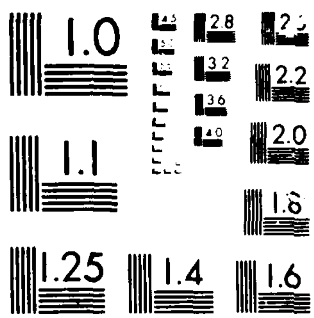
AD-A177 099 INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATOR OF THE VARIANCE-COVARIA..(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS B K SINHA NOV 86 TR-86-35 1/1

UNCLASSIFIED AFOSR-TR-87-0096 F49620-85-C-0008 F/G 12/1 NL

END
DATE
FILMED
4 87


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 87-0096	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Inadmissibility of the best equivariant estimators of the variance-covariance matrix, the precision matrix, and the generalized variance: A survey		5. TYPE OF REPORT & PERIOD COVERED Technical - November 1986
		6. PERFORMING ORG. REPORT NUMBER 86-33
7. AUTHOR(s) Bimal K. Sinha		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis Fifth Floor Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Department of the Air Force Bolling Air Force Base, DC 20332		12. REPORT DATE November 1986
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE FEB 2 4 1987 S D E

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Co

Best equivariant estimator, squared error loss, entropy loss, generalized variance, MANOVA test, Roy's maximum root test, testimator, minimax estimator, Wishart distribution, Wishart identity.

20. ABSTRACT (Co

DTIC FILE COPY

Based on a data matrix $X = (X_1,\ldots,X_k):p \times k$ with independent columns $X_i \sim N_p(\xi_i,\Sigma)$, and an independent $p \times p$ Wishart matrix $S \sim W_p(n,\Sigma)$, procedures to obtain estimators dominating the best equivariant estimators of Σ, Σ^{-1} and $|\Sigma|$ under various loss functions are reviewed.

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A177 099

AFOSR-TR- 87-00[illegible]0

INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATORS OF THE VARIANCE-COVARIANCE MATRIX, THE PRECISION MATRIX, AND THE GENERALIZED VARIANCE: A SURVEY

Bimal K. Sinha[1]
University of Maryland Baltimore County
and
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

87 2 20 211

INADMISSIBILITY OF THE BEST EQUIVARIANT
ESTIMATORS OF THE VARIANCE-COVARIANCE MATRIX,
THE PRECISION MATRIX, AND THE GENERALIZED VARIANCE:
A SURVEY

Bimal K. Sinha[1]
University of Maryland Baltimore County
and
University of Pittsburgh

November 1986

Technical Report No. 86-35

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

[1] Research supported by the Air Force Office of Scientific Research (AFOSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATORS OF THE VARIANCE-COVARIANCE MATRIX, THE PRECISION MATRIX, AND THE GENERALIZED VARIANCE:

A SURVEY

Bimal K. Sinha[1]
University of Maryland Baltimore County

and

University of Pittsburgh

ABSTRACT

Based on a data matrix $X = (X_1,\dots,X_k): p \times k$ with independent columns $X_i \sim N_p(\xi_i,\Sigma)$, and an independent $p \times p$ Wishart matrix $S \sim W_p(n,\Sigma)$, procedures to obtain estimators dominating the best equivariant estimators of Σ, Σ^{-1} and $|\Sigma|$ under various loss functions are reviewed.

AMS 1980 Subject Classification. Primary 62F10; Secondary 62H99.

Key Words: Best equivariant estimator, squared error loss, entropy loss, generalized variance, MANOVA test, Roy's maximum root test, testimator, minimax estimator, Wishart distribution, Wishart identity.

[1]Research supported by the Air Force Office of Scientific Research (AFOSR) under Contract F49620-85-C-0008.

1. Introduction

A multivariate normal linear model in its canonical form consists of a data matrix $X = (X_1,\dots,X_k):p \times k$ with independent columns $X_i \sim N_p(\xi_i,\Sigma)$, $i = 1,\dots,k$, and an independent $p \times p$ Wishart matrix $S \sim W_p(n,\Sigma)$. Here ξ_i's and Σ are unknown. This kind of data arises, for example, if we are sampling from k different multivariate normal populations $N_p(\mu_i,\Sigma)$, $i = 1,\dots,k$, with a sample of size n_i from the ith population. If $\bar{X}_i$ and S_i *denote respectively the sample mean vector and the sample sum of* squares and products matrix, we may write $X_i = \sqrt{n_i}\,\bar{X}_i$, $i = 1,\dots,k$, and $S = S_1 +\dots+ S_k$ to get the above setup where $n = n_1 +\dots+ n_k - k$ and $\xi_i = \sqrt{n_i}\,\mu_i$, $i = 1,\dots,k$. Under this formulation we consider the problems of estimating Σ, the common variance-covariance matrix, Σ^{-1}, the common precision matrix, and $|\Sigma|$, the common generalized variance. We assume that $n > p + 1$.

