Sankhyā : The Indian Journal of Statistics
1991, Volume 53, Series A, Pt. 3, pp. 389-411.

SHRINKAGE ESTIMATION IN A RESTRICTED PARAMETER SPACE

By DEBAPRIYA SENGUPTA
Iowa State University
and
PRANAB KUMAR SEN
University of North Carolina

SUMMARY. For the multivariate normal mean model when the parameter is restricted to a positively homogeneous cone, Stein-rule or shrinkage versions of restricted maximum likelihood estimators RSMLE are considered, and under suitable quadratic loss, their dominance properties are studied systematically. Some applications to linear models are also considered.

1. Introduction

The past three decades have witnessed a phenomenal growth of research on improved estimation (under quadratic loss) based on Stein-rule (or shrinkage) estimators (SRE). Consider a $p(\geqslant 3)$-variate normal distribution with mean vector $\boldsymbol{\theta} \in \Theta$ where $\boldsymbol{\theta}$ is a restricted subset of R^p. In the unrestricted case (i.e., for $\Theta = R^p$), the SRE of $\boldsymbol{\theta}$ dominate the usual maximum likelihood estimators (MLE). In the restricted case, the pivot $(\boldsymbol{\theta}_0)$ on which a SRE is based may not be an inner point of Θ, and hence, the relative dominance picture can be quite different. In fact, restricted (R) MLE (derived under the parameter restraints) may not possess all the asymptotic optimality properties of the classical MLE (although the RMLE generally perform better than the MLE when $\boldsymbol{\theta} \in \Theta$). One may theorefore wonder whether the RMLE can be dominated in quadratic risk by suitable SRE in the same manner as the unrestricted (U)MLE is dominated by a SRE ? A comprehensive study of this dominance picture needs to focus on the risk of all the UMLE, RMLE, SRE and their restricted versions (RSRE). The object of the present investigation is to present a systematic and unified account of this relative dominance picture in a variety of restricted models.

To motivate, we may mention some typical problems where the RMLE are advocated.

Paper received. November 1988 ; revised May 1990.

AMS (1980) *subject classifications.* 62C15, 62F10, 62H12.

Key words and phrases. Admissibility, positively homogeneous cone, positive-rule estimators, quadratic loss, restricted MLE ; Stein-rule estimators, Stein-type identity.

(i) *Orthant alternative problem.* Here $\Theta = \{\boldsymbol{\theta} \in R^p : \boldsymbol{\theta} \geqslant \boldsymbol{\theta}_0\}$, for some specified $\boldsymbol{\theta}_0 \in \boldsymbol{R}^p$; translating the observations by $\boldsymbol{\theta}_0$, we may set $\boldsymbol{\theta}_0 = \boldsymbol{0}$.

(ii) *Ordered alternative problem.* Here $\Theta = \{\boldsymbol{\theta} \in \boldsymbol{R}^p : \theta_1 \leqslant \ldots \leqslant \theta_p\}$ so that the pivot (or null hypothesis) relates to the line of equality : $\boldsymbol{\theta} = \theta\boldsymbol{1}, \theta \in \boldsymbol{R}$.

Other models include the so called *umbrella* alternatives, *tree* alternatives, *loop* alternatives, etc., and we may refer to Robertson et al. (1988) for some details. The RMLE computed under the restricted setup are often termed the *isotonic* MLE (viz., Barlow *et al.* (1972). Kudo (1963), Nüesch (1966), Perlman (1969) and others have studied various properties of the usual RMLE for some of these models. In this context, one may introduce a *positively homogeneous cone* Θ by setting that $\boldsymbol{\theta} \in \Theta$ implies that for every $M > 0$, $M\boldsymbol{\theta} \in \Theta$, so that Θ is a proper subset of R^p, restricted by some inequalities. Generally, a RSMLE on a positively homogeneous cone Θ performs better than the usual UMLE, although an opposite picture may emerge on the complementary space $R^p \setminus \Theta$. In the simplest orthant model (when the covariance matrix is assumed to be diagonal and known), Chang (1981, 1982) has proposed some shrinkage estimators. His formulation has mostly been on heuristic grounds, and the full impact of shrinkage has not been incorporated in the estimators considered by him. Further, the formulation becomes ineffective when the covariance matrix is not diagonal. For the case of a single inequality constraint, Judge and Yancey (1986) have considered some SRE ; their formulation also encounters considerable difficulties in the case of multiple constraints when the covariance matrix is arbitrary. For both the cases, for a completely arbitrary covariance matrix, or in general, for a positively homogeneous cone Θ, an explicit formulation of the RMLE is a precursor for the construction of suitable RSMLE which would have better dominance properties. With this objective in mind, we shall consider here a formulation of RMLE and RSMLE in a unified manner, and then incorporate them in our desired dominance picture study.

In Section 2 we start with an explicit formulation of the RMLE, and this in turn provides a clear motivation for the construction of the RSMLE, treated in Section 3. The dominance of the RSMLE over the RMLE is established in Section 4 ; a general class of restricted minimax estimators is also considered there. Section 5 deals with some restricted parameter spaces in some common linear models. Section 6 is devoted to the relative risk picture for the proposed estimators, and the concluding section deals with some general comments with special attention to the estimators considered by Chang (1981, 1982) and others.

2. Preliminary notions

We shall see in Section 5 that a general class of restricted alternative models can be reduced to the positive orthant model for which

$$\boldsymbol{\Theta} = \boldsymbol{\Theta}_+ = \{\boldsymbol{\theta} \in \boldsymbol{R}^p : \boldsymbol{\theta} \geqslant \boldsymbol{0},\ \|\boldsymbol{\theta}\| \geqslant 0\} = R_{+p}. \qquad \ldots \quad (2.1)$$

Hence, for the sake of simplicity, in Sections 2 through 4, we shall consider specifically the case of an orthant restriction model with an arbitrary covariance matrix. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be n independent and identically distributed (i.i.d.) random vectors (r.v.) having a p-variate normal distribution with mean vector $\boldsymbol{\theta}$ and dispersion matrix $\boldsymbol{\Sigma}$ assumed to be positive definite (p.d.). Let us denote by

$$\overline{\boldsymbol{X}}_n = n^{-1} \sum_{i=1}^{n} \boldsymbol{X}_i \text{ and } S_n = \sum_{i=1}^{n} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}_n)(\boldsymbol{X}_i - \overline{\boldsymbol{X}}_n)' \ ; \ \hat{\boldsymbol{\Sigma}}_n = (n-1)^{-1}\boldsymbol{S}_n.$$

Then $\overline{\boldsymbol{X}}_n$ is the UMLE of $\boldsymbol{\theta}$; it is unbiased for $\boldsymbol{\theta}$ but not admissible for $p \geqslant 3$ The formulation of the RMLE of $\boldsymbol{\theta}$ (restricted to Θ_+) is due to Nuesch (1966). We shall find it convenient to express this in the following compact form. Later on, we also provide suitable simplifications in some special cases.

For every $p \geqslant 1$, let $N_p = \{1, \ldots, p\}$. By $a \subseteq N_p$, we mean a subset of N_p ordered by natural ordering, $|a|$ stands for the cardinality of a, and $a' = N_p \setminus a$ stands for the complement of a. For a p-vector $\boldsymbol{x}$, we define for each $a \subseteq N_p$, $\boldsymbol{x}_a$ as the $|a|$—vector consisting of the components with indices $i \in a$; $\boldsymbol{x}_\phi = 0$, conventionally. For a $p \times p$ p.d. matrix $\boldsymbol{Q}$ and for every $a \subseteq N_p$, $b \subseteq N_p$, let

$$\boldsymbol{x}_{a:b}(\boldsymbol{Q}) = \boldsymbol{x}_a - \boldsymbol{Q}_{ab}\boldsymbol{Q}_{bb}^{-1}\boldsymbol{x}_b \text{ and } \boldsymbol{Q}_{a:b} = \boldsymbol{Q}_{aa} - \boldsymbol{Q}_{ab}\boldsymbol{Q}_{bb}^{-1}\boldsymbol{Q}_{ba}, \qquad \ldots \quad (2.3)$$

where $\boldsymbol{Q}_{ab}$ denotes the $|a| \times |b|$ submatrix of $\boldsymbol{Q}$ consisting of the rows in a and columns in b ; $\boldsymbol{Q}_{aa}$ and $\boldsymbol{Q}_{bb}$ are defined analogously. Let then

$$\mathscr{X}_a(\boldsymbol{Q}) = \{\boldsymbol{x} \in \boldsymbol{R}^p = \boldsymbol{Q}_{a'a'}^{-1}\boldsymbol{x}_{a'} \leqslant \boldsymbol{0}, \boldsymbol{x}_{a:a'}(\boldsymbol{Q}) > 0\}, \phi \subseteq a \subseteq N_p. \qquad \ldots \quad (2.4)$$

Then [viz., Kudo (1963)] the $\mathscr{X}_a(\boldsymbol{Q})$, $a \in N_p$ are disjoint and $\bigcup_{a \in N_p} \mathscr{X}_a(\boldsymbol{Q}) = \boldsymbol{R}^p$.

Finally, for every $a \subseteq N_p$, such that $|a| = r$ $(0 \leqslant r \leqslant p)$, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are r- and $(p-r)$—vectors respectively, we denote by

$$[\boldsymbol{\mathscr{P}}_a(\boldsymbol{u}, \boldsymbol{v})]_i = u_i \text{ or } v_i \text{ according as } i \in a \text{ or } i \in a'. \qquad \ldots \quad (2.5)$$

The RMLE of $\boldsymbol{\theta}$ (restricted to Θ_+) is given by

$$\hat{\boldsymbol{\theta}}^*_{RM} = \sum_{\{\phi \subseteq a \subseteq N_p\}} \mathscr{P}_a(\overline{\boldsymbol{X}}_{na:a'}(\boldsymbol{S}_n), \boldsymbol{0})\, 1\, (\overline{\boldsymbol{X}}_n \in \mathscr{X}_a(\boldsymbol{S}_n)), \qquad \ldots \quad (2.6)$$

where $1(A)$ stands for the indicator function of the set A. If $\mathbf{\Sigma}$ is known, in (2.6) we replace $\mathbf{S}_n$ by $\mathbf{\Sigma}$, and this RMLE is denoted by $\hat{\mathbf{\theta}}_{RM}$. We may further note that only one of the 2^p indicators functions $1(\bar{X}_n \epsilon \mathscr{X}_a(S_n))$, $a \epsilon N_p$ is non-null, so that (2.6) is actually expressible in terms of a single (random) term in this closed form. We may remark that whereas the UMLE $\bar{\mathbf{X}}_n$ is equivariant under any nonsingular transformation of the $\mathbf{X}_i$ the restricted space Θ_+ is not invariant with respect to such transformations, and also the RMLE in (2.6) may not enjoy this equivariance property in a general setup. If however, $\mathbf{\Sigma}$ is a diagonal matrix, we may need only the diagonal terms of $\mathbf{S}_n$ to estimate it (denote this diagonal matrix by $\mathbf{S}_n^0$), then the RMLE in (2.6) with $\mathbf{S}_n$ replaced by $\mathbf{S}_n^0$ remains equivariant under any coordinatewise scalar transformation. In such a case, a is'the set of all'coordinates of $\bar{\mathbf{X}}_n$ having positive elements (easy to determine instantly), and that simply yields the unique partition a (ϵN_p) for which $1(\bar{\mathbf{X}}_n \epsilon \mathscr{X}_a(\mathbf{S}_n^0)) = 1$. This simplification is not generally tenable for an arbitrary $\mathbf{\Sigma}$ (even when it is known) For p not so large, the computational algothithm for (2.6) has been discussed in various places (viz., Robertson *et al.* (1988)), although as p increases, the task becomes highly laborious. However, in actual practice, generally p is not large, and hence, (2.6) does not pose any threatening task for its computation. The beauty of (2.6) is that it provides a natural motivation for the SRE which we consider in the next section. However, to study the related dominance results, we need to introduce the relavant risk function. For this, consider the quadratic loss function (for an estimator $\mathbf{T}$ of $\mathbf{\theta}$) :

$$L(\mathbf{T}, \mathbf{\theta}) = (\mathbf{T}-\mathbf{\theta})' \mathbf{\Sigma}^{-1}(\mathbf{T}-\mathbf{\theta}) = ||\mathbf{T}-\mathbf{\theta}||^2_{\mathbf{\Sigma}}, \quad \ldots \quad (2.7)$$

so that the risk is given by

$$R(\mathbf{T}, \mathbf{\theta}) = E_{\mathbf{\theta}}L(\mathbf{T}, \mathbf{\theta}) = Tr\{\mathbf{\Sigma}^{-1}E_{\mathbf{\theta}}[(\mathbf{T}-\mathbf{\theta})(\mathbf{T}-\mathbf{\theta})'], \mathbf{\theta} \epsilon \Theta_+. \quad \ldots \quad (2.8)$$

