Estimating A Covariance Matrix in MANOVA Model

Hisayuki Hara

METR 98-07 September 1998

Communicated by Professor C.Hirotsu

Estimating A Covariance Matrix in MANOVA Model

Hisayuki Hara *

Department of Mathematical Engineering and Information physics

University of Tokyo

Summary

On the estimation of the covariance matrix in the framework of multivariate analysis of variance(MANOVA) model, Sinha and Ghosh(1987) proposed a Stein type truncated estimater improving on the uniformly minimum variance unbiased(UMVU) estimator under entropy loss. However the estimator is discontinuous. This article obtains some other continuous estimators which dominate the UMVU estimator and, furthermore, one of which is shown analytically to improve on Sinha-Ghosh's estimator.

Key words and phrases: Conditional risk difference, continuity, entropy loss, inadmissibility, Wishart matrix.

AMS 1985 subject classifications: primary 62F10; secondary 62H99.

*Address: Department of Mathematical Engineering and Information Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

1 Introduction

Consider the MANOVA model,

$$\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k) : \ \ p \times k \sim \mathrm{N}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}_k),$$

$$\boldsymbol{S} : \ \ p \times p \sim \mathrm{W}_p(n, \boldsymbol{\Sigma}) \quad (n \geq p), \tag{1}$$

$$\boldsymbol{X} \text{ and } \boldsymbol{S} \text{ are independent,}$$

where $\mathrm{N}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}_k)$ denotes that $\boldsymbol{X}_i$ are independently distributed to the multivariate normal distribution with mean $\boldsymbol{\mu}_i$, the corresponding column of $\boldsymbol{M} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k)$, and covariance matrix $\boldsymbol{\Sigma}$. $\mathrm{W}_p(n, \boldsymbol{\Sigma})$ denotes the Wishart distribution with the parameter $\boldsymbol{\Sigma}$ and degrees of freedom n. Let $\boldsymbol{M}$ and $\boldsymbol{\Sigma}$ be unknown. Throughout this paper, for two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of the same order, let $\boldsymbol{A} \geq \boldsymbol{B}$ implies that $\boldsymbol{A} - \boldsymbol{B}$ is nonnegative definite.

Here we treat the estimation of $\boldsymbol{\Sigma}$ under the entropy loss,

$$\mathrm{L}(\hat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \mathrm{tr}(\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}) - \log|\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}| - p. \tag{2}$$

This problem remains invariant under the full affine group acting as

$$(\boldsymbol{X}, \boldsymbol{S}) \rightarrow (\boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}, \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}'), \quad (\boldsymbol{M}, \boldsymbol{\Sigma}) \rightarrow (\boldsymbol{A}\boldsymbol{M} + \boldsymbol{B}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}')$$

for $p \times p$ nonsingular matrix $\boldsymbol{A}$ and $p \times k$ matrix $\boldsymbol{B}$. The unbiased estimator

$$\delta_0 = \frac{1}{n}\boldsymbol{S} \tag{3}$$

is known to be the best affine equivariant and the uniformly minimum variance unbiased estimator for $\boldsymbol{\Sigma}$.

In the univariate case, where $p = 1$,

$$X_i \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\mu_i, \sigma^2), \quad S \sim \sigma^2\chi^2,$$

Stein(1964) presented a truncated estimator for dominating S/n by using the information of $\boldsymbol{X} = (X_1, \ldots, X_p)$,

$$\delta^{ST} = \min\left\{ \frac{S}{n}, \quad \frac{S + \|\boldsymbol{X}\|^2}{n+k} \right\}.$$

As a multivariate extention, Sinha and Ghosh(1987) derived the Stein type truncated estimator

$$\delta_k^{SG} = \begin{cases} \dfrac{\boldsymbol{S} + \boldsymbol{X}\boldsymbol{X}'}{n+k}, & \text{if} \quad \dfrac{\boldsymbol{S} + \boldsymbol{X}\boldsymbol{X}'}{n+k} \le \dfrac{\boldsymbol{S}}{n} \\ \\ \dfrac{\boldsymbol{S}}{n}, & \text{otherwise.} \end{cases} \tag{4}$$

which improves δ_0 under the loss (2). However δ_k^{SG} is discontinuous where the largest eigenvalue of $\boldsymbol{X}\boldsymbol{X}' - (k/n)\boldsymbol{S}$ equals to 0.

For $k = 1$ the estimator corresponding to (4) is

$$\delta_1^{SG} = \begin{cases} \dfrac{1}{n}\boldsymbol{S}, & \text{if } \boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} \ge \dfrac{1}{n} \\ \\ \dfrac{1}{n+1}(\boldsymbol{S} + \boldsymbol{X}\boldsymbol{X}'), & \text{otherwise.} \end{cases} \tag{5}$$

Perron(1990) proposed the continuous estimator

$$\delta^{PR} = \begin{cases} \dfrac{1}{n+1}(\boldsymbol{S} + \dfrac{1}{n}\dfrac{1}{\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}}\boldsymbol{X}\boldsymbol{X}'), & \text{if } \boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} \ge \dfrac{1}{n} \\ \\ \dfrac{1}{n+1}(\boldsymbol{S} + \boldsymbol{X}\boldsymbol{X}'), & \text{otherwise.} \end{cases} \tag{6}$$

which dominates δ_1^{SG} under (2). Kubokawa et al.(1992) and Kubokawa et al.(1993) proposed the estimator,

$$\delta^{EB} = \begin{cases} \dfrac{1}{n+1}(\boldsymbol{S} + \dfrac{p}{n+1-p}\dfrac{1}{\boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}}\boldsymbol{X}\boldsymbol{X}'), & \text{if } \boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X} \le \dfrac{p}{n+1-p} \\ \\ \dfrac{1}{n+1}(\boldsymbol{S} + \boldsymbol{X}\boldsymbol{X}'), & \text{otherwise.} \end{cases} \tag{7}$$

δ^{EB} was also shown to dominate δ_1^{SG} under (2). They also pointed out that δ^{EB} can be interpreted as an empirical Bayes estimator.