From the classical point of view without any decision-theoretic consideration, one would simply ignore X (since ξ_i's are unknown) and propose multiples of S, S^{-1} and $|S|$ which are unbiased for Σ, Σ^{-1} and $|\Sigma|$ respectively. Even from a decision-theoretic point of view with the availability of a suitable loss function for each of these problems, one would normally be tempted to use very simple estimates like best multiples of S, S^{-1} and $|S|$ especially for invaiant loss functions. To illustrate this point clearly, let us first write down the various loss functions which are commonly used in this context. We denote by $L(\hat{\Sigma},\Sigma)$ the loss is estimating Σ by $\hat{\Sigma}$. A similar notation is used for $L(\hat{\Sigma}^{-1},\Sigma^{-1})$ and $L(|\hat{\Sigma}|,|\Sigma|)$.

(1.1)
$$L_1(\hat{\Sigma},\Sigma) = \mathrm{tr}(\hat{\Sigma} - \Sigma)^2$$
$$L_2(\hat{\Sigma},\Sigma) = \mathrm{tr}(\hat{\Sigma}\,\Sigma^{-1} - I)^2$$
$$L_3(\hat{\Sigma},\Sigma) = \mathrm{tr}\,\hat{\Sigma}\,\Sigma - \ln|\hat{\Sigma}\,\Sigma^{-1}| - p$$
$$L_4(\hat{\Sigma},\Sigma|Q) = \sum_{i\le j}(\hat{\sigma}_{ij} - \sigma_{ij})^2 q_{ij},\quad q_{ij} \ge 0 \text{ (weights)}$$

$$(1.2)\qquad L_1(\hat{\Sigma}^{-1},\Sigma^{-1}) = \mathrm{tr}(\hat{\Sigma}^{-1} - \Sigma^{-1})^2$$

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

$$L_3(\hat{\Sigma}^{-1},\Sigma^{-1}) = \mathrm{tr}\ \hat{\Sigma}^{-1}\Sigma - \ln|\hat{\Sigma}^{-1}\Sigma| - p\ .$$

$$L_4(\hat{\Sigma}^{-1},\Sigma^{-1}|Q) = \mathrm{tr}(\hat{\Sigma}^{-1} - \Sigma^{-1})^2 Q,\quad Q \text{ arbitrary p.d. matrix}$$

$$(1.3)\qquad L_1(|\hat{\Sigma}|,\ |\Sigma|) = (|\hat{\Sigma}| - |\Sigma|)^2$$

$$L_2(|\hat{\Sigma}|,\ |\Sigma|) = (|\hat{\Sigma}|/|\Sigma| - 1)^2$$

$$L_3(|\hat{\Sigma}|,\ |\Sigma|) = (|\hat{\Sigma}|/|\Sigma|) - \ln(|\hat{\Sigma}|/|\Sigma|) - 1$$

$$L_4(|\hat{\Sigma}|,\ |\Sigma|) = (|\Sigma|/|\hat{\Sigma}| - 1)^2$$

$$L_5(|\hat{\Sigma}|,\ |\Sigma|) = (|\Sigma|/|\hat{\Sigma}|) - \ln(|\Sigma|/|\hat{\Sigma}|) - 1$$

The loss function $L_1(\cdot)$ in (1.1) and (1.2) is a generalization of squared error loss and is noninvariant. The loss function $L_4(\cdot)$ in (1.1), due to Perlman (1972), is also noninvariant and is a generalization of squared error loss with constant weights q_{ij}'s. The loss functions $L_2(\cdot)$ and $L_3(\cdot)$ in (1.1) and (1.2) are invariant, $L_3(\cdot)$ in (1.1) being the entropy loss introduced in James and Stein (1961) while $L_2(\cdot)$ in (1.1) is proposed in Selliah (1964). $L_3(\cdot)$ in (1.2) is introduced and justified in Sinha and Ghosh (1986). All the four loss functions in (1.2) are completely analogous to those in (1.1). $L_4(\cdot)$ in (1.2) appears in Haff (1979a). The loss functions in (1.3) are all invariant, $L_1(\cdot)$ and $L_2(\cdot)$ being equivalent as far as risk is concerned. We mention in passing that the loss functions $L_1(\cdot)$, $L_2(\cdot)$ and $L_3(\cdot)$ in (1.1) and (1.2) can be used interchangeably. For example, $L_2(\hat{\Sigma}^{-1},\Sigma^{-1})$ in (1.2) can also be used as a genuine loss function for Σ because it satisfies $L_2(\hat{\Sigma}^{-1},\Sigma^{-1}) \geq 0$ and '$= 0$' if and only if $(\hat{\Sigma}^{-1})^{-1} = \Sigma$. [Loss functions developed by Efron and Morris (1976) from an empirical Bayes argument are not considered here.]