We may recall that an estimator $\mathbf{T}^*$ dominantes another one ($\mathbf{T}$) in quadratic risk if

$$R(\mathbf{T}^*, \mathbf{\theta}) \leqslant R(\mathbf{T}, \mathbf{\theta}), \forall\ \mathbf{\theta} \epsilon \Theta_+, \text{ with strict inequality, for some } \mathbf{\theta}. \quad \ldots \quad (2.9)$$

3. The RSMLE for Θ_+

Note that $\{\mathscr{X}_a(\mathbf{S}_n) ; \phi \subseteq a \subseteq N_p\}$ is a partitioning of R^p into 2^p disjoint subsets. Also, on $\mathscr{X}_a(\mathbf{S}_n)$, the problem of estimating $\mathbf{\theta}$ under the constraint that $\mathbf{\theta} \geqslant \mathbf{0}$ essentially reduces to that of estimating $\mathbf{\theta}_a$ under the constraint $\mathbf{\theta}_{a'} = 0$, and the RMLE coincides with the UMLE on this $|a|$-dimensional subspace. This suggests that while adopting the James and Stein (1961) shrinkage methodology, it may be wiser to adapt it to the particular subset (i.e., $\mathscr{X}_a(\boldsymbol{S}_n)$) where the sample statistic belongs to. Thus, we allow the shrinkage factor to be dependent on the (random) subset $a : \phi \subseteq a \subseteq N_p$, and propose the following RSMLE of $\boldsymbol{\theta}$(on Θ_+) :

$$\hat{\boldsymbol{\theta}}^*_{\mathrm{RSM}} = \sum_{\{\phi \subseteq a \subseteq N_p\}} 1(\overline{\boldsymbol{X}}_n \in \mathscr{X}_a(\boldsymbol{S}_n)) \{1-c_a(n\|\hat{\boldsymbol{\theta}}^*_{\mathrm{RM}}\|^2_{\boldsymbol{S}_n})^{-1}\} \hat{\boldsymbol{\theta}}^*_{\mathrm{RM}}, \quad \dots \quad (3.1)$$

where the RMLE $\hat{\boldsymbol{\theta}}^*_{\mathrm{RM}}$ is defined by (2.6), so that

$$\|\hat{\boldsymbol{\theta}}^*_{\mathrm{RM}}\|^2_{\boldsymbol{S}_n} = [\overline{\boldsymbol{X}}_{na:a'}(\boldsymbol{S}_n)]' \boldsymbol{S}^{-1}_{na:a'} (\overline{\boldsymbol{X}}_{na:a'}(\boldsymbol{S}_n)] \text{ on } \mathscr{X}_a(\boldsymbol{S}_n), \phi \subseteq a \subseteq N_p, \dots \quad (3.2)$$

and the shrinkage factors c_a, $\phi \subseteq a \subseteq N_p$ are nonnegative and they satisfy :

$$0 \leqslant c_a \leqslant 2(|a|-2)^+/(n-p+2) \text{ where } q^+ = \max(0, q), \text{ for real } q. \quad \dots \quad (3.3)$$

It may be remarked that $\boldsymbol{S}_n$ follows the Wishart distribution $W(p, n-1, \boldsymbol{\Sigma})$ with $n-1$ degrees of freedom (DF). In general if $\boldsymbol{S}$ were $W(p, m, \boldsymbol{\Sigma})$, independently of $\overline{\boldsymbol{X}}_n$, then in (3.1) and (3.2), we would have replaced $\boldsymbol{S}_n$ by $\boldsymbol{S}$, and in (3.3), for the upper bound of c_a, $n-p+2$ by $(m-p+3)$, $a \subseteq N_p$. Also, if $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{V}$, $\boldsymbol{V}$ known, and if there exists an S^2, independent of $\overline{\boldsymbol{X}}_n$, such that $mS^2/\sigma^2 \sim \chi^2_m$, for some $m \geqslant 1$, then letting $s^2 = m(m+2)^{-1}S^2$ and $\hat{\boldsymbol{\Sigma}} = s^2\boldsymbol{V}$, in (3.1)—(3.2), we would replace $\boldsymbol{S}_n$ by $\hat{\boldsymbol{\Sigma}}$ and in (3.3), the upper bound for c_a would be simply $2(|a|-2)^+$, $a \subseteq N_p$. In particular, if $\boldsymbol{\Sigma}$ is known, then we may further replace $\hat{\boldsymbol{\Sigma}}$ by $\boldsymbol{\Sigma}$ without any further change in the upper bound for c_a, $a \subseteq N_p$. For the case of a diagonal $\boldsymbol{\Sigma}$, the computation of (3.1) becomes much simpler (vide Section 2 for the corresponding RMLE). In this case, Chang (1981, 1982) has considered some alternaive SRE where the shrinkage factor is not made to depend on the particular (random) set a for which $\overline{\boldsymbol{X}}_n \in \mathscr{X}_a(\boldsymbol{I})$. We shall make a detailed comparison of our proposed estimators and the ones by Chang (1981, 1982) in Section 7. In passing, however, we may remark that (3.1) besides being applicable in a more general situation is also more efficient than the Chang estimators when $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$.

$$\hat{\boldsymbol{\theta}}^{+*}_{\mathrm{RSM}} = \sum_{\{\phi \subseteq a \subseteq N_p\}} 1(\overline{\boldsymbol{X}}_n \in \mathscr{X}_a(\boldsymbol{S}_n)) \{1-c_a(n\|\hat{\boldsymbol{\theta}}^*_{\mathrm{RM}}\|^2_{\boldsymbol{S}_n})^{-1}\}^+ \hat{\boldsymbol{\theta}}^*_{\mathrm{RM}}. \quad \dots \quad (3.4)$$

Here also, the (lower) turncation of the shrinkage factor is made to depend on the partitioning $\mathscr{X}_a(\boldsymbol{S}_n)$, $\phi \subseteq a \subseteq N_p$. Some analogues of the Baranchik (1970) and Strawderman (1971) estimators for the restricted parameter space under consideration will be considered in the next section.

4. Dominance results for the RSMLE

Our main contention is to prove the following :

Theorem 4.1. *Under the quadratic loss in* (2.7), *for every* $p \geqslant 3$,

$$R(\boldsymbol{\theta}^*_{\text{RSM}}, \boldsymbol{\theta}) \leqslant R(\hat{\boldsymbol{\theta}}^*_{\text{RSM}}, \boldsymbol{\theta}) \leqslant R(\hat{\boldsymbol{\theta}}^*_{\text{RM}}, \boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \Theta_+, \quad \ldots \quad (4.1)$$

where strict inequality holds in a neighborhood of the pivot $\mathbf{0}$ ($\in \Theta_+$). *Thus, the positive rule RSMLE dominates the RSMLE which in turn dominates the RMLE when* θ *is restricted to the positive orthant space* Θ_+.

The proof of this theorem and some related lemmas is relegated to the Appendix.

Remarks. First, we may note that in adapting c_a to the set $\mathscr{X}_a(\boldsymbol{S}_n)$, $a \in N_p$, maximal gain is achieved when $c_a = (|a|-2)^+$, $a \in N_p$. This is very similar to that in the unrestricted case. Secondly, as expected, the risk-reduction due to shrinkage is a maximum at the pivot (i.e., $\boldsymbol{\theta} = \mathbf{0}$). Thirdly, in the unrestricted case, the shrinkage factor is c_p while in our case, it depends on the c_a, $a \in N_p$. Since the c_a are generally montone increasing in $|a|$, and our risk reduction involves an average of the reductions over the variaous sectors $\mathscr{X}_a(\boldsymbol{S}_n)$, $a \in N_p$, we would have a comparatively smaller reduction of the risk of the RSMLE over RMLE than in the unrestricted case (i.e., SMLE over UMLE). This is not surprising, as the RMLE are themselves adjusted estimators with due considerations on the set restraints. This raises the issue of comparing the SMLE and RSMLE, and this will be addressed in Section 6. Fourthly, for any $\boldsymbol{\theta} \neq \mathbf{0}$, the reduction in the risk due to shrinkage depends on the sample size n through the noncentralities $n^{\frac{1}{2}} \boldsymbol{\theta}_a$, $a \in N_p$. It is also known [viz., Sen (1984)] that in the unrestricted case, as the noncentrality parameter increases, the risk-reduction becomes smaller, and in the asymptotic case, it achieves the value 0. A similar situation holds here. As n increases, the risk-reduction of the RSMLE over the RMLE becomes smaller and approaces the asymptote 0 as $n \to \infty$. For this reason, in the asymptotic case, it has been suggested [Sen(1984)] that one should use Pitman alternatives (i.e., $\boldsymbol{\theta} = n^{-\frac{1}{2}} \boldsymbol{\lambda}$, $\boldsymbol{\lambda} \in \boldsymbol{\Theta}_+$), for which we would have the same picture as in Theorem 4.1 where $\boldsymbol{\theta}$ has to be replaced by $\boldsymbol{\lambda}$. Finally, the relative dominance picture in Theorem 4.1 is adapted to the particular loss function in (2.7). If instead of $\boldsymbol{\Sigma}^{-1}$ we use an arbitrary $\boldsymbol{W}$ (p.d.) then (4.1) may not hold. This situation is quite comparable to the unrestricted case where the dominance of the SMLE rests on the adoptation of a particular loss function, and for different loss functions, shrinkage estimators having the desired dominance property may differ.