In this paper we consider the extention of Perron and Kubokawa's result for $k = 1$ to $k \geq 2$. In section 2, we will present the Perron type continuous estimator which dominates δ_k^{SG} and a sufficient condition to improve δ_0 is given. In section 3, we obtain another estimator which is similar to (7) and also improves δ_0. Section 4 gives the simulation results about the risk performance of the estimators proposed here against that of δ_k^{SG} and δ_0.

2 Perron type continuous estimator

Let $\boldsymbol{T}$ be any $p \times p$ matrix satisfying $\boldsymbol{T}\boldsymbol{T}' = \boldsymbol{S}$ and the spectrum decomposition of $\boldsymbol{T}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{T}'^{-1}$ be expressed as $\boldsymbol{T}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{T}'^{-1} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, $\lambda_1 \geq \cdots \geq \lambda_p$ and $\boldsymbol{H}$ is an orthogonal matrix.

Now we consider the following estimator,

$$\delta^{TR} = \frac{1}{n+k}(\boldsymbol{S} + \boldsymbol{T}\boldsymbol{H}\boldsymbol{\Lambda}^{TR}\boldsymbol{H}'\boldsymbol{T}') \tag{8}$$

where $\boldsymbol{\Lambda}^{TR} = \mathrm{diag}\{\min(k/n, \lambda_i)\}$. This estimator is continuous. It is easy to show that this estimator for $k = 1$ coincides with (6) using the fact that the only eigenvalue of $\boldsymbol{T}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{T}'^{-1}$ is $\mathrm{tr}\boldsymbol{T}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{T}'^{-1} = \boldsymbol{X}'\boldsymbol{S}^{-1}\boldsymbol{X}$. Then we can say that (8) for $k \geq 2$ is an extention of (6).

First we consider to show the dominance of δ^{TR} over δ_0 under the criterion (2).

Let $\boldsymbol{W}$ be $p \times p$ random matrix whose probability density with respect to the Lebesgue measure $\mathrm{d}\boldsymbol{W}$ is

$$\mathrm{Const.}|\boldsymbol{W}\boldsymbol{W}'|^{(n-p)/2} \exp\left(-\frac{1}{2}\mathrm{tr}\boldsymbol{W}\boldsymbol{W}'\right). \tag{9}$$

It is easy to show that both $\boldsymbol{W}\boldsymbol{W}'$ and $\boldsymbol{W}'\boldsymbol{W}$ follow $\mathrm{W}_p(n, \boldsymbol{I}_p)$. We can set $\boldsymbol{W}\boldsymbol{W}' = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{S}\boldsymbol{\Sigma}'^{-1/2} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{T}\boldsymbol{T}'\boldsymbol{\Sigma}'^{-1/2}$, where $\boldsymbol{\Sigma}^{-1/2}$ is a $p \times p$ constatnt matrix which satisfies $\boldsymbol{\Sigma}^{-1/2'}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1}$. Letting $\boldsymbol{P}$ be $\boldsymbol{P} = \boldsymbol{T}^{-1}\boldsymbol{\Sigma}^{1/2}\boldsymbol{W}$ with

$\boldsymbol{\Sigma}^{1/2} = (\boldsymbol{\Sigma}^{-1/2})^{-1}$, we can rewrite $\boldsymbol{W}$ as

$$\boldsymbol{W} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{T}\boldsymbol{P}. \tag{10}$$

We can easily see that $\boldsymbol{P}$ is an orthogonal matrix. Then we have

$$\begin{aligned}
\mathrm{tr}\delta^{TR}\boldsymbol{\Sigma}^{-1} &= \frac{1}{n+k}\mathrm{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{T}\boldsymbol{P}\boldsymbol{P}'\boldsymbol{H}(\boldsymbol{I}_p + \boldsymbol{\Lambda}^{TR})\boldsymbol{H}'\boldsymbol{P}\boldsymbol{P}'\boldsymbol{T} \\
&= \frac{1}{n+k}\mathrm{tr}\boldsymbol{W}'\boldsymbol{W}\boldsymbol{P}'\boldsymbol{H}(\boldsymbol{I}_p + \boldsymbol{\Lambda}^{TR})\boldsymbol{H}'\boldsymbol{P},
\end{aligned}$$

$$\mathrm{tr}\delta_0\boldsymbol{\Sigma}^{-1} = n^{-1}\mathrm{tr}\boldsymbol{S}\boldsymbol{\Sigma}^{-1} = n^{-1}\mathrm{tr}\boldsymbol{W}'\boldsymbol{W}.$$

Let $\tilde{\boldsymbol{X}}$ be $\tilde{\boldsymbol{X}} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{X}$. We note that the spectrum decomposition of $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}'\boldsymbol{W}'^{-1}$ can be expressed as

$$\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}'\boldsymbol{W}'^{-1} = \boldsymbol{P}'\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'\boldsymbol{P},$$

which indicates that $\boldsymbol{P}'\boldsymbol{H}(\boldsymbol{I}_p + \boldsymbol{\Lambda}^{TR})\boldsymbol{H}'\boldsymbol{P}$ is constant with $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ fixed.