When a loss function is invariant, an affine (translation and multiplication by nonsingular matrices) equivariant estimator turns out to be of the form cS for Σ, dS^{-1} for Σ^{-1} and $e|S|$ for $|\Sigma|$ where c, d and e are constants. These estimators ignore X and, due to the invariant nature of the loss function, have constant risks, independent of Σ. It is then possible to find the best choice of these constants, resulting in the best affine equivariant estimators. For example, it is easy to show that (assuming that $n > p + 4$)

$$(1.4)\quad c_2 = \frac{1}{n+p+1},\ c_3 = \frac{1}{n},\ d_2 = (n-p)(n-p-3)/(n-1),\ d_3 = n-p-1,$$

$$e_1 = e_2 = e_3 = (n-p)!/n!,\quad e_4 = (n-p-4)!/(n-4)!,$$

$$e_5 = (n-p-2)!/(n-2)!$$

where c_2 is the best choice of the constant c for the loss function $L_2(\cdot)$ in (1.1) and so on.

The question then arises whether one should be content with these simple estimators or look for improved estimators. For a decision-theorist, it is, of course, essential to know if these best multiples of S, S^{-1} and $|S|$ are admissible for the respective problems. Unfortunately, it turns out none of these estimators is admissible and, in fact, there are many ways to improve over them. The improved estimators, however, lack simplicity and, unlike for the above simple estimators, their frequency properties are extremely difficult to study analytically. Numerical computations (Lin and Perlman (1985), Dey and Srinivasan (1984), Sinha and Ghosh (1986)) show that risk improvements are marginal in many cases but substantial in some cases.

The various methods leading to improved estimators of Σ, Σ^{-1} and $|\Sigma|$ are reviewed in this paper. Broadly speaking, there are three methods: (i) minimax estimators due to James and Stein (1961) and recently modified and improved by Dey and Srinivasan (1985); (ii) empirical Bayes estimators developed by Haff (1977, 1979a, 1979b, 1980); and (iii) Stein's testimators (Stein (1964)) as developed by Shorrock and Zidek (1976), Sinha (1976), and Sinha and Ghosh (1986).

Some open problems are mentioned in the concluding section.

2. Minimax Estimators.

To improve over the best affine equivariant estimators, James and Stein (1961) considered a somewhat smaller group of translation and multiplication by $p \times p$ lower triangular matrices with positive diagonal elements. The component translation of this group eliminates X and the estimators which are equivariant under the triangular group G_T^+ of the above type turn out to be of the form $K\Delta K'$ where $K \varepsilon G_T^+$ satisfies $S = KK'$ and Δ is a diagonal matrix. For an invariant loss function, the risk of such an estimator is independent of Σ and depends only on the diagonal elements of Δ (apart from n and p). To compute the risk, one has to use the simple fact that if $S \sim W_p(n,I)$ and $S = KK'$ with $K \varepsilon G_T^+$, then $K_{ii}^2 \sim \chi^2_{n-i+1}$, $K_{ji} \sim N(0,1)$, $j > i$, and all variables in K are independent (see Kshirsagar (1978)). It is then possible to find the best choice of the diagonal elements of Δ, thus resulting in the best G_T^+-equivariant estimator which, if different from the best affine equivariant estimator, is certainly an improvement over it. Moreover, because of the solvability of the group G_T^+ (Kiefer (1957)), the resultant estimator with constant risk is automatically minimax. This procedure can be followed to get improved estimators for the invariant loss functions $L_2(\cdot)$ and $L_3(\cdot)$ in (1.1) and (1.2). For example, for the loss function $L_2(\cdot)$ in (1.1), the minimax estimator dominating the best affine equivariant estimator c_2S is given by $K\Delta_0K'$ where the diagonal elements $\delta_1,\dots,\delta_p$ of Δ_0 are the solutions of the equations