Motivated by the positive-rule shrinkage estimators and a general class of minimax estimators (in the unrestricted case) introduced by Baranchik (1970) and Strawderman (1971), we consider here some more general minimax estimators in the restricted case. In the unrestricted case, the MLE and its shrinkage versions are all equivariant under nonsingular transformations $\boldsymbol{X} \rightarrow \boldsymbol{Y} = \boldsymbol{BX}, \boldsymbol{B}$ nonsingular. However, this equivariance is not generally true for Θ_+ (when $\boldsymbol{\Sigma}$ is arbitrary p.d.) or for a positively homogeneous cone in $\boldsymbol{R}^p$. This makes it difficult to use a canonical reduction (on $\boldsymbol{\theta}, \boldsymbol{\Sigma}$) to establish the desired results for an arbitrary $\boldsymbol{\Sigma}$. However, under an additional condition on $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, such a general dominance result may be obtained. Let us define

$$< \boldsymbol{x}, \boldsymbol{y} >_{\boldsymbol{\Sigma}} = \boldsymbol{x}' \boldsymbol{\Sigma}^{-1} \boldsymbol{y}, < \boldsymbol{x}, \boldsymbol{x} >_{\boldsymbol{\Sigma}} = \| \boldsymbol{x} \|^2_{\boldsymbol{\Sigma}} ; \quad \dots \quad (4.2)$$

$$\psi_m (\boldsymbol{\theta}, \boldsymbol{\Sigma}) = (2\pi)^{p/2} |\boldsymbol{\Sigma}|^{-1/2} \int_{\boldsymbol{z} > \boldsymbol{0}} \dots \int \|z\|^{-m}_{\boldsymbol{\Sigma}} < \boldsymbol{\theta}, z >^{m}_{\boldsymbol{\Sigma}} \exp \left\{ -\frac{1}{2} \|z\|^2_{\boldsymbol{\Sigma}} \right\} dz, m \geqslant 0; \quad \dots \quad (4.3)$$

$$\Theta^*_+ = \{\boldsymbol{\theta} \in \Theta_+ : \psi_m (\boldsymbol{\theta}_{a:a'}, \boldsymbol{\Sigma}_{a:a'}) \geqslant 0, \forall\, m \geqslant 0 \text{ and } a \subseteq N_p\}. \quad \dots \quad (4.4)$$

Note that a sufficient condition for Θ^*_+ to be non-empty is that $\boldsymbol{\Sigma}^{-1}_{a:a'} \boldsymbol{\theta}_{a:a'} \geqslant \boldsymbol{0}$, for every $a \in N^*_p$ $(= \{a \in N_p : |a| > 2\}), p \geqslant 3$, and this is true in particular when $\boldsymbol{\Sigma}$ is diagonal. Thus, the results to follow hold for Θ_+ when $\boldsymbol{\Sigma}$ is diagonal and $p \geqslant 3$.

For every $a\, \phi \subseteq a \subseteq N_p$, let $t_a(y) : \boldsymbol{R}^+ \rightarrow (0, 2(|a|-2)^+)$ be a nondecreasing, non-negative and bounded function of y, and for $p \geqslant 3$, let $\boldsymbol{X} \sim N_p(\boldsymbol{\theta}, \boldsymbol{\Sigma})$,

$$\hat{\boldsymbol{\theta}}^* = \sum_{\{\phi \subseteq a \subseteq N_n\}} 1(\boldsymbol{X} \in \mathscr{X}_a(\boldsymbol{\Sigma})) \left\{ 1 - 1(|a| > 2) [y^{-1} t_a(y)]_{y = \|\hat{\boldsymbol{\theta}}_{\text{RM}}\|^2_{\boldsymbol{\Sigma}}} \right\} \hat{\boldsymbol{\theta}}^*_{\text{RM}}, \quad \dots \quad (4.5)$$

where $\hat{\boldsymbol{\theta}}^*_{\text{RM}}$ refers to the RMLE in the case of known Σ (and $n = 1$). Note that the positive-rule estimator in (3.4) [with $\boldsymbol{S}_n$ replaced by $\boldsymbol{\Sigma}$] is a member of this class.

Theorem 4.2. *Let $\boldsymbol{X} \sim N_p(\theta, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is such that $\boldsymbol{\theta} \in \Theta_+$. Also assume that $t_a(y)$ is nonnegative, monotone nondecreeasing and and bounded from above by $2(|a|-2)$, for every $a \subseteq N^*_p$. Then, for $p > 3$, any estimator of the form* (4.5) *dominates $\hat{\boldsymbol{\theta}}^*_{\text{RM}}$ (over $\boldsymbol{\Theta}_+$), and hence, is minimax (over Θ_+).*

The proof of the theorem is sketched in the Appendix. In passing, we may remark that when $\boldsymbol{\Sigma}$ is diagonal, the $\mathscr{X}_a(\boldsymbol{\Sigma}) = \mathscr{X}_a(\boldsymbol{I})$ and $\hat{\boldsymbol{\theta}}_{\text{RM}}$ do not depend on $\boldsymbol{\Sigma}$, but the factor $[y^{-1}t_a(y)]_{y = \|\hat{\boldsymbol{\theta}}_{\text{RM}}\|^2_{\boldsymbol{\Sigma}}}$ may depend on $\boldsymbol{\Sigma}$, and hence, (4.5) depends on $\mathbf{\Sigma}$, even if it is diagonal. However, for a diagonal $\mathbf{\Sigma}$, $\Theta_+^* = \Theta_+$, and the conclusion holds for the entire domain Θ_+. This may hold for some non-diagonal $\mathbf{\Sigma}$ too, and some of these cases will be treated in the next section.

5. Some extensions and important applications

An extension of the positive orthant model is considered here and some specific applications in linear models are presented along with.

5.1. *Sub-orthant model.* As an extension of (2.1), we consider the following :

$$X(m+p \text{ -vector}) = (X_1', X_2')' \sim N_{m+p}(\boldsymbol{\theta}, \mathbf{\Sigma}) \; ; \boldsymbol{\theta} = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2')' \quad \dots \quad (5.1)$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are m and p-vectors respectively, and $\mathbf{\Sigma}$ is a $(m+p)\times(m+p)$ matrix (p.d.). In this setup, $\boldsymbol{\mu}_1$ is unrestricted while $\boldsymbol{\mu}_2$ belongs to the positive orthant $\boldsymbol{R}^{+p}$, i.e.,

$$\Theta_+^0 = \{\boldsymbol{\theta} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) : \boldsymbol{\mu}_1 \in \boldsymbol{R}^m \text{ and } \boldsymbol{\mu}_2 \in \boldsymbol{R}^{+p}\}, \quad \dots \quad (5.2)$$

and the *pivot* for $\boldsymbol{\mu}_2$ is 0. We denote by $N_p^0 = \{m+1, \dots, m+p\}$ and for every $a : \phi \subseteq a \subseteq N_p^0$, the complemetnary subset (a') as well as the $\boldsymbol{X}_{1:a}(\mathbf{\Sigma})$, $\boldsymbol{X}_{a:a'}(\Sigma)$ and $\underset{a:a'}{\Sigma}$ etc., are defined as in earlier sections. It can be shown (viz., Sengupta and Sen (1987)) that for this sub-orthant model, the RMLE of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_{\text{RM}} = \sum_{\phi \subseteq a \subseteq N_p^0} \mathscr{P}_{N_m, a}(\boldsymbol{X}_{1:a'}, \boldsymbol{X}_{a:a'}, \mathbf{0}) \; 1(\mathbf{\Sigma}_{a'a'}^{-1}\boldsymbol{X}_{a'} \leqslant \mathbf{0}, \boldsymbol{X}_{a:a'} > \mathbf{0}), \dots \quad (5.3)$$

where $N_m = \{1, \dots, m\}$. Actually, for the case of unknown $\mathbf{\Sigma}$, whenever we have a Wishart matrix $\boldsymbol{S}$ (with M DF), independent of $\boldsymbol{X}$, the RMLE of $\boldsymbol{\theta}$ is also given by (5.3) provided in the definition of the $\boldsymbol{X}_{1:a}$, $\boldsymbol{X}_{a:a'}$ etc., we replace $\mathbf{\Sigma}$ by $\boldsymbol{S}$.

Motivated by the RMLE in (5.3) and the Stein-rule estimators in Sections 3 and 4, we consider the following theorem on improved estimation for this sub-orthant model.

Theorem 5.1. *The shrinkage estimator*

$$\hat{\boldsymbol{\theta}}_{\text{RSM}} = \sum_{\phi \subseteq a \subseteq N_p^0} 1(\boldsymbol{X}_2 \in (\mathscr{X}_a(\mathbf{\Sigma}_2)) \{1 - c_a \|\hat{\boldsymbol{\theta}}_{\text{RM}}\|_{\mathbf{\Sigma}}^{-2}\} \hat{\boldsymbol{\theta}}_{\text{RM}} \quad \dots \quad (5.4)$$

dominates the RMLE in (5.3) [*over* Θ_+^0] *whenever* $p+m \geqslant 3$ *and*

$$0 \leqslant c_a \leqslant 2(|a|+m-2)^+, \textit{for every } a : \phi \subseteq a \subseteq N_p^0. \quad \dots \quad (5.5)$$

The proof is relegated to the Appendix.

For the case of unknown $\boldsymbol{\Sigma}$, under the provision of a Wishart matrix $\boldsymbol{S} \sim W(m+p, M, \boldsymbol{\Sigma})$, in (5.4), we need to replace $\boldsymbol{\Sigma}$ by $\boldsymbol{S}$, while in (5.5), the upper bound for c_a is to be taken as $2(|a|+m-2)/(M-m-p+3)$, for $a : \phi \subseteq a \subseteq N_p^0$. With these changes, the dominance result in Theorem 5.1 remains in tact. In passing, we may remark that a particular case of Theorem 5.1 [viz., $p = 1$] dealing with a single inequality restraint has been considered in Judge and Yancey (1986) [see also Chang (1982)]. However, their heuristic approach may run into considerable difficulties in the non-orthogonal case (where $\boldsymbol{\Sigma}$ may not be diagonal), while in the orthogonal case, under (4.4), we have a more general class of estimators (which also extends readily for this sub-orthant model). We may also remark that the modifications for the case of $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{V}$, $\boldsymbol{V}$ known, can be made as in after (3.2)—(3.3) : If S^2/σ^2 is $\sim \chi^2_m$, independently of $\boldsymbol{X}$, then letting $s^2 = (M+2)^{-1} M S^2$ and $\hat{\boldsymbol{\Sigma}} = s^2 \boldsymbol{V}$, we may allow $0 < c_a < 2(|a|+m-2)^+$, $\forall\, a : \phi \subseteq a \subseteq N_p^0$. This special case is of considerable importance in the context of some useful linear models.

5.2. *Ordered alternative model.* Suppose that $X_{ij}, j = 1, \ldots, n_i$ are *i.i.d.r.v.*'s with the normal distribution $N(\mu_i, \sigma^2)$, $i = 1, \ldots, r$; all these r samples being independent. The ordered alternative model relates to the following positively homogeneous subsepace of $\boldsymbol{R}^r : \Theta^{>} = \{(\mu_1, \ldots, \mu_r) \in \boldsymbol{R}^r : \mu_1 \leqslant \ldots \leqslant \mu_r\}$, so that the pivot is the line of equality $\mu_1 = \ldots = \mu_r = \mu_1 \in \boldsymbol{R}$. If we write

$$\mu_i = \mu_1 \text{ for } i = 1 \text{ and } \mu_i = \mu_1 + \beta_2 + \ldots + \beta_i,\ i = 2, \ldots, r\ ;\ \boldsymbol{\beta}_2 = (\beta_2, \ldots, \beta_r)', \quad \ldots \quad (5.6)$$

then $\boldsymbol{\theta} = (\mu_1, \boldsymbol{\beta}_2)$ relates to the sub-orthant model with $m = 1$ and $p = r-1$. We may also consider a two-way layout : $X_{ij} = \mu_i + \tau_j + e_{ij}$, $1 \leqslant j \leqslant n$, $i = 1, \ldots, n$ and characterize the order alternative model for the treatment effects $\tau_1, \ldots, \tau_r$ as a sub-orthant model. Since these are both particular cases of some linear models, we consider the latter in details.