Calculating the conditional risk difference between δ^{TR} and δ_0 with $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ given,

$$\begin{aligned}
\mathrm{E}[\mathrm{tr}(\delta^{TR} - \delta_0)\boldsymbol{\Sigma}^{-1}|\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}] &= (n+k)^{-1}\mathrm{E}\left[\mathrm{tr}\boldsymbol{W}'\boldsymbol{W}\boldsymbol{P}'\boldsymbol{H}(\boldsymbol{\Lambda}^{TR} - (k/n)\boldsymbol{I}_p)\boldsymbol{H}'\boldsymbol{P}|\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}\right] \\
&= (n+k)^{-1}\mathrm{tr}\left\{\mathrm{E}[\boldsymbol{W}'\boldsymbol{W}|\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}]\boldsymbol{P}'\boldsymbol{H}(\boldsymbol{\Lambda}^{TR} - (k/n)\boldsymbol{I}_p)\boldsymbol{H}'\boldsymbol{P}\right\} \\
&\leq \mathrm{tr}\boldsymbol{H}'\boldsymbol{P}(\boldsymbol{I}_p + \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}'\boldsymbol{W}'^{-1})^{-1}\boldsymbol{P}'\boldsymbol{H}(\boldsymbol{\Lambda}^{TR} - (k/n)\boldsymbol{I}_p) \\
&= \mathrm{tr}(\boldsymbol{I}_p + \boldsymbol{\Lambda})^{-1}(\boldsymbol{\Lambda}^{TR} - (k/n)\boldsymbol{I}_p) \\
&= \sum_{i=1}^{p}\min\left(\frac{\lambda_i - k/n}{1+\lambda_i},\ 0\right).
\end{aligned} \tag{11}$$

The inequality in (11) is by Lemma 4 in Sinha and Ghosh(1987),

$$\mathrm{E}[\boldsymbol{W}'\boldsymbol{W}|\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}] \geq (n+k)(\boldsymbol{I}_p + \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}'\boldsymbol{W}'^{-1})^{-1}$$

and the non-positive definiteness of $\boldsymbol{\Lambda}^{TR} - (k/n)\boldsymbol{I}_p$.

On the other hand,

$$\begin{aligned}
\log|\delta^{TR}\boldsymbol{\Sigma}^{-1}| - \log|\delta_0\boldsymbol{\Sigma}^{-1}| &= \log\left\{|\boldsymbol{W}'\boldsymbol{W}|\prod_{i=1}^{p}\frac{\min(1+\lambda_i, 1+k/n)}{n+k}\right\} - \log\left|\frac{\boldsymbol{S}\boldsymbol{\Sigma}^{-1}}{n}\right| \\
&= \sum_{i=1}^{p}\log\left(\frac{\min(n+n\lambda_i, n+k)}{n+k}\right) \\
&= \sum_{i=1}^{p}\log\left(\min\left(\frac{n(1+\lambda_i)}{n+k}, 1\right)\right).
\end{aligned} \tag{12}$$

Combining (11) and (12),

$$\begin{aligned}
&\mathrm{E}[\mathrm{L}(\delta^{TR}, \boldsymbol{\Sigma}) - \mathrm{L}(\delta_0, \boldsymbol{\Sigma}) | \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}] \\
&\quad \le \sum_{i=1}^{p} \left\{ \min\left(0, \frac{\lambda_i - k/n}{1+\lambda_i}\right) - \log\left(\min\left(\frac{n(1+\lambda_i)}{n+k}, 1\right)\right)\right\} \qquad (13) \\
&\quad = \sum_{i=1}^{p} R_i, \qquad (14)
\end{aligned}$$

where R_i is

$$R_i = \begin{cases} 0, & \lambda_i \ge k/n \\ \dfrac{\lambda_i - k/n}{1+\lambda_i} - \log\left(\dfrac{n(1+\lambda_i)}{n+k}\right) & \lambda_i < k/n \end{cases}$$

With respect to R_i such that $\lambda_i < k/n$,

$$\frac{\mathrm{d}R_i}{\mathrm{d}\lambda_i} = \frac{1+k/n}{(1+\lambda_i)^2} - \frac{1}{1+\lambda_i} > 0.$$

Since $R_i = 0$ at $\lambda_i = k/n$, we can say that $\mathrm{E}[\mathrm{L}(\delta^{TR}, \boldsymbol{\Sigma}) - \mathrm{L}(\delta_0, \boldsymbol{\Sigma}) | \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}] \le 0$ for any $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$.

Proposition 2.1 *δ^{TR} dominates δ_0 under (2).*

Next we consider the inadmissibility of δ_k^{SG}. With respect to $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ such that δ_k^{SG} takes $(\boldsymbol{S}+\boldsymbol{X}\boldsymbol{X}')/(n+k)$, δ^{TR} also takes $(\boldsymbol{S}+\boldsymbol{X}\boldsymbol{X}')/(n+k)$. Since δ^{TR} improves δ_0 under the conditional risk for any $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$, the inadmissibility of δ_k^{SG} is proved.

Theorem 2.1 *δ^{TR} dominates δ_k^{SG}.*

Using the above argument, we can also obtain a sufficient condition to improve δ_0. We consider the class of estimators as follows,

$$\delta_\phi = \boldsymbol{T}' \left(\frac{\boldsymbol{I}_p + \boldsymbol{H}\boldsymbol{\Lambda}^{\phi}\boldsymbol{H}'}{n+k}\right)\boldsymbol{T} \qquad (15)$$

where $\boldsymbol{\Lambda}_{\phi_i} = \mathrm{diag}\{\phi_i(\boldsymbol{\lambda})\}$ and $\boldsymbol{\lambda} = (\lambda_1, \dots \lambda_p)'$. We suppose that $\phi_i(\boldsymbol{\lambda}) \le k/n$ for all i. In the same way as (13), we have

$$\mathrm{E}[\mathrm{L}(\delta_\phi, \boldsymbol{\Sigma}) - \mathrm{L}(\delta_0, \boldsymbol{\Sigma}) | \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}] \le \sum_{i=1}^{p} \left\{ \frac{\phi_i(\boldsymbol{\lambda}) - k/n}{1+\lambda_i} - \log\left(\frac{n(1+\phi_i(\boldsymbol{\lambda}))}{n+k}\right)\right\} \qquad (16)$$

Therefore we get the following theorem.

