ON THE SECOND ORDER APPROXIMATION OF THE RISK OF SEQUENTIAL POINT ESTIMATION IN A MULTIVARIATE ANALYSIS

HISAO NAGAO

Received July 23, 2002

ABSTRACT. This paper deals with the problem of sequential point estimator of the mean under the general multivariate continuous distribution. As a loss function, we consider the square loss plus sample size times cost. Under this loss, the second term of the regret can be shown as the cost goes to zero. Also the negative regret is discussed.

1. Introduction

Let $p \times 1$ vectors $X, X_1, X_2, \cdots$ be independent observations from a population which has a continuous known distribution form with unknown parameters. We denote μ and Σ by mean vector and covariance matrix ($|\Sigma| \neq 0$). Given sample size n, we want to estimate μ by a sample $\bar{X}_n$ and we suffer a loss

$$L_n = (\bar{X}_n - \mu)'(\bar{X}_n - \mu) + cn, \tag{1.1}$$

where $c > 0$ is a cost par one observation. The risk is

$$R_n = \mathrm{E}L_n = \frac{1}{n}\mathrm{tr}\Sigma + cn, \tag{1.2}$$

which is minimized by using the optimal fixed sample size $n_c = \sqrt{\dfrac{\mathrm{tr}\Sigma}{c}}$, when Σ is known. Then the corresponding minimum fixed sample size risk is $R_{n_c} = 2cn_c$. When Σ is unknown, the optimal sample size n_c cannot be used, and there is no fixed sample size rule that will achieve the risk R_{n_c}. For this case, we consider the following stopping rule

$$N_c = N = \inf\left\{ n \geq m \mid n \geq \sqrt{\frac{\mathrm{tr}S_n}{c}} \right\}, \tag{1.3}$$

where $S_n = \sum_{i=1}^n (X_i - \bar{X}_n)(X_i - \bar{X}_n)'/n$ and m may depend on c. We define the risk for the stopping time N.

$$R_c^* = \mathrm{E}\left\{ (\bar{X}_N - \mu)'(\bar{X}_N - \mu) + cN \right\} \tag{1.4}$$

and the regret $\omega(c) = R_c^* - R_{n_c}$.

In this paper, we shall give the second order approximation of the regret $\omega(c)$ as $c \to 0$. This problem has been dealt by many authors. Many of them considered the loss $L_n = A(\bar{X}_n - \mu)'(\bar{X}_n - \mu) + n$ instead of L_n in (1.1) and treated the problem as $A \to \infty$ instead of $c \to 0$. This model in this paper seems to be more natural. When $p = 1$ and a normal distribution, Robbins [12] gave a numerical example of the expectation of N by showing that the sample mean $\bar{X}_n$ and $S_m, \cdots, S_n (n = m, m+1, \cdots)$ are independent. Also under the same situation, Starr [14] has given $R_c^*/R_{n_c} \to 1$ and the much stronger result $\omega(c) = O(1)$

2000 *Mathematics Subject Classification.* 62L12, 62H12.
Key words and phrases. Stopping rule, risk, regret, uniformly integrable, nagative regret, t-distribution.

has been derived by Starr and Woodroofe [15]. Furthermore, Woodroofe [19] has given the second order approximation of the regret of the procedure as $c \to 0$

$$\omega(c) = \frac{1}{2}c + o(c). \tag{1.5}$$

In these papers, the delay m does not depend on c. In a multivariate normal population, Khan [7] and Rohatgi and O'Neill [13] considered the sequential estimation problems under the assumption that the covariance matrix is a diagonal matrix with unknown elements. For the trivial extension of diagonal matrix, Ghosh, Sinha and Mukhopadyay [6] and Wang [18] have given that the regret is $\omega(c) = O(c)$, and for the stopping time, Ghosh et al. have given numerical values by simulation and Wang has given exact distribution. For non-normal case, Starr and Woodroofe [16] treated the mean of the exponential density and the formula of their stopping time is different from (1.3) and they showed that the regret $\omega(c) \leq O(c)$. For the problem that the univariate distribution is unspecified, Ghosh and Mukhopadhyay [4] and Martinsek [8] considered the same problem as one in this paper. Ghosh et al. showed that $R^*(c)/R_{n_c} \to 1$ as $c \to 0$ and Martinsek has given the second order approximation of $\omega(c)$ by using Chow and Martinsek [2] under a little different loss mentioned above. See Ghosh , Mukhopadyay and Sen [5].

Also Takada and Nagao [17] and Nagao [10] considered the linex loss under a multivariate normal distribution and regression model. For the fixed width confidene region, Nagao and Srivastava [11] considered for multivariate normal case. The problem in this paper treats the non-normal case.

Let $U_c = N(\mathrm{tr}\Sigma/\mathrm{tr}S_N)^{1/2} - n_c$ which is called the excess and the random variable U to which U_c converges in law as $c \to 0$ has distribution defined in terms of the first time that the random walk $W_n = \frac{1}{2}\sum_{i=1}^{n}(3 - (\mathrm{tr}\Sigma)^{-1}(X_i - \mu)'(X_i - \mu))$ is positive.

2. The expectation of stopping time

We shall give the proof of the following theorem. We can assume that the mean of X is zero vector.

Let $Z_n = n(\mathrm{tr}\Sigma/\mathrm{tr}S_n)^{1/2}$, then we have $Z_n = W_n + \xi_n$, where with $* \in (\mathrm{tr}\Sigma, \mathrm{tr}S_n)$

$$\begin{aligned} W_n &= \frac{1}{2}\sum_{i=1}^{n}(3 - (\mathrm{tr}\Sigma)^{-1}X_i'X_i) \\ \text{and} \qquad & \\ \xi_n &= \frac{n}{2}(\mathrm{tr}\Sigma)^{-1}\bar{X}_n'\bar{X}_n + \frac{3}{8} *^{-5/2}\, n(\mathrm{tr}(S_n - \Sigma))^2(\mathrm{tr}\Sigma)^{1/2}. \end{aligned} \tag{2.1}$$

Then by a routine consideration, we have Theorem 2.1.