5.3. *RSMLE for univariate linear models.* Consider the usual linear model

$$\boldsymbol{X}_n = (X_1, \ldots, X_n)' = \boldsymbol{C\beta} + \boldsymbol{e}_n\ ;\ \boldsymbol{e}_n \sim N_n(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n),\ 0 < \sigma^2 < \infty; \quad \ldots \quad (5.7)$$

where $\boldsymbol{C}$ is a known matrix (of order $n \times p^*$) of regression constants, $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$ is a p^*-vector of unknown resression parameters, $\boldsymbol{\beta}_j$ is a p_j-vector, $j = 1, 2$, and σ^2 is unknown. We consider the following sub-orthant model :

$$\boldsymbol{\beta} \in \Theta^{>} = \{\boldsymbol{\beta} : \boldsymbol{\beta}_1 \in \boldsymbol{R}^{p_1}, \boldsymbol{\beta}_2 \in \boldsymbol{R}^{+p}\},\ p_2 = p \text{ and } p^* = p_1 + p_2. \quad \ldots \quad (5.8)$$

Without any loss of generality (and allowing reparameterization if necessary), we may assume that $\boldsymbol{C}$ is of full rank $p^* < n$, and we desire to construct improved estimators of $\boldsymbol{\beta}$ under the set of restraints on $\boldsymbol{\beta}_2$ in (5.8). This model includes the K-sample location model as a special case of $p = K-1$.

Note that the classical MLE for $(\boldsymbol{\beta}, \sigma^2)$ are

$$\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1', \hat{\boldsymbol{\beta}}_2')' = (\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{X}_n \text{ and } \hat{\sigma}^2 = (n-p^*)^{-1}\|\boldsymbol{X}_n - \boldsymbol{C}\hat{\boldsymbol{\beta}}\|^2, \quad \ldots \quad (5.9)$$

and these are jointly sufficient for $(\boldsymbol{\beta}, \sigma^2)$. Moreover, $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ are independent with

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\boldsymbol{C}'\boldsymbol{C}^{-1}) \text{ and } (n-p^*)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n-p^*}. \quad \ldots \quad (5.10)$$

In this case, the RMLE of $\boldsymbol{\beta}$ is denoted by $\hat{\boldsymbol{\beta}}^*_{\mathrm{RM}}$ and is defined by (5.3) with $\boldsymbol{X} = \hat{\boldsymbol{\beta}}$ and $\boldsymbol{\Sigma} = \hat{\sigma}^2(\boldsymbol{C}'\boldsymbol{C})^{-1}$, and as in Theorem 5.1, we obtain the RSMLE as

$$\hat{\boldsymbol{\theta}}_{\mathrm{RSM}} = \sum_{\phi \subseteq a \subseteq N_p^0} 1(\hat{\boldsymbol{\beta}}_2 \in \mathscr{X}_a(\hat{\sigma}^2(\boldsymbol{C}'\boldsymbol{C})^{-1}))\,\{1 - c_a\hat{\sigma}^2\,[\hat{\boldsymbol{\beta}}^{*\prime}_{\mathrm{RM}}(\boldsymbol{C}'\boldsymbol{C})\hat{\boldsymbol{\beta}}^*_{\mathrm{RM}}]^{-1}\}\hat{\boldsymbol{\beta}}^*_{\mathrm{RM}} \ldots \quad (5.11)$$

where

$$0 \leqslant c_a \leqslant 2(|a| + p_1 - 2)^+(n-p)/(n-p+2), \ \forall\ a : \phi \subseteq a \subseteq N_p^0. \ \ldots \quad (5.12)$$

Then, under the quadratic error loss $(\boldsymbol{Y}-\boldsymbol{\beta})'(\boldsymbol{C}'\boldsymbol{C})(\boldsymbol{Y}-\boldsymbol{\beta})/\sigma^2$, we have

$$R(\hat{\boldsymbol{\beta}}_{\mathrm{RSM}}, \boldsymbol{\beta}) \leqslant R(\hat{\boldsymbol{\beta}}_{\mathrm{RM}}, \boldsymbol{\beta}), \ \forall\ \boldsymbol{\beta} \in \boldsymbol{R}^{p_1} \times \boldsymbol{R}^{+p}. \quad \ldots \quad (5.13)$$

We may also consider a positive-rule version of the RSMLE (as in (3.4) with the obvious modifications) and verify that (4.1) holds as well in this model. Parallel dominance results hold for a general class of multivariate linear models too.

6. Relative risk picture of the proposed estimators

Inspite of having some similarity, there is a basic difference in the SMLE and RSMLE. For the normal mean problem, positive orthant model, the SMLE is invariant under orthogonal transformations : $\boldsymbol{X} \to \boldsymbol{Y} = \boldsymbol{U}\boldsymbol{X}$, $\boldsymbol{U}'\boldsymbol{U} = I$, but the RSMLE is not so. Thus the risk function of the SMLE is constant on the countours $\delta = \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}$, and this characterization can be incorporated in the simplification of the risk picture of the SMLE. The risk function of of the RMLE and RSMLE may depend on the unknown $\boldsymbol{\theta}$ in a much more involved manner. This picture may actually depend on whether $\boldsymbol{\theta}$ belongs to the interior of Θ_+ or to its lower dimensional faces (edges). In Section 6.1, we shall study the differential picture of the risk of RSMLE in different subspaces of Θ_+ and incorporate this in the next sub-section to draw the relative picture for the SMLE and RSMLE.

6.1. *Directional variation of the risk of the RSMLE and RMLE over subspaces of* Θ_+. We may observe that at the pivot $\boldsymbol{\theta} = \mathbf{0}$,

$$R(\hat{\boldsymbol{\theta}}_{\text{RMS}}, \boldsymbol{\theta}) = p/2-(p-4)/2-2^{-p}(p+2) = 2-(p+2)/2^p. \quad \dots \quad (6.1)$$

The growth of the risk of the RSMLE (or RMLE) is very much dependent on the direction of deviation of $\boldsymbol{\theta}$ from the pivot. To study this directional variation, we proceed in two steps. First, we study the directional variation of the risk of the RMLE. Then by using (A.11), we draw the picture for the RSMLE. Also, for simplicity, we consider the model $\boldsymbol{X} \sim N(\boldsymbol{\theta}, \boldsymbol{I})$, $\boldsymbol{\theta} \in \Theta_+$ with the pivot $\mathbf{0}$, so that

$$R(\hat{\boldsymbol{\theta}}_{\text{RM}}, \boldsymbol{\theta}) = \sum_{\phi \subseteq a \subseteq N_p} [P_\theta\{\boldsymbol{X}_{a'} \leqslant \mathbf{0}\}E_{\boldsymbol{\theta}}\{\|\boldsymbol{X}_a-\boldsymbol{\theta}_a\|^2\, 1(\boldsymbol{X}_a > \mathbf{0})\}$$

$$+P_{\boldsymbol{\theta}}\{\boldsymbol{X}_a > \mathbf{0}\}P_{\boldsymbol{\theta}}\{\boldsymbol{X}_{a'} \leqslant \mathbf{0}\}.\ \|\boldsymbol{\theta}_{a'}\|^2]. \quad \dots \quad (6.2)$$

Let $\Phi(x)$ and $\phi(x)$ be respectively the standard normal d.f. and p.d.f. Then,

$$P_{\boldsymbol{\theta}}\{X_{a'} \leqslant \mathbf{0}\} = \prod_{j \in a'} \Phi(-\theta_j),\ P_{\boldsymbol{\theta}}\{\boldsymbol{X}_a > \mathbf{0}\} = \prod_{j \in a} \Phi(\theta\ ),\ \phi \subseteq a \subseteq N_p \quad \dots \quad (6.3)$$

$$E_{\boldsymbol{\theta}}\{\|X_a-\boldsymbol{\theta}_a\|^2 1(\boldsymbol{X}_a > \mathbf{0})\} = P_{\boldsymbol{\theta}}\{\boldsymbol{X}_a \geqslant 0\} \sum_{j \in a} \{1-\theta_j\phi(\theta_j)/\Phi(\theta_j)\}, \quad \dots \quad (6.4)$$

for every $a : \phi \subseteq a \subseteq N_p$, so that by (6.2), (6.3) and (6.4), we have

$$R(\hat{\boldsymbol{\theta}}_{\text{RM}}, \boldsymbol{\theta}) = \sum_{\phi \subseteq a \subset N_p} \sum_{j \in a} \prod_{l \in a'} \Phi(\theta_j)\Phi\{-\theta_l)\}\,|a| - \sum_{j \in a} \theta_j\phi(\theta_j)/\Phi(\theta_j) + \sum_{l \in a'} \theta_l^2\}. \quad \dots \quad (6.5)$$

It is clear from (6.5) that the risk of the RMLE depends on the individual $\theta_1, \dots, \theta_p$ in an involved manner (and not simply on $\|\boldsymbol{\theta}\|^2$). For the case of the SMLE or the classical MLE, by virtue of the canonical reduction, it sufficies to consider the case where $\theta = (\delta, 0, \dots, 0)'$ and $\delta^2 = \|\boldsymbol{\theta}\|^2$. Hence, we study first the risk of the RMLE in this case (relating to a one dimensional face of Θ_+). In this case, letting $\boldsymbol{\theta}^0 = (\delta.\ 0')'$, we obtain from (6.5) by some routine steps that

$$(R\hat{\boldsymbol{\theta}}_{\text{RM}}, \boldsymbol{\theta}^0) = p/2+[\Phi(\delta)-1/2]-\{\delta\phi(\delta)-\delta^2[1-\Phi(\delta)]\}. \quad \dots \quad (6.6)$$

Note that by virtue of the Mill's ratio, $\delta\phi(\delta)-\delta^2[1-\phi\delta)]$ is nonnegative for every $\delta \geqslant 0$ and it converges to 0 as $\delta \to \infty$. Thus, at $\boldsymbol{\theta} = \mathbf{0}$, (6.6) is equal to $p/2$ and it is bounded away from above by $(p+1)/2$, for every $\delta \geqslant 0$; the upper bound is attained as $\delta \to \infty$. By virtue of (A.11) and (6.6), we have no letting $c_a = (|a|-2)^+$, for $a : \phi \subseteq a \subseteq N_p$, that under $\boldsymbol{\theta}^0$,

$$R(\hat{\boldsymbol{\theta}}_{\text{RSM}}, \boldsymbol{\theta}^0) = R(\hat{\boldsymbol{\theta}}_{\text{RM}}, \boldsymbol{\theta}^0) - \Sigma_{a \in N_p^*}(|a|-2)^2 E_{\boldsymbol{\theta}^0}\{1(\boldsymbol{X}_a > \mathbf{0}, \boldsymbol{X}_{a'} \leqslant \boldsymbol{\theta})\,|\,|\boldsymbol{X}_a|\,|^2\}. \quad \dots \quad (6.7)$$

Note that if $\{1\} \notin a$,

$$E_{\boldsymbol{\theta}^0}\{1(\boldsymbol{X}_a > \boldsymbol{\theta}, \boldsymbol{X}_{a'} \leqslant \boldsymbol{0})\|\boldsymbol{X}_a\|^2\} = \Phi(-\delta)2^{-p+1}(|a|-2)^{-1}, \quad \ldots \quad (6.8)$$

while on letting $g_m(v)$ be the central chi square p.d.f. with m DF, we have for $\{1\} \epsilon a$,

$$E_{\boldsymbol{\theta}^0}\{1(\boldsymbol{X}_a > 0, \boldsymbol{X}_{a'} \leqslant \boldsymbol{0})\|\boldsymbol{X}_a\|^2\} = 2^{-p+1} \int_0^\infty \int_0^\infty (v+y^2)^{-1} \phi(y-\delta) g_{|a|-1}(v) dv dy \quad \ldots \quad (6.9)$$

Therefore, from (6.6) through (6.9), we obtain that

$$\begin{aligned} R(\hat{\boldsymbol{\theta}}_{RMS}, \boldsymbol{\theta}^0) = \; & p/2 + [\Phi(\delta) - 1/2] - \{\delta\phi(\delta) - \delta^2[1-\Phi(\delta)]\} \quad \ldots \quad (6.10) \\ & -2^{-(p-1)}[\Phi(-\delta) \sum_{r=3}^{p} (r-2) \binom{p-1}{r} \\ & - \sum_{r=3}^{p} (r-2)^2 \binom{p-1}{r-1} \int_0^\infty \int_0^\infty (y^2+v)^{-1} g_{r-1}(v) . \phi(y-\delta) dv dy]. \end{aligned}$$

It may be noted that by the usual expansion of $\phi(y-\delta)$ in a Taylor series expansion in δ and identifying the Hermite polynomials in y in these successive terms, we may as well use Lemma A.1 to provide a power series representation for the last term in (6.10) in δ with the coefficients depending on the $\psi_m(\boldsymbol{\theta}^0, \boldsymbol{1})$. However, for intended brevity, we refrain ourself from this formidable task.