Theorem 2.2 *If $\phi_i(\boldsymbol{\lambda})$ satisties*

1. $\phi_i(\boldsymbol{\lambda}) \le k/n$
2.
$$\frac{\phi_i(\boldsymbol{\lambda}) - k/n}{1+\lambda_i} - \log\left(\frac{n(1+\phi_i(\boldsymbol{\lambda}))}{n+k}\right) \le 0 \quad \text{for all } i,$$

δ_ϕ *improves* δ_0.

If we put $\phi_i(\boldsymbol{\lambda})$ as

$$\phi_i(\boldsymbol{\lambda}) = \begin{cases} \lambda_i, & \text{if} \ \ \lambda_1 \le k/n \\ \\ \dfrac{k}{n}, & \text{otherwise,} \end{cases}$$

we can easily see that δ_k^{SG} also satisties this condition and their discontinuity of ϕ_i.

3 Another Class of Improved estimators

In this section we give another class of estimators which dominates δ_0. We consider here the class of estimators,

$$\delta_{a,b} = a\boldsymbol{S} + b\boldsymbol{G}, \quad \boldsymbol{G} = \textstyle\sum_{i=1}^k \dfrac{\boldsymbol{X}_i\boldsymbol{X}_i'}{\boldsymbol{X}_i'\boldsymbol{S}^{-1}\boldsymbol{X}_i}, \quad a > 0, b \ge 0. \tag{17}$$

We calculate the risk of this estimator to find a suitable value of a and b by the argument in Shorrock and Zidek(1974).

Write $\tilde{\boldsymbol{X}}_i = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{X}_i$, $\tilde{\boldsymbol{S}} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{S}\boldsymbol{\Sigma}^{-1/2}$. Set $\boldsymbol{U}_i = \boldsymbol{Q}_i\tilde{\boldsymbol{S}}\boldsymbol{Q}_i'$ for an orthogonal matrix $\boldsymbol{Q}_i$ such that $\boldsymbol{Q}_i\tilde{\boldsymbol{X}}_i = (\|\tilde{\boldsymbol{X}}_i\|, 0, \ldots, 0)'$. The conditional distribution of $\boldsymbol{U}_i$ given $\tilde{\boldsymbol{X}}_i$ is $\mathrm{W}(n, \boldsymbol{I}_p)$, which remains independent of $\tilde{\boldsymbol{X}}_i$. Therefore $\boldsymbol{U}_i$ and $\tilde{\boldsymbol{X}}_i$ are mutually independent. Define $\tilde{x}_i$, v_i, $\boldsymbol{R}_i$ by

$$\tilde{x}_i = \|\tilde{\boldsymbol{X}}_i\|^2, \quad v_i = u_{11,i} - \boldsymbol{U}_{12,i}\boldsymbol{U}_{22,i}^{-1}\boldsymbol{U}_{21,i}, \quad \boldsymbol{R}_i = \boldsymbol{U}_i - v_i\boldsymbol{E}_{11},$$

where $\boldsymbol{U}_i$ is partitioned as

$$\boldsymbol{U}_i = \begin{pmatrix} u_{11.i} & \boldsymbol{U}_{12,i} \\ \boldsymbol{U}_{21,i} & \boldsymbol{U}_{22,i} \end{pmatrix}$$

with a 1×1 matrix $u_{11,i}$ and $\boldsymbol{E}_{11}$ is a $p \times p$ matrix such that $(1,1)$ element is 1 and the other elements are all 0. Then $\tilde{x}_i$, u_i and $\boldsymbol{R}_i$ are mutually independent and $v_i \sim \chi^2_{n-p+1}$. Leting η_i, $i = 1, \ldots, p$, be the eigenvalues of $\boldsymbol{T}'^{-1}\boldsymbol{G}\boldsymbol{T}^{-1}$, we have

$$\begin{aligned} \mathrm{tr}\delta_{a,b}\boldsymbol{\Sigma}^{-1} &= \mathrm{tr}\left(a\tilde{\boldsymbol{S}} + b\sum_{i=1}^{k} \frac{\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{X}}_i'}{\tilde{\boldsymbol{X}}_i'\tilde{\boldsymbol{S}}^{-1}\tilde{\boldsymbol{X}}_i}\right) \\ &= a\mathrm{tr}\tilde{\boldsymbol{S}} + b\sum_{i=1}^{k} v_i. \end{aligned} \tag{18}$$

$$|\delta_{a,b}\boldsymbol{\Sigma}^{-1}| = |\boldsymbol{S}\boldsymbol{\Sigma}^{-1}| \prod_{i=1}^{k}(a + b\eta_i). \tag{19}$$

Noting that

$$\begin{aligned} \mathrm{tr}\boldsymbol{T}'^{-1}\boldsymbol{G}\boldsymbol{T}^{-1} &= \mathrm{tr}\left(\sum_{i=1}^{k} \frac{\boldsymbol{T}'^{-1}\boldsymbol{X}_i\boldsymbol{X}_i'\boldsymbol{T}^{-1}}{\boldsymbol{X}_i'\boldsymbol{S}^{-1}\boldsymbol{X}_i}\right) \\ &= k \end{aligned}$$

and $\eta_i \geq 0$ for all i, we have

$$|\delta_{a,b}\boldsymbol{\Sigma}^{-1}| \geq |\boldsymbol{S}\boldsymbol{\Sigma}^{-1}|a^{k-1}(a + kb).$$

As a result we get

$$\mathrm{E}[\mathrm{L}(\delta_{a,b}, \boldsymbol{\Sigma})] \leq apn + bk(n-p+1) - (k-1)\log a - \log(a+kb) - \mathrm{E}[\log|\boldsymbol{S}\boldsymbol{\Sigma}^{-1}|] - p$$

According to proposition 2.1 in Kubokawa et al.(1993), if we choose (a, b) as,

$$(a_0, b_0) = \left(\frac{1}{n+1}, \frac{p}{(n+1)(n-p+1)k}\right), \tag{20}$$

the risk difference between δ_{a_0,b_0} and δ_0 is

$$\mathrm{E}[\mathrm{L}(\delta_{a_0,b_0}, \boldsymbol{I}_p) - \mathrm{L}(\delta_0, \boldsymbol{I}_p)] \leq \log\left[\left(\frac{n+1}{n}\right)^p \frac{n+1-p}{n+1}\right] < 0.$$

Then we can obtain the following proposition.