$$
\begin{aligned}
(2.1)\qquad & (n+p-1)(n+p+1)\delta_1 + (n+p-3)\delta_2 + \dots + (n-p+1)\delta_p = n + p - 1 \\
& (n+p-3)\delta_1 + (n+p-3)(n+p-1)\delta_2 + \dots + (n-p+1)\delta_p = n+p-3 \\
& \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
& (n-p+1)\delta_1 + (n-p+1)\delta_2 + \dots + (n-p+1)(n-p+3)\delta_p = n-p+1
\end{aligned}
$$

Similarly, the minimax estimator dominating the best affine equivariant estimator c_3S for the loss function $L_3(\cdot)$ in (1.1) is given by $K\Delta_1K'$ with the ith diagonal element δ_i' of Δ_1 being equal to $(n+p-2i+1)^{-1}, i = 1,\dots,p$. Analogous results can be obtained for the loss functions $L_2(\cdot)$ in $L_3(\cdot)$ in (1.2).

Remark 2.1. The above procedure does not work for the loss functions in (1.3) relevant for estimating the generalized variance because $|K\Delta K'| = e|S|$ for some e, which is also affine equivariant. A method for improving over $e|S|$ is discussed in Section 4.

Remark 2.2. A drawback of the minimax estimators presented above is their dependence on the coordinate system. Although this can be overcome by averaging the minimax estimator over the group of $p \times p$ orthogonal matrices with respect to Haar measure, the resultant estimator does not have a simple form for $p \geq 3$ (see Takemura (1984)). Another drawback is that the amount of risk improvement is very marginal. Stein (1975, 1977a, 1977b) instead considered the class of orthogonally invariant estimators of the form $\hat{\Sigma} = R\,\phi(L)\,R'$ where $S = RLR'$ with R the matrix of normalized eigen vectors ($RR' = R'R = I$), $L = \mathrm{diag}(\ell_1,\ldots,\ell_p)$ is the diagonal matrix of eigen values of S with $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_p$, and $\phi(L) = (\phi_1(L),\ldots,\phi_p(L))$. For estimating Σ under the entropy loss function $L_1(\cdot)$ in (1.1), Stein (1975) proposed using

$$\phi_i^S(L) = \ell_i/[(n-p+1) + 2\ell_i \sum_{j\neq i} 1/(\ell_i - \ell_j)], \quad i = 1,\ldots,p. \tag{2.2}$$

However, the risk function of the resultant estimator is very complicated and *it has not been theoretically determined that this estimator dominates* c_3S, the best affine equivariant estimator for this loss, although Monte Carlo simulation results of Lin and Perlman (1985) indicate this to be the case. Recently, for the loss function $L_3(\cdot)$ in (1.1), Dey and Srinivasan (1985), by successfully applying Stein's technique, obtained orthogonally invariant minimax estimators which are improvements over c_3S as well as over the minimax estimator presented before for $p \geq 3$. The improved estimators which are obtained by solving a certain differential inequality have very simple forms. Also the percentage risk improvements over the risk of c_3S and over the minimax risk are both significant, except for very large n (Dey and Srivivasan (1984)). We may mention that Haff's (1982) modification of Stein's estimator $\hat{\Sigma}$ also performs quite well. Dey and Srinivasan's (1985) improved estimator over c_3S is given by $R\,\phi^*(L)R'$ and that over the minimax estimator $K\Delta_1 K'$ by $R\,\phi^1(L)R'$. The components of $\phi^*(L)$ are obtained as

(2.3) $\phi_i^o(L) = \ell_i/n - (\ell_i \log \ell_i)\tau_o(u)/(b_o + u)$, $i = 1,\dots,p$,

where $u = \sum_1^p \log^2 l_i$, $b_o > 144(p-2)^2/(25n^2)$ is a constant, and $\tau_o(u)$ is a function satisfying (i) $0 < \tau_o(u) < 2(p-2)/n^*$, $n^* = 5n^2/6$; (ii) $\tau_o(u)$ monotone nondecreasing in u and $E[\tau_o'(u)] < \infty$. On the other hand, the components of $\phi^1(L)$ are given by