Theorem 2.1. For stopping time N, we have

$$\mathrm{E}(N - n_c) = \nu - 0.5 - \frac{3}{8}(\mathrm{tr}\Sigma)^{-2}\mathrm{Var}((X-\mu)'(X-\mu)) + o(1), \tag{2.2}$$

where $\nu = \mathrm{E}(U)$.

3.The derivation of the regret $\omega(c)$.

The proof is mainly based on Chow, Robbins and Teicher [3], Chow and Martinsek [2] and Martinsek [8]. Next we shall give the approximation of $\omega(c)$ as $c \to 0$. When $p = 1$, Martinsek [8] derived the expression of the risk. In this case,

$$\begin{aligned} \omega(c) &= \mathrm{E}(\bar{X}_N'\bar{X}_N) + c\mathrm{E}(N) - 2cn_c = \mathrm{E}(\bar{X}_N'\bar{X}_N - cN) \\ &+2c\mathrm{E}(N - n_c) = c\mathrm{E}(\frac{1}{c}\bar{X}_N'\bar{X}_N - N) + 2c\mathrm{E}(N - n_c). \end{aligned} \tag{3.1}$$

The lemma in Chow, Robbins and Teicher is helpful to give the derivation of $\omega(c)$. So we mention their lemma.

Let $\{Y_n, \mathcal{F}_n, 1 \leq n < \infty\}$ be a stochastic process, where Y_n is $\mathcal{F}_n$-measurable and put $\mathcal{F}_0 = \{\phi, \Omega\}$, $Y_0 = 0$, $Z_n = Y_n - Y_{n-1}$. Then we have

Lemma 3.1. (Chow, Robbins and Teicher.) If $Y_n \geq 0$ $(n = 1, 2, \cdots)$, then for any stopping time t,

$$\mathrm{E}(Y_t) = \mathrm{E}(\sum_{i=1}^{t} \mathrm{E}\left(Z_i | \mathcal{F}_{i-1}\right)),$$

where $Y_n = \sum_{i=1}^{n} Z_i$.

By Lemma 3.1,

Lemma 3.2. Let $S_n^* = \sum_{i=1}^{n} X_i$ with $S_0^* = 0$. If $\mathrm{E}(N) < \infty$, then we have

$$\mathrm{E}(S_N^{*'} S_N^*) = (\mathrm{tr}\Sigma)\mathrm{E}(N). \tag{3.2}$$

Thus we have

$$\mathrm{E}(\frac{1}{c}\bar{X}_N' \bar{X}_N - N) = \mathrm{E}(\frac{1}{cN^2} S_N^{*'} S_N^* - N) = \mathrm{E}(S_N^{*'} S_N^* (\frac{1}{cN^2} - \frac{1}{\mathrm{tr}\Sigma})) = I + II, \tag{3.3}$$

where

$$\begin{aligned} I &= \mathrm{E}\left(S_N^{*'} S_N^* ((cN^2)^{-1} - (\mathrm{tr}S_N)^{-1})\right), \\ \text{and} \qquad II &= \mathrm{E}\left(S_N^{*'} S_N^* ((\mathrm{tr}S_N)^{-1} - (\mathrm{tr}\Sigma)^{-1})\right). \end{aligned} \tag{3.4}$$

Then we have

$$I = -\mathrm{E}\left(S_N^{*'} S_N^* \frac{(\mathrm{tr}\Sigma)^{-1/2}}{N} U_c (\frac{1}{\sqrt{c}N} + \frac{1}{\sqrt{\mathrm{tr}S_N}})\right). \tag{3.5}$$

After some calculation, we have

$$U_c \leq \frac{3}{2} n_c (N-1)^{-1} + \frac{3}{8} n_c (N-1)^{-3/2}. \tag{3.6}$$

So we have to show that $\frac{n_c}{N}$ is u.i.(uniformly integrable). Also we need u.i. of some statistics. So we summarize here as lemma. See Chow and Martinsek [2].

Lemma 3.3. Let $Q_n = \sum_{i=1}^{n} X_i' X_i$ with $Q_0 = 0$. We assume that there is positive number δ such that $\delta c^{-1/4} \leq m = o(c^{-1/2})$.

(1)

$$\text{If } \ \mathrm{E}(X'X) < \infty, \ \text{ then } \ (\frac{n_c}{N})^q \ \text{ is u.i. for any } \ q > 0. \tag{3.7}$$

(2)

$$\text{If } \ \mathrm{E}(X'X)^t < \infty \ \text{ for } \ t \geq 1, \ \text{ then } \ (\frac{N}{n_c})^t \ \text{ is u.i.} \tag{3.8}$$

(3)

$$\text{If } \ \mathrm{E}(X'X)^t < \infty \ \text{ for } \ t \geq 1, \ \text{ then } \ |\frac{1}{n_c} S_N^{*'} S_N^*|^t \ \text{ is u.i.} \tag{3.9}$$

(4)

$$\text{If } \mathrm{E}(X'X)^t < \infty \quad \text{for} \quad t \geq 2, \quad \text{then} \quad |\frac{1}{\sqrt{n_c}}(Q_N - N\mathrm{tr}\Sigma)|^t \quad \text{is u.i.} \tag{3.10}$$

