TSUKUBA J. MATH.
Vol. 6, No. 1 (1982), 103–126

ASYMPTOTIC RISK COMPARISON OF IMPROVED ESTIMATORS FOR NORMAL COVARIANCE MATRIX

By

Nariaki SUGIURA and Masahiro FUJIMOTO

Asymptotic risks of the empirical Bayes estimators $\hat{\Sigma}_H$ by Haff [5] for a covariance matrix Σ in a p-dimensional normal distribution are computed and compared with that of James and Stein's minimax estimators $\hat{\Sigma}_{JS}$. For $p\geqq 6$, it is shown that $\hat{\Sigma}_{JS}$ are always better than $\hat{\Sigma}_H$ asymptotically, though the leading terms are the same. New estimators which dominate $\hat{\Sigma}_{JS}$ for some Σ in any p asymptotically are proposed. Some numerical comparisons are given. Exact risks for ordinary estimators $\hat{\Sigma}_O$ and minimax estimators $\hat{\Sigma}_{JS}$ are also computed and compared with asymptotic ones for which the approximations are shown to be excellent.

1. Introduction

Let S have a Wishart distribution with unknown scale matrix Σ and n degrees of freedom, for which we shall write $S: W_p(n, \Sigma)$ and assume $n>p+1$. Let $\hat{\Sigma}$ be an estimator of Σ. The loss function is taken to be

$$L_1(\hat{\Sigma}, \Sigma)=\mathrm{tr}\,\hat{\Sigma}\Sigma^{-1}-\log|\hat{\Sigma}\Sigma^{-1}|-p \tag{1.1}$$

or

$$L_2(\hat{\Sigma}, \Sigma)=\frac{1}{2}\mathrm{tr}(\hat{\Sigma}\Sigma^{-1}-I)^2. \tag{1.2}$$

The L_1 loss is equivalent to the likelihood ratio statistic for testing the hypothesis $\Sigma=\Sigma_0$ against all alternatives. The L_2 loss can also be used as a test statistic for the same problem as in Nagao [10]. The factor 1/2 in the L_2 loss is not essential. However we wish to retain it, since L_1 loss tends to $\mathrm{tr}(\hat{\Sigma}\Sigma^{-1}-I)^2/2$, when $\hat{\Sigma}$ is close to Σ. The risk function is given by $R_i(\hat{\Sigma}, \Sigma)=E[L_i(\hat{\Sigma}, \Sigma)]$ for $i=1$ or 2. Haff [5] proved that among the scalar multiples of S, the best estimator under L_1 is $\hat{\Sigma}_O^{(1)}=S/n$ and that under L_2 it is given by $\hat{\Sigma}_O^{(2)}=S/(n+p+1)$, which we call ordinary estimators. Then he considered the posterior mean of Σ for a prior distribution $W_p[n', (\gamma C)^{-1}]$ for Σ^{-1} with unknown scalar $\gamma>0$ and known p. d. matrix

Received October 13, 1981. Revised February 15, 1982.

C. It is given by $E[\Sigma|S,\gamma]=(S+\gamma C)/(n+n'-p-1)$. In the process of estimating γ by maximizing approximate marginal likelihood of S, he obtained $ut(u)$ for $u=1/\mathrm{tr}(S^{-1}C)$ as an estimator for γ, where $t(\cdot)$ is nonincreasing. He then proved that under L_1 the estimator

$$\hat{\Sigma}_H^{(1)}=\frac{1}{n}[S+ut(u)C] \tag{1.3}$$

for $0\leqq t(u)\leqq 2(p-1)/n$, dominates $\hat{\Sigma}_O^{(1)}=S/n$ for any $n>p+1$ and under L_2 the estimator

$$\hat{\Sigma}_H^{(2)}=\frac{1}{n+p+1}(S+utC) \tag{1.4}$$

for $0\leqq t\leqq 2(p-1)/(n-p+3)$, dominates $\hat{\Sigma}_O^{(2)}=S/(n+p+1)$ for any $n>p+1$. It was also shown that if $t(u)$ in (1.3) is constant, the best choice of $t(u)$ is $(p-1)/n$ and that the best choice of t in (1.4) is $(p-1)/(n-p+3)$. In this paper we always take these optimal values for t and call them Haff's estimators $\hat{\Sigma}_H^{(1)}$ and $\hat{\Sigma}_H^{(2)}$ respectively.

A minimax estimator for Σ was earlier obtained by James and Stein [7], giving

$$\hat{\Sigma}_{JS}^{(i)}=K\Delta^{(i)}K' \tag{1.5}$$

for the loss L_i ($i=1$ or 2), where the lower triangular matrix K with positive diagonal elements is obtained from $S=KK'$ and $\Delta^{(i)}=diag[\Delta_1^{(i)},\cdots,\Delta_p^{(i)}]$. For the L_1 loss, they proved that $\Delta_j^{(1)}=1/(n+p+1-2j)$ and reported that they were unable to get explicit form of $\Delta_j^{(2)}$. Sharma [13] derived the linear equations for $\Delta_j^{(2)}$, from which numerical values are computed for given n and p. They were also obtained earlier by Selliah [12].

The primary purpose of this paper is to compare the asymptotic risk of Haff's estimator $\hat{\Sigma}_H^{(i)}$ with that of James and Stein's estimator $\hat{\Sigma}_{JS}^{(i)}$ under L_i for $i=1$ or 2. Under L_2, we have derived an asymptotic form of $\Delta_j^{(2)}$ for large n. It is shown that the leading terms of the asymptotic risks for $\hat{\Sigma}_H^{(i)}$ and $\hat{\Sigma}_{JS}^{(i)}$ are the same and that the next term for $\hat{\Sigma}_H^{(i)}$ is less than that of $\hat{\Sigma}_{JS}^{(i)}$ only for $2\leqq p\leqq 5$ and for some Σ. If $p\geqq 6$, the second term of the asymptotic expansion of $R_i(\hat{\Sigma}_H^{(i)},\Sigma)$ is always larger than that of $R_i(\hat{\Sigma}_{JS}^{(i)},\Sigma)$ for all Σ.

Secondly we shall propose new estimators for Σ by minimizing risks empirically, which are given by

$$\hat{\Sigma}^{(1)}=\frac{1}{n}\left[S+b\frac{\mathrm{tr}\,CS^{-1}}{\mathrm{tr}(CS^{-1})^2}C\right],\quad 0\leqq b\leqq\frac{2(p-1)}{n} \tag{1.6}$$

for L_1 loss and

$$\hat{\Sigma}^{(2)}=\frac{1}{n+p+1}\left[S+b\frac{\mathrm{tr}\,CS^{-1}}{\mathrm{tr}(CS^{-1})^2}C\right],\quad 0\leqq b\leqq\frac{2(p-1)}{n} \tag{1.7}$$

for L_2 loss. It is shown that our new estimator $\hat{\Sigma}^{(1)}$ dominates $\hat{\Sigma}_O^{(1)}$ for all $n>p+1$ and that $\hat{\Sigma}^{(2)}$ dominates $\hat{\Sigma}_O^{(2)}$ asymptotically. The result also holds for more general form of $\hat{\Sigma}^{(1)}$, that is, the constant b in (1.6) can be replaced by $t(\cdot)$ in (1.3) for $u=\operatorname{tr} CS^{-1}/\operatorname{tr}(CS^{-1})^2$. However we prefer to (1.6) to simplify later discussions. The leading term of the asymptotic risk is the same as that of $\hat{\Sigma}_{JS}^{(i)}$ and the second term is less than that of $\hat{\Sigma}_{JS}^{(i)}$ for some Σ and for all $p>1$. Eliminating the leading term, the range of $R_i(\hat{\Sigma}^{(i)}, \Sigma)$ is much wider below than $R_i(\hat{\Sigma}_H^{(i)}, \Sigma)$ asymptotically. However the absolute difference $R_i(\hat{\Sigma}^{(i)}, \Sigma)-R_i(\hat{\Sigma}_{JS}^{(i)}, \Sigma)$ or $R_i(\hat{\Sigma}_H^{(i)}, \Sigma)-R_i(\hat{\Sigma}_{JS}^{(i)}, \Sigma)$ is not so large.

To get some idea for the errors of asymptotic approximations, the terms of order n^{-3} (third terms) are computed for $R_i(\hat{\Sigma}_H^{(i)}, \Sigma)$ and $R_i(\hat{\Sigma}^{(i)}, \Sigma)$. The exact risks of $\hat{\Sigma}_{JS}^{(i)}$ are computed and asymptotic values up to order n^{-3} are compared. For $2\leqq p\leqq 6$ and $n\geqq 16$, asymptotic values for $\hat{\Sigma}_{JS}^{(i)}$ are accurate for three (two) significant digits for L_1 (L_2) loss in most cases examined. The rates of the reduction of the risks of $\hat{\Sigma}_H^{(i)}(\hat{\Sigma}^{(i)})$ with respect to $\hat{\Sigma}_O^{(i)}$ are shown to be the highest 8%(20%) for $i=1$, $n\geqq 16$ and 4%(11%) for $i=2$, $n\geqq 32$ respectively within our examples computed in Tables.

2. Derivation of new estimators

Since our goal is to find an estimator $\hat{\Sigma}$ which minimizes the risk, we shall look for a solution in a form $\hat{\Sigma}^{(1)}=(S+\gamma C)/n$ for L_1 or $\hat{\Sigma}^{(2)}=(S+\gamma C)/(n+p+1)$ for L_2. The risk for L_1 is given by

$$R_1(\hat{\Sigma}^{(1)}, \Sigma)=\frac{\gamma}{n}\operatorname{tr} C\Sigma^{-1}-E[\log|\frac{1}{n}(S+\gamma C)\Sigma^{-1}|]. \tag{2.1}$$

Hence the derivative with respect to γ is

$$\frac{1}{n}\operatorname{tr} C\Sigma^{-1}-E[\operatorname{tr}(\gamma I+SC^{-1})^{-1}], \tag{2.2}$$

where the expectation is taken by S having $W_p(n, \Sigma)$ distribution. At $\gamma=0$, the derivative has a negative value $-(p+1)\operatorname{tr} C\Sigma^{-1}/\{n(n-p-1)\}$, since $E(S^{-1})=\Sigma^{-1}/(n-p-1)$, by Kshirsagar [9], for example. This shows that the risk will be smaller if we take γ positive near zero. Assume that γ is small and put the derivative (2.2) equal to zero. We get an equation for γ, an approximate solution of which is given by

$$\gamma=(p+1)\frac{\operatorname{tr} C\Sigma^{-1}}{\operatorname{tr}(C\Sigma^{-1})^2}, \tag{2.3}$$

which yields the estimator (1.6). The estimator (1.7) for L_2 is similarly derived.

The constant factor b is restricted so that it dominates ordinary estimator $\hat{\Sigma}_O^{(i)}$, which will be discussed later.