(2.4) $\phi_i^1(L) = \ell_i\delta_i^1 - (\ell_i \log \ell_i)\tau_1(u)/(b_1 + u)$, $i = 1,\dots,p$,

where $b_1 > 144(p-2)^2/\{25(n+p-1)^2\}$ is a constant, and $\tau_1(u)$ is a function satisfying (i) $0 < \tau_1(u) < 12(p-2)/(5(n+p-1)^2)$, (ii) $\tau(u)$ monotone non-decreasing in u and $E[\tau_1'(u)] < \infty$. Incidentally, another very simple estimator $R\,\phi^2(L)R'$ with $\phi_i^2(L) = \ell_i\delta_i^1$, $i = 1,\dots,p$, is also minimax and is an improvement over both c_3S and $K\Delta_1K'$ for the entropy loss function $L_3(\cdot)$ in (1.1) (Stein (1982), Dey and Srivivasan (1985)).

3. Empirical Bayes Estimators

A unified approach to constructing estimators of Σ and Σ^{-1} substantially better (in terms of amount of risk improvement) than the best affine equivariant estimators for invariant loss functions in (1.1) and (1.2) and otherwise better than the unbiased estimators has been successfully developed in a series of papers by Haff (1977, 1979a, 1979b, 1980). Assuming an Wishart conjugate prior for Σ^{-1}, $\Sigma^{-1} \sim W_p(n',C^{-1}/\gamma)$, $n' > p$, $\gamma > 0$, C p.d., which results in the Wishart posterior for Σ^{-1}, $\Sigma^{-1}|S \sim W_p(n + n', (S + \gamma C)^{-1})$, it turns out that for most loss functions in (1.1) and (1.2), the estimator minimizing the Bayesian expected loss is of the form $\hat{\Sigma} = a(S + \gamma C)$ for some constant $a > 0$. An empirical Bayes estimator is then obtained by pretending that γ is unknown and suitably estimating it using the marginal density of S which is proportional to $\gamma^{pn'/2}|S|^{(n-p)/2}\,|S + \gamma C|^{-(n+n')/2}$. Quite generally, for estimating Σ, estimators of the form

(3.1) $\hat{\Sigma} = a\{S + ut(u)I_p\}$

are proposed where for simplicity, C is taken as identity, a is a positive constant, u is an average eigen value of S, and t(u) is nonnegative, bounded and nonincreasing. On the other hand, for estimating the precision matrix Σ^{-1}, estimators of the form

$$\hat{\Sigma}^{-1} = b\{S^{-1} + v\ t(v)I_p\} \tag{3.2}$$

are proposed where $b > 0$ is a constant, v is an average eigen value of S^{-1}, and $t(v)$ is nonnegative, bounded and nonincreasing. In applications, depending on the loss function, $u(v)$ maybe the arithmetic, geometric or harmonic mean of the eigen values of $S(S^{-1})$. By (repeatedly) applying the extremely powerful Wishart Identity (Stein (1975), Haff (1979b)), Haff was able to construct estimators of the forms given in (3.1) and (3.2) better than $c_{opt}S$ and $d_{opt}S^{-1}$ for all the loss functions in (1.1) and (1.2) respectively (except for the loss $L_3(\cdot)$ in (1.2) which is new). To get a flavor of Haff-type improved estimators, we mention below a few results. Note that the loss function $L_1(\cdot)$ in (1.1) or (1.2) is a special case of corresponding $L_4(\cdot)$ when Q is taken as the identity matrix. We denote by $\hat{\Sigma}_i^{(H)}$ and $\hat{\Sigma}_i^{-1(H)}$ the improved estimators due to Haff when the loss function is $L_i(\cdot)$ in (1.1) and (1.2) respectively.

(3.3) $\hat{\Sigma}_2^{(H)} = a\{S + u\ t(u)I_p\}$, $a = (n+p+1)^{-1}$, $u = 1/tr(S^{-1})$,

$0 < t \leq 2(p-1)/(n-p+3)$, t a constant, dominates c_2S.

(3.4) $\hat{\Sigma}_3^{(H)} = a\{S + u\ t(u)I_p\}$, $a = n^{-1}$, $u = 1/tr(S^{-1})$,

$0 < t(u) \leq 2(p-1)/n$, an absolutely continuous and nonincreasing function, dominates c_3S. (If t is chosen as a constant, the optimal value of t is $(p-1)/n$)

(3.5) $\hat{\Sigma}_4^{(H)} = a\{S + u\ t(u)I_p\}$, $u = |S|^{1/p}$, $t(u)$ satisfying

(i) $(4q^*/pn^2)u\ t'(u) + 2q^*[a - (pn-2)/pn^2]t(u) + a^2t^2(u) \leq 0$

(ii) $u\ t''(u) + 2t'(u) \geq 0$

where $q^* = p(\prod_1^p q_{ii})^{1/p}/\sum_1^p q_{ii}$, dominates aS, whatever Q. (If t is chosen as a constant, $0 < t \leq 2q^*[(pn-2)/pn^2-a]/a^2$ will do.)