Since U_c is u.i., we have

$$I = -2\nu + o(1). \tag{3.11}$$

For II, we have $II = II_a + II_b$, where with $V_n = \sum_{i=1}^n (X_i - \bar{X}_n)(X_i - \bar{X}_n)'$,

$$II_a = \frac{(\mathrm{tr}\Sigma)^{-1}}{n_c} \mathrm{E}\left\{(S_N^{*'} S_N^* - \mathrm{tr}V_N)(N - (\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N)\right\} \tag{3.12}$$

and

$$\begin{aligned} II_b = \quad & \frac{1}{n_c}\mathrm{E}\left\{(S_N^{*'} S_N^* - \mathrm{tr}V_N)(n_c - (\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N)(\mathrm{tr}V_N)^{-1}(N \right. \\ & \left. -(\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N)\right\} + (\mathrm{tr}\Sigma)^{-1}\mathrm{E}(\frac{1}{N} S_N^{*'} S_N^*). \end{aligned} \tag{3.13}$$

Then we have

$$\begin{aligned} II_a \quad = & \frac{(\mathrm{tr}\Sigma)^{-1}}{n_c}\left\{\mathrm{E}(S_N^{*'} S_N^* - Q_N)(N - (\mathrm{tr}\Sigma)^{-1}Q_N) + (\mathrm{tr}\Sigma)^{-1} \right. \\ & \times\mathrm{E}(\frac{1}{N} S_N^{*'} S_N)^2 \; + \; (\mathrm{tr}\Sigma)\mathrm{E}(N) - 2(\mathrm{tr}\Sigma)^{-1}\mathrm{E}(\frac{1}{N} Q_N S_N^{*'} S_N^*) \\ & \left. +(\mathrm{tr}\Sigma)^{-1}\mathrm{E}(\frac{1}{N}(S_N^{*'} S_N^*)^2)\right\}. \end{aligned} \tag{3.14}$$

Lemma 3.4.

$$\frac{1}{n_c}\mathrm{E}(\frac{1}{N} S_N^{*'} S_N^*)^2 = o(1), \tag{3.15}$$

$$\frac{1}{n_c}\mathrm{E}(\frac{1}{N}(S_N^{*'} S_N^*)^2) = (\mathrm{tr}\Sigma)^2 + 2\mathrm{tr}\Sigma^2 + o(1), \tag{3.16}$$

$$\frac{1}{n_c}\mathrm{E}(\frac{1}{N} Q_N(S_N^{*'} S_N^*)) = (\mathrm{tr}\Sigma)^2 + o(1). \tag{3.17}$$

Proof. By Hölder inequality, we have for any set A,

$$\frac{1}{n_c}\mathrm{E}(\frac{1}{N} S_N^{*'} S_N^*)^2 I_A \leq \frac{1}{n_c}\left\{\mathrm{E}(\frac{n_c}{N})^{2+\epsilon} I_A\right\}^{2/(2+\epsilon)} \left\{\mathrm{E}(\frac{1}{n_c} S_N^{*'} S_N^*)^{2(2+\epsilon)/\epsilon} I_A\right\}^{\epsilon/(2+\epsilon)}, \tag{3.18}$$

where I_A is an indicator function. Thus $\frac{1}{n_c}\mathrm{E}(\frac{1}{N} S_N^{*'} S_N^*)^2$ is u.i. Thus by Lemma 3.3, we have (3.15). Similarly we have for (3.16)

$$\frac{1}{n_c}\mathrm{E}\left(\frac{1}{N}(S_N^{*'} S_N^*)^2\right) = \mathrm{E}(\frac{n_c}{N})(\frac{1}{n_c} S_N^{*'} S_N^*)^2.$$

Also for (3.17), we have the L.H.S. of (3.17)$= \frac{1}{n_c}\mathrm{E}(\frac{1}{N})(Q_N - N\mathrm{tr}\Sigma)(S_N^{*'} S_N^*) + \frac{(\mathrm{tr}\Sigma)^2}{n_c}\mathrm{E}(N)$.

Lemma 3.5. We have

$$\mathrm{E}\sum_{i=1}^{N} S^*_{i-1} = \mathrm{E}NS^*_N. \tag{3.19}$$

Proof. Let a be a $p\times 1$ fixed vector. We put $a'S^*_N = U_N = \sum_{i=1}^N u_i$ with $u_i = a'X_i$. Let $Y_n = nU_n$, then we have with $U_0 = 0$

$$Y_n - Y_{n-1} = U_{n-1} + nu_n.$$

Thus we obtain $Y_n = \sum_{j=1}^n (U_{j-1} + ju_j)$. Therefore we have

$$\mathrm{E}NU_N = \mathrm{E}\sum_{j=1}^{N}(U_{j-1} + ju_j) = \mathrm{E}\sum_{j=1}^{N} U_{j-1} + \mathrm{E}\sum_{j=1}^{N} ju_j.$$

To prove (3.19), we must show that $\mathrm{E}\sum_{j=1}^N U_{j-1}$ and $\mathrm{E}\sum_{j=1}^N ju_j$ are absolute convergences. Thus we have

$$\mathrm{E}|\sum_{j=1}^{N} ju_j| \le \mathrm{E}(N\sum_{j=1}^{N}|u_j|) \le (\mathrm{E}N^2)^{1/2}\{\mathrm{E}(\sum_{j=1}^{N}|u_j|)^2\}^{1/2}.$$