3. Risks of ordinary and James and Stein's minimax estimators

Using the Bartlett's decomposition (Giri [3], page 126) of Wishart matrix S when $\Sigma=I$, we get

$$R_1(\hat{\Sigma}_O^{(1)}, \Sigma)=p\log n-\sum_{j=1}^{p} E[\log \chi^2_{n-j+1}], \tag{3.1}$$

where χ^2_m denotes the χ^2 variate with m degrees of freedom. Using digamma function $\psi(x)=d\log\Gamma(x)/dx$, we can rewrite it

$$p\log\frac{n}{2}-\sum_{j=1}^{p}\psi\left(\frac{n-j+1}{2}\right). \tag{3.2}$$

If n is an integer larger than one, we know that

$$\psi(n)=1+\frac{1}{2}+\cdots+\frac{1}{n-1}-\gamma \tag{3.3}$$

for Euler's constant $\gamma=0.57721\ 56649\ 01532\ 9\cdots$ (Abramowitz and Stegun [1]). For half integer argument ($n\geqq 1$),

$$\psi\left(n+\frac{1}{2}\right)=-\gamma-2\log 2+2\left(1+\frac{1}{3}+\cdots+\frac{1}{2n-1}\right). \tag{3.4}$$

These are sufficient for the computation of $R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$. If n is large, an asymptotic formula for ψ is available, which is derived from Stirling's formula (Kendall [8], page 245)

$$\psi(x+h)=\log x+\frac{h-1/2}{x}+\sum_{r=1}^{n}\frac{(-1)^r B_{r+1}(h)}{x^{r+1}(r+1)}+O\left(\frac{1}{x^{n+2}}\right), \tag{3.5}$$

where $B_r(h)$ are the Bernoulli polynomials given by $B_2(h)=h^2-h+1/6$, $B_3(h)=h^3-(3/2)h^2+(1/2)h$. This yields

$$R_1(\hat{\Sigma}_O^{(1)}, \Sigma)=\frac{p(p+1)}{2n}+\frac{p(2p^2+3p-1)}{12n^2}+\frac{p(p^2-1)(p+2)}{12n^3}+O(n^{-4}). \tag{3.6}$$

Some numerical values of $R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$ are computed based on (3.2)~(3.4) and compared with the asymptotic values (3.6) for $p=2\sim 6$ and $n=8\sim 128$. They are shown in Table 1. We can see that the asymptotic approximations are excellent, namely, for $n\geqq 16$ and $p\leqq 6$, the values are accurate with three significance digits.

Under L_2 loss, Haff [5] noted that

$$R_2(\hat{\Sigma}_O^{(2)}, \Sigma)=\frac{p(p+1)}{2(n+p+1)}, \tag{3.7}$$

Table 1. Values of $R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$

		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$	$O(n^{-1})$	.37500	.187500	.093750	.046875	.023438
	$O(n^{-2})$	.03385	.008464	.002116	.000529	.000132
	$O(n^{-3})$	.00391	.000488	.000061	.000008	.000001
	approx.	.4128	.19645	.095927	.047412	.023571
	exact	.413314	.196484	.095929	.047412	.023571
$p=3$	$O(n^{-1})$	.75000	.37500	.187500	.093750	.046875
	$O(n^{-2})$	.10156	.02539	.006348	.001587	.000397
	$O(n^{-3})$	.01953	.00244	.000305	.000038	.000005
	approx.	.871	.4028	.19415	.095375	.047276
	exact	.876824	.403141	.194171	.095376	.047277
$p=4$	$O(n^{-1})$	1.2500	.62500	.312500	.156250	.078125
	$O(n^{-2})$	.2240	.05599	.013997	.003499	.000875
	$O(n^{-3})$	.0586	.00732	.000916	.000114	.000014
	approx.	1.533	.6883	.32741	.159864	.079014
	exact	1.559962	.689672	.327490	.159868	.079015
$p=5$	$O(n^{-1})$	1.8750	.9375	.46875	.234375	.117188
	$O(n^{-2})$	.4167	.1042	.02604	.006510	.001628
	$O(n^{-3})$	.1367	.0171	.00214	.000267	.000033
	approx.	2.43	1.059	.4969	.24115	.118848
	exact	2.52347	1.06300	.497161	.241166	.118849
$p=6$	$O(n^{-1})$	2.6250	1.3125	.65626	.328125	.164063
	$O(n^{-2})$	.6953	.1738	.04346	.010864	.002716
	$O(n^{-3})$	.2734	.0342	.00427	.000534	.000067
	approx.	3.59	1.521	.7040	.33952	.166845
	exact	3.87328	1.53134	.704554	.339557	.166847

which is asymptotically the same as $R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$ for large n. This is the reason why we prefer multiplier 1/2 in the definition of L_2 loss in (1.2). Unlike the simple form of (3.7), the asymptotic approximations

$$R_2(\hat{\Sigma}_O^{(2)}, \Sigma)=\frac{p(p+1)}{2n}-\frac{p(p+1)^2}{2n^2}+\frac{p(p+1)^3}{2n^3}+O(n^{-4}) \tag{3.8}$$

are not so excellent as $R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$. For example, the exact value of $R_2(\hat{\Sigma}_O^{(2)}, \Sigma)$ in (3.7) for $p=2$ and $n=16$ is 0.15789, while the asymptotic value of (3.8) gives 0.15894 which is accurate for three significant digits. From Table 1, the corresponding exact value of $R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$ is 0.19648 and the asymptotic value is 0.19645 which is accurate for one more digit than $R_2(\hat{\Sigma}_O^{(2)}, \Sigma)$. This is the case with other values of parameters n and p.

Next we shall evaluate the risks of the minimax estimators by James and Stein [7]. By considering a best equivariant estimator $\phi(LSL')=L\phi(S)L'$ for the transformation group of lower triangular matrices L with positive diagonal elements, they obtained a minimax estimator of (1.5) under L_1 loss and derived

$$R_1(\hat{\Sigma}_{JS}^{(1)}, \Sigma)=\sum_{j=1}^{p}\log(n+p-2j+1)-\sum_{j=1}^{p}E[\log\chi^2_{n-j+1}]. \tag{3.9}$$

Using digamma function $\psi(x)$, this can be simplified as

$$\sum_{j=1}^{p} \log \frac{1}{2}(n+p-2j+1) - \sum_{j=1}^{p} \psi\left(\frac{n-j+1}{2}\right), \tag{3.10}$$

which is useful for numerical computations. The asymptotic form of (3.10) is obtained by (3.5), giving

$$R_1(\hat{\Sigma}^{(1)}_{JS}, \Sigma) = \frac{p(p+1)}{2n} + \frac{p(3p+1)}{12n^2} + \frac{p(p^2-1)(p+2)}{12n^3} + O(n^{-4}). \tag{3.11}$$

In Table 2 exact and asymptotic values of $R_1(\hat{\Sigma}^{(1)}_{JS}, \Sigma)$ are compared. It is found that for $n \geqq 16$ and $p \leqq 6$, the asymptotic values are accurate for three significant digits, which is the same conclusion as for $R_1(\hat{\Sigma}^{(1)}_{O}, \Sigma)$. Since equivariant estimators contain best scalar multiple of S, namely, $\hat{\Sigma}^{(1)}_{O}$, inequality $R_1(\hat{\Sigma}^{(1)}_{JS}, \Sigma) < R_1(\hat{\Sigma}^{(1)}_{O}, \Sigma)$ holds as a matter of fact. If we take difference of the risks by asymptotic form, we get

$$R_1(\hat{\Sigma}^{(1)}_{JS}, \Sigma) - R_1(\hat{\Sigma}^{(1)}_{O}, \Sigma) = -\frac{p(p^2-1)}{6n^2} + O(n^{-4}), \tag{3.12}$$

which is negative for $p \geqq 2$, neglecting the higher order terms. This suggests the

Table 2. Exact and asymptotic values of $R_1(\hat{\Sigma}^{(1)}_{JS}, \Sigma)$

		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$	$O(n^{-1})$	.37500	.187500	.093750	.046875	.023438
	$O(n^{-2})$	.01823	.004557	.001139	.000285	.000071
	$O(n^{-3})$	.00391	.000488	.000061	.000008	.000001
	approx.	.3971	.19255	.094950	.047167	.023510
	exact	.39757	.19257	.094952	.047168	.023510
$p=3$	$O(n^{-1})$	.75000	.37500	.187500	.093750	.046875
	$O(n^{-2})$	.03906	.00977	.002441	.000610	.000153
	$O(n^{-3})$	.01953	.00244	.000305	.000038	.000005
	approx.	.809	.3872	.19025	.094398	.047033
	exact	.81229	.38739	.190257	.094399	.047032
$p=4$	$O(n^{-1})$	1.2500	.62500	.312500	.156250	.078125
	$O(n^{-2})$	.0677	.01693	.004232	.001058	.000265
	$O(n^{-3})$	.0586	.00732	.000916	.000114	.000014
	approx.	1.376	.6493	.31765	.157422	.078404
	exact	1.3927	.64997	.31768	.157425	.078404
$p=5$	$O(n^{-1})$	1.8750	.9375	.46875	.234375	.117188
	$O(n^{-2})$	.1042	.0260	.00651	.001628	.000407
	$O(n^{-3})$	.1367	.0171	.00214	.000267	.000033
	approx.	2.12	.981	.4774	.236270	.117628
	exact	2.1713	.98271	.47750	.236275	.117628
$p=6$	$O(n^{-1})$	2.6250	1.3125	.65625	.328125	.164063
	$O(n^{-2})$	.1484	.0371	.00928	.002319	.000580
	$O(n^{-3})$	.2734	.0342	.00427	.000534	.000067
	approx.	3.05	1.384	.6698	.33098	.164709
	exact	3.2107	1.3889	.67003	.330991	.164710

validity of the asymptotic comparisons.

Under L_2 loss, the exact $\Delta^{(2)}$ is not available. However Selliah [12] and Sharma [13] show that $\Delta=[\Delta_1^{(2)}, \cdots, \Delta_p^{(2)}]'$, satisfies linear equations $A\Delta=b$, where $p\times p$ matrix A and p-vector b are given by

$$A=\begin{pmatrix} (n+p-1)(n+p+1) & n+p-3 & \cdots\cdots & n-p+1 \\ n+p-3 & (n+p-3)(n+p-1) & \cdots\cdots & n-p+1 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ n-p+1 & n-p+1 & \cdots\cdots & (n-p+1)(n-p+3) \end{pmatrix} \tag{3.13}$$

$$b=(n+p-1,\ n+p-3, \cdots, n-p+1)'.$$

With this Δ, the risk is given by

$$R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)=\frac{1}{2}p-\frac{1}{2}\sum_{j=1}^{p}(n-2j+p+1)\Delta_j^{(2)}. \tag{3.14}$$

We can see by checking the exact values of $\Delta^{(1)}$ and $\Delta^{(2)}$ that the choice of $\Delta_j^{(1)}$ is always larger than $\Delta_j^{(2)}$ and the risks of $\hat{\Sigma}_{JS}^{(1)}$ are larger than that of $\hat{\Sigma}_{JS}^{(2)}$. The best scalar multiple $1/n$ for L_1 loss and $1/(n+p+1)$ for L_2 loss lie always smaller than the middle of $\Delta_1, \cdots, \Delta_p$. Sharma [13] gives the values of $R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)$ for $p=2$ and $n=5(5)30$. Using (3.13), we can evaluate Δ for large n, giving

$$\begin{aligned}\Delta_j^{(2)}=&\frac{1}{n}-\frac{2}{n^2}(p+1-j)+\frac{1}{n^3}[4(p+1)^2-(8p+9)j+5j^2]\\ &+\frac{1}{3n^4}[-2(p+1)(11p^2+22p+12)+(66p^2+150p+85)j\\ &-3(28p+33)j^2+38j^3]+O(n^{-5})\end{aligned} \tag{3.15}$$

and

$$R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)=\frac{p(p+1)}{2n}-\frac{p(p+1)(2p+1)}{3n^2}+\frac{p^2(p+1)^2}{n^3}+O(n^{-4}). \tag{3.16}$$

Note that optimal scalar multiplier for S is $1/n$ under L_1 loss and $1/(n+p+1)$ under L_2 loss. Asymptotic expansion of $\Delta_j^{(1)}=1/(n+p+1-2j)$ replaced n by $n+p+1$ yields the same terms as in (3.15) up to order n^{-2}. The difference of the risks, $R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)-R_2(\hat{\Sigma}_{O}^{(2)}, \Sigma)$ in the asymptotic form is exactly the same as (3.12) up to $O(n^{-2})$. In Table 3, exact and asymptotic values of $R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)$ are shown based on (3.14) and (3.16). We can see that the asymptotic approximations are worse than $R_1(\hat{\Sigma}_{JS}^{(1)}, \Sigma)$ and are comparative for $R_2(\hat{\Sigma}_{O}^{(2)}, \Sigma)$. This suggests that the loss L_1 is favourable for the asymptotic approximations. The maximum rate of reduction of risks for $\hat{\Sigma}_{JS}^{(1)}$ with respect to $\hat{\Sigma}_{O}^{(1)}$ within Tables 1 and 2 is given by 17% for $n=8$ and $p=6$. However the corresponding rate for L_2 loss in Table 3 is only 5%.