Remark 3.1. The result for the loss function $L_1(\cdot)$ in (1.1) is obtained by putting $Q = I_p$ i.e., $q^* = 1$ in (3.5). In this case a is usually taken as n^{-1}. There are, of course, simpler improved estimators (improvement over $n^{-1}S$) for this loss function. Perlman (1972) proved that $\beta\, n^{-1}S$ dominates $n^{-1}S$ for any β, $(np-2)/(np+2) \leq \beta < 1$. Such a simple dominance result holds for the loss function $L_4(\cdot)$ in (1.1) as well. It is proved in Haff (1979b) that a S dominates $n^{-1}S$ for any a, $(n-1)/(n(n+1)) \leq a < n^{-1}$, whatever be Q. However, unlike for the estimator in (3.5), the amount of risk improvement for these simple estimators is marginal.

Remark 3.2. The identity matrix I_p in (3.3) and (3.4) can be replaced by any p.d. matrix C in which case u need be redefined as $u = 1/(\mathrm{tr}S^{-1}C)$.

(3.6) $$\hat{\Sigma}_2^{-1(H)} = b\{S^{-1} + vt(v)I_p\},\ b = (n-p-3)(n-p)/(n-1),$$

$$v = 1/\mathrm{tr}(S),\ 0 < t(v) \leq 2(p-1)/(n-p)$$

and $t'(v) \leq 0$, dominates d_2S^{-1}.

(3.7) $$\hat{\Sigma}_4^{-1(H)} = b\{S^{-1} + vt(v)I_p\},\ v = |S|^{-1/p},\ t(v)\ \text{satisfying}$$

$$\{(4/p)[vt'(v) + t(v)] + 2b^*t(v)\}q^* + bt^2(v) \leq 0,\ \text{where } b^* = b - (n-p-1),$$

$$q^* = p(\prod_1^p q_{ii})^{1/p}/\sum_1^p q_{ii},$$

dominates b S^{-1}, whatever Q.

Remark 3.3. The result for the loss function $L_1(\cdot)$ in (1.2) is obtained from (3.7) by putting $Q = I_p$ i.e., $q^* = 1$. In this case b is usually taken as (n-p-1). Again, for this loss function also, there are simpler improved estimators. Haff (1977) showed that for any b, $n-p-3 \leq b < n-p-1$, b S^{-1} dominates $(n-p-1)S^{-1}$. However, the amount of risk improvement is marginal.

Remark 3.4. For the loss function $L_4(\cdot)$ in (1.2), Haff (1979a) obtained another class of more improved estimators of Σ^{-1}. This is not discussed here.

<u>Remark 3.5.</u> So far no Haff-type improved estimator is available which dominates $e_{opt}|S|$ for estimating the generalized variance $|\Sigma|$ under the loss functions in (1.3). Also, no Haff-type dominance result is known for the loss function $L_3(\cdot)$ in (1.2) which improves over d_3S^{-1} for estimating Σ^{-1}. Dominance results for these problems are presented in the next section.

4. <u>Stein's testimators.</u>

The methods described in Sections 2 and 3 above for improving over the best affine equivariant estimators for the loss functions in (1.1) and (1.2) are based on different functions of S, but ignoring X all the time. Clearly these methods do not work when p = 1. This is because for estimating the variance σ^2 of a normal population with unknown mean μ, if the sample mean is ignored, the best multiple of the sample variance is admissible under both entropy loss and squared error loss. Stein (1964) came up with an ingenius idea to use the sample mean along with the sample variance in order to obtain an improved estimator of σ^2. For the squared error loss, Stein's improved estimator has the form $\hat{\sigma}^2 = \min[(n+1)^{-1}\sum_1^n (y_i-\bar{y})^2,\ (n+2)^{-1}\sum_1^n (y_i-\mu_0)^2]$ where μ_0 is any fixed constant. Such an estimator is also called a testimator because of its obvious interpretation as the result of a suitable test procedure.