Proposition 3.1 δ_{a_0,b_0} *dominates* δ_0.

With the argument in the previous section, we can also get the following results.

Theorem 3.1

$$\delta_{a_0,b_0}^{TR} = \begin{cases} \dfrac{\boldsymbol{S}+\boldsymbol{X}\boldsymbol{X}'}{n+k}, & \text{if } \dfrac{\boldsymbol{S}+\boldsymbol{X}\boldsymbol{X}'}{n+k} \leq \delta_{a_0,b_0}, \\ \\ \delta_{a_0,b_0}, & \text{otherwise}, \end{cases} \tag{21}$$

dominates δ_{a_0,b_0}.

The estimator δ_{a_0,b_0}^{TR} is also discontinuous. So it seems to be not still preferable. But we can present a continuous estimator which dominates $\delta_{a,b}$ by the argument in the derivation of Theorem 2.1. We give the proof of Theorem 3.1 simultaniously.

Let $\boldsymbol{\psi}_0$, $\boldsymbol{\psi}$, $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{\Lambda}}$ be defined as

$$\boldsymbol{\psi}_0 = \frac{\boldsymbol{I}_p + \boldsymbol{T}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{T}'^{-1}}{n+k} = \tilde{\boldsymbol{T}}\tilde{\boldsymbol{T}},'$$

$$\boldsymbol{\psi} = \frac{1}{n+1}\left(\boldsymbol{I}_p + \frac{p}{(n-p+1)k}\boldsymbol{T}^{-1}\boldsymbol{G}\boldsymbol{T}'^{-1}\right) = \tilde{\boldsymbol{T}}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{T}}',$$

$$\tilde{\boldsymbol{\Lambda}} = \mathrm{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p), \quad \tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_p$$

where $\tilde{\lambda}_i$, $i = 1, \ldots, p$, are the characteristic roots of $\boldsymbol{\psi}$ in the metric of $\boldsymbol{\psi}_0$, that is, the roots of

$$|\boldsymbol{\psi} - \tilde{\lambda}_i \boldsymbol{\psi}_0| = 0,$$

and $\tilde{\boldsymbol{T}}$ consists of the corresponding characteristic vector. We consider the class of estimator with a function $\phi_i(\cdot)$,

$$\delta_{a_0,b_0}^{\phi} = \boldsymbol{T}\tilde{\boldsymbol{T}}\boldsymbol{\Lambda}_\phi\tilde{\boldsymbol{T}}'\boldsymbol{T}',$$

where $\boldsymbol{\Lambda}_\phi$ is $\mathrm{diag}\{\phi_i(\tilde{\boldsymbol{\lambda}})\}$ and $\tilde{\boldsymbol{\lambda}} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p)'$. Then

$$\begin{aligned} \mathrm{tr}(\delta_{a_0,b_0}^{\phi} - \delta_{a_0,b_0})\boldsymbol{\Sigma}^{-1} &= \mathrm{tr}\tilde{\boldsymbol{T}}'\boldsymbol{P}\boldsymbol{P}'\boldsymbol{T}'\boldsymbol{\Sigma}^{-1}\boldsymbol{T}\boldsymbol{P}\boldsymbol{P}'\tilde{\boldsymbol{T}}(\boldsymbol{\Lambda}_\phi - \tilde{\boldsymbol{\Lambda}}) \\ &= \mathrm{tr}\boldsymbol{W}'\boldsymbol{W}\boldsymbol{P}'\tilde{\boldsymbol{T}}'(\boldsymbol{\Lambda}_\phi - \tilde{\boldsymbol{\Lambda}})\tilde{\boldsymbol{T}}\boldsymbol{P}, \end{aligned}$$

$$\log|\delta^{\phi}_{a_0,b_0}\boldsymbol{\Sigma}^{-1}| - \log|\delta_{a_0,b_0}\boldsymbol{\Sigma}^{-1}| = \log\prod_{i=1}^{p}\frac{\phi_i(\tilde{\boldsymbol{\lambda}})}{\tilde{\lambda}_i} = \sum_{i=1}^{p}\log\left(\frac{\phi_i(\tilde{\boldsymbol{\lambda}})}{\tilde{\lambda}_i}\right)$$

We note that if $\boldsymbol{P}\boldsymbol{T}'^{-1}\boldsymbol{X}_i = \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}_i\ i = 1,\ldots,p$ are all given, which is equivalent to $\boldsymbol{P}\boldsymbol{T}'^{-1}\boldsymbol{X} = \boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ is given, $\boldsymbol{P}\tilde{\boldsymbol{T}}'(\boldsymbol{\Lambda}_\phi - \tilde{\boldsymbol{\Lambda}})\tilde{\boldsymbol{T}}\boldsymbol{P}'$ is constant. Then $\tilde{\boldsymbol{\lambda}}$ is also constant with $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ given.