Table 3. Exact and asymptotic values of $R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)$

		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$	$O(n^{-1})$	.37500	.18750	.093750	.046875	.023438
	$O(n^{-2})$	—.15625	—.03906	—.009766	—.002441	—.000610
	$O(n^{-3})$	.07031	.00879	.001099	.000137	.000017
	approx.	.289	.1572	.0851	.04457	.022844
	exact	.26697	.15559	.084970	.044563	.022844
$p=3$	$O(n^{-1})$	.75000	.37500	.18750	.093750	.046875
	$O(n^{-2})$	—.43750	—.10938	—.02734	—.006836	—.001709
	$O(n^{-3})$	.28125	.03516	.00440	.000549	.000069
	approx.	.59	.301	.1646	.08746	.045235
	exact	.48250	.29211	.16393	.087422	.045232
$p=4$	$O(n^{-1})$	1.2500	.62500	.31250	.15625	.078125
	$O(n^{-2})$	—.9375	—.23438	—.05859	—.01465	—.003662
	$O(n^{-3})$	.7813	.09766	.01221	.00153	.000191
	approx.	1.09	.488	.266	.1431	.07465
	exact	.73548	.45918	.26397	.14298	.074644
$p=5$	$O(n^{-1})$	1.8750	.9375	.46875	.23438	.117188
	$O(n^{-2})$	—1.7188	—.4297	—.10742	—.02686	—.006714
	$O(n^{-3})$	1.7578	.2197	.02747	.00343	.000429
	approx.	1.9	.73	.389	.2110	.11090
	exact	1.0189	.65233	.38311	.21056	.11088
$p=6$	$O(n^{-1})$	2.625	1.3125	.65625	.32813	.164063
	$O(n^{-2})$	—2.844	—.7109	—.17773	—.04443	—.011108
	$O(n^{-3})$	3.445	.4307	.05383	.00673	.000841
	approx.	3.2	1.03	.532	.2904	.15380
	exact	1.3283	.86807	.51965	.28952	.15374

4. Risks under L_1 loss

4.1. Risk of Haff's estimator. As Sharma [13] noted, the exact values of the risks of Haff's estimators are difficult to compute. Asymptotic evaluation of them gives some useful information. We shall put $C=I$ in (1.3) without loss of generality and assume that $t(u)=b=$constant, namely, the estimator

$$\hat{\Sigma}_H^{(1)}=\frac{1}{n}\left(S+\frac{b}{\operatorname{tr} S^{-1}}I\right) \tag{4.1}$$

is considered for L_1 loss. The difference of risks can be written by

$$\begin{aligned} &R_1(\hat{\Sigma}_H^{(1)}, \Sigma)-R_1(\hat{\Sigma}_O^{(1)}, \Sigma) \\ &=\frac{b}{n}E\left[\frac{\operatorname{tr}\Sigma^{-1}}{\operatorname{tr} S^{-1}}\right]-E\left[\log\left|I+\frac{b}{\operatorname{tr} S^{-1}}S^{-1}\right|\right], \end{aligned} \tag{4.2}$$

which is bounded from above by

$$\frac{b}{n}E\left[\frac{\operatorname{tr}\Sigma^{-1}}{\operatorname{tr} S^{-1}}\right]-b+\frac{b^2}{2}E\left[\frac{\operatorname{tr} S^{-2}}{(\operatorname{tr} S^{-1})^2}\right]. \tag{4.3}$$

By the Wishart identity due to Haff [5], we get

$$E\left[\frac{\operatorname{tr}\Sigma^{-1}}{\operatorname{tr}S^{-1}}\right]=n-p-1+2E\left[\frac{\operatorname{tr}S^{-2}}{(\operatorname{tr}S^{-1})^2}\right]. \tag{4.4}$$

This yields an upper bound of (4.2)

$$\frac{b}{n}\left(-p-1+2+\frac{nb}{2}\right), \tag{4.5}$$

which is negative if and only if $0\leqq b\leqq 2(p-1)/n$, and the minimum value is attained by $b=(p-1)/n$. This is the special case of Theorem 4.3 by Haff [5]. We impose this restriction on b. Note that $b=O(n^{-1})$ and $Y=\sqrt{n}(S/n-\Sigma)$ converges in law to a $p(p+1)/2$ variate normal distribution with mean zero. We can evaluate (4.2) asymptotically as

$$\frac{b}{n}\left\{E\left[\frac{\operatorname{tr}\Sigma^{-1}}{\operatorname{tr}S^{-1}}\right]-n+\frac{nb}{2}E\left[\frac{\operatorname{tr}S^{-2}}{(\operatorname{tr}S^{-1})^2}\right]-\frac{b^2n}{3}\frac{\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-1})^3}\right\}+O(n^{-4}). \tag{4.6}$$

In getting the last term of (4.6), we should take $E[\operatorname{tr}S^{-3}/(\operatorname{tr}S^{-1})^3]$, which can be evaluated by writing $S/n=\Sigma+Y/\sqrt{n}$ and noting that $E(Y)=0$ and $Y=O_p(1)$, giving $\operatorname{tr}\Sigma^{-3}/(\operatorname{tr}\Sigma^{-1})^3+O(n^{-1})$. Now we need the following lemma to complete our asymptotic expansion.

LEMMA 4.1. *Let S have a Wishart distribution $W_p(n,\Sigma)$. Then*

$$E\left[\frac{\operatorname{tr}S^{-2}}{(\operatorname{tr}S^{-1})^2}\right]=\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}+\frac{1}{n}\left\{6\frac{(\operatorname{tr}\Sigma^{-2})^2}{(\operatorname{tr}\Sigma^{-1})^4}-8\frac{\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-1})^3}+\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}+1\right\}+O(n^{-2}). \tag{4.7}$$

PROOF. From the Wishart identity, we get

$$E\left[\frac{\operatorname{tr}S^{-2}}{(\operatorname{tr}S^{-1})^2}\operatorname{tr}\Sigma^{-1}\right]=4E\left[\frac{(\operatorname{tr}S^{-2})^2}{(\operatorname{tr}S^{-1})^3}-\frac{\operatorname{tr}S^{-3}}{(\operatorname{tr}S^{-1})^2}\right]+(n-p-1)E\left[\frac{\operatorname{tr}S^{-2}}{\operatorname{tr}S^{-1}}\right]. \tag{4.8}$$

$$E\left[\frac{\operatorname{tr}S^{-2}}{\operatorname{tr}S^{-1}}\operatorname{tr}\Sigma^{-1}\right]=2E\left[\frac{(\operatorname{tr}S^{-2})^2}{(\operatorname{tr}S^{-1})^2}-2\frac{\operatorname{tr}S^{-3}}{\operatorname{tr}S^{-1}}\right]+(n-p-1)E[\operatorname{tr}S^{-2}]. \tag{4.9}$$

By Haff [4], we know that

$$\begin{aligned}E[\operatorname{tr}S^{-2}]&=\frac{(\operatorname{tr}\Sigma^{-1})^2}{(n-p)(n-p-1)(n-p-3)}+\frac{\operatorname{tr}\Sigma^{-2}}{(n-p)(n-p-3)}\\&=\frac{1}{n^2}\operatorname{tr}\Sigma^{-2}+\frac{2p+3}{n^3}\operatorname{tr}\Sigma^{-2}+\frac{1}{n^3}(\operatorname{tr}\Sigma^{-1})^2+O(n^{-4}).\end{aligned} \tag{4.10}$$

Combined with these formulas, we get the desired result (4.7).
Substituting (4.4) and (4.7) into (4.6) and using (3.12) we get

THEOREM 4.1. *An asymptotic expansion of the difference of risks between Haff's estimator* $\hat{\Sigma}_H^{(1)}$ *defined by* (4.1) *with* $b=(p-1)/n$ *and James and Stein's minimax estimator* $\hat{\Sigma}_{JS}^{(1)}$ *for* L_1 *loss is given by*

$$R_1(\hat{\Sigma}_H^{(1)}, \Sigma)-R_1(\hat{\Sigma}_{JS}^{(1)}, \Sigma)=\frac{p-1}{6n^2}\left\{(p+1)(p-6)+3(p+3)\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}\right\}$$
$$(4.11)\qquad +\frac{(p-1)(p+3)}{2n^3}\left\{6\frac{(\operatorname{tr}\Sigma^{-2})^2}{(\operatorname{tr}\Sigma^{-1})^4}-8\frac{\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-1})^3}+\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}+1\right\}$$
$$-\frac{(p-1)^3}{3n^3}\frac{\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-1})^3}+O(n^{-4}).$$

We can see that the term of $O(n^{-2})$ in (4.11) is always positive, if $p\geqq 6$. This shows that the risk of $\hat{\Sigma}_H^{(1)}$ is always larger than that of $\hat{\Sigma}_{JS}^{(1)}$ asymptotically, if $p\geqq 6$. Note that

$$(4.12)\qquad \frac{1}{p}\leqq\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}\leqq 1.$$

The lower and upper bounds of $O(n^{-2})$ in (4.11) are given by

$$(4.13)\qquad \frac{1}{6}(p-1)\left(p^2-5p-3+\frac{9}{p}\right)\quad\text{and}\quad\frac{1}{6}(p-1)(p^2-2p+3).$$

Some numerical values are given in the following:

Ranges of $O(n^{-2})$ in (4.11).