Recently, Stein's (1964) technique has been successfully generalized in the multivariate case to deal with all the loss functions in (1.3) and the entropy losses in (1.1) and (1.2). We now have testimators which are better than $e_{opt}|S|$ for estimating $|\Sigma|$, than c_3S for estimating Σ when the loss is $L_3(\cdot)$ in (1.1), and than d_3S^{-1} for estimating Σ^{-1} when the loss is $L_3(\cdot)$ in (1.2). See Sinha (1976), Sinha and Ghosh (1986), and also Shorrock and Zidek (1976). However, simulation results show that the amount of risk improvement is marginal in most cases except when the loss is $L_3(\cdot)$ in (1.2). A brief description of these testimators is given below. We use the notation X, S as given in the Introduction.

For estimating $|\Sigma|$ under the loss function $L_1(\cdot)$ or $L_2(\cdot)$ or $L_3(\cdot)$ in (1.3), the testimator better than $e_1|S|$ is given by $|\hat{\Sigma}|_1 = \min\{\frac{(n-p+k)!}{(n+k)!}\ |S + XX'|,$

$\frac{(n-p)!}{n!}|S|\}$. Under the loss functions $L_4(\cdot)$ and $L_5(\cdot)$ in (1.3), the improved testimators are respectively given by $|\hat{\Sigma}|_4 = \min\{\frac{(n-p-4+k)!}{(n-4+k)!}|S + XX'|, \frac{(n-p-4)!}{(n-4)!}|S|\}$ and $|\hat{\Sigma}|_5 = \min\{\frac{(n-p-2+k)!}{(n-2+k)!}|S + XX'|, \frac{(n-p-2)!}{(n-2)!}|S|\}$. On the other hand, for estimating Σ when the loss is $L_3(\cdot)$ in (1.1), the testimator $\hat{\Sigma} = (S + XX')/(n+k)$ if $\lambda_{max}(X'S^{-1}X) \leq k/n$, $= S/n$ otherwise, dominates the best affine invariant estimator S/n. For the other loss function $L_3(\cdot)$ in (1.2), the improved testimator $\hat{\Sigma}^{-1} = (n-p-1+k)(S + XX')^{-1}$ if $\lambda_{max}(X'S^{-1}X) \leq k/(n-p-1)$, $= (n-p-1)S^{-1}$ otherwise, dominates d_3S^{-1} for estimating Σ^{-1}. Here $\lambda_{max}(\cdot)$ stands for the largest eigen value of $(\cdot)$.

A general procedure to obtain improved testimators is to consider estimators equivariant under a nonnormal subgroup H of the full affine group G. Typically, the group G acts on (X,S) and (ξ,Σ) as $X \to AX + B$, $\xi \to A\xi + B$, $S \to ASA'$, $\Sigma \to A\Sigma A'$ where $A:p \times p$ is nonsingular and $B:p \times k$ is any matrix. Here $\xi:p \times k = (\xi_1,\ldots,\xi_k)$. The subgroup H is obtained from G by putting $B = 0$. It turns out that estimators of Σ equivariant under H are of the form $\hat{\Sigma} = W\psi W'$ where $S = WW'$, $W:p \times p$ is nonsingular, and $\psi \equiv \psi(UU'):p \times p$ with $U = W^{-1}X$. For a specified loss function, the best choice of the function $\psi(\cdot)$ usually involves the unknown parameters ξ and Σ besides being computationally intractable. This is also the case for estimating Σ^{-1} and $|\Sigma|$ for which estimators of the form $\hat{\Sigma}^{-1} = W'^{-1}\psi^{-1}W^{-1}$ and $|\hat{\Sigma}| = |S|\psi(X'S^{-1}X)$ respectively are considered where ψ in the second case is a scalar. However, for most problems, an upper bound of the optimum ψ, say ψ_0, is available. Because of the convex nature of the loss functions, it then follows that given any estimator with a specified ψ^*, putting $\tilde{\psi} = \min(\psi_0,\psi^*)$ results in improvement of risk. When ψ^* and ψ_0 are matrices, $\tilde{\psi}$ is defined as $\tilde{\psi} = \psi_0$ if $\psi_0 \leq \psi^*$, $= \psi^*$ otherwise, where $\psi_0 \leq \psi^*$ means $\psi^* - \psi_0$ is n.n.d. This is the basis of derivation of all the testimators discussed in this section.

Remark 4.1. It is interesting to observe that the testimators of Σ, Σ^{-1} of $|\Sigma|$ defined above are based on different test statistics. For Σ and Σ^{-1}, a test based on the maximum root of $X'S^{-1}X$ is used, while for $|\Sigma|$ it is the likelihood ratio test. In all the cases, the underlying null hypothesis is $\xi = 0$.

