r \text{tr} Q \Sigma), \quad U_N = b(\text{tr} Q \Sigma)(\text{tr} Q T)/(v_N \text{tr} Q \Psi).
\end{equation}

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

\begin{equation}
\frac{2}{1+\theta_N U_N} - \frac{1}{(1+\theta_N U_N)^2} \geq \frac{a}{1+\theta_N a} (2U_N^4 - a U_N^z),
\end{equation}

which yields that $\Delta \geq 0$ for all $\omega$ if

\begin{equation}
E\left[ g(N) N^{-1} v_N \left\{ 2b \frac{\text{tr} Q \Psi}{\text{tr} Q T} - a \left( \frac{\text{tr} Q \Psi}{\text{tr} Q T} \right)^z \frac{v_N}{\sigma} \right\} \right] \geq 0
\end{equation}

for all $\omega$, where $g(n) = (1+\theta_N a)^{-1}$ and $\sigma = \text{tr} Q \Sigma$. Similar to (2.5),

\begin{equation}
\frac{E[\text{tr} Q \Psi / \text{tr} Q T]}{E[\{\text{tr} Q \Psi / \text{tr} Q T\}^z]} = \frac{E[\sum_{i=1}^p \eta_i w_i^{-1}]}{E[\sum_{i=1}^p \eta_i w_i^{-z}]} \geq \min_{1 \leq i \leq p} \{E[w_i^{-1}]\},
\end{equation}

where $w_1, \ldots, w_p$ are mutually independent random variables, each $w_i$ having $\chi_{1-r}$ and $\eta_1, \ldots, \eta_p$ are parameters satisfying $\sum_{i=1}^p \eta_i = 1$ and $\eta_i > 0, i=1, \ldots, p$. Here the inequality in (3.6) follows from theorem 2.2 of Bhattacharya (1984).

From the condition $a \leq 2(l-r-4)b$ and the fact that $E[w_i^{-1}]/E[w_i^{-z}] = l-r-4$, the inequality (3.5) holds if $E[g(N) N^{-1} v_N (1-v_N/\sigma)] \geq 0$ for all $\omega$, which is rewritten as

\begin{equation}
\sum_{n=m}^\infty g(n) E[\{n^{-1} v_n (1-v_n/\sigma) I_{\{n \geq n_0\}}\}] \geq 0
\end{equation}

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 $(\geq m)$ such that $n(n-r)/(3r(n-1)) \geq \sigma$. It should be noted that $n_0$ is uniquely determined. Then we write the l.h.s of (3.7)

\begin{equation}
= \sum_{n=m}^{n_0} g(n) E\left[ \frac{1}{n} v_n (1-v_n/\sigma) I_{\{n \geq n_0\}} \right] + g(n_0) E\left[ \frac{1}{n_0} v_{n_0} (1-v_{n_0}/\sigma) I_{\{n \geq n_0\}} \right]
\end{equation}
\[ + \sum_{n=n_0}^{\infty} \left\{ g(n+1)E \left[ \frac{1}{n+1} v_{n+1} (1-v_{n+1}/\sigma) I_{\{N\geq n+1\}} \right] - g(n) E \left[ \frac{1}{n} v_n (1-v_n/\sigma) I_{\{N\geq n+1\}} \right] \right\}, \]

where the first term in the rhs of (3.8) should be interpreted as zero if \( n_0 = m \).

Note that for \( n \geq n_0 \), on the set \( \{ N \geq n+1 \} \), \( v_n > \varepsilon n(n-r)/\{3r(n-1)\} \geq \sigma \). Since \( g(n) \) is nonincreasing,

third term in the rhs of (3.8)

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

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 \( u = \text{tr} QT_n \). Then,

(3.9) \[ E \left[ \left\{ \frac{1}{n+1} v_{n+1} (1-v_{n+1}/\sigma) - \frac{1}{n} v_n (1-v_n/\sigma) \right\} I_{\{N\geq n+1\}} \right] \]

\[ = I_{\{N\geq n+1\}} \left[ \frac{1}{n+1} \left\{ \frac{n-1}{n} v_n + \frac{1}{3n} E[u^2] \right\} - \frac{1}{n+1} \left\{ \left( \frac{n-1}{n} \right)^2 v_n^2 + \frac{2(n-1)}{3n^2} v_n E[u^2] + \frac{1}{9n^4} E[u^4] \right\} \right] \frac{1}{\sigma} - \frac{1}{n} v_n + \frac{1}{n\sigma} v_n^2 \]

\[ \geq I_{\{N\geq n+1\}} \left[ \frac{3n-1}{n^2(n+1)} v_n^2 / \sigma - \frac{8n-2}{3n^2(n+1)} v_n + \frac{n-1}{3n^2(n+1)} \sigma \right] \]

since \( E[u] = \sigma \) and \( E[u^2] \leq 3\sigma^2 \). 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 \( \{ N \geq n+1 \} \), \( v_n > \sigma \), it follows that

extreme rhs of (3.9)

\[ \geq I_{\{N\geq n+1\}} \left[ \frac{3n-1}{n^2(n+1)} - \frac{8n-2}{3n^2(n+1)} + \frac{n-1}{3n^2(n+1)} \right] \sigma \]

\[ \geq 0. \]