$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
$\left(-\frac{3}{4}, \frac{1}{2}\right)$	$(-2, 2)$	$\left(-\frac{19}{8}, \frac{11}{2}\right)$	$\left(-\frac{4}{5}, 12\right)$	$\left(\frac{15}{4}, \frac{45}{2}\right)$

The risk is unchanged for any scalar multiple of Σ. Some numerical values based on (4.11) are given in Table 4. The term of $O(n^{-3})$ gives some idea for the error of our asymptotic approximation. For $\Sigma^{-1}=\lambda\, diag(1, 1, \cdots, 1)$, the lower bound of (4.12) is attained and for $\Sigma^{-1}\to\lambda\, diag(1, 0, \cdots, 0)$, the upper bound is approached. In Table 4 we write $\Sigma^{-1}=\lambda(1, \cdots, 1)$ instead of $\Sigma^{-1}=\lambda\, diag(1, \cdots, 1)$ for abbreviation. Inspection of Table 4 shows that for $p\geqq 6$, the risk differences are positive and that for $p=5$ and $\Sigma^{-1}=\lambda\, diag(1, \cdots, 1)$, the values are positive for $n=8$ and $n=16$, while they are negative for $n\geqq 32$. Precisely speaking they are positive for $n\leqq 21$ and negative for $n\geqq 22$. Whether this is due to the poor accuracy of the asymptotic approximation for small n is not clear. For $p\leqq 4$ and $\Sigma^{-1}=\lambda\, diag(1, \cdots, 1)$, the values are all negative. Thus $p=5$ is the boundary. $\hat{\Sigma}_H^{(1)}$ is better than $\hat{\Sigma}_{JS}^{(1)}$ for these type of Σ if $p\leqq 5$. For $0\leqq b\leqq 2(p-1)/n$, inequality $R_1(\hat{\Sigma}_H^{(1)}, \Sigma)<R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$ holds

exactly. This can be verified also by the asymptotic consideration, namely, we have

$$R_1(\hat{\Sigma}_H^{(1)}, \Sigma)-R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$$
$$=\frac{p-1}{n^2}\left[-(p+1)+\frac{1}{2}(p+3)\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}\right]+O(n^{-3}). \tag{4.14}$$

The term of $O(n^{-2})$ is always negative because of (4.12). This gives again a weak support as in (3.12) for the usefulness of the asymptotic comparison, when exact inequality between risks is not known. From Tables 1 and 4, we can compute the rates of the reduction of the risks of Haff's estimator $\hat{\Sigma}_H^{(1)}$ with respect to the

Table 4. Asymptotic values of $R_1(\hat{\Sigma}_H^{(1)}, \Sigma)-R_1(\hat{\Sigma}_{JS}^{(1)}, \Sigma)$

Σ^{-1}		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$ $\lambda(1,1)$	$O(n^{-2})$	—.011719	—.002930	—.000732	—.000183	—.000046
	$O(n^{-3})$	.004720	.000590	.000074	.000009	.000001
	approx.	—.0070	—.00234	—.000659	—.000174	—.000045
$\lambda(1,2)$	$O(n^{-2})$	—.009549	—.002387	—.000597	—.000149	—.000037
	$O(n^{-3})$	.003400	.000425	.000053	.000007	.000001
	approx.	—.0061	—.00196	—.000544	—.000143	—.000036
$\lambda(1,10)$	$O(n^{-2})$	.001356	.000339	.000085	.000021	.000005
	$O(n^{-3})$	—.000496	—.000062	—.000008	—.000001	—.000000
	approx.	.00086	.000277	.000077	.000020	.000005
$\lambda(1,0)$	$O(n^{-2})$	.007813	.001953	.000488	.000122	.000031
	$O(n^{-3})$	—.000651	—.000081	—.000010	—.000001	—.000000
	approx.	.00716	.001872	.000478	.000121	.000030
$p=3$ $\lambda(1,1,1)$	$O(n^{-2})$	—.031250	—.007813	—.001953	—.000488	—.000122
	$O(n^{-3})$	.012442	.001555	.000194	.000024	.000003
	approx.	—.019	—.0063	—.00176	—.000464	—.000119
$\lambda(1,2,3)$	$O(n^{-2})$	—.026042	—.006510	—.001628	—.000407	—.000102
	$O(n^{-3})$	.010417	.001302	.000163	.000020	.000003
	approx.	—.016	—.0052	—.00146	—.000387	—.000099
$\lambda(1,10,10^2)$	$O(n^{-2})$	.014358	.003590	.000897	.000224	.000056
	$O(n^{-3})$	—.003847	—.000481	—.000060	—.000008	—.000001
	approx.	.0105	.00311	.000837	.000217	.000055
$\lambda(1,0,0)$	$O(n^{-2})$	.031250	.007813	.001953	.000488	.000122
	$O(n^{-3})$	—.005208	—.000651	—.000081	—.000010	—.000001
	approx.	.0260	.00716	.001872	.000478	.000121
$p=4$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	—.037109	—.009277	—.002319	—.000580	—.000145
	$O(n^{-3})$	.021973	.002747	.000343	.000043	.000005
	approx.	—.015	—.0065	—.00198	—.000537	—.000140
$\lambda(1,2,3,4)$	$O(n^{-2})$	—.028906	—.007227	—.001807	—.000452	—.000113
	$O(n^{-3})$	.019570	.002446	.000306	.000038	.000005
	approx.	—.009	—.0048	—.00150	—.000413	—.000108
$\lambda(1,10,10^2,10^3)$	$O(n^{-2})$	.056135	.014034	.003508	.000877	.000219
	$O(n^{-3})$	—.012895	—.001612	—.000201	—.000025	—.000003
	approx.	.043	.0124	.00331	.000852	.000216
$\lambda(1,0,0,0)$	$O(n^{-2})$	.085938	.021484	.005371	.001343	.000336
	$O(n^{-3})$	—.017578	—.002197	—.000275	—.000034	—.000004
	approx.	.068	.0193	.00510	.001308	.000331

Table 4. (continued)

Σ^{-1}		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=5$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	$-.012500$	$-.003125$	$-.000781$	$-.000195$	$-.000049$
	$O(n^{-3})$	.033333	.004167	.000521	.000065	.000008
	approx.	.021	.0010	$-.00026$	$-.000130$	$-.000041$
$\lambda(1,2,\cdots,5)$	$O(n^{-2})$	$-.001389$	$-.000347$	$-.000087$	$-.000022$	$-.000005$
	$O(n^{-3})$	.030648	.003831	.000479	.000060	.000007
	approx.	.029	.0035	.00039	.000038	.000002
$\lambda(1,10,\cdots,10^4)$	$O(n^{-2})$	.142050	.035512	.008878	.002220	.000555
	$O(n^{-3})$	$-.030504$	$-.003813$	$-.000477$	$-.000060$	$-.000007$
	approx.	.112	.0317	.00840	.002160	.000547
$\lambda(1,0,\cdots,0)$	$O(n^{-2})$	.187500	.046875	.011719	.002930	.000732
	$O(n^{-3})$	$-.041667$	$-.005208$	$-.000651$	$-.000081$	$-.000010$
	approx.	.146	.0417	.01107	.002848	.000722
$p=6$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	.058594	.014648	.003662	.000916	.000229
	$O(n^{-3})$	.046568	.005821	.000728	.000091	.000011
	approx.	.105	.0205	.00439	.001006	.000240
$\lambda(1,2,\cdots,6)$	$O(n^{-2})$	.072545	.018136	.004534	.001134	.000283
	$O(n^{-3})$	.043624	.005453	·000682	.000085	.000011
	approx.	.116	.0236	.00522	.001219	.000294
$\lambda(1,10,\cdots,10^5)$	$O(n^{-2})$	.287643	.071911	.017978	.004494	.001124
	$O(n^{-3})$	$-.059523$	$-.007440$	$-.000930$	$-.000116$	$-.000015$
	approx.	.228	.0645	.01705	.00438	.001109
$\lambda(1,0,\cdots,0)$	$O(n^{-2})$	.351563	.087891	.021973	.005493	.001373
	$O(n^{-3})$	$-.081380$	$-.010173$	$-.001272$	$-.000159$	$-.000020$
	approx.	.270	.078	.0207	.00533	.001353

maximum likelihood estimator $\hat{\Sigma}_O^{(1)}$, namely $100\times\{R_1(\hat{\Sigma}_O^{(1)},\Sigma)-R_1(\hat{\Sigma}_H^{(1)},\Sigma)\}/R_1(\hat{\Sigma}_O^{(1)},\Sigma)$, which range above to 8% for $n\geqq 16$. The rates of the reduction of the risks of $\hat{\Sigma}_H^{(1)}$ with respect to $\hat{\Sigma}_{JS}^{(1)}$ range only from -5.6% to 1.6% for $n\geqq 16$ in Table 4.

4.2. Risk of new estimator. Now we shall consider the risk of a new estimator $\hat{\Sigma}^{(1)}$ given in (1.6). We can write the risk difference

$$R_1(\hat{\Sigma}^{(1)},\Sigma)-R_1(\hat{\Sigma}_O^{(1)},\Sigma) \tag{4.15}$$

$$=\frac{b}{n}(\operatorname{tr}\Sigma^{-1})E\left[\frac{\operatorname{tr}S^{-1}}{\operatorname{tr}S^{-2}}\right]-E\left[\log\left|I+\frac{b\operatorname{tr}S^{-1}}{\operatorname{tr}S^{-2}}S^{-1}\right|\right].$$

By the Wishart identity, we get

$$E\left[\frac{\operatorname{tr}S^{-1}}{\operatorname{tr}S^{-2}}\operatorname{tr}\Sigma^{-1}\right]=4E\left[\frac{\operatorname{tr}S^{-3}\operatorname{tr}S^{-1}}{(\operatorname{tr}S^{-2})^2}\right]-2 \tag{4.16}$$

$$+(n-p-1)E\left[\frac{(\operatorname{tr}S^{-1})^2}{\operatorname{tr}S^{-2}}\right].$$

Using (4.16), the risk difference is bounded from above by

$$\frac{b}{n}\left\{4E\left[\frac{\operatorname{tr}S^{-3}\operatorname{tr}S^{-1}}{(\operatorname{tr}S^{-2})^2}\right]-2+\left(\frac{bn}{2}-p-1\right)E\left[\frac{(\operatorname{tr}S^{-1})^2}{\operatorname{tr}S^{-2}}\right]\right\}. \tag{4.17}$$

Note that

$$2\frac{\operatorname{tr} S^{-3}\operatorname{tr} S^{-1}}{(\operatorname{tr} S^{-2})^2}\leqq 1+\frac{(\operatorname{tr} S^{-1})^2}{\operatorname{tr} S^{-2}}, \tag{4.18}$$

where the equality holds if and only if $S^{-1}=\lambda\, diag(1, 0, \cdots, 0)$ except for permutation of the diagonal elements. The upper bound (4.17) is further simplified as

$$\frac{b}{n}\left(\frac{bn}{2}-p+1\right)E\left[\frac{(\operatorname{tr} S^{-1})^2}{\operatorname{tr} S^{-2}}\right]. \tag{4.19}$$

Hence $\hat{\Sigma}^{(1)}$ dominates $\hat{\Sigma}_O^{(1)}$ if $0\leqq b\leqq 2(p-1)/n$ and the minimum of (4.19) is attained by $b=(p-1)/n$. The choice of b is the same as for the Haff's estimator.

To get asymptotic expansion of the risk difference (4.15), we can rewrite it as in (4.6) by

$$\begin{aligned} &\frac{b}{n}\left\{\left(\frac{nb}{2}-p-1\right)E\left[\frac{(\operatorname{tr} S^{-1})^2}{\operatorname{tr} S^{-2}}\right]-2+4E\left[\frac{\operatorname{tr} S^{-3}\operatorname{tr} S^{-1}}{(\operatorname{tr} S^{-2})^2}\right]\right\} \\ &\quad -\frac{b^3}{3}\frac{(\operatorname{tr}\Sigma^{-1})^3\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-2})^3}+O(n^{-4}). \end{aligned} \tag{4.20}$$

To evaluate each expectation asymptotically, we need the following lemma.