Suppose that $\phi_i(\tilde{\boldsymbol{\lambda}}) \le \tilde{\lambda}_i$. Let $\tilde{\boldsymbol{X}}_i$ be $\tilde{\boldsymbol{X}}_i = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{X}$. Then with respect to the conditional risk difference between $\delta^{\phi}_{a_0,b_0}$ and δ_{a_0,b_0} when $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ is given, we have

$$\begin{aligned}
&\mathrm{E}[\mathrm{L}(\delta^{\phi}_{a_0,b_0},\boldsymbol{\Sigma}) - \mathrm{L}(\delta_{a_0,b_0},\boldsymbol{\Sigma})|\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}] \\
&= \mathrm{tr}\tilde{\boldsymbol{T}}'\boldsymbol{P}'\mathrm{E}[\boldsymbol{W}'\boldsymbol{W}|\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}]\boldsymbol{P}\tilde{\boldsymbol{T}}(\boldsymbol{\Lambda}_\phi - \tilde{\boldsymbol{\Lambda}}) - \sum_{i=1}^{p}\log\left(\frac{\phi_i(\tilde{\boldsymbol{\lambda}})}{\tilde{\lambda}_i}\right) \\
&\le \mathrm{tr}(n+k)\tilde{\boldsymbol{T}}'(\boldsymbol{I}_p + \boldsymbol{T}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{T}'^{-1})^{-1}\tilde{\boldsymbol{T}}(\boldsymbol{\Lambda}_\phi - \tilde{\boldsymbol{\Lambda}}) - \sum_{i=1}^{p}\log\left(\frac{\phi_i(\tilde{\boldsymbol{\lambda}})}{\tilde{\lambda}_i}\right) \\
&= \sum_{i=1}^{p}\left\{\phi_i(\tilde{\boldsymbol{\lambda}}) - \tilde{\lambda}_i - \log\left(\frac{\phi_i(\tilde{\boldsymbol{\lambda}})}{\tilde{\lambda}_i}\right)\right\}.
\end{aligned}$$

Then we obtain a sufficient condition to improve δ_0 as

$$\phi_i(\boldsymbol{\lambda}) - \lambda_i - \log\left(\frac{\phi_i(\boldsymbol{\lambda})}{\lambda_i}\right) \le 0 \quad \text{for all} \;\; i$$

Setting $\phi_i^{TR}(\boldsymbol{\lambda})$ as

$$\phi_i^{TR}(\boldsymbol{\lambda}) = \begin{cases} 1, & \text{if} \;\; \tilde{\lambda}_p \ge 1 \\ \tilde{\lambda}_i, & \text{otherwise,} \end{cases}$$

it is easy to show that $\phi_i^{TR}(\boldsymbol{\lambda})$ satisfies the above condition. The corresponding estimator is $\delta^{TR}_{a_0,b_0}$. Then we proved the Theorem 3.1.

$\phi_i(\boldsymbol{\lambda}) = \min(\lambda_i, 1) \equiv \phi_0(\boldsymbol{\lambda})$ for all i also satisfies this condition. Let the corresponding estimator be denoted as $\delta^{\phi_0}_{a_0,b_0}$.

For $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ such that $\delta^{TR}_{a_0,b_0}$ takes $(\boldsymbol{S} + \boldsymbol{X}'\boldsymbol{X})/(n+k)$, $\delta^{\phi_0}_{a_0,b_0}$ also takes $(\boldsymbol{S} + \boldsymbol{X}'\boldsymbol{X})/(n+k)$. On the other hand for any $\boldsymbol{W}^{-1}\tilde{\boldsymbol{X}}$ such that $\delta^{TR}_{a_0,b_0}$ takes δ_{a_0,b_0}, $\delta^{\phi_0}_{a_0,b_0}$

improves δ_{a_0,b_0} under the conditional risk. Then the domination of $\delta^{\phi_0}_{a_0,b_0}$ against $\delta^{TR}_{a_0,b_0}$ is proved.

Theorem 3.2 $\delta^{\phi_0}_{a_0,b_0}$ *improves* $\delta^{TR}_{a_0,b_0}$.

The comparison between $\delta^{\phi_0}_{a_0,b_0}$ and δ^{SG}_k is interesting but we have not been able to obtain analytical results about it. We give some Monte Carlo studies in the following section.

4 Monte Carlo Study

We study the risk performance of the proposed estimators with some Monte Carlo studies. We compare the average losses of δ^{SG}_k, δ^{TR}, δ_{a_0,b_b}, $\delta^{TR}_{a_0,b_b}$ and $\delta^{\phi_0}_{a_0,b_b}$. We assume here that $\boldsymbol{\Sigma} = \boldsymbol{I}_p$. The average risk gains of the proposed estimators against δ_0,

$$\mathrm{E}[\mathrm{L}(\delta_0, \boldsymbol{I}_p) - \mathrm{L}(\delta, \boldsymbol{I}_p)]$$

over 100,000 replications for several combinations of $(p, n, k, \|\boldsymbol{\mu}_i\|^2)$ are given in Table 1 to 5.

The summary of this experiment is as follows.

- We can see the dominance δ^{TR} over δ^{SG}_k. When the degrees of freedom of $\boldsymbol{S}$ is rather small, the dominance relationship $\delta^{\phi_0}_{a_0,b_0} \geq \delta^{TR} \geq \delta^{SG}_k$ may hold. Taking into account that $\delta^{\phi_0}_{a_0,b_0}$ is the estimator which replace $\boldsymbol{S}/n$ in δ^{TR} with the improved estimator δ_{a_0,b_0} and that the analytical results of Kubokawa et al.(1993) in the case of $k = 1$, this relationship is expected intuitively, but have not been able to show analytically.

- These estimators all use the information of $\boldsymbol{X}$. Then as k increase for fixed p, n, $\boldsymbol{\mu}$, the risk gain is expected to increase. δ^{SG}_k may not always hold this relationship. In this sense δ^{SG}_k does not use the information of $\boldsymbol{X}$ effectively.

- For the trace of noncentral parameter matrix, the improvement of all estimators may be monotone nonincreasing.
- For $p = 5$, the risk gain of δ_k^{SG} nearly zero. But the gains of the other estimators is still considerable.

Acknowledgement

The author is deeply grateful to Professor Chihiro Hirotsu of University of Tokyo for his helpful comments and encouragement.

References

[1] Kubokawa, T., Robert, C. and Saleh, A.K.MD.E.(1992). Empirical Bayes estimation of the covariance matrix of a normal distribution with unknown mean under an entropy loss, *Sankhya*, A, **54**, 402-410.

[2] Kubokawa, T., Honda, T., Morita, K. and Saleh, A.K.MD.E.(1993). Estimating a covariance matrix of a normal distribution with unknown mean, *J. Japan Statist. Soc.*, **23**, 131-144.