Since

$$\begin{aligned}\mathrm{E}(\textstyle\sum_{j=1}^N |u_j|)^2 = &\ \mathrm{E}\{(\textstyle\sum_{j=1}^N(|u_j| - \mathrm{E}|u_j|)) + N\mathrm{E}|u_1|\}^2 \\ &\le 2\{\mathrm{E}(\textstyle\sum_{j=1}^N(|u_j| - \mathrm{E}|u_j|))^2\} + (\mathrm{E}N^2)(\mathrm{E}|u_1|)^2\}\end{aligned}$$

and $\mathrm{E}(\sum_{j=1}^N(|u_j| - \mathrm{E}(|u_j|))^2 = (\mathrm{E}N)\mathrm{Var}(|u_1|) \le (\mathrm{E}N)(\mathrm{E}(u_1^2))$, $\mathrm{E}\sum_{j=1}^N ju_j$ is absolute convergence. Next we consider $\sum_{j=1}^N U_{j-1}$. Thus we have $|U_{j-1}| \le \sum_{i=1}^j |u_i|$. Then we have $\sum_{j=1}^N |U_{j-1}| \le \sum_{i=1}^N (N-i+1)|u_i| \le N\sum_{i=1}^N |u_i|$. Thus similar caluculation as the proof of Lemma 3.1 yields the desired conclusion.

$$\begin{aligned}II_a = &\ \frac{(\mathrm{tr}\Sigma)^{-1}}{n_c}\{\mathrm{E}(S^{*'}_N S^*_N - N\mathrm{tr}\Sigma)(N - (\mathrm{tr}\Sigma)^{-1}Q_N) + (\mathrm{tr}\Sigma) \\ &\times\mathrm{E}(N - (\mathrm{tr}\Sigma)^{-1}Q_N)^2\} + 2(\mathrm{tr}\Sigma^2)(\mathrm{tr}\Sigma)^{-2} + o(1).\end{aligned} \tag{3.20}$$

We note that

$$\begin{aligned}&2\,\mathrm{E}((\mathrm{tr}\Sigma)^{-1}S^{*'}_N S^*_N - N)((\mathrm{tr}\Sigma)^{-1}Q_N - N) = -(\mathrm{tr}\Sigma)^{-2}\mathrm{E}(S^{*'}_N S^*_N - Q_N)^2 \\ &\quad+\mathrm{E}((\mathrm{tr}\Sigma)^{-1}S^{*'}_N S^*_N - N)^2 + \mathrm{E}((\mathrm{tr}\Sigma)^{-1}Q_N - N)^2.\end{aligned} \tag{3.21}$$

By Lemma 3.1, we have the following lemma.

Lemma 3.6.

$$\mathrm{E}(S^{*'}_N S^*_N - Q_N)^2 = 4\mathrm{E}(\textstyle\sum_{k=1}^N S^{*'}_{k-1}\Sigma S^*_{k-1}), \tag{3.22}$$

$$\begin{aligned}&\mathrm{E}((\mathrm{tr}\Sigma)^{-1}S^{*'}_N S^*_N - N)^2 = 4(\mathrm{tr}\Sigma)^{-2}\mathrm{E}(\textstyle\sum_{k=1}^N S^{*'}_{k-1}\Sigma S^*_{k-1}) + 4(\mathrm{tr}\Sigma)^{-2} \\ &\times\mathrm{E}((X'X)X')\mathrm{E}\textstyle\sum_{i=1}^N S^*_{i-1} + \{(\mathrm{tr}\Sigma)^{-2}\mathrm{E}(X'X)^2 - 1\}\mathrm{E}(N),\end{aligned} \tag{3.23}$$

$$\mathrm{E}((\mathrm{tr}\Sigma)^{-1}Q_N - N)^2 = ((\mathrm{tr}\Sigma)^{-2}\mathrm{E}(X'X)^2 - 1)\mathrm{E}(N). \tag{3.24}$$

Therefore we have

$$\begin{aligned}\mathrm{E}((\mathrm{tr}\Sigma)^{-1}S_N^{*'}S_N^*-N)\,((\mathrm{tr}\Sigma)^{-1}Q_N-N)&=2(\mathrm{tr}\Sigma)^{-2}\mathrm{E}((X'X)X')\mathrm{E}(N-n_c)S_N^*\\&\quad+((\mathrm{tr}\Sigma)^{-2}\mathrm{E}(X'X)^2-1)\mathrm{E}(N).\end{aligned}\tag{3.25}$$

Thus we obtain

$$II_a=-2\frac{(\mathrm{tr}\Sigma)^{-2}}{n_c}\mathrm{E}(X'XX')\mathrm{E}(N-n_c)S_N^*+2(\mathrm{tr}\Sigma^2)(\mathrm{tr}\Sigma)^{-2}+o(1).\tag{3.26}$$

We shall consider the expectation $\frac{2}{n_c}\mathrm{E}a'(N-n_c)S_N^*$, where a is a $p\times 1$ fixed vector. Since $N-n_c=U_c+\frac{1}{2}((\mathrm{tr}\Sigma)^{-1}Q_N-N)-\xi_N$, we have

$$\frac{2}{n_c}\mathrm{E}a'(N-n_c)S_N^*=\frac{1}{n_c}\mathrm{E}\left\{(\mathrm{tr}\Sigma)^{-1}Q_N-N+2(U_c-\xi_N)\right\}a'S_N^*.\tag{3.27}$$

By Schwarz's inequality, we have

$$\frac{1}{n_c}\mathrm{E}|(U_c-\xi_N)a'S_N^*|\le\frac{||a||}{n_c}\mathrm{E}|(U_c-\xi_N)||S_N^{*'}S_N^*|^{1/2}.\tag{3.28}$$