LEMMA 4.2. *Let S have a Wishart distribution $W_p(n, \Sigma)$. Then*

$$\begin{aligned} &E\left[\frac{(\operatorname{tr} S^{-1})^2}{\operatorname{tr} S^{-2}}\right] \\ &\quad =\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}+\frac{1}{n}\left[8\frac{\operatorname{tr}\Sigma^{-4}(\operatorname{tr}\Sigma^{-1})^2}{(\operatorname{tr}\Sigma^{-2})^3}-\frac{(\operatorname{tr}\Sigma^{-1})^4}{(\operatorname{tr}\Sigma^{-2})^2}\right. \\ &\quad \left.-8\frac{\operatorname{tr}\Sigma^{-3}\operatorname{tr}\Sigma^{-1}}{(\operatorname{tr}\Sigma^{-2})^2}-\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}+2\right]+O(n^{-2}), \end{aligned} \tag{4.21}$$

$$\begin{aligned} &E\left[\frac{\operatorname{tr} S^{-1}\operatorname{tr} S^{-3}}{(\operatorname{tr} S^{-2})^2}\right] \\ &\quad =\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-2})^2}+\frac{1}{n}\left[24\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}\operatorname{tr}\Sigma^{-4}}{(\operatorname{tr}\Sigma^{-2})^4}\right. \\ &\quad -\frac{2}{(\operatorname{tr}\Sigma^{-2})^3}\{(\operatorname{tr}\Sigma^{-1})^3\operatorname{tr}\Sigma^{-3}+12\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-5} \\ &\quad +4(\operatorname{tr}\Sigma^{-3})^2\}+\frac{1}{(\operatorname{tr}\Sigma^{-2})^2}\{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}+6\operatorname{tr}\Sigma^{-4}\} \\ &\quad \left.+\frac{3(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}\right]+O(n^{-2}). \end{aligned} \tag{4.22}$$

Unlike Lemma 4.1, it seems to be impossible to prove Lemma 4.2 from the Wishart identity only. We obtained it by another method used by Ito [6], Siotani [14], Okamoto [11], Sugiura [15], Fujikoshi [2] and others, that is, for analytic function $f(S)$, it holds

$$E\left[f\left(\frac{1}{n}S\right)\right]=f(\Sigma)+\frac{1}{n}\operatorname{tr}(\Sigma\partial)^2 f(\Lambda)|_{\Lambda=\Sigma}+O(n^{-2}), \tag{4.23}$$

where ∂ is a matrix of differential operators and its (i, j) element is given by $(1/2)(1+\delta_{ij})(\partial/\partial\lambda_{ij})$ for $\Lambda=(\lambda_{ij})$. The following lemma is useful for the repeated application of (4.23).

LEMMA 4.3. *Let E_{ij} $(i\neq j)$ be $p\times p$ matrix having $1/2$ at the (i, j) and (j, i) positions and zero at other positions. Let E_{ii} be diagonal matrix having 1 at i-th diagonal and zero otherwise. Then for any symmetric matrices $A=(a_{ij})$ and $B=(b_{ij})$,*

$$\begin{aligned}&\sum_{i,j}\lambda_i\lambda_j \operatorname{tr} AE_{ij}\operatorname{tr} BE_{ij}=\sum_{i,j}\lambda_i\lambda_j a_{ij}b_{ij}\\ &\sum_{i,j}\lambda_i\lambda_j \operatorname{tr} AE_{ij}BE_{ij}=\frac{1}{2}\sum_{i,j}\lambda_i\lambda_j a_{ij}b_{ij}+\frac{1}{2}\sum_i\lambda_i a_{ii}\sum_j\lambda_j b_{jj}.\end{aligned} \tag{4.24}$$

Applying Lemma 4.2 to (4.20), we get

THEOREM 4.2. *An asymptotic expansion of the difference of risks between new estimator $\hat{\Sigma}^{(1)}$ defined by* (1.6) *with $b=(p-1)/n$ and James and Stein's minimax estimator $\hat{\Sigma}^{(1)}_{JS}$ for L_1 loss is given by*

$$\begin{aligned}R_1(\hat{\Sigma}^{(1)}, \Sigma)-R_1(\hat{\Sigma}^{(1)}_{JS}, \Sigma)=&\frac{p(p^2-1)}{6n^2}+\frac{p-1}{n^2}\left[-2+4\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-2})^2}\right.\\ &\left.-\frac{p+3}{2}\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}\right]+\frac{p-1}{n^3}\left[96\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}\operatorname{tr}\Sigma^{-4}}{(\operatorname{tr}\Sigma^{-2})^4}\right.\\ &-\frac{1}{(\operatorname{tr}\Sigma^{-2})^3}\left\{\left(8+\frac{(p-1)^2}{3}\right)(\operatorname{tr}\Sigma^{-1})^3\operatorname{tr}\Sigma^{-3}+96\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-5}\right.\\ &\left.+32(\operatorname{tr}\Sigma^{-3})^2+4(p+3)(\operatorname{tr}\Sigma^{-1})^2\operatorname{tr}\Sigma^{-4}\right\}\\ &+\frac{1}{(\operatorname{tr}\Sigma^{-2})^2}\left\{4(p+4)\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}+24\operatorname{tr}\Sigma^{-4}+\frac{p+3}{2}(\operatorname{tr}\Sigma^{-1})^4\right\}\\ &\left.+\left(12+\frac{p+3}{2}\right)\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}-p-3\right]+O(n^{-4}).\end{aligned} \tag{4.25}$$

By the inequalities (4.12) and (4.18), the term of $O(n^{-2})$ in (4.25) ranges from

$$-\frac{1}{3}(p-1)(p^2+4p-6) \text{ to } \frac{1}{6}(p-1)(p^2-2p+3). \tag{4.26}$$

The lower bound is obtained by noting that $(\mathrm{tr}\,\Sigma^{-1})^2/\mathrm{tr}\,\Sigma^{-2} \leqq p$ and $\mathrm{tr}\,\Sigma^{-1}\,\mathrm{tr}\,\Sigma^{-3}/(\mathrm{tr}\,\Sigma^{-2})^2 \geqq 1$, where both equalities are satisfied by $\Sigma^{-1}=\lambda I$. The upper bound is the same as for $\hat{\Sigma}_H^{(1)}$ given in (4.13), while the lower bound is smaller than that of $\hat{\Sigma}_H^{(1)}$, and is always negative. Some numerical values are given below. The lower bound is considerably smaller than (4.13).

Ranges of $O(n^{-2})$ in (4.25).

$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
$\left(-2, \frac{1}{2}\right)$	$(-10, 2)$	$\left(-26, \frac{11}{2}\right)$	$(-52, 12)$	$\left(-90, \frac{45}{2}\right)$

The upper bound is approached as $\Sigma^{-1} \to \lambda\, diag(1, 0, \cdots, 0)$ or any permutation of the diagonal elements of it. This shows that $\hat{\Sigma}^{(1)}$ is better than $\hat{\Sigma}_{JS}^{(1)}$ for $\Sigma^{-1}=\lambda I$ and worse for $\Sigma^{-1}=\lambda\, diag(1, 0, \cdots, 0)$, which is the same conclusion as in Haff's estimator $\hat{\Sigma}_H^{(1)}$. However the lower bound is always negative for $\hat{\Sigma}^{(1)}$ and it is not dominated by $\hat{\Sigma}_{JS}^{(1)}$ for any p if n is large.

Some numerical values based on Theorem 4.2 are given in Table 5, in contrast to Table 4. For $n=8$ and $\Sigma^{-1}=\lambda I$, the positive risk differences are observed, which is probably due to the error of asymptotic approximation for small n. It is found that for $\Sigma^{-1}=\lambda I$ and $\lambda\, diag(1, 2, \cdots, p)$, $\hat{\Sigma}^{(1)}$ is better than $\hat{\Sigma}_H^{(1)}$; for $\Sigma^{-1}=\lambda\, diag(1, 10, \cdots, 10^{p-1})$, $\hat{\Sigma}^{(1)}$ is slightly worse than $\hat{\Sigma}_H^{(1)}$; for $\Sigma^{-1}=\lambda\, diag(1, 0, \cdots, 0)$, the asymptotic differences are consistent up to $O(n^{-3})$. The last statement can be confirmed by putting $\Sigma^{-1}=\lambda\, diag(1, 0, \cdots, 0)$ in Theorems 4.1 and 4.2. From Tables 1, 2 and 5, we can compute the rates of the reduction of the risks of $\hat{\Sigma}^{(1)}$ with respect to $\hat{\Sigma}_O^{(1)}$, namely, $100 \times \{R_1(\hat{\Sigma}_O^{(1)}, \Sigma) - R_1(\hat{\Sigma}^{(1)}, \Sigma)\}/R_1(\hat{\Sigma}_O^{(1)}, \Sigma)$ which range above to 20% for $n \geqq 16$. This may be compared with 8% for $\hat{\Sigma}_H^{(1)}$. If we compare the rates of $\hat{\Sigma}^{(1)}$

Table 5. Asymptotic values of $R_1(\hat{\Sigma}^{(1)}, \Sigma) - R_1(\hat{\Sigma}_{JS}^{(1)}, \Sigma)$

Σ^{-1}			$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$	$\lambda(1, 1)$	$O(n^{-2})$	$-.031250$	$-.007813$	$-.001953$	$-.000488$	$-.000122$
		$O(n^{-3})$	.033854	.004232	.000529	.000066	.000008
		approx.	.003	$-.0036$	$-.00142$	$-.000422$	$-.000114$
	$\lambda(1, 2)$	$O(n^{-2})$	$-.018438$	$-.004609$	$-.001152$	$-.000288$	$-.000072$
		$O(n^{-3})$	.008778	.001097	.000137	.000017	.000002
		approx.	$-.0097$	$-.0035$	$-.00102$	$-.000271$	$-.000070$
	$\lambda(1, 10)$	$O(n^{-2})$	.005040	.001260	.000315	.000079	.000020
		$O(n^{-3})$	$-.001753$	$-.000219$	$-.000027$	$-.000003$	$-.000000$
		approx.	.0033	.00104	.000288	.000075	.000019
	$\lambda(1, 0)$	$O(n^{-2})$	.007813	.001953	.000488	.000122	.000031
		$O(n^{-3})$	$-.000651$	$-.000081$	$-.000010$	$-.000001$	$-.000000$
		approx.	.00716	.001872	.000478	.000121	.000030

Table 5. (continued)

Σ^{-1}		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=3$ $\lambda(1,1,1)$	$O(n^{-2})$	−.156250	−.039063	−.009766	−.002441	−.000610
	$O(n^{-3})$	.153646	.019206	.002401	.000300	.000038
	approx.	−.003	−.020	−.0074	−.00214	−.000573
$\lambda(1,2,3)$	$O(n^{-2})$	−.103316	−.025829	−.006457	−.001614	−.000404
	$O(n^{-3})$	.069561	.008695	.001087	.000136	.000017
	approx.	−.034	−.0171	−.0054	−.00148	−.000387
$\lambda(1,10,10^2)$	$O(n^{-2})$	.021771	.005443	.001361	.000340	.000085
	$O(n^{-3})$	−.009179	−.001147	−.000143	−.000018	−.000002
	approx.	.0126	.0043	.00122	.000322	.000083
$\lambda(1,0,0)$	$O(n^{-2})$	.031250	.007813	.001953	.000488	.000122
	$O(n^{-3})$	−.005208	−.000651	−.000081	−.000010	−.000001
	approx.	.0260	.00716	.001872	.000478	.000121
$p=4$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	−.406250	−.101563	−.025391	−.006348	−.001587
	$O(n^{-3})$	.404297	.050537	.006317	.000790	.000099
	approx.	−.002	−.051	−.0191	−.00556	−.001488
$\lambda(1,2,3,4)$	$O(n^{-2})$	−.276042	−.069010	−.017253	−.004313	−.001078
	$O(n^{-3})$	.204965	.025621	.003203	.000400	.000050
	approx.	−.07	−.043	−.0140	−.00391	−.001208
$\lambda(1,10,10^2,10^3)$	$O(n^{-2})$	.066391	.016598	.004149	.001037	.000259
	$O(n^{-3})$	−.027263	−.003408	−.000426	−.000053	−.000007
	approx.	.039	.0132	.00372	.000984	.000253
$\lambda(1,0,0,0)$	$O(n^{-2})$	.085938	.021484	.005371	.001343	.000336
	$O(n^{-3})$	−.017578	−.002197	−.000275	−.000034	−.000004
	approx.	.068	.0193	.00510	.001308	.000331
$p=5$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	−.812500	−.203125	−.050781	−.012695	−.003174
	$O(n^{-3})$	.841667	.105208	.013151	.001644	.000205
	approx.	.03	−.10	−.038	−.0111	−.00297
$\lambda(1,2\cdots,5)$	$O(n^{-2})$	−.556302	−.139075	−.034769	−.008692	−.002173
	$O(n^{-3})$	.435419	.054427	.006803	.000850	.000106
	approx.	−.12	−.085	−.0280	−.00784	−.00207
$\lambda(1,10,\cdots,10^4)$	$O(n^{-2})$	.154470	.038618	.009654	.002414	.000603
	$O(n^{-3})$	−.061226	−.007653	−.000957	−.000120	−.000015
	approx.	.093	.0310	.00870	.00229	.000588
$\lambda(1,0,\cdots,0)$	$O(n^{-2})$	.187500	.046875	.011719	.002930	.000732
	$O(n^{-3})$	−.041667	−.005208	−.000651	−.000081	−.000010
	approx.	.146	.0417	.01107	.002848	.000722
$p=6$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	−1.406250	−.351563	−.087891	−.021973	−.005493
	$O(n^{-3})$	1.529948	.191243	.023905	.002988	.000374
	approx.	.1	−.16	−.064	−.0190	−.00512
$\lambda(1,2,\cdots,6)$	$O(n^{-2})$	−.963619	−.240905	−.060226	−.015057	−.003764
	$O(n^{-3})$	.782396	.097799	.012225	.001528	.000191
	approx.	−.18	−.143	−.048	−.0135	−.00357
$\lambda(1,10,\cdots,10^5)$	$O(n^{-2})$	.301591	.075398	.018849	.004712	.001178
	$O(n^{-3})$	−.116287	−.014536	−.001817	−.000227	−.000028
	approx.	.19	.061	.0170	.00449	.001150
$\lambda(1,0,\cdots,0)$	$O(n^{-2})$	.351563	.087891	.021973	.005493	.001373
	$O(n^{-3})$	−.081380	−.010173	−.001272	−.000159	−.000020
	approx.	.270	.078	.0207	.00533	.001353