Next, we consider another extreme case where $\boldsymbol{\theta} = \theta\boldsymbol{1}$, $\theta > 0$, so that $\Delta = p\theta^2$. Thus, here $\boldsymbol{\theta}$ lies on the line of equality $\theta_1 = \ldots = \theta_p$ and is an interior point of Θ_+. If we let $\alpha = \Phi(\theta)$, $\beta = 1-\theta\phi(\theta)/\Phi(\theta)$ and $\gamma = \phi(-\theta)/\Phi(\theta)$ (so that $\alpha(1+\gamma) = 1$), then (6.5) simplifies to

$$\begin{aligned} & \alpha^p\{p(1-\beta)(1+\gamma)^{p-1} + p\,\theta^2\gamma(1+\gamma)^{p-1}\} \\ & = p\{\Phi(\theta) - \theta\Phi(\theta) + \theta^2[1-\phi(\theta)]. \quad \ldots \quad (6.11) \end{aligned}$$

In passing, we may remark that the SMLE has the risk equal to 2 at $\boldsymbol{\theta} = \boldsymbol{0}$, while for any $\boldsymbol{\theta} \neq \boldsymbol{0}$, this risk is greater than 2 (but bounded from above by p). Also, we may note that at $\boldsymbol{\theta} = \boldsymbol{0}$, both (6.6) and (6.11) reduce to $p/2$. Thus, for $p \geqslant 5$, the RMLE may not perform better than the SMLE (particularly near the pivot). Given this discouraging feature of the RMLE, the need for RSMLE is felt even more for larger values of p. Next, to compare (6.6) and (6.11), we set $h(x) = p\{\Phi(x) - x\phi(x) + x^2[1-\Phi(x)]\} - [\Phi(x\sqrt{p}) - x\sqrt{p}\phi(x\sqrt{p})] + px^2[1-\Phi(x\sqrt{p})[$, and note that

$$h'(x) = 2\sqrt{p}x\,[\Phi(x\sqrt{p}) - \Phi(x)] \geqslant 0, \text{ for every } x > 0. \quad \ldots \quad (6.12)$$

Thus, corresponding to a given Δ ($= \delta^2$ in Case (i) and $p\theta^2$ in Case (ii)), the differenence between (6.11) and (6.6) is equal to 0 at $\Delta = 0$ and is a monotone nondecreasing function of Δ with the upper asymptote equal to $(p-1)/2(\geqslant 1)$. This clearly indicates the non-uniformity of the risk of the RMLE over the Δ-contours : The risk is smaller on the boundaries of Θ_+ while it may be considerably larger as $\boldsymbol{\theta}$ moves away to the interior. By setting $c_a = (|a|-2)^+$, for $\phi \subsetneq a \subsetneq N_p$ and using (A.1), we may simplify the right hand side of (A.11) to

$$\sum_{r=3}^{p} \binom{p}{r} (r-2)^2[\Phi(-\theta)]^{p-r} (\exp\{-(r/2)\theta^2\}) \{ \sum_{k \geqslant 0} (k\,!)^{-1} 2^{(k-1)/2}$$

$$. \psi_k(\theta\, 1_r, \boldsymbol{I}_r)\Gamma((r+k-1)/2)/\Gamma((r/2))\}, \qquad \dots \quad (6.13)$$

where $\mathbf{1}_r$ and $\boldsymbol{I}_r$ stand for the r-dimensional vector $(1, \dots, 1)$, and identity matrix, respectively. Thus, for $\boldsymbol{\theta} = \boldsymbol{\theta}^{=} = \theta\,\mathbf{1}_p$, $\Delta = p\theta^2$, we have

$$R(\hat{\boldsymbol{\theta}}_{\text{RMS}}, \boldsymbol{\theta}^{=}) = p\{\Phi(\theta) - \theta\phi(\theta) + \theta^2[1-\Phi(\theta)] - \sum_{r=3}^{p} (r-2)^2 \binom{p}{r} [\Phi(-\theta)]^{p-r}.$$

$$\{(\exp\{-(r/2)\}^2) \sum_{k \geqslant 0} (k\,!)^{-1} 2^{(k-1)/2} \psi_k(\theta \mathbf{1}_r, \boldsymbol{I}_r)\, \Gamma((r+k-1)/2)/\Gamma((r/2)\}. \qquad \dots \quad (6.14)$$

Note that by (4.3), for $\boldsymbol{\theta} = \boldsymbol{\theta}^{=}$, $\psi_k(\theta\mathbf{1}_r, \boldsymbol{I}_r)$ is $d_k\, \theta^k = d_k\, \Delta^k/p^{k/2}$, where the d_k are positive constants depending on the k and hence the second termon the right hand side of (6.14) can equivalently be expressed as a sole function of Δ, although in an infinite series form.

In a similar manner, we may consider a lower dimensional face of Θ_+ by setting

$$\boldsymbol{\theta} = \boldsymbol{\theta}_{(s)} = (\Delta/s)^{\frac{1}{2}}(\mathbf{1}'_s, \mathbf{0}'_{p-s})' \; ; \Delta \geqslant 0, \text{ for } s = 1, \dots, p. \qquad \dots \quad (6.15)$$

Simplification of (6.5) is indeed possible under (6.15), and it would then be a function of Δ. It can be shown that for $\Delta > 0$, the risk of the RMLE under (6.15) is a monotone nondecreasing function of $s (0 \leqslant s \leqslant p)$. This picture reveals that the performance of the RMLE is not uniform over the Δ-contours; rather the more $\boldsymbol{\theta}$ is close to the boundary of Θ_+, the better may be the performance. Also, using (A.1), under (6.15), the right hand side of (A.11) may also be expressed in a form (somewhat) similar to that in (6.13), although for $s < p$, the form will be more complicated. Actually, for an arbitrary $\boldsymbol{\theta} \in \Theta_+$, we may use (A.1) to express a typical term in (A.11) [when $\boldsymbol{\Sigma} = \boldsymbol{I}_p$] in terms of homogeneous functions of $\boldsymbol{\theta}$ of degrees k ; $k \geqslant 0$.

6.2. *Dominance properties of the RSMLE.* By virtue of the discussions made in Section 6.1., we draw the following conclusions :

For $p = 3$, if $\boldsymbol{\theta} = (\delta, 0, 0)'$, the RMLE dominates the SMLE, and hence, by Theorem 4.1, the RSMLE dominates the SMLE. In this context observe that by (6.6), $R(\hat{\boldsymbol{\theta}}_{SM}, \boldsymbol{\theta}) = 2$ for $\boldsymbol{\theta} = \boldsymbol{\theta}^0$, and it monotonically increases as $\delta(= \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta})$ increases, and finally, its upper asymptote is equal to $p\ (= 3)$ (viz., James and Stein, 1961).

Therefore, in the sequel, we confine ourselves to the case of $p \geqslant 4$.

Theorem 6.1. *Let* $\boldsymbol{X} \sim N_p(\boldsymbol{\theta}, \boldsymbol{I})$ *where* $p \geqslant 3$ *and* $\boldsymbol{\theta} \in \Theta_+$. *Then*

$$R(\hat{\boldsymbol{\theta}}_{RSM}, \boldsymbol{\theta}) \leqslant R(\hat{\boldsymbol{\theta}}_{SM}, \boldsymbol{\theta}), \textit{for all } \boldsymbol{\theta} \in \{\sqrt{\delta}\boldsymbol{e}_1 ; \delta > 0\}, \quad \ldots \quad (6.16)$$

where $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$ *is the basis vector and* $\delta = \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}$.

Proof. Note that the relevance of $\{\sqrt{\delta}\boldsymbol{e}_1, \delta > 0\}$ follows from the invariance of the risk of the SMLE under rotation, although other choices may be important for the RSMLE. We only consider the case of $p \geqslant 4$, as for $p = 3$, the result has already been proved earlier.

First, consider the case of $p = 4$. Since $R(\hat{\boldsymbol{\theta}}_{SM}, \boldsymbol{\theta}) = p - (p-2)^2 E(\chi^{-2}_{p,\delta})$, by (A.11) and (6.6), we obtain that

$$R(\hat{\boldsymbol{\theta}}_{SM}, \boldsymbol{\theta}) - R(\hat{\boldsymbol{\theta}}_{RMS}, \boldsymbol{\theta}) = p/2 - (p-2)^2 E(\chi^{-2}_{p,\delta}) - [\phi(\delta^{\frac{1}{2}}) - \tfrac{1}{2}] + \{\delta^{\frac{1}{2}}\phi(\delta^{\frac{1}{2}}) - \delta(1-\phi(\delta^{\frac{1}{2}}))\}$$
$$+ \sum_{a \subseteq N_p^*} (|a|-2)^2 E_{\boldsymbol{\theta}}\{1(\boldsymbol{X} \in \mathcal{X}_a)\|\boldsymbol{X}_a\|^{-2}\} \quad \ldots \quad (6.17)$$

The right hand side of (6.17) reduces to

$$2 - 4E(\chi^{-2}_{4,\delta}) - [\phi(\delta^{\frac{1}{2}}) - \tfrac{1}{2}\delta^{\frac{1}{2}}\phi(\delta^{\frac{1}{2}}) + (1-\phi(\delta^{\frac{1}{2}}))] + \sum_{a \subseteq N_p^*} (|a|-2)^2 E_{\boldsymbol{\theta}}\{1(\boldsymbol{X} \in X_a)\|\mathcal{X}_a\|^{-2}\}$$
$$= A_4(\delta) + B_4(\delta), \text{ say.} \quad \ldots \quad (6.18)$$

Note that $B_4(\delta)$ is nonnegative for all $\delta \geqslant 0$. Hence, it suffices to show that $A_4(\delta)$ is $\geqslant 0$, for every $\delta \geqslant 0$. For this, note that

$$E(X^{-2}_{4,\delta}) = \exp(-\delta/2) \sum_{r \geqslant 0} (\delta/2)^r (r\,!)^{-1}(2+2r)^{-1}, \text{ for every } \delta \geqslant 0. \quad \ldots \quad (6.19)$$