Since ξ_n is given by (3.1), we have to show that $(\frac{n_c}{N})^\alpha(\frac{1}{n_c}S_N^{*'}S_N^*)^{1/2}$, $(\frac{1}{n_c}S_N^{*'}S_N^*)^{3/2}$ and $*^{-5/2}N(\mathrm{tr}(S_N-\Sigma))^2(\frac{1}{n_c}S_N^{*'}S_N^*)^{1/2}$. are u.i. for any $\alpha>0$. Since $*^{-1}\le(\mathrm{tr}\Sigma)^{-1}+(\mathrm{tr}S_N)^{-1}\le(\mathrm{tr}\Sigma)^{-1}+(\mathrm{const.})(\frac{n_c}{N})^5$, $*^{-\alpha}$ is u.i. for any $\alpha>0$. Since $N(\mathrm{tr}(S_N-\Sigma))^2=\frac{1}{N}(\mathrm{tr}(V_N-N\Sigma))^2$ and $\mathrm{tr}(V_N-N\Sigma)=(Q_N-N\mathrm{tr}\Sigma)-\frac{1}{N}S_N^{*'}S_N^*$, we can show their uniformly integrability by Hölder inequality. Therefore we have

$$\frac{2}{n_c}\mathrm{E}a'(N-n_c)S_N^*=\frac{1}{n_c}\mathrm{E}\left((\mathrm{tr}\Sigma)^{-1}Q_N-N\right)a'S_N^*+O(n_c^{-1/2}).\tag{3.29}$$

Since

$$\begin{aligned}\mathrm{E}((\mathrm{tr}\Sigma)^{-1}Q_N-N)a'S_N^*&=\tfrac{1}{2}\{\mathrm{E}((\mathrm{tr}\Sigma)^{-1}Q_N-N+a'S_N^*)^2\\&\quad-\mathrm{E}((\mathrm{tr}\Sigma)^{-1}Q_N-N)^2-\mathrm{E}(a'S_N^*)^2\},\end{aligned}\tag{3.30}$$

by Lemma 3.1, we have

$$\mathrm{E}((\mathrm{tr}\Sigma)^{-1}Q_N-N)a'S_N^*=(\mathrm{tr}\Sigma)^{-1}\mathrm{E}(N)a'\mathrm{E}(XX'X).\tag{3.31}$$

Thus we have

$$\frac{2}{n_c}\mathrm{E}a'(N-n_c)S_N^*=(\mathrm{tr}\Sigma)^{-1}a'\mathrm{E}(XX'X)+o(1).\tag{3.32}$$

Therefore we have

$$II_a=-\{\mathrm{E}(X'XX')\}\{\mathrm{E}(XX'X)\}(\mathrm{tr}\Sigma)^{-3}+2(\mathrm{tr}\Sigma^2)(\mathrm{tr}\Sigma)^{-2}+o(1).\tag{3.33}$$

Next we shall calculate II_b given by (3.13). Then we have $II_b = II_{b_1} + II_{b_2}$, where

$$II_{b_1} = \frac{1}{n_c}\mathrm{E}\left\{(S_N^{*'}S_N^* - \mathrm{tr}V_N)(N - (\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N)\left((\mathrm{tr}\Sigma)^{-1}Q_N - U_c\right.\right.$$
$$\left.\left.+\xi_N - (\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N)(\mathrm{tr}V_N\right)^{-1}\right\} + (\mathrm{tr}\Sigma)^{-1}\mathrm{E}(\frac{1}{N}S_N^{*'}S_N) \quad (3.34)$$

$$II_{b_2} = \frac{3}{2}\frac{1}{n_c}\mathrm{E}\left\{(S_N^{*'}S_N^* - \mathrm{tr}V_N)(N - (\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N)(N - (\mathrm{tr}\Sigma)^{-1}Q_N)\right.$$
$$\left.\times(\mathrm{tr}V_N)^{-1}\right\}. \quad (3.35)$$

Since $(\mathrm{tr}V_N)^{-1} \le (\mathrm{const.})\dfrac{n_c^2}{N^3}$, by Hölder inequality, $II_{b_1} = 1 + o(1)$. Since $N - (\mathrm{tr}\Sigma)^{-1}\mathrm{tr}V_N = N - (\mathrm{tr}\Sigma)^{-1}Q_N + \dfrac{(\mathrm{tr}\Sigma)^{-1}}{N}S_N^{*'}S_N^*$, we have

$$\begin{aligned}II_{b_2} = \; & \frac{3}{2n_c}\mathrm{E}\{S_N^{*'}S_N^*(N - (\mathrm{tr}\Sigma)^{-1}Q_N)^2(\mathrm{tr}V_N)^{-1}\} - \frac{3}{2n_c}\mathrm{E}(N \\ & -(\mathrm{tr}\Sigma)^{-1}Q_N)^2 + \frac{3}{2n_c}\mathrm{E}\{\frac{(\mathrm{tr}\Sigma)^{-1}}{N}(S_N^{*'}S_N^*)^2(N - (\mathrm{tr}\Sigma)^{-1}Q_N) \\ & \times(\mathrm{tr}V_N)^{-1}\} - \frac{3}{2n_c}(\mathrm{tr}\Sigma)^{-1}\mathrm{E}(\frac{1}{N}S_N^{*'}S_N^*(N - (\mathrm{tr}\Sigma)^{-1}Q_N)).\end{aligned} \quad (3.36)$$