with respect to $\hat{\Sigma}^{(1)}_{JS}$, we get the range from -5.6% to 12% in Table 5 for $n\geqq 16$. The rates for $\hat{\Sigma}^{(1)}$ with respect to $\hat{\Sigma}^{(1)}_{H}$ range from -0.4% to 12% for $n\geqq 16$.

5. Risks under L_2 loss

5.1. Risk of Haff's estimator. We shall now consider the estimator

$$\hat{\Sigma}_H^{(2)}=\frac{1}{n+p+1}\left[S+\frac{b}{\operatorname{tr} S^{-1}}I\right] \tag{5.1}$$

proposed by Haff [5], where C is taken to be I in (1.4) without loss of generality. The loss function is given by (1.2), throughout Section 5. It is known by Haff [5] that the best scalar multiple of S is given by $\hat{\Sigma}_O^{(2)}=S/(n+p+1)$. The difference of risks can be written by

$$\begin{aligned} &R_2(\hat{\Sigma}_H^{(2)},\Sigma)-R_2(\hat{\Sigma}_O^{(2)},\Sigma)\\ &=\frac{b}{2(n+p+1)^2}E\left[\frac{2}{\operatorname{tr} S^{-1}}\operatorname{tr}\{S\Sigma^{-2}-(n+p+1)\Sigma^{-1}\}+\frac{b\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr} S^{-1})^2}\right]. \end{aligned} \tag{5.2}$$

To evaluate each expectation, we need the following equations due to Haff [5] derived from the Wishart identity.

$$E\left[\frac{\operatorname{tr} S\Sigma^{-2}}{\operatorname{tr} S^{-1}}\right]=nE\left[\frac{\operatorname{tr}\Sigma^{-1}}{\operatorname{tr} S^{-1}}\right]+2E\left[\frac{\operatorname{tr} S^{-1}\Sigma^{-1}}{(\operatorname{tr} S^{-1})^2}\right]. \tag{5.3}$$

$$E\left[\frac{\operatorname{tr} S^{-1}\Sigma^{-1}}{(\operatorname{tr} S^{-1})^2}\right]=(n-p-2)E\left[\frac{\operatorname{tr} S^{-2}}{(\operatorname{tr} S^{-1})^2}\right]+4E\left[\frac{\operatorname{tr} S^{-3}}{(\operatorname{tr} S^{-1})^3}\right]-1. \tag{5.4}$$

$$E\left[\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr} S^{-1})^2}\right]=4E\left[\frac{\operatorname{tr} S^{-2}\Sigma^{-1}}{(\operatorname{tr} S^{-1})^3}\right]+(n-p-1)E\left[\frac{\operatorname{tr} S^{-1}\Sigma^{-1}}{(\operatorname{tr} S^{-1})^2}\right]. \tag{5.5}$$

Together with (4.4) and Lemma 4.1, we can rewrite (5.2) as

$$\begin{aligned} &\frac{b}{(n+p+1)^2}\Bigg[-n(p+1)+\left\{2n-4p-4-bn(p+1)+\frac{bn^2}{2}\right\}\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}+(p+1)^2\\ &\qquad -8\frac{\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-1})^3}+3(bn+4)\frac{(\operatorname{tr}\Sigma^{-2})^2}{(\operatorname{tr}\Sigma^{-1})^4}\Bigg]+O(n^{-4}). \end{aligned} \tag{5.6}$$

Assuming that $b=O(1/n)$, the term of $O(n^{-2})$ in (5.6) is

$$-n(p+1)+2n\left(1+\frac{bn}{4}\right)\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}\leqq -n(p+1)+2n\left(1+\frac{bn}{4}\right). \tag{5.7}$$

The condition that the R. H. S. of (5.7) is negative is given by $b\leqq 2(p-1)/n$ which is in contrast with the exact result $b\leqq 2(p-1)/(n-p+3)$ in Haff [5]. The equality in (5.7) is attained by $\Sigma^{-1}=\lambda\, diag(1,0,\cdots,0)$, for which the value of (5.6) is minimized by

$$b=\frac{(n-p+1)(p-1)}{n^2-2(p-2)n}=\frac{1}{n}(p-1)\left(1+\frac{p-3}{n}\right)+O(n^{-3}). \tag{5.8}$$

Again the result is the same as the optimal choice $b=(p-1)/(n-p+3)$ by Haff [5] asymptotically. Note that

$$
\begin{aligned}
(5.9)\qquad & R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)-R_2(\hat{\Sigma}_{O}^{(2)}, \Sigma) \\
&= -\frac{p(p^2-1)}{6n^2}+\frac{p(p+1)^2(p-1)}{2n^3}+O(n^{-4}).
\end{aligned}
$$

We get

THEOREM 5.1. *An asymptotic expansion of the difference of risks between Haff's estimator* $\hat{\Sigma}_H^{(2)}$ *defined by* (5.1) *and James and Stein's minimax estimator* $\hat{\Sigma}_{JS}^{(2)}$ *for* L_2 *loss is given by*

$$
\begin{aligned}
& R_2(\hat{\Sigma}_H^{(2)}, \Sigma)-R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)=\frac{p-1}{6n^2}\left[(p+1)(p-6)+3(p+3)\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}\right] \\
(5.10)\qquad & +\frac{p-1}{n^3}\left[\frac{1}{2}(p+1)^2(6-p)-\Delta(p+1)+3(p+3)\frac{(\operatorname{tr}\Sigma^{-2})^2}{(\operatorname{tr}\Sigma^{-1})^4}\right. \\
& \left.+(p+1)(\Delta-2p-6)\frac{\operatorname{tr}\Sigma^{-2}}{(\operatorname{tr}\Sigma^{-1})^2}-8\frac{\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-1})^3}\right]+O(n^{-4}),
\end{aligned}
$$

where $b=(p-1)(1+\Delta/n)/n$ *and an optimal choice of* Δ *is* $p-3$.

The term of $O(n^{-2})$ in (5.10) is the same as that of $R_1(\hat{\Sigma}_H^{(1)}, \Sigma)-R_1(\hat{\Sigma}_{JS}^{(1)}, \Sigma)$ in Theorem 4.1. However the term of $O(n^{-3})$ is different which yields poor asymptotic approximations as can be seen in Table 6 compared with Table 4. For instance, when $n=16$, $p=6$ and $\Sigma^{-1}=\lambda I$, the approximate value of $R_2(\hat{\Sigma}_H^{(2)}, \Sigma)-R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)$ is equal to -0.032. However we can not say that this is negative, because of the error that may arise in the asymptotic approximations. The corresponding value for $\hat{\Sigma}_H^{(1)}$ is 0.0205 from Table 4 and we are certain that this is positive. One might think that an asymptotic expansion with respect to $n+p+1$ is better for $\hat{\Sigma}_H^{(2)}$, because of (3.7). We can easily rewrite (5.10) in terms of powers of $n+p+1$ instead of n. For the above example we get the term of order $(n+p+1)^{-2}$ is equal to 0.007089 and the term of order $(n+p+1)^{-3}$ is equal to -0.011290. The approximate value is -0.004201, which is different from -0.032. However still the second term is larger than the first in absolute value. If we increase $n=128$ in this example, the approximate value is 0.000150, the corresponding value in Table 6 is 0.000138. Hence these values are reliable. The fact that the asymptotic approximations are better for L_1 loss than for L_2 loss, is ascertained again. From Tables 3 and 6, the rates of the reduction of the risks of $\hat{\Sigma}_H^{(2)}$ with respect to $\hat{\Sigma}_O^{(2)}$ can be computed, the range of which is given by 0%~4% for $n\geqq 32$ in Table 6.

Table 6. Asymptotic values of $R_2(\hat{\Sigma}_H^{(2)}, \Sigma)-R_2(\hat{\Sigma}_{JS}^{(2)}, \Sigma)$