Also,

$$[1 - \Phi(\delta^{\frac{1}{2}})] \leqslant \delta^{-\frac{1}{2}}\phi(\delta^{\frac{1}{2}}) = (2\pi\delta)^{-\frac{1}{2}} \exp(-\delta/2), \ \forall\ \delta \geqslant 0, \quad \ldots \quad (6.20)$$

while, by standard results,

$$4E(X^{-4}_{6,\delta}) = \exp(-\delta/2) \sum_{r \geqslant 0} (\delta/2)^r (r\,!)^{-1} 4/[(2+2r)\,(4+2r)] \geqslant_{\frac{1}{2}} \exp(-\delta/2), \delta \geqslant 0. \quad \ldots \quad (6.21)$$

Hence, by (6.19), (6.20) and (6.21), we have

$$A_4(0) = 0, \quad \ldots \quad (6.22)$$

$$(\partial/\partial u)A_4(u)\,|_{u=\delta} = 4E(X^{-4}_{6,\delta}) - [1-\Phi(\delta^{\frac{1}{2}})] > 0, \ \forall\ \delta > (2/\pi)\ (\simeq 0.636). \quad \ldots \quad (6.23)$$

At $\delta = 0$, (6.23) is equal to 0, while for $0 < \delta \leqslant 2/\pi$, we note that

$$(\partial/\partial\theta)\,\{4E(X_{6,\delta}^{-4})-[1-\Phi(\delta^{\frac{1}{2}})]\} = \phi(\delta^{\frac{1}{2}})-16\delta^{\frac{1}{2}}E(\chi_{8,\delta}^{-2})$$

$$= \exp(-\delta/2)\,\{(2\pi)^{-\frac{1}{2}}-\sum_{r\geqslant 0}(\delta/2)^r(r\,!)^{-1}[(2+2r)\,(4+2r)\,(6+2r)]^{-1}\,\delta^{\frac{1}{2}}\}$$

$$\geqslant \exp(-\delta/2)\,\{(2\pi)^{-\frac{1}{2}}-(\delta^{\frac{1}{2}}3)\sum_{r\geqslant 0}(\delta/2)^r(r\,!)^{-1}\}$$

$$= \phi(\delta^{\frac{1}{2}})\,\{1-(2\pi)^{\frac{1}{2}}(\delta^{\frac{1}{2}}/3)\exp(\delta/2)\}$$

$$\geqslant \phi(\delta^{\frac{1}{2}})\,\{1-(2/3)\exp(\pi^{-1})\} > 0,\ 0 < \delta \leqslant (2/\pi), \quad \ldots (6.24)$$

as $\delta^{\frac{1}{2}}\exp(\delta/2) \leqslant (2/\pi)\exp(\pi^{-1})$ and $\exp(\pi^{-1}) < 3/2$. Thus, (6.23) is $\geqslant 0$, for all $0 \leqslant \delta \leqslant (2/\pi)$, and hence, (6.23) and (6.24) imply that $A_4(\delta)$ is $\geqslant 0$, for every $\delta \geqslant 0$. Hence for $p = 4$, the RSMLE dominates the SMLE.

Consider now the general case of $p \geqslant 5$. Note that we have here $X_1 \sim N(\delta^{\frac{1}{2}},1)$ and $X_i \sim N(0, 1)$, for $i = 2, \ldots, p$. We denote the last term on the right hand side of (6.17) by $C_p(\delta)$ and write it as

$$C_p(\delta) = \sum_{a\subseteq N_p^*}(|a|-2)^2E_\theta\{1(\boldsymbol{X}\in\mathscr{X}_a)\,\|\boldsymbol{X}_a\|^2\,I(1\notin a)\}$$

$$+\sum_{a\subseteq N_p^*}(|a|-2)^2E_\theta\,\{1(\boldsymbol{X}\in\mathscr{X}_a)\,\|\boldsymbol{X}_a\|^2\,1(1\in a)$$

$$= C_p^{(1)}(\delta)+C_p^{(2)}(\delta),\ \text{say.} \quad \ldots (6.25)$$

Now, by direct evaluation, it follows that

$$C_p^{(1)}(\delta) = \Phi(-\delta^{\frac{1}{2}})\,\{(p-5)/2+(p+1)2^{-p+1}\}\,; \quad \ldots (6.26)$$

$$C_p^{(2)}(0) = (p-3)/4+2^{-p}\ \text{and}\ C_p(0) = (p-4)/2+(p+2)2^{-p}. \quad \ldots (6.27)$$

For every $r > 2$, let us denote by

$$A_r(\delta) = E_\theta\{(X_1^2+\ldots+X_r^2)^{-1}\,1(X_j \geqslant 0, j = 1, \ldots, r)\}. \quad \ldots (6.28)$$

Then, we may write

$$C_p^{(2)}(\delta) = \sum_{r=2}^{p}\binom{p-1}{r-1}2^{-(p-r)}(r-2)^2\,A_r(\delta). \quad \ldots (6.29)$$

Noting that the joint density of $X_1, \ldots, X_r$ is given by $\phi(x-\delta^{\frac{1}{2}})\phi(x_2)\ldots\phi(x_r)$, we obtain that

$$(\partial/\partial u)A_r(u)\,|_{u=\delta} = (2\delta^{1/2})^{-1}\,E_\theta\{(X_1-\delta)\|\boldsymbol{X}_r\|^{-2}\,1(\boldsymbol{X}_r \geqslant 0)\},\ r > 2. \quad \ldots (6.30)$$

Since in (6.30) X_1 is confined to $\boldsymbol{R}^+$, the derivative is positive at $\delta = 0$ and it continues to be positive for all $\delta \geqslant \delta_0^{(r)}$, for some $\delta_0^{(r)} > 0$, while it may be negative for $\delta > \delta_0^{(r)}$. Moreover, by (6.30), we obtain that even for $\delta > \delta_0^* = \min_{r > z} \delta_0^{(r)}$,

$$(\partial/\partial u)\, A_r(u)\,|_{u=\delta} > -\frac{1}{2}\, A_r(\delta), \qquad \ldots \quad (6.31)$$

and actually, (6.31) holds for all $\delta \geqslant 0$. Hence, (6.29) and (6.31) ensure that

$$(\partial/\partial\delta) C_p^{(2)}(\delta) > -\frac{1}{2}\, C_p^{(2)}(\delta), \text{ for all } \delta, \qquad \ldots \quad (6.32)$$

which in turn implies that

$$C_p^{(2)}(\delta) \geqslant \exp(-\delta/2) C_p^{(2)}(0), \text{ for all } \delta \geqslant 0. \qquad \ldots \quad (6.33)$$

Thus, by (6.25), (6.26), (6.27) and (6.33), we have

$$C_p(\delta) \geqslant \Phi(-\delta^{\frac{1}{2}})\, \{(p-5)/2+(p+1)2^{-p+1}\}+\exp(-\delta/2)\, \{(p-3)/2+2^{-p}\}, \forall \delta \geqslant 0. \qquad \ldots \quad (6.34)$$

Consequently, the right hand sidejof (6.17) is bounded from below by

$$(p/2)-(p-2)^2 E(\chi_{p,\delta}^{-2}+\{\delta^{\frac{1}{2}}\phi(\delta^{\frac{1}{2}})-[1-\Phi(\delta^{\frac{1}{2}})]\}-(1/2\ +\Phi(-\delta^{\frac{1}{2}})$$

$$\{(p-3^{\frac{1}{2}}/2+(p+1)2^{-p+1}\}+\exp(-\delta/2\{(p-3)/2+2^{-p}\}, \forall\ \delta \geqslant 0. \qquad \ldots \quad (6.35)$$

At $\delta = 0$, (6.35) reduces to $(p+2)/2^p > 0$, and proceeding as in (6.19) through (6.21) but replacing 4 by p], it follows that the derivative of (6.35) with respect to δ is indeed nonnegative for all δ, and hence, (6.35) is nonnegative for all $\delta > 0$. This completes the proof of the theorem.

Let us now consider the relative risk picture in the *least favourable direction* : $\boldsymbol{\theta} = \delta^{\frac{1}{2}}\, \mathbf{1}$ where $\mathbf{1} = (1, \ldots, 1)'$ and $\delta > 0$. In this case, (4.4) reduces to

$$\sum_{r=3}^{p} \binom{p}{r} (r-2)^2\, [\Phi(-\delta^{\frac{1}{2}})]^{p-r}\, A_r^0(\delta) = C_p^*(\delta), \text{ say}, \qquad \ldots \quad (6.36)$$

where

$$A_r^0(\delta) = \int \ldots \int_{R^{r+}} ||\boldsymbol{x}||^{-2} \prod_{j=1}^{r} \phi(x_j-\delta^{\frac{1}{2}}) d\boldsymbol{x}, \text{ for } r = 3, \ldots, p. \qquad \ldots \quad (6.37)$$

Proceeding as in (6.30) and (6.31), we obtain that for every $\delta > 0$,

$$(\partial/\partial u) A_r^0(u)\,|_{u=\delta} > -(r/2)\, A_r^0(\delta) \geqslant (-p/2)\, A_r^0(\delta), \text{ for } r = 3, \ldots, p. \qquad \ldots \quad (6.38)$$

Consequently, proceeding as in (6.32) through (6.33), me obtain from (6.36)—(6.38) that

$$C_p^*(\delta) \geqslant \exp(-p\delta/2)C_p^*(0) = \exp(-p\delta/2)\{(p-4)/2+(p+2)2^{-p}\}, \forall\ \delta \geqslant 0. \quad \ldots \quad (6.39)$$

As such, noting that for $\boldsymbol{\theta} = \delta^{\frac{1}{2}}\mathbf{1}$,

$$R(\hat{\boldsymbol{\theta}}_{SM}, \boldsymbol{\theta}) - R(\hat{\boldsymbol{\theta}}_{RSM}, \boldsymbol{\theta}) = p-(p-2)^2E(X_{p,p\delta}^{-2}) - p\{\Phi(\delta^{\frac{1}{2}}) - \delta^{\frac{1}{2}}\phi(\delta^{\frac{1}{2}}) + [1-\Phi(\delta^{\frac{1}{2}})]\} + C_p^*(\delta),$$

we obtain by using (6.39) and similar manipulations as in the proof of Theorem 6.1 that the following result holds.

Theorem 6.2. *If* $\boldsymbol{\theta} = \delta^{\frac{1}{2}}\mathbf{1}, \delta > 0$, *the RSMLE dominates the SMLE for* $p = 3$ *and* 4.. *For* $p \geqslant 5$, *the same conclusion holds at least for large values of* δ.

We may refer to Sengupta and Sen (1987) for the proof of Theorem 6.2 and some other related results.

7. Comparison with the change estimators

We have mentioned earlier that Chang (1981, 1982) has considered some versions of restricted shrinkage estimators. If $X \sim N(\mu, 1), \mu > 0$, the Katz (1961) showed that

$$\hat{\mu}(X) = X + \phi_0(X);\ \phi_0(x) = \exp\left(-\frac{1}{2}x^2\right) \Big/ \left\{\int_{-\infty}^{x} \exp\left(-\frac{1}{2}u^2\right)du\right\},\ x \in R \quad \ldots \quad (7.1)$$

is an admissible estimator of μ. State by side, consider the model $\boldsymbol{X} \sim N(\boldsymbol{\theta}, \boldsymbol{I})$' Chang (1982) considered a Katz-type estimator $\boldsymbol{\delta}_2 = (\delta_1^2, \ldots, \delta_p^2)'$ where

$$\delta_i^2(\boldsymbol{X}) = X_i + \phi_0(X_i) - c(X_i + \phi_0(X_i))/\{\Sigma(X_j + \phi_0(X_j))^2\},\ i = 1, \ldots, p, \quad \ldots \quad (7.2)$$

where $0 < c < 2(p-2)$ and $p \geqslant 3$; here also $\boldsymbol{\theta} \in \Theta_+$. Under quadratic error loss, $\boldsymbol{\delta}_2$ dominates the classical MLE ($\boldsymbol{X}$). Further, defining $t(.)$ as in (4.5), we may also introduce the other restricted shrinkage estimator due to Chang (1982, 1982), which is denoted by $\boldsymbol{\delta}_1 = (\delta_1^1, \ldots, \delta_p^1)$,. Here

$$\delta_i^1(X_i) = X_i + t(X_i) - cX_i/(\Sigma X_j^2),\ \text{if all the } X_j \text{ are} \geqslant 0, \quad \ldots \quad (7.3)$$

and $X_i + t(X_i)$, otherwise, for $i = 1, \ldots, p$.