Since, by Hölder inequality,

$$\begin{aligned}&\frac{1}{n_c}\mathrm{E}\left\{\frac{1}{N}(S_N^{*'}S_N^*)^2(N - (\mathrm{tr}\Sigma)^{-1}Q_N)(\mathrm{tr}V_N)^{-1}\right\} \le (\mathrm{const})\frac{1}{n_c}\mathrm{E}\left\{(\frac{n_c}{N})^4\right. \\ &\left.\times\, |(\frac{1}{n_c}S_N^{*'}S_N^*)^2(N - (\mathrm{tr}\Sigma)^{-1}Q_N)|\right\} \le (\mathrm{const})\frac{1}{\sqrt{n_c}}\mathrm{E}^{1/4}\left\{(\frac{n_c}{N})^{16}\right\} \\ &\quad \times\mathrm{E}^{1/2}\left(\frac{1}{n_c^2}(S_N^{*'}S_N^*)^4\right)\mathrm{E}^{1/4}\left(\frac{1}{\sqrt{n_c}}(N - (\mathrm{tr}\Sigma)^{-1}Q_N)\right)^4 \\ &\quad = O(\frac{1}{\sqrt{n_c}})\end{aligned} \quad (3.37)$$

and since, by similar calculation,

$$\frac{1}{n_c}\mathrm{E}|\frac{1}{N}S_N^{*'}S_N^*(N - (\mathrm{tr}\Sigma)^{-1}Q_N)| = O(\frac{1}{\sqrt{n_c}}), \quad (3.38)$$

we have

$$\begin{aligned}II_{b_2} = \; & \frac{3}{2n_c}\mathrm{E}\{S_N^{*'}S_N^*(N - (\mathrm{tr}\Sigma)^{-1}Q_N)^2(\mathrm{tr}V_N)^{-1}\} - \frac{3}{2n_c}\mathrm{E}(N \\ & -(\mathrm{tr}\Sigma)^{-1}Q_N)^2 + o(1).\end{aligned} \quad (3.39)$$

From (3.24), we have

$$\mathrm{E}(N - (\mathrm{tr}\Sigma)^{-1}Q_N)^2 = \mathrm{E}\left\{(\mathrm{tr}\Sigma)^{-2}(X'X)^2 - 1\right\}\mathrm{E}(N). \quad (3.40)$$

We shall evaluate the remainder term in II_{b_2}.

At first we have

$$\frac{1}{n_c}\mathrm{E}\left\{S_N^{*'}S_N^*(N-(\mathrm{tr}\Sigma)^{-1}Q_N)^2(\mathrm{tr}V_N)^{-1}\right\}=\frac{1}{n_c^2}\mathrm{E}\left\{S_N^{*'}S_N^*(N\right.$$
$$\left.-(\mathrm{tr}\Sigma)^{-1}Q_N)^2\ (n_c-(\mathrm{tr}\Sigma)^{-1}(\mathrm{tr}V_N))(\mathrm{tr}V_N)^{-1})\right\}+\frac{(\mathrm{tr}\Sigma)^{-1}}{n_c^2} \tag{3.41}$$
$$\times\mathrm{E}\left\{S_N^{*'}S_N^*(N-(\mathrm{tr}\Sigma)^{-1}Q_N)^2\right\}.$$

For the first term in (3.41), by $U_c=N-n_c-\frac{1}{2}\{(\mathrm{tr}\Sigma)^{-1}Q_N-N\}+\xi_N$ and Hölder inequality with the assumption $\mathrm{E}(X'X)^{4+\epsilon}<\infty$ for some $\epsilon>0$, we can get that the first term of the R.H.S. in (3.41) is $o(1)$. By Lemma 3.1, we have

$$\mathrm{E}S_N^{*'}S_N^*(N-(\mathrm{tr}\Sigma)^{-1}Q_N)^2=\mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)^2\mathrm{E}\textstyle\sum_{\alpha=1}^N S_{\alpha-1}^{*'}S_{\alpha-1}^*$$
$$+4\mathrm{E}\textstyle\sum_{\alpha=1}^N \mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)(X'S_{\alpha-1}^*)(\alpha-1-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1})$$
$$+2\mathrm{E}\textstyle\sum_{\alpha=1}^N \mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)^2(X'S_{\alpha-1}^*)+(\mathrm{tr}\Sigma)\mathrm{E}\sum_{\alpha=1}^N(\alpha-1 \tag{3.42}$$
$$-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1})^2+2\mathrm{E}(X'X)(1-(\mathrm{tr}\Sigma)^{-1}X'X)$$
$$\times\mathrm{E}\textstyle\sum_{\alpha=1}^N(\alpha-1-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1})+\mathrm{E}(X'X)(1-(\mathrm{tr}\Sigma)^{-1}X'X)^2\mathrm{E}(N).$$

Here we give the lemma.

Lemma 3.7. For fixed $p\times 1$ vector a, we have the following formulas.

$$\frac{1}{n_c^2}\mathrm{E}\sum_{\alpha=1}^N S_{\alpha-1}^{*'}S_{\alpha-1}^*=\frac{1}{2}\mathrm{E}(X'X)+o(1). \tag{3.43}$$

$$\frac{1}{n_c^2}\mathrm{E}\sum_{\alpha=1}^N a'S_{\alpha-1}^*(\alpha-1-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1})=\frac{1}{2}\mathrm{E}a'X$$
$$\times(1-(\mathrm{tr}\Sigma)^{-1}X'X)+o(1). \tag{3.44}$$

$$\frac{1}{n_c^2}\mathrm{E}a'\sum_{\alpha=1}^N S_{\alpha-1}^*=\frac{1}{n_c^2}\mathrm{E}a'NS_N^*=o(1) \tag{3.45}$$

$$\frac{1}{n_c^2}\mathrm{E}\sum_{\alpha=1}^N(\alpha-1-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1})^2=\frac{1}{2}\mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)^2+o(1), \tag{3.46}$$

$$\frac{1}{n_c^2}\mathrm{E}\sum_{\alpha=1}^N(\alpha-1-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1})=o(1). \tag{3.47}$$

Proof. By Lemma 3.1, we have

$$\mathrm{E}NS_N^{*'}S_N^*=\mathrm{E}(X'X)\mathrm{E}\frac{1}{2}N(N+1)+\mathrm{E}\sum_{\alpha=1}^N S_{\alpha-1}^{*'}S_{\alpha-1}^*.$$