Σ^{-1}		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$ $\lambda(1,1)$	$O(n^{-2})$	—.011719	—.002930	—.000732	—.000183	—.000046
	$O(n^{-3})$	.012207	.001526	.000191	.000024	.000003
	approx.	.0005	—.0014	—.00054	—.000159	—.000043
$\lambda(1,2)$	$O(n^{-2})$	—.009549	—.002387	—.000597	—.000149	—.000037
	$O(n^{-3})$	.009042	.001130	.000141	.000018	.000002
	approx.	—.0005	—.0013	—.00046	—.000132	—.000035
$\lambda(1,10)$	$O(n^{-2})$	.001356	.000339	.000085	.000021	.000005
	$O(n^{-3})$	—.004123	—.000515	—.000064	—.000008	—.000001
	approx.	—.0028	—.00018	.000020	.000013	.000004
$\lambda(1,0)$	$O(n^{-2})$	.007813	.001953	.000488	.000122	.000031
	$O(n^{-3})$	—.009766	—.001221	—.000153	—.000019	—.000002
	approx.	—.0020	.0007	.00034	.000103	.000028
$p=3$ $\lambda(1,1,1)$	$O(n^{-2})$	—.031250	—.007813	—.001953	—.000488	—.000122
	$O(n^{-3})$	.035590	.004449	.000556	.000070	.000009
	approx.	.004	—.0034	—.00134	—.000419	—.000113
$\lambda(1,2,3)$	$O(n^{-2})$	—.026042	—.006510	—.001628	—.000407	—.000102
	$O(n^{-3})$	.026259	.003282	.000410	.000051	.000006
	approx.	.0002	—.0032	—.00122	—.000356	—.000095
$\lambda(1,10,10^2)$	$O(n^{-2})$	.014358	.003590	.000897	.000224	.000056
	$O(n^{-3})$	—.035581	—.004448	—.000556	—.000069	—.000009
	approx.	—.021	—.0009	.00034	.000155	.000047
$\lambda(1,0,0)$	$O(n^{-2})$	.031250	.007813	.001953	.000488	.000122
	$O(n^{-3})$	—.054688	—.006836	—.000854	—.000107	—.000013
	approx.	—.023	.0010	.00110	.00038	.000109
$p=4$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	—.037109	—.009277	—.002319	—.000580	—.000145
	$O(n^{-3})$	.026733	.003342	.000418	.000052	.000007
	approx.	—.010	—.0059	—.00190	—.000528	—.000138
$\lambda(1,2,3,4)$	$O(n^{-2})$	—.028906	—.007227	—.001807	—.000452	—.000113
	$O(n^{-3})$	.009316	.001165	.000146	.000018	.000002
	approx.	—.0196	—.0061	—.00166	—.000433	—.000111
$\lambda(1,10,10^2,10^3)$	$O(n^{-2})$	.056135	.014034	.003508	.000877	.000219
	$O(n^{-3})$	—.146300	—.018288	—.002286	—.000286	—.000036
	approx.	—.09	—.004	.0012	.00059	.000184
$\lambda(1,0,0,0)$	$O(n^{-2})$	.085938	.021484	.005371	.001343	.000336
	$O(n^{-3})$	—.187500	—.023438	—.002930	—.000366	—.000046
	approx.	—.10	—.002	.0024	.00098	.000290
$p=5$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	—.012500	—.003125	—.000781	—.000195	—.000049
	$O(n^{-3})$	—.079375	—.009922	—.001240	—.000155	—.000019
	approx.	—.092	—.0130	—.0020	—.00035	—.000068
$\lambda(1,2,\cdots,5)$	$O(n^{-2})$	—.001389	—.000347	—.000087	—.000022	—.000005
	$O(n^{-3})$	—.106505	—.013313	—.001664	—.000208	—.000026
	approx.	—.11	—.014	—.0018	—.00023	—.000031
$\lambda(1,10,\cdots,10^4)$	$O(n^{-2})$	.142050	.035512	.008878	.002220	.000555
	$O(n^{-3})$	—.410155	—.051269	—.006409	—.000801	—.000100
	approx.	—.27	—.016	.0025	.00142	.00046
$\lambda(1,0,\cdots,0)$	$O(n^{-2})$	.187500	.046875	.011719	.002930	.000732
	$O(n^{-3})$	—.484375	—.060547	—.007568	—.000946	—.000118
	approx.	—.30	—.014	.0042	.00198	.00061

Table 6. (continued)

Σ^{-1}		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=6$ $\lambda(1,\cdots,1)$	$O(n^{-2})$	.058594	.014648	.003662	.000916	.000229
	$O(n^{-3})$	—.370822	—.046353	—.005794	—.000724	—.000091
	approx.	—.31	—.032	—.0021	.00019	.000138
$\lambda(1,2,\cdots,6)$	$O(n^{-2})$	.072545	.018136	.004534	.001134	.000283
	$O(n^{-3})$	—.409160	—.051145	—.006393	—.000799	—.000100
	approx.	—.38	—.033	—.0019	.00033	.00018
$\lambda(1,10,\cdots,10^5)$	$O(n^{-2})$	.287643	.071911	.017978	.004494	.001124
	$O(n^{-3})$	—.924538	—.115567	—.014446	—.001806	—.000226
	approx.	—.64	—.04	.004	.0027	.00090
$\lambda(1,0,\cdots,0)$	$O(n^{-2})$	.351563	.087891	.021973	.005493	.001373
	$O(n^{-3})$	—1.044922	—.130615	—.016327	—.002041	—.000255
	approx.	—.7	—.04	.006	.0035	.00112

5.2. Risk of new estimator. Finally we shall consider the estimator (1.7) for $C=I$ without loss of generality, namely,

$$\hat{\Sigma}^{(2)}=\frac{1}{n+p+1}\left(S+\frac{b\operatorname{tr}S^{-1}}{\operatorname{tr}S^{-2}}I\right). \tag{5.11}$$

The risk difference can be written by

$$\begin{aligned}R_2(\hat{\Sigma}^{(2)},\Sigma)-R_2(\hat{\Sigma}^{(2)}_0,\Sigma)=\frac{b}{(n+p+1)^2}E\Big[&\frac{\operatorname{tr}S^{-1}}{\operatorname{tr}S^{-2}}\operatorname{tr}\{S\Sigma^{-1}-(n+p+1)I\}\Sigma^{-1}\\&+\frac{b}{2}\left(\frac{\operatorname{tr}S^{-1}}{\operatorname{tr}S^{-2}}\right)^2\operatorname{tr}\Sigma^{-2}\Big].\end{aligned} \tag{5.12}$$

Each expectation can be computed by the following relations obtained from the Wishart identity in Haff [5].

$$\begin{aligned}E\left[\frac{\operatorname{tr}S^{-1}\operatorname{tr}S^{-2}\Sigma^{-1}}{(\operatorname{tr}S^{-2})^2}\right]=&2E\left[4\frac{\operatorname{tr}S^{-1}\operatorname{tr}S^{-5}}{(\operatorname{tr}S^{-2})^3}-\frac{\operatorname{tr}S^{-4}}{(\operatorname{tr}S^{-2})^2}\right]\\&-2E\left[\frac{(\operatorname{tr}S^{-1})^2}{\operatorname{tr}S^{-2}}\right]+(n-p-3)E\left[\frac{\operatorname{tr}S^{-1}\operatorname{tr}S^{-3}}{(\operatorname{tr}S^{-2})^2}\right].\end{aligned} \tag{5.13}$$

$$E\left[\frac{\operatorname{tr}S^{-1}\Sigma^{-1}}{\operatorname{tr}S^{-2}}\right]=n-p-2-E\left[\frac{(\operatorname{tr}S^{-1})^2}{\operatorname{tr}S^{-2}}\right]+4E\left[\frac{\operatorname{tr}S^{-4}}{(\operatorname{tr}S^{-2})^2}\right]. \tag{5.14}$$

$$\begin{aligned}E\left[\frac{(\operatorname{tr}S^{-1})^2}{(\operatorname{tr}S^{-2})^2}\operatorname{tr}S^{-1}\Sigma^{-1}\right]=&(n-p-2)E\left[\frac{(\operatorname{tr}S^{-1})^2}{\operatorname{tr}S^{-2}}\right]-E\left[\frac{(\operatorname{tr}S^{-1})^4}{(\operatorname{tr}S^{-2})^2}\right]\\&+8E\left[\frac{(\operatorname{tr}S^{-1})^2\operatorname{tr}S^{-4}}{(\operatorname{tr}S^{-2})^3}\right]-4E\left[\frac{\operatorname{tr}S^{-1}\operatorname{tr}S^{-3}}{(\operatorname{tr}S^{-2})^2}\right].\end{aligned} \tag{5.15}$$

For example, the first term of the expectation in the R. H. S. of (5.12) can be expressed by the Whisart identity as

$$nE\left[\frac{\operatorname{tr} S^{-1}}{\operatorname{tr} S^{-2}}\operatorname{tr}\Sigma^{-1}\right]-(n-p-1)(n+p+1)E\left[\frac{(\operatorname{tr} S^{-1})^2}{\operatorname{tr} S^{-2}}\right]+4E\left[\frac{\operatorname{tr} S^{-1}\operatorname{tr} S^{-2}\Sigma^{-1}}{(\operatorname{tr} S^{-2})^2}\right]$$

$$-2E\left[\frac{\operatorname{tr} S^{-1}\Sigma^{-1}}{\operatorname{tr} S^{-2}}\right]-4(n+p+1)E\left[\frac{\operatorname{tr} S^{-1}\operatorname{tr} S^{-3}}{(\operatorname{tr} S^{-2})^2}\right]+2(n+p+1),$$

which can be reduced further by (5.13), (5.14) and (4.16). Assuming that $b=O(n^{-1})$, we can finally rewrite (5.12) as

$$(5.16)\quad \frac{b}{(n+p+1)^2}\Big[-2n+n\Big(\frac{b}{2}n-p-1\Big)E\left[\frac{(\operatorname{tr} S^{-1})^2}{\operatorname{tr} S^{-2}}\right]+4nE\left[\frac{\operatorname{tr} S^{-1}\operatorname{tr} S^{-3}}{(\operatorname{tr} S^{-2})^2}\right]$$
$$+4p+6+\left\{(p+1)^2-6-\frac{b}{2}n(2p+3)\right\}\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}-16\frac{\operatorname{tr}\Sigma^{-4}}{(\operatorname{tr}\Sigma^{-2})^2}$$
$$-4(bn+2p+4)\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-2})^2}+32\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-5}}{(\operatorname{tr}\Sigma^{-2})^3}$$
$$+8bn\frac{(\operatorname{tr}\Sigma^{-1})^2\operatorname{tr}\Sigma^{-4}}{(\operatorname{tr}\Sigma^{-2})^3}-\frac{bn}{2}\frac{(\operatorname{tr}\Sigma^{-1})^4}{(\operatorname{tr}\Sigma^{-2})^2}\Big]+O(n^{-4}).$$

By (4.18) the term of $O(n^{-2})$ in (5.16) is bounded from above by

$$(5.17)\quad \left\{\frac{1}{2}bn^2-n(p+1)+2n\right\}\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}},$$

which is negative only if $b\leqq 2(p-1)/n$. The upper bound (5.17) is attained for $\Sigma^{-1}=\lambda\, diag(1,0,\cdots,0)$ or any permutation of the diagonal elements of it. For this Σ^{-1}, the risk difference (5.16) can be written by

$$(5.18)\quad \frac{b}{(n+p+1)^2}\left\{\frac{1}{2}bn^2-n(p-1)+(p-1)^2-(p-2)bn\right\}+O(n^{-4}),$$

which is minimized by $b=(p-1)(1+\Delta/n)/n$ for $\Delta=p-3$ asymptotically. This optimal choice of b is the same as for $\hat{\Sigma}_H^{(2)}$. Using (5.9), we get

THEOREM 5.2. *An asymptotic expansion of the difference of risks between estimator* $\hat{\Sigma}^{(2)}$ *defined by* (5.11) *with* $b=(p-1)(1+\Delta/n)/n$ *and James and Stein's estimator* $\hat{\Sigma}_{JS}^{(2)}$ *for* L_2 *loss is given by*

$$R_2(\hat{\Sigma}^{(2)},\Sigma)-R_2(\hat{\Sigma}_{JS}^{(2)},\Sigma)$$
$$(5.19)\quad =\frac{p-1}{6n^2}\left[(p-3)(p+4)-3(p+3)\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}+24\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-2})^2}\right]$$
$$+\frac{p-1}{n^3}\Big[-\frac{1}{2}(p+1)(p^2+p-14)-2\Delta+(p^2+6p+13-2\Delta)\frac{(\operatorname{tr}\Sigma^{-1})^2}{\operatorname{tr}\Sigma^{-2}}$$
$$+4(\Delta-4p-1)\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}}{(\operatorname{tr}\Sigma^{-2})^2}+8\frac{\operatorname{tr}\Sigma^{-4}}{(\operatorname{tr}\Sigma^{-2})^2}+2\frac{(\operatorname{tr}\Sigma^{-1})^4}{(\operatorname{tr}\Sigma^{-2})^2}$$

$$-\frac{4}{(\operatorname{tr}\Sigma^{-2})^3}\{-(p-5)(\operatorname{tr}\Sigma^{-1})^2\operatorname{tr}\Sigma^{-4}+16\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-5}$$

$$+2(\operatorname{tr}\Sigma^{-1})^3\operatorname{tr}\Sigma^{-3}+8(\operatorname{tr}\Sigma^{-3})^2\}$$

$$\left.+96\frac{\operatorname{tr}\Sigma^{-1}\operatorname{tr}\Sigma^{-3}\operatorname{tr}\Sigma^{-4}}{(\operatorname{tr}\Sigma^{-2})^4}\right]+O(n^{-4}).$$

An optimal choice of Δ is given by $p-3$.