Remark 4.2. Sequential testimators of Σ, Σ^{-1} and $|\Sigma|$, based on suitable tests of parts of ξ, all of which dominate the corresponding best affine equivariant estimators are also available. See Sinha (1976) and Sinha and Ghosh (1986) for detail.

5. Open Problems

The following is a list of the open problems in the area discussed in the paper.

(a) It will be interesting to develop Haff-type improved estimators for estimating the precision matrix Σ^{-1} under the loss function $L_3(\cdot)$ in (1.2).

(b) For estimating the generalized variance $|\Sigma|$ under the loss functions in (1.3), so far no Haff-type improved estimator is available. Is it true that the estimators $e_{opt}|S|$ are admissible when X is ignored? A partial answer is available in Das Gupta (1982).

(c) So far, Stein's testimators are available only for the loss functions $L_3(\cdot)$ in (1.1) and (1.2). What happens for other loss functions?

REFERENCES

DasGupta, A. (1982). Is the best equivariant estimator of the generalized variance admissible? Technical Report, Stat-Math Division, Indian Statistical Institute.

Dey, D. K. and Srinivasan, C. (1984). Estimation of covariance matrix under Stein's loss. Technical Report, Texas Tech. University.

Dey, D. K. and Srinivasan, C. (1985). Estimation of a covariance matrix under Stein's loss. Ann. Statist., 13, 1581-1591.

Efron, B. and Morris, C. (1976). Multivariate empirical Bayes estimation of covariance matrices. Ann. Statist., 4, 22-32.

Haff, L. R. (1977). Minimax estimators for a multinormal precision matrix. Jour. Multivariate Anal., 7, 374-385.

Haff, L. R. (1979a). Estimation of the inverse covariance matrix: Random mixtures of the inverse Wishart matrix and the identity. Ann. Statist., 7, 1264-1276.

Haff, L. R. (1979b). An identity for the Wishart distribution with applications. J. Multivariate Anal., 9, 531-542.

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

Haff, L. R. (1982). Solutions of the Euler-Lagrange equations for certain multivariate normal estimation problems. Unpublished manuscript.

James, W. and Stein, C. (1961). Estimation with quadratic loss. Proc. Fourth Berkeley Symposium Math. Statist. Prob. 1, 361-380.

Khirsagar, A. M. (1978). Multivariate Analysis. Marcel Dekker, Inc. New York and Basel.

Kiefer, J. (1957). Invariance, minimax sequential estimation and continuous time process. Ann. Math. Statist., 28, 573-601.

Lin, S. P. and Perlman, M. D. (1985). A Monte Carlo comparison of four estimators for a covariance matrix. Multivariate Anal., 6, 411-429. Edited by Krishnaiah. North Holland, Amsterdam.

Perlman, M. D. (1972). Reduced mean square error estimation for several parameters. Sankhya (B), 34, 89-92.

Selliah, J. B. (1964). Estimation and testing problems in a Wishart distribution. Tech. Reprot No. 10, Dept. of Statistics, Stanford University.

Shorrock, R. W. and Zidek, J. V. (1976). An improved estimator of the generalized variance. Ann. Statist., 4, 629-638.

Sinha, B. K. (1976). On improved estimators of the generalized variance. Jour. Multivariate. Anal., 6, 617-625.

Sinha, B. K. and Ghosh, M. (1986). Inadmissibility of the best equivariant estimators of the variance-covariance matrix, the precision matrix, and the generalized variance under entropy loss. To appear in Statistics and Decisions.

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Inst. Stat. Math., 16, 155-160.

Stein, C. (1975). Reitz Lecture. 39th Annual Meeting IMS, Atlanta, Georgia.

Stein, C. (1977a). Unpublished notes on estimating the covariance matrix.

Stein, C. (1977b). Lectures on the theory of estimation of many parameters. (In Russian). In: Studies in the Statistical Theory of Estimation, Part I (Ibragimov, I. A. and Nikulin, M. S. eds.), Proceedings of Scientific Seminars of the Steklov Institute, Leningrad Division, 74, 4-65.

Stein, C. (1982). Unpublished notes. University of Washington, Seattle.

Takemura, A. (1984). An orthogonally invariant minimax estimator of the covariance matrix of a multivariate normal population Tsukuba J. Math., 8, 367-376.

ATE
LMED
– 8