Then we have the first formula in Lemma 3.7. For the next formula, we have $\frac{1}{n_c^2}\mathrm{E}\sum_{\alpha=1}^{N} a'S_{\alpha-1}^{*}$ $\times(\alpha-1-(\mathrm{tr}\Sigma)^{-1}Q_{\alpha-1}) = \frac{1}{n_c^2}\mathrm{E}Na'S_N^*(N-(\mathrm{tr}\Sigma)^{-1}Q_N) - \frac{1}{2}\mathrm{E}a'X(1-(\mathrm{tr}\Sigma)^{-1}X'X) + o(1)$. Let $Y_i = (a'X_i, 1-(\mathrm{tr}\Sigma)^{-1}X_i'X_i)'$ $(i=1,2,\cdots)$, then we have that the limiting distribution of $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$ has a normal with covariance $\mathrm{E}a'X(1-(\mathrm{tr}\Sigma)^{-1}X'X)$. Thus by Hölder inequality, we have the desired one. For third one, $\frac{1}{n_c^2}\mathrm{E}a'\sum_{\alpha=1}^{N} S_{\alpha-1}^{*} = \frac{1}{n_c^2}\mathrm{E}a'NS_N^* = o(1)$. Thus we have Lemma 3.7. Thus we have

$$\begin{aligned}\frac{1}{n_c^2}\,\mathrm{E}S_N^{*'}S_N^{*'}(N-(\mathrm{tr}\Sigma)^{-1}Q_N)^2 &= (\mathrm{E}X'X)\mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)^2\\ &+2\left\{\mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)X'\right\}\left\{\mathrm{E}X(1-(\mathrm{tr}\Sigma)^{-1}X'X)\right\}+o(1).\end{aligned}\tag{3.48}$$

Therefore we have

$$II_b = 3(\mathrm{tr}\Sigma)^{-1}\{\mathrm{E}(1-(\mathrm{tr}\Sigma)^{-1}X'X)X'\}\{\mathrm{E}X(1-(\mathrm{tr}\Sigma)^{-1}X'X)\}+1+o(1).$$

Thus from (3.33), we have

$$II = 2(\mathrm{tr}\Sigma)^{-3}\mathrm{E}(X'XX')\mathrm{E}(XX'X)+2(\mathrm{tr}\Sigma)^{-2}\mathrm{tr}\Sigma^2+1+o(1).$$

Hence the regret $\omega(c)$ is given by

$$\begin{aligned}\omega(c) &= c\,(I+II+2\mathrm{E}(N-n_c))\\ &= c\,\big(2(\mathrm{tr}\Sigma)^{-3}\{\mathrm{E}(X'XX')\}\{\mathrm{E}(XX'X)\}+2(\mathrm{tr}\Sigma^2)(\mathrm{tr}\Sigma)^{-2}\\ &-\frac{3}{4}(\mathrm{tr}\Sigma)^{-2}\mathrm{Var}(X'X)\Big)+o(c).\end{aligned}\tag{3.49}$$

Theorem 3.1. If $\mathrm{E}\left((X-\mu)'(X-\mu)\right)^{4+\epsilon}<\infty$ for some $\epsilon>0$ and there is positive number δ such that $\delta c^{-1/4}\le m = o(c^{-1/2})$, the regret $\omega(c)$ is given by

$$\begin{aligned}\omega(c) = \;& c\left\{2(\mathrm{tr}\Sigma)^{-3}\{\mathrm{E}(X-\mu)'(X-\mu)(X-\mu)'\}\{\mathrm{E}(X-\mu)\right.\\ &\times(X-\mu)'(X-\mu)\}+2(\mathrm{tr}\Sigma)^{-2}\mathrm{tr}\Sigma^2\\ &\left.-\frac{3}{4}(\mathrm{tr}\Sigma)^{-2}\mathrm{Var}((X-\mu)'(X-\mu))\right\}+o(c).\end{aligned}\tag{3.50}$$

Instead of (1.3), we define another stopping time

$$N^* = \inf\left\{n\ge m \;\mid\; n\ge \ell_n\sqrt{\frac{\mathrm{tr}S_n}{c}}\right\},$$

where $\ell_n = 1+\frac{\ell_0}{n}+o(n^{-1})$. Let $\omega^*(c)$ be the regret for N^*. Then by similar caluculation, we have

$$\mathrm{E}(N^*-n_c) = \mathrm{E}(N-n_c)+\ell_0+o(1)$$

and

$$\omega^*(c) = \omega(c)+o(c).$$

When we use unbiased estimator $V_n/(n-1)$ of Σ, we have $\ell_0 = -\frac{1}{2}$. From this comment, when we adapt unbiased estimator instead of sample covariance matrix for Σ, we find that the method deceases the sample size, but increases MSE since both methods asymptotically have the same regret.

4. Applications

We here only explain the case of a p-variate normal distribution. Then we have

$$\frac{c}{2p} + o(c) \leq \omega(c) = \frac{c}{2}\frac{\mathrm{tr}\Sigma^2}{(\mathrm{tr}\Sigma)^2} + o(c) \leq \frac{c}{2} + o(c).$$

Thus for any p and covariance matrix Σ the regret is bounded by amounts independent of Σ as $c \to 0$.