Note that the term of $O(n^{-2})$ for $\hat{\Sigma}^{(2)}$ in (5.19) is the same as the corresponding term of Theorem 4.2 for $\hat{\Sigma}^{(1)}$. Also the term of $O(n^{-2})$ for $\hat{\Sigma}_H^{(2)}$ in Theorem 5.1 is the same as that of Theorem 4.1 for $\hat{\Sigma}_H^{(1)}$. Hence the ranges of $O(n^{-2})$ in (4.13) and (4.26) hold also for $\hat{\Sigma}_H^{(2)}$ and $\hat{\Sigma}^{(2)}$. Asymptotically, the range for $\hat{\Sigma}^{(2)}$ is wider below than that for $\hat{\Sigma}_H^{(2)}$. Some numerical values of the risk differences for $\hat{\Sigma}^{(2)}$ are shown in Table 7. Comparing with Table 6, we can see that for $\Sigma^{-1}=\lambda I$ and $\lambda\, diag(1,2,\cdots,p)$, $\hat{\Sigma}^{(2)}$ is better considerably; for $\Sigma^{-1}=\lambda\, diag(1,10,\cdots,10^{p-1})$, $\hat{\Sigma}_H^{(2)}$ is better and for $\Sigma^{-1}=\lambda\, diag(1,0,\cdots,0)$, they are the same. The last statement can be checked by putting $\Sigma^{-1}=\lambda\, diag(1,0,\cdots,0)$ in (5.10) and (5.19). Comparing with Table 5, we can see that the asymptotic approximations are poor for $\hat{\Sigma}^{(2)}$. Again the positive values for $\Sigma^{-1}=\lambda I$ and negative values for $\Sigma^{-1}=$

Table 7. Asymptotic values of $R_2(\hat{\Sigma}^{(2)},\Sigma)-R_2(\hat{\Sigma}_{JS}^{(2)},\Sigma)$

Σ^{-1}			$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=2$		$O(n^{-2})$	—.031250	—.007813	—.001953	—.000488	—.000122
	$\lambda(1,1)$	$O(n^{-3})$	.039063	.004883	.000610	.000076	.000010
		approx.	.008	—.0029	—.00134	—.000412	—.000113
		$O(n^{-2})$	—.018438	—.004609	—.001152	—.000288	—.000072
	$\lambda(1,2)$	$O(n^{-3})$	.014372	.001796	.000225	.000028	.000004
		approx.	—.004	—.0028	—.00093	—.000260	—.000069
		$O(n^{-2})$	.005040	.001260	.000315	.000079	.000020
	$\lambda(1,10)$	$O(n^{-3})$	—.007077	—.000885	—.000111	—.000014	—.000002
		approx.	—.0020	.00038	.00020	.000065	.000018
		$O(n^{-2})$	.007813	.001953	.000488	.000122	.000031
	$\lambda(1,0)$	$O(n^{-3})$	—.009766	—.001221	—.000153	—.000019	—.000002
		approx.	—.0020	.0007	.00034	.000103	.000028
$p=3$		$O(n^{-2})$	—.156250	—.039063	—.009766	—.002441	—.000610
	$\lambda(1,1,1)$	$O(n^{-3})$	.236979	.029622	.003703	.000463	.000058
		approx.	.08	—.009	—.0061	—.00198	—.000552
		$O(n^{-2})$	—.103316	—.025829	—.006457	—.001614	—.000404
	$\lambda(1,2,3)$	$O(n^{-3})$	.128827	.016103	.002013	.000252	.000031
		approx.	.26	—.010	—.0044	—.00136	—.000372
		$O(n^{-2})$	.021771	.005443	.001361	.000340	.000085
	$\lambda(1,10,10^2)$	$O(n^{-3})$	—.042109	—.005264	—.000658	—.000082	—.000010
		approx.	—.020	.0002	.00070	.000258	.000075
		$O(n^{-2})$	.031250	.007813	.001953	.000488	.000122
	$\lambda(1,0,0)$	$O(n^{-3})$	—.054688	—.006836	—.000854	—.000107	—.000013
		approx.	—.023	.0010	.00110	.00038	.000109

Table 7. (continued)

Σ^{-1}		$n=8$	$n=16$	$n=32$	$n=64$	$n=128$
$p=4$ $\lambda(1, \cdots, 1)$	$O(n^{-2})$	$-.406250$	$-.101563$	$-.025391$	$-.006348$	$-.001587$
	$O(n^{-3})$	$.708984$	$.088623$	$.011078$	$.001385$	$.000173$
	approx.	$.30$	$-.013$	$-.014$	$-.0050$	$-.00141$
$\lambda(1, 2, 3, 4)$	$O(n^{-2})$	$-.276042$	$-.069010$	$-.017253$	$-.004313$	$-.001078$
	$O(n^{-3})$	$.419957$	$.052495$	$.006562$	$.000820$	$.000103$
	approx.	$.14$	$-.017$	$-.0107$	$-.00349$	$-.00098$
$\lambda(1, 10, 10^2, 10^3)$	$O(n^{-2})$	$.066391$	$.016598$	$.004149$	$.001037$	$.000259$
	$O(n^{-3})$	$-.155830$	$-.019479$	$-.002435$	$-.000304$	$-.000038$
	approx.	$-.09$	$-.003$	$.0017$	$.00073$	$.000221$
$\lambda(1, 0, 0, 0)$	$O(n^{-2})$	$.085938$	$.021484$	$.005371$	$.001343$	$.000336$
	$O(n^{-3})$	$-.187500$	$-.023438$	$-.002930$	$-.000366$	$-.000046$
	approx.	$-.10$	$-.002$	$.0024$	$.00098$	$.000290$
$p=5$ $\lambda(1, \cdots, 1)$	$O(n^{-2})$	$-.812500$	$-.203125$	$-.050781$	$-.012695$	$-.003174$
	$O(n^{-3})$	1.590625	$.198828$	$.024854$	$.003107$	$.000388$
	approx.	$.8$	$-.004$	$-.026$	$-.0096$	$-.00279$
$\lambda(1, 2, \cdots, 5)$	$O(n^{-2})$	$-.556302$	$-.139075$	$-.034769$	$-.008692$	$-.002173$
	$O(n^{-3})$	$.968478$	$.121060$	$.015132$	$.001892$	$.000236$
	approx.	$.41$	$-.02$	$-.020$	$-.0068$	$-.00194$
$\lambda(1, 10, \cdots, 10^4)$	$O(n^{-2})$	$.154470$	$.038618$	$.009654$	$.002414$	$.000603$
	$O(n^{-3})$	$-.421797$	$-.052725$	$-.006591$	$-.000824$	$-.000103$
	approx.	$-.27$	$-.014$	$.0031$	$.00159$	$.00050$
$\lambda(1, 0, \cdots, 0)$	$O(n^{-2})$	$.187500$	$.046875$	$.011719$	$.002930$	$.000732$
	$O(n^{-3})$	$-.484375$	$-.060547$	$-.007568$	$-.000946$	$-.000118$
	approx.	$-.30$	$-.014$	$.0042$	$.00198$	$.00061$
$p=6$ $\lambda(1, \cdots, 1)$	$O(n^{-2})$	-1.406250	$-.351563$	$-.087891$	$-.021973$	$-.005493$
	$O(n^{-3})$	3.040365	$.380046$	$.047506$	$.005938$	$.000742$
	approx.	1.6	$.03$	$-.040$	$-.0160$	$-.00475$
$\lambda(1, 2, \cdots, 6)$	$O(n^{-2})$	$-.963619$	$-.240905$	$-.060226$	$-.015057$	$-.003764$
	$O(n^{-3})$	1.865664	$.233208$	$.029151$	$.003644$	$.000455$
	approx.	$.9$	$-.01$	$-.031$	$-.0114$	$-.00331$
$\lambda(1, 10, \cdots, 10^5)$	$O(n^{-2})$	$.301591$	$.075398$	$.018849$	$.004712$	$.001178$
	$O(n^{-3})$	$-.936962$	$-.117120$	$-.014640$	$-.001830$	$-.000229$
	approx.	$-.64$	$-.04$	$.004$	$.0029$	$.00095$
$\lambda(1, 0, \cdots, 0)$	$O(n^{-2})$	$.351563$	$.087891$	$.021973$	$.005493$	$.001373$
	$O(n^{-3})$	-1.044922	$-.130615$	$-.016327$	$-.002041$	$-.000255$
	approx.	$-.7$	$-.04$	$.006$	$.0035$	$.00112$

$\lambda\, diag(1, 0, \cdots, 0)$ when $n=8$ or 16 in Table 7 are doubtful. From Tables 3 and 7, we can compute the rates of the reduction of the risks for $\hat{\Sigma}^{(2)}$ with respect to $\hat{\Sigma}^{(2)}_O$, which range above to 11% for $n \geqq 32$. This may be compared with 4% for $\hat{\Sigma}^{(2)}_H$. Comparing the rates for $\hat{\Sigma}^{(2)}$ with respect to $\hat{\Sigma}^{(2)}_H$, the range is given by $-0.2\% \sim 7\%$ for $n \geqq 32$ in Table 7. Also the rates for $\hat{\Sigma}^{(2)}$ with respect to $\hat{\Sigma}^{(2)}_{JS}$ range $-1.2\% \sim 8\%$ while the rates for $\hat{\Sigma}^{(2)}_H$ with respect to $\hat{\Sigma}^{(2)}_{JS}$ range only $-1.2\% \sim 0.8\%$ for $n \geqq 32$.

Acknowledgment. The authors wish to express their gratitude to the referee for his careful reading the manuscript and useful comments.

References

[1] Abramowitz, M. and Stegun, I. A., Handbook of Mathematical Functions. NBS App. Math. Ser. 55 1964.

[2] Fujikoshi, Y., Asymptotic expansions of the distributions of test statistics in multivariate analysis. J. Sci. Hiroshima Univ. Ser. A-I 34 (1970) 73-144.

[3] Giri, N., Multivariate Statistical Inference. Academic, 1977.

[4] Haff, L. R., An identity for the Wishart distribution with applications. J. Multivariate Anal. 9 (1979) 531-544.

[5] Haff, L. R., Empirical Bayes estimation of the multivariate normal covariance matrix. Ann. Statist. 8 (1980) 586-597.

[6] Ito, K., Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. Ann. Math. Statist. 27 (1956) 1091-1105.

[7] James, W. and Stein, C., Estimation with quadratic loss. Fourth Berkeley Symp. Math. Statist. Probability, Univ. California Press Berkeley 1961.

[8] Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics. Vol. 2 3rd Edition, Griffin 1973.

[9] Kshirsagar, A. M., Multivariate Analysis. Marcel Dekker 1972.

[10] Nagao, H., On some test criteria for covariance matrix. Ann. Statist. 1 (1973) 700-709.

[11] Okamoto, M., An asymptotic expansion for the distribution of the linear discriminant function. Ann. Math. Statist. 34 (1963) 1286-1301.

[12] Selliah, J., Estimation and testing problems in a Wishart distribution. Ph. D. thesis, Dept. Statist. Stanford Univ. 1964.

[13] Sharma, D., An estimator of normal covariance matrix. Calcutta Statist. Assoc. Bulletin 29 (1980) 161-167.

[14] Siotani, M., On the distribution of the Hotelling's T^2-statistic. Ann. Inst. Statist. Math. 8 (1957) 1-14.

[15] Sugiura, N., Derivatives of the characteristic root of a symmetric or a hermitian matrix with two applications in multivariate analysis. Commun. Statist. 1 (1973) 393-417.

Institute of Mathematics
University of Tsukuba
Ibaraki 305 Japan