Let us also introduce $\boldsymbol{\delta}_0 = (\delta_1^0, \ldots, \delta_p^0)$ where

$$\delta_i^0(X_i) = X_i + t(X_i),\quad i = 1, \ldots, p. \quad \ldots \quad (7.4)$$

Then Chang (1982) has shown that for the simultaneous estimation of $\theta \epsilon \Theta_+$), under the quadratic error loss, (7.3) dominates (7.4). However, we may remark that it is possible to construct alternative estimators (similar to those in Section 3) which dominate (7.3) and/or (7.4). Corresponding to (7.3), our proposed estimator is

$$\delta_1^{*i}(\boldsymbol{X}) = X_i+t(X_i)-c_a X_i^+\|\boldsymbol{X}_a\|^{-2} \text{ on } \mathscr{X}_a(I),\ a \subsetneq N_p. \quad \ldots \quad (7.5)$$

Note then

$$R(\boldsymbol{\delta}_1(\boldsymbol{X}), \boldsymbol{\theta})-R(\boldsymbol{\delta}_1^*(\boldsymbol{X}), \boldsymbol{\theta}) = \sum_{\phi \subseteq a \subseteq N} E_{\boldsymbol{\theta}}\Big[\sum_{i \epsilon a} (X_i+t(X_i)-\theta_i)^2$$

$$- \sum_{i \epsilon a} (X_i+t(X_i)-c_a X_i^+\|\boldsymbol{X}_a\|^{-2}-\theta_i)^2\Big] \,.\, 1(\boldsymbol{X} \epsilon \mathscr{X}_a(I)))$$

$$= \sum_{a \subset N_p^*} E_{\boldsymbol{\theta}}\Big[2 \sum_{i \epsilon a} c_a X_i(X_i+t(X_i)-\theta_i)-c_a^2\|\boldsymbol{X}\|^{-2}\Big] 1(\boldsymbol{X} \epsilon \mathscr{X}_a(I)))$$

$$= \sum_{a \subseteq N_p^*} E_{\boldsymbol{\theta}}[1(\boldsymbol{X} \epsilon \mathscr{X}_a(I))\|X\|^{-2}] > 0, \quad \ldots \quad (7.6)$$

where the penultimate step follows from the nonnegativity of $t(.)$ and Theorem 4.1. In a similar manner, corresponding to $\boldsymbol{\delta}_2(\boldsymbol{X})$ in (7.2), we introduce the following :

$$\delta_2^{*i}(\boldsymbol{X}) = \begin{cases} X_i+\phi_0(X_i)-c_a(X_i+\phi_0(X_i)) [\sum_{j \epsilon a} (X_j+\phi_0(X_j))^2]^{-1}, & i \epsilon a \\ X_i+\phi_0(X_i), \text{ if } i \notin a, & \end{cases} \quad \ldots \quad (7.7)$$

on $\mathscr{X}_a(I)$, for every $a \subsetneq N_p$. Here, $c_a \epsilon [0, 2(|a|-2)^+]$, for every $a \subsetneq N_p$. Then proceeding as in Chang (1982) but employing the results developed in this paper, it follows that

$$R(\boldsymbol{\delta}_0(\boldsymbol{X}), \boldsymbol{\theta})-R(\boldsymbol{\delta}_2^*(\boldsymbol{X}), \boldsymbol{\theta}) \geqslant 0, \ \forall\ \boldsymbol{\theta} \epsilon \Theta_+, \quad \ldots \quad (7.8)$$

so that (7.7) dominates (7.4). It can also be shown that both $\boldsymbol{\delta}_2$ and $\boldsymbol{\delta}_2^*$ are equivalent when $\boldsymbol{\theta}$ moves away from the pivot $\boldsymbol{0}$ (inside Θ_+) and their difference in risk is exponentially in $\|\boldsymbol{\theta}\|$ negligible. For small departures from the pivot, the relative risk picture of these two estimators depend very much on the direction of the shift, and no general conclusion for their relative superiority can be drawn. However, from the motivational point of view, looking at our earlier sections, we are more in favor of $\boldsymbol{\delta}_2^*$ than $\boldsymbol{\delta}_2$. We conclude this section with a remark that for the single inequality contraint alternative, Judge and Yancey (1986) have considered some RSMLE, although their findings rest on a restrictive regularity condition which is not generally met in the case of a positively homogeneous subset of R_p, and hence, may not be applicable here.

Appendix

Proofs of Theorems 4.1, 4.2 and 5.1

We need a basic lemma in these derivations, and this is presented first.

Lemma A.1. *Let* $X \sim N_p(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ *p.d. Then, for every* $(r, s) : s \geqslant 0$, $p+r+s > 0$.

$$E_{\boldsymbol{\theta}}\{\|\boldsymbol{X}\|^r_{\boldsymbol{\Sigma}} < \boldsymbol{X}, \boldsymbol{\theta} >^s_{\boldsymbol{\Sigma}} 1(\boldsymbol{X} > \boldsymbol{0})\} = (\exp\{-\tfrac{1}{2}\|\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}}\}) \sum_{k \geqslant 0} (k\,!)^{-1}\psi_{k+s}(\boldsymbol{\theta}, \boldsymbol{\Sigma}).$$

$$2^{(r+s+k)/2}\Gamma((r+s+k+p)/\Gamma(p/2). \quad \ldots \quad (A.1)$$

Proof of Lemma A.1. Note that the left hand side of (A.1) is equal to

$$\int_{\boldsymbol{R}^{+p}} \cdots\cdots \int \|\boldsymbol{X}\|^r_{\boldsymbol{\Sigma}} < \boldsymbol{X}, \boldsymbol{\theta} >^s_{\boldsymbol{\Sigma}} \exp\left\{-\frac{1}{2}\|\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}} + < \boldsymbol{\theta}, \boldsymbol{X} >_{\boldsymbol{\Sigma}} - \frac{1}{2}\|\boldsymbol{X}\|^2_{\boldsymbol{\Sigma}}\right\} d\boldsymbol{X} \ldots \quad (A.2)$$

By the use of the Cauchy-Schwarz inequality (on $< \boldsymbol{X}, \boldsymbol{\theta} >_{\boldsymbol{\Sigma}}$) and the Dominated Convergence Theorem, it can be shown that (A.2) is finite if $s \geqslant 0$ and $r+s+p \geqslant 0$. Also, the left hand side of (A.1) may be written as

$$\left(\exp\left\{-\frac{1}{2}\|\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}}\right\}\right) \sum_{k \geqslant 0} (k\,!)^{-1} \int_{\boldsymbol{R}^{+p}} \cdots\cdots \int \|\boldsymbol{X}\|^r_{\boldsymbol{\Sigma}} < \boldsymbol{X}, \boldsymbol{\theta} >^{s+k}_{\boldsymbol{\Sigma}} \exp\left\{-\frac{1}{2}\|\boldsymbol{X}\|^2_{\boldsymbol{\Sigma}}\right\} d\boldsymbol{X}$$

$$= \left(\exp\left\{-\frac{1}{2}\|\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}}\right\}\right) \sum_{k \geqslant 0} (k\,!)^{-1} E_{\boldsymbol{0},\boldsymbol{\Sigma}}\|\boldsymbol{X}\|^r_{\boldsymbol{\Sigma}} < \boldsymbol{X}, \boldsymbol{\theta} >^{s+k}_{\boldsymbol{\Sigma}} 1(\boldsymbol{X} \geqslant 0)\Big\}, \ldots \quad (A.3)$$

where $E_{\boldsymbol{0},\boldsymbol{\Sigma}}$ denotes the expectation under $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$. Now, it is known that under $N_p(0, \boldsymbol{\Sigma})$, $(\|X\|^{-2}_{\boldsymbol{\Sigma}}\boldsymbol{X}, 1(\boldsymbol{X} \geqslant \boldsymbol{0}))$ is (jointly) independent of $\|\boldsymbol{X}\|_{\boldsymbol{\Sigma}}$ so that the right hand side of (4.5) can be written as

$$\left(\exp\left\{-\frac{1}{2}\|\boldsymbol{\theta}\|^2_{\boldsymbol{\Sigma}}\right\}\right) \Sigma_{k \geqslant 0}(k\,!)^{-1}E_{0,\boldsymbol{\Sigma}}\left[\|\boldsymbol{X}\|^{-(k+s)}_{\boldsymbol{\Sigma}} < \boldsymbol{X}\,\boldsymbol{\theta} >^{s+k,}_{\boldsymbol{\Sigma}} 1(\boldsymbol{X} \geqslant 0)\right]$$

$$E_{0,\boldsymbol{\Sigma}}\ \|\boldsymbol{X}\|^{r+s+k}_{\boldsymbol{\Sigma}}\Big\}$$

$$= \left(\exp\left\{\frac{1}{2}\|\boldsymbol{\theta}\|^2\right\}\right] \sum_{k \geqslant 0} (k\,!)^{-1}\psi_{k+s}(\boldsymbol{\theta},\boldsymbol{\Sigma})2^{(r+s+k)/2}\Gamma((r+s+k+p)/2)/\Gamma(p/2), \quad \ldots \quad (A.4)$$

where the last identity follows from (4.3) and the central moments of $\|\boldsymbol{X}\|_{\boldsymbol{\Sigma}}$ Q.E.D.