Next, note that \( I_{\{N\geq n\}} = 1 \) with probability 1 and that

(3.10) \[ E[v_m (1-v_m/\sigma)] = E[v_m] - E[v_m^2] / \sigma \]

\[ \geq \frac{m-r}{3(m-1)^2} \sigma - \frac{(m-r)(m-r+2)}{\{3(m-1)\}^2} \sigma \]

\[ = (m-r)(2m+r-5)\sigma / \{9(m-1)^2\} \]

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

the first two terms in the rhs of (3.8)

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

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

$$\begin{align*}
(3.12) & E\left[ n_0^{-1} v_{n_0} (1-v_{n_0}/\sigma) I_{\{N=n_0\}} \right] \geq \left\{ a_{n_0} v_{n_0-1} - b_{n_0} v_{n_0-1}^2/\sigma + \sigma c_{n_0} \right\} I_{\{N=n_0\}} \\
& \geq b_{n_0} v_{n_0-1} (1-v_{n_0-1}/\sigma) I_{\{N=n_0\}},
\end{align*}$$

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

the rhs of (3.11) $\geq g(n_0) \left\{ \sum_{n=m}^{n_0-2} E\left[ n^{-1} v_n (1-v_n/\sigma) I_{\{N=n\}} \right] \right\} + E\left[ (n_0-1)^{-1} v_{n_0-1} (1-v_{n_0-1}/\sigma) I_{\{N=n_0\}} \right] \geq 0$.

Proceed inductively to get

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

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

4. Asymptotic risk expansion

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

From (3.2), the risk difference is written as

$$\begin{align*}
\Delta = -\varepsilon^2 a^2 (r tr Q)\Sigma^{-1} E\left[ (n^*/N)^2 P_N \right],
\end{align*}$$

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

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

**Proof.** Put $Z_N = N(1+\theta_N)(1+\theta_N U_N)^{-1}$. Observe that for $d, \delta > 0$,
(4.2) \[ E[(n^* / N)^{\rho} Z_N I_{(n^* / N)^{\rho} Z_N > s_\varepsilon}] \leq d^{-\frac{\varepsilon}{2}} E[(n^* / N)^{\rho (1 + \varepsilon)} Z_N^{\rho}] \leq d^{-\frac{\varepsilon}{2}} E[(n^* / N)^{\rho (1 + \varepsilon)}] E[Z_N^{\rho (1 + \varepsilon)}] \]\n
Since \((n^* / N)^{\rho (1 + \varepsilon)}\) is uniformly integrable under the condition \(p(m-r) > 8\) by Lemma 2.2, there exists some constant \(M_1\) independent of \(\varepsilon\) such that \(E[(n^* / N)^{\rho (1 + \varepsilon)}] < M_1\) for \(0 < \varepsilon < \varepsilon_0\). Also,

\[ Z_N = \frac{N}{1 + \theta_N} \left( \frac{1 + \theta_N}{1 + \theta_N U_N} \right)^{\rho} \leq \frac{N}{1 + \theta_N} \left( 1 + \frac{1}{\theta_N} \right)^{\rho} U_N^\rho. \]

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

\[ n(1 + \theta_n)^{-1} (1 + \theta_n)^{-\rho} \leq M_2 \text{ for all } n \geq m. \]

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

\[ E[Z_N^{\rho (1 + \varepsilon)}] \leq \sum_{j=1}^{n} C_j E[(tr QT)^{-\rho (1 + \varepsilon)}] E[\sup_{n \geq m} Q_j^{\nu} / (n-r)] \]

for constants \(C_j\) independent of \(\varepsilon\). 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 \((n^* / N)^{\rho} Z_N\) holds. Similarly we can show the uniform integrability of \((n^* / N)^{\rho} N(1 + \theta_n U_N)^{-1}\).

Note that \((n^* / N)^{\rho} \to 1\) a.s. and

\[ P_N \to \frac{r \ tr \ Q \Sigma}{(3b)^{\rho} \ tr \ Q(C)\Sigma^{-1} \ tr \ Q \Psi} \left\{ a \left[ \frac{tr \ Q \Psi}{tr \ QT} \right] - 6b \left[ \frac{tr \ Q \Psi}{tr \ QT} \right] \right\} \text{ a.s.} \]

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

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

\[ R(\omega, \xi_N(a, b)) = R(\omega, \xi_N) = R(\omega, \xi_N) \]

\[ + \varepsilon a \left[ \frac{tr \ Q \Psi}{tr \ QT} \right] - 6b \left[ \frac{tr \ Q \Psi}{tr \ QT} \right] + o(\varepsilon^2). \]

From Theorem 4.1 and the inequality (3.6), we can see that \(\xi_N(a, b)\) asymptotically dominates \(\xi_N\) if \(a \leq 6(l-r-4)b\). In the univariate case \(p=r=1, a_i=1, Q=1, \Sigma=\sigma^2, C=(1, \ldots, 1), \Psi=\phi^2\). Theorems 2.2 and 4.1 give that

\[ R(\omega, \xi_N) = R(\omega, \xi_N) = R(\omega, \xi_N) = R(\omega, \xi_N) + \varepsilon a \left[ \frac{a - 6(l-5)b}{9b^2(l-3)(l-5)b^2} \right] + o(\varepsilon^2). \]
Acknowledgement

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

References

Mukhopadhyay, N. (1985), A note on three-stage and sequential point estimation procedures for a normal mean, Sequential Analysis 4, 311-319.
Takada, Y. (1990), Sequential point estimation rule with bounded risk for the mean of a multivariate normal distribution, Unpublished.

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