[3] Perron, F.(1990). Equivariant estimators of the covariance matrix, *Canadian J. Statist.*, **18**, 179-182.

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

[5] Sinha, B. K. and Ghosh, M.(1989). Inadmissibility of the best equivariant estimators of the variance-covariance matrix, the precision matrix, and the generalized variance under entropy loss, *Statist. & Decisions*, **5**, 201-227.

[6] 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.

Table 1. Risk gains for δ_k^{SG}, δ^{TR}, δ_{a_0,b_0}, δ_{a_0,b_0}^{TR} and $\delta_{a_0,b_0}^{\phi_0}$ for $p = 3$ and $n = 4$.

$\mathrm{E}[\mathrm{L}(\delta_0, \boldsymbol{I}_p)] = 2.1966$.

	$\|\boldsymbol{\mu}_i\|^2$	0.0	1.0	2.0	5.0	10.0
$k = 3$	δ_k^{SG}	0.0057	0.0029	0.0013	0.0001	0.0000
	δ^{TR}	0.1774	0.1740	0.1695	0.1608	0.1568
	δ_{a_0,b_0}	0.3890	0.3778	0.3774	0.3506	0.3198
	δ_{a_0,b_0}^{TR}	0.4208	0.4000	0.3924	0.3544	0.3202
	$\delta_{a_0,b_0}^{\phi_0}$	0.5033	0.4874	0.4776	0.4252	0.3755
$k = 5$	δ_k^{SG}	0.0054	0.0022	0.0007	0.0000	0.0000
	δ^{TR}	0.2470	0.2392	0.2291	0.2124	0.2064
	δ_{a_0,b_0}	0.4205	0.4165	0.4033	0.3714	0.3373
	δ_{a_0,b_0}^{TR}	0.4487	0.4322	0.4116	0.3726	0.3373
	$\delta_{a_0,b_0}^{\phi_0}$	0.5965	0.5815	0.5521	0.4851	0.4297
$k = 7$	δ_k^{SG}	0.0050	0.0017	0.0005	0.0000	0.0000
	δ^{TR}	0.2888	0.2766	0.2614	0.2423	0.2345
	δ_{a_0,b_0}	0.4305	0.4251	0.4141	0.3846	0.3440
	δ_{a_0,b_0}^{TR}	0.4554	0.4372	0.4198	0.3853	0.3440
	$\delta_{a_0,b_0}^{\phi_0}$	0.6470	0.6258	0.5924	0.5236	0.4593
$k = 10$	δ_k^{SG}	0.0050	0.0010	0.0002	0.0001	0.0000
	δ^{TR}	0.3301	0.3111	0.2914	0.2684	0.2601
	δ_{a_0,b_0}	0.4457	0.4398	0.4258	0.3873	0.3513
	δ_{a_0,b_0}^{TR}	0.4697	0.4493	0.4295	0.3876	0.3514
	$\delta_{a_0,b_0}^{\phi_0}$	0.7028	0.6737	0.6322	0.5493	0.4880

Table 2. Risk gains for δ_k^{SG}, δ^{TR}, δ_{a_0,b_0}, δ_{a_0,b_0}^{TR} and $\delta_{a_0,b_0}^{\phi_0}$ for $p=3$ and $n=7$.

$\mathrm{E}[\mathrm{L}(\delta_0, \boldsymbol{I}_p)] = 1.0285$.

	$\|\boldsymbol{\mu}_i\|^2$	0.0	1.0	2.0	5.0	10.0
$k=3$	δ_k^{SG}	0.0046	0.0020	0.0008	0.0001	0.0000
	δ^{TR}	0.0716	0.0701	0.0673	0.0635	0.0620
	δ_{a_0,b_0}	0.1141	0.1148	0.1092	0.1022	0.0917
	δ_{a_0,b_0}^{TR}	0.1257	0.1216	0.1131	0.1028	0.0917
	$\delta_{a_0,b_0}^{\phi_0}$	0.1578	0.1561	0.1461	0.1301	0.1141
$k=5$	δ_k^{SG}	0.0045	0.0015	0.0005	0.0000	0.0000
	δ^{TR}	0.1100	0.1050	0.0996	0.0922	0.0891
	δ_{a_0,b_0}	0.1275	0.1208	0.1180	0.1089	0.0959
	δ_{a_0,b_0}^{TR}	0.1387	0.1256	0.1199	0.1090	0.0959
	$\delta_{a_0,b_0}^{\phi_0}$	0.2057	0.1933	0.1827	0.1596	0.1390
$k=7$	δ_k^{SG}	0.0047	0.0015	0.0003	0.0000	0.0000
	δ^{TR}	0.1377	0.1293	0.1212	0.1113	0.1075
	δ_{a_0,b_0}	0.1293	0.1262	0.1236	0.1102	0.0987
	δ_{a_0,b_0}^{TR}	0.1402	0.1303	0.1248	0.1103	0.0987
	$\delta_{a_0,b_0}^{\phi_0}$	0.2348	0.2219	0.2081	0.1776	0.1567
$k=10$	δ_k^{SG}	0.0048	0.0010	0.0002	0.0000	0.0000
	δ^{TR}	0.1667	0.1542	0.1420	0.1306	0.1260
	δ_{a_0,b_0}	0.1313	0.1274	0.1241	0.1155	0.0998
	δ_{a_0,b_0}^{TR}	0.1416	0.1302	0.1250	0.1155	0.0998
	$\delta_{a_0,b_0}^{\phi_0}$	0.2655	0.2475	0.2287	0.1999	0.1735

Table 3. Risk gains for δ_k^{SG}, δ^{TR}, δ_{a_0,b_0}, δ_{a_0,b_0}^{TR} and $\delta_{a_0,b_0}^{\phi_0}$ for $p=3$ and $n=10$.

$\mathrm{E}[\mathrm{L}(\delta_0, \boldsymbol{I}_p)] = 0.6778$.