This theorem contains the results of many models in a multivariate analysis.
Khan [7] and Rohatagi and O'Neill [13] considered the problem of the mean vector in a p-variate normal distribution and they assumed that the covariance matrix $\Sigma = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_p^2)$. Then from the above Theorem, the regret is given by

$$\omega(c) = \frac{c}{2}\frac{\sum_{i=1}^p \sigma_i^4}{(\sum_{i=1}^p \sigma_i^2)^2} + o(c).$$

Next we consider the intraclass correlation model under the p-variate normal distribution. Then the covariance matrix is given by, with $-\frac{1}{p-1} < \rho < 1$, $\Sigma = \sigma^2[(1-\rho)I_p + \rho(1, \cdots, 1)'(1, \cdots, 1)]$. Then the regret is

$$\omega(c) = c(\frac{1 + \overline{p-1}\rho^2}{2p}) + o(c).$$

From this expression, we find that the second approximation of $\omega(c)$ does not depend on σ^2.

Finally we consider a multivariate t-distribution with k degrees of freedom as an error distribution. Let U and V be a random vector and variable, respectively, which are independent. The distribution of U is a p-variate normal distribution with mean zero and covariance matrix Σ and V has chi-square distribution with k degrees of freedom. The distribution of $Z = U/\sqrt{V/k}$ is called a multivariate t-distribution with k degrees of freedom. See Anderson (1984). Then

$$\mathrm{Var}(Z'Z) = 2(\frac{k}{k-2})^2\{\frac{k-2}{k-4}\mathrm{tr}\Sigma^2 + \frac{1}{k-4}(\mathrm{tr}\Sigma)^2\}.$$

Then we have, since $\mathrm{E}(Z'ZZ') = 0$,

$$\begin{aligned}\omega(c) = \quad & c\{\frac{(k-10)}{2(k-4)}\frac{\mathrm{tr}\Sigma^2}{(\mathrm{tr}\Sigma)^2} - \frac{3}{2(k-4)}\} + o(c) \\ \leq & \frac{c(k-13)}{2(k-4)} + o(c) \quad \text{ if } \ 10 < k < 13.\end{aligned}$$

Also we have when $4 < k \leq 10$,

$$\omega(c) \leq \frac{c}{2(k-4)}\{\frac{(k-10)}{p} - 3\} + o(c) = \frac{c}{2(k-4)}(\frac{k-10-3p}{p}) + o(c).$$

Therefore if we choose k as $4 < k < 13$, the regret is asymptotically negative for all p. This shows that the sequential consideration is better even if a covariance matrix is known.

When $p = 1$, Martinsek [9] dealt with the negative regret through numerical examples. Also when $p = 1$, Ghosh, Mukhopadhyay and Sen [5] claimed that the regret may be negative with t-distribition with $k \leq 7$.

Acknowledgement. The author is grateful to the referee for his careful reading.

References

[1] Anderson, T.W.(1984) An introduction to Multivariate Statistical Analysis. 2nd Edition. John Wiley & Sons, New York.

[2] Chow, Y.S. and Martinsek, A.T.(1982) Bounded regret of a sequential procedure for estimation of the mean. Ann. Statist., 10, 909-914.

[3] Chow, Y.S., Robbins, H. and Teicher, H.(1965). Moments of randomly stopped sums. Ann. Math. Statist. 36, 789-799.

[4] Ghosh, M. and Mukhopadhyay, N.(1979). Sequential point estimation of the mean when the distribution is unspecified. Commun. Statist. Theory Methods A., 8, 637-652.

[5] Ghosh, M., Mukhopadhyay, N. and Sen, P.K.(1997). Sequential Estimation. John Wiley & Sons, New York.

[6] Ghosh, M., Sinha, B.K. and Mukhopadhyay, N.(1976). Multivariate sequential point estimation. Jour. Multivar. Anal., 6, 281-294.

[7] Khan, R.A.(1968). Sequential estimation of the mean vector of a multivariate normal distribution. Sankhyā Ser. A. 30, 331-342.

[8] Martinsek, A.T.(1983). Second order approximation to the risk of a sequential procedure. Ann. Statist., 11, 827-836; Correction (1986), Ann. Statist., 14, 359.

[9] Martinsek, A.T.(1988). Negative regret, optional stopping, and the elimination of outliers. Jour. Amer. Statist. Assoc. 83, 160-163.

[10] Nagao, H.(2002) Sequential estimations of some vector in linear regression model under a linex loss. Scien. Math. Japon., 56, 493-499.

[11] Nagao, H. and Srivastava, M.S.(2002). Fixed width confidence region for the mean of a multivariate normal distribution. Jour. Multivar. Anal., 81, 259-273.

[12] Robbins, H.(1959). Sequential estimation of the mean of a normal population. In Probability and Statistics, H. Cramér Volume (ed. by U. Grenander), 235-245. Uppsala: Almquist and Wiksell.

[13] Rohatgi, V.K. and O'Neill, R.T.(1973). On sequential estimation of the mean of a multinormal population. Ann. Inst. Statist. Math., 25, 321-325.

[14] Starr, N.(1966). On the asymptotic efficiency of a sequential procedure for estimating the mean. Ann. Math. Statist., 36, 1173-1185.

[15] Starr, N. and Woodroofe, M.(1969). Remarks on sequential point estimation. Proc. Nat. Acad. Sci. USA, 63, 285-288.

[16] Starr, N. and Woodroofe, M.(1972). Further remarks on sequential estimation: The exponential case. Ann. Math. Statist., 43, 1147-1154.

[17] Takada, Y. and Nagao, H. (2001). Sequential point estimation of a multivariate normal mean vector under a linex loss function. Submitted for publication.

[18] Wang, Y.H.(1980). Sequential estimation of the mean of a multinormal population. Jour. Amer. Statist. Assoc., 75, 977-983.

[19] Woodroofe, M.(1977). Second order approximations for sequential point and interval estimation. Ann. Statist., 5, 984-995.

Department of mathematical sciences
Graduate school of Engineering
Osaka Prefecture University
Sakai, Osaka Japan