	$\lVert\boldsymbol{\mu}_i\rVert^2$	0.0	1.0	2.0	5.0	10.0
$k=3$	δ_k^{SG}	0.0033	0.0014	0.0005	0.0001	0.0000
	δ^{TR}	0.0389	0.0378	0.0366	0.0344	0.0331
	δ_{a_0,b_0}	0.0561	0.0520	0.0539	0.0485	0.0440
	δ_{a_0,b_0}^{TR}	0.0627	0.0555	0.0559	0.0487	0.0440
	$\delta_{a_0,b_0}^{\phi_0}$	0.0798	0.0743	0.0741	0.0638	0.0562
$k=5$	δ_k^{SG}	0.0037	0.0013	0.0004	0.0000	0.0000
	δ^{TR}	0.0628	0.0598	0.0562	0.0524	0.0506
	δ_{a_0,b_0}	0.0593	0.0552	0.0540	0.0513	0.0476
	δ_{a_0,b_0}^{TR}	0.0662	0.0582	0.0551	0.0514	0.0476
	$\delta_{a_0,b_0}^{\phi_0}$	0.1048	0.0971	0.0912	0.0811	0.0732
$k=7$	δ_k^{SG}	0.0039	0.0009	0.0002	0.0000	0.0000
	δ^{TR}	0.0816	0.0761	0.0709	0.0649	0.0625
	δ_{a_0,b_0}	0.0587	0.0585	0.0577	0.0528	0.0468
	δ_{a_0,b_0}^{TR}	0.0658	0.0606	0.0584	0.0528	0.0467
	$\delta_{a_0,b_0}^{\phi_0}$	0.1227	0.1160	0.1085	0.0935	0.0823
$k=10$	δ_k^{SG}	0.0043	0.0008	0.0002	0.0000	0.0000
	δ^{TR}	0.1030	0.0941	0.0860	0.0790	0.0765
	δ_{a_0,b_0}	0.0618	0.0617	0.0597	0.0537	0.0478
	δ_{a_0,b_0}^{TR}	0.0689	0.0632	0.0600	0.0537	0.0478
	$\delta_{a_0,b_0}^{\phi_0}$	0.1467	0.1367	0.1246	0.1068	0.0952

Table 4. Risk gains for δ_k^{SG}, δ^{TR}, δ_{a_0,b_0}, δ_{a_0,b_0}^{TR} and $\delta_{a_0,b_0}^{\phi_0}$ for $p=5$ and $n=6$. $\mathrm{E}[\mathrm{L}(\delta_0, \boldsymbol{I}_p)] = 3.9778$.

	$\|\boldsymbol{\mu}_i\|^2$	0.0	1.0	5.0	10.0
$k=5$	δ_k^{SG}	0.0000	0.0000	0.0000	0.0000
	δ^{TR}	0.3311	0.3301	0.3210	0.3190
	δ_{a_0,b_0}	0.8249	0.8351	0.7955	0.7751
	δ_{a_0,b_0}^{TR}	0.8249	0.8377	0.7960	0.7751
	$\delta_{a_0,b_0}^{\phi_0}$	1.0440	1.0481	0.9805	0.9424
$k=7$	δ_k^{SG}	0.0000	0.0000	0.0000	0.0000
	δ^{TR}	0.4073	0.4023	0.3829	0.3821
	δ_{a_0,b_0}	0.9018	0.8800	0.8153	0.7652
	δ_{a_0,b_0}^{TR}	0.9046	0.8812	0.8153	0.7652
	$\delta_{a_0,b_0}^{\phi_0}$	1.1892	1.1608	1.0561	0.9879
$k=10$	δ_k^{SG}	0.0000	0.0000	0.0000	0.0000
	δ^{TR}	0.4846	0.4741	0.4498	0.4389
	δ_{a_0,b_0}	0.9006	0.9207	0.8514	0.7667
	δ_{a_0,b_0}^{TR}	0.9027	0.9210	0.8514	0.7667
	$\delta_{a_0,b_0}^{\phi_0}$	1.2661	1.2716	1.1522	1.0400

Table 5. Risk gains for δ_k^{SG}, δ^{TR}, δ_{a_0,b_0}, δ_{a_0,b_0}^{TR} and $\delta_{a_0,b_0}^{\phi_0}$ for $p = 5$ and $n = 10$. $\mathrm{E}[\mathrm{L}(\delta_0, \boldsymbol{I}_p)] = 1.8698$.

	$\|\boldsymbol{\mu}_i\|^2$	0.0	1.0	5.0	10.0
$k = 5$	δ_k^{SG}	0.0001	0.0001	0.0000	0.0000
	δ^{TR}	0.1484	0.1482	0.1408	0.1394
	δ_{a_0,b_0}	0.2344	0.2514	0.2196	0.2322
	δ_{a_0,b_0}^{TR}	0.2352	0.2518	0.2196	0.2322
	$\delta_{a_0,b_0}^{\phi_0}$	0.3307	0.3469	0.3022	0.3086
$k = 7$	δ_k^{SG}	0.0001	0.0000	0.0000	0.0000
	δ^{TR}	0.1945	0.1927	0.1789	0.1772
	δ_{a_0,b_0}	0.2628	0.2707	0.2159	0.2011
	δ_{a_0,b_0}^{TR}	0.2635	0.2708	0.2159	0.2011
	$\delta_{a_0,b_0}^{\phi_0}$	0.4029	0.4078	0.3325	0.3100
$k = 10$	δ_k^{SG}	0.0000	0.0000	0.0000	0.0000
	δ^{TR}	0.2479	0.2417	0.2230	0.2211
	δ_{a_0,b_0}	0.2676	0.2665	0.2292	0.2275
	δ_{a_0,b_0}^{TR}	0.2677	0.2664	0.2292	0.2275
	$\delta_{a_0,b_0}^{\phi_0}$	0.4599	0.4515	0.3860	0.3753