Let us now return to the proof of Theorem 4.1. For simplicity of presentation, we shall consider only the case of $\boldsymbol{\Sigma}$ being known ; a similar but more complicated proof works out as well for the case of totally unknown $\boldsymbol{\Sigma}$ or the case of $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{V}$ with $\boldsymbol{V}$ known, and for these details, we may refer to Sengupta and Sen (1987). Note that by virtue of (3.1) [with S_n replaced by Σ] and (2.7)—(2.8), for every $p \geqslant 3$,

$$R(\hat{\theta}^*_{RSM}, \theta) = \sum_{\phi \subseteq a N_p} E_\theta\{\|\theta_{RM} - \theta - c_a(\hat{\theta}'_{RM}\Sigma^{-1}\hat{\theta}_{RM})^{-1}\hat{\theta}_{RM}\|^2_\Sigma \, 1(X \in \mathscr{X}_a(\Sigma))\}$$

$$= R(\hat{\theta}^*_{RM}, \theta) - 2 \sum_{a \subseteq N^*_p} E_\theta\{c_a 1(X \in \mathscr{X}_a(\Sigma))$$

$$+2 \sum_{a \subseteq N^*_p} E_\theta \left\{ c_a 1(X \in \mathscr{X}_a(\Sigma)) \frac{<\theta, \hat{\theta}^*_{RM}>_\Sigma}{\|\hat{\theta}^*_{RM}\|_\Sigma} \right\} + \sum_{a \subseteq N^*_p} c_a^2 E\hat{\theta}\{\|\hat{\theta}^!_{RM}\|^{-2}_\Sigma 1(X \in \mathscr{X}_a(\Sigma))\}, \quad \ldots \quad (A.5)$$

where X stands for X_n and by an appeal to the sufficiency of X_n, we may set, without any loss of generality, $n = 1$. Thus, it suffices to show that for every $\theta \in \Theta$, the net contribution of the last three terms in (A.5) is negative. Towards this, note that

$$\sum_{a \subseteq N^*_p} E_\theta\{c_a 1(X_a \in \mathscr{X}(\Sigma))\} = \sum_{a \subseteq N^*_p} c_a P_\theta\{X_{a:a'}(\Sigma) > 0, \Sigma^{-1}_{a'a'} X_{a'} \leqslant 0\}$$

$$= \sum_{a \subseteq N^*_p} P_\theta\{X_{a:a'}(\Sigma) > 0\}. \quad P_\theta\{\Sigma^{-1}_{a'a'} X_{a'} \leqslant 0\}, \quad \ldots \quad (A.6)$$

as $X_{a:a'}(\Sigma)$ and $X_{a'}$ are mutually independent, for every $a \subseteq N_p$. Thus, by (A.1), reduces to ... (A.6)

$$\sum_{a \subseteq N^*_p} c_a P_\theta\{ \Sigma^{-1}_{a'a'} X_{a'} \leqslant 0\}. \left(\exp\left\{-\frac{1}{2} \|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a'}}\right\}\right)$$

$$\left[\sum_{k \geqslant 0} 2^{k/2}(k\,!)^{-1} \psi_k(\theta_{a:a'}(\Sigma), \Sigma_{a:a'}) \Gamma((k+|a|)/2)/\Gamma((|a|)/2)\right]. \quad \ldots \quad (A.7)$$

Similarly, using the identities that $\Sigma^{aa} = \Sigma^{-1}_{a:a'}$ and $\Sigma^{a'a} = -\Sigma^{-1}_{a'a'} \Sigma_{a'a} \Sigma^{aa}$ along with (3.2) (with S_n and $\bar{X}_n$ being replaced by Σ and X respectively), we have

$$< \theta, \hat{\theta}_{RM} >_\Sigma = (\theta'_a \Sigma^{a'a} + \theta'_{a'} \Sigma^{a'a}) X_{a:a'}(\Sigma) = \theta'_{a:a} \Sigma^{-1}_{a:a'} X_{a:a'}(\Sigma), \ldots \quad (A.8)$$

so that

$$\sum_{a \subseteq N^*_p} E_\theta\{c_a \|\hat{\theta}_{RM}\|^{-2}_\Sigma < \theta, \hat{\theta}_{RM} >_\Sigma 1(X \in \mathscr{X}_a(\Sigma))\}$$

$$= \sum_{a \subseteq N^*_p} c_a E_\theta\{1(X_{a:a'}(\Sigma) > 0, \ \Sigma^{-1}_{a'a'} X_{a'} \leqslant 0) < \theta, \hat{\theta}_{RM} >_\Sigma . \ \|\hat{\theta}_{RM}\|^{-2}_\Sigma\}$$

$$= \sum_{a \subseteq N^*_p} c_a P_\theta\{ \Sigma_{a'a'} X_{a'} \leqslant 0\}. \left(\exp\left\{-\frac{1}{2}\|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a}}\right\}\right).$$

$$\left[\sum_{k \geqslant 0} (k\,!)^{-1} 2^{(k-1)/2} \psi_{k+1}(\theta_{a:a'}(\Sigma), \ \Sigma^{a:a'}(\Gamma((k+|a|)/2)/\Gamma((|a|)/2)\right]$$

$$= \sum_{a \subseteq N^*_p} c_a P_\theta\{\Sigma^{-1}_{a'a'} X_{a'} \leqslant 0\}. \ \left(\exp\left\{-\frac{1}{2}\|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a'}}\right\}\right).$$

$$\left[\sum_{k \geqslant 1} k(k!)^{-1} 2^{(k-2)/2} \psi_k(\theta_{a:a'}(\Sigma)), \Sigma_{a:a'}) \Gamma((k-2+|a|)/2)/\Gamma(|a|)/2)\right]. \ldots \quad (A.9)$$

Similarly, using (A.6) (with $s = 0$ and $r = -2$, $p = |a|$), the last term in (A.5) reduces to

$$\sum_{a \subseteq N_p^*} c_a^2 P_{\theta}\{\Sigma_{a'a'}^{-1} X_{a'} \leqslant 0\}. \left(\exp\left\{-\frac{1}{2}\|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a'}}\right\}\right).$$

$$\left[\sum_{k \geqslant 0} (k\,!)^{-1} 2^{(k-2)/2} \psi_k(\theta_{a:a'}(\Sigma), \Sigma_{a:a'}) \Gamma((k-2+|a|)/2)/\Gamma((|a|)/2)\right]. \quad \ldots \quad (A.10)$$

Therefore, from (A.5) through (A.10), we obtain that

$$R(\hat{\theta}^*_{RM}, \theta) - R(\hat{\theta}^*_{RSM}, \theta)$$

$$= \sum_{a \subseteq N_p^*} c_a P_{\theta}\{\Sigma_{a'a'}^{-1} X_{a'} \leqslant 0\}. \left(\exp\left\{-\frac{1}{2}\|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a'}}\right)\right\}.$$

$$\{\psi_0(\theta_{a:a'}(\Sigma), \Sigma_{a:a'}) [2 - c_a \Gamma((|a|)/2-1)/(2\Gamma((|a|)/2))]$$

$$+ \sum_{k \geqslant 1} (k\,!)^{-1} \psi_k(\theta_{a:a'}(\Sigma), \Sigma_{d:a'}) [2^{k/2}/\Gamma((|a|)/2)]$$

$$[2\Gamma((k+|a|)/2) - k\Gamma((k-2+|a|)/2) - (c_a/2)((k+|a|-2)/2)]\}$$

$$= \sum_{a \subseteq N_p^*} c_a(2(|a|-2)-c_a) P_{\theta}\{\Sigma_{a'a'}^{-1} X_{a'} \leqslant 0\}. \left(\exp\left\{-\frac{1}{2}\|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a'}}\right\}\right).$$

$$\left[\sum_{k \geqslant 0} \psi_k(\theta_{a:a'}(\Sigma), \Sigma_{a:a'})(k\,!)^{-1} 2^{(k-2)/2} \Gamma((k-2+|a|)/2)/\Gamma((|a|)/2)\right]$$

$$= \sum_{a \subseteq N_p^*} c_a(2(|a|-2)-c_a) E_{\theta}\{1(X \in \mathscr{X}_a(\Sigma))\|\hat{\theta}_{RM}\|^{-2}_{\Sigma}\} \geqslant 0, \quad \ldots \quad (A.11)$$

as on N_p^*, $|a| > 2$ and $0 < c_a < 2(|a|-2)$. This proves the second inequality in (4.1). To establish the first inequality (i.e., the dominance of the positive-rule RSME), we may write $\{1 - c_a\|\hat{\theta}_{RM}\|^{-2}_{\Sigma}\}^+ = 1(\|\hat{\theta}_{RM}\|^2_{\Sigma} > c_a)\{1 - c_a\|\hat{\theta}_{RM}\|^{-2}_{\Sigma}\}$ and virtually repeat the same proof as in (A.5) through (A.11). Hence, for intended brievity, we omit the details. This completes the proof of Theorem 4.1.

Consider next the proof of Theorem 4.2. We very much follow the lines of the proof of Theorem 4.1. Parallel to (A.5), we have here

$$R(\hat{\theta}^*, \theta) = R(\hat{\theta}_{RM}, \theta) - 2(I) + 2(II) + (III), \quad \ldots \quad (A.12)$$

where compared to (A.6), (A.9) and (A.10), we have here

$$(I) = \sum_{a \subseteq N_p^*} P_{\theta}\left\{\Sigma_{a'a'}^{-1} X_a \leqslant 0\right\}\left[\exp\left\{-\frac{1}{2}\|\theta_{a:a}(\Sigma)\|^2_{\Sigma_{a:a'}}\right\}\right].$$

$$\left(\sum_{\geqslant 0} \frac{1}{k\,!} \psi_k(\theta_{a:a'} \Sigma_{a:a'})\right) E_0[\|X_{a:a'}\|^k_{\Sigma_{a:a'}} t_a(\|X_{a:a'}\|^2_{\Sigma_{a:a'}})]), \quad \ldots \quad (A.13)$$

$$(II) = \sum_{a \subseteq N_p^*} P_{\theta}\{\Sigma_{a:a'}^{-1} X_{a'} \leqslant 0\}\Big[\exp\Big\{-\frac{1}{2}\|\theta_{a:a'}(\Sigma)\|^2_{\Sigma_{a:a'}}\Big\}\Big].$$

$$\Big(\sum_{k\geqslant 0}\frac{1}{k!}\psi_k(\theta_{a:a'}\,\Sigma_{a:a'})E_0[\|X_{a:a'}\|^k_{\Sigma_{a:u'}} t_a\|X_{a:a'}\|^2_{\Sigma_{a:a'}})]\Big), \quad \ldots \quad (A.14)$$

$$(III) = \sum_{a \subseteq N_p^*} P_{\theta}\Big\{\Sigma_{a'a'}^{-1} X_a \leqslant 0\Big\}\Big[\exp\Big\{-\frac{1}{2}\|\theta_{a:a'}\|^2_{\Sigma_{a:a'}}\Big\}\Big].$$

$$\Big(\sum_{k\geqslant 0}\frac{1}{k!}\psi_k(\theta_{a:a'},\Sigma_{a:a'})E_0[\|X_{a:a'}\|^{k-2}\Sigma_{a:a}\, t_a^2\|X_{a:a'}\|^2\Sigma, \quad \ldots \quad (A.15)$$

Thus, from (A.12) through (A.15), we obtain that

$$R(\hat{\theta}_{RM},\theta)-R(\hat{\theta}^*,\theta) = \sum_{a \subseteq N_p^*} P_{\theta}\{\Sigma_{a'a'}^{-1} x_{a'} \leqslant 0\}\Big[\exp\Big\{-\frac{1}{2}\|\theta_{a:a'}\|\Sigma_{a:a'}\Big\}\Big].$$

$$\cdot\Big[\sum_{k\geqslant 0}\frac{1}{k!}\psi_k(\theta_{a:a'},\Sigma_{a:a'})E_0[t_a(y)\{2y^{k/2}-2ky^{k/2-1}-y^{k/2-1}t_a(y)\}]_{y=\|X_{a:a'}\|^2_{\Sigma_{a:a'}}}\Big] \quad \ldots \quad (A.16)$$

Now, under $\theta = 0$, $\|X_{a:a'}\|^2\Sigma_{a:a'} \sim X^2_{|a|}$, for every $a : \phi \subsetneq a \subsetneq N_p$, so that writing $U = \|X_{a:a'}\|\Sigma_{a:a'}$, we obtain that for every a $\subsetneq N_p^*$ and $k \geqslant 0$,

$$E_0\{t_a(U)[2U^{k/2}-2kU^{k/2-1}-U^{k/2-1}t_a(U)]\}$$

$$= \Big\{2^{\frac{1}{2}|a|}\Gamma\Big(\frac{1}{2}|a|\Big)\Big\}^{-1}\int_0^{\infty} e^{-\frac{1}{2}u}u^{\frac{1}{2}|a|-1}[t_a(u)\{2u^{k/2}-2ku^{k/2-1}-u^{k/2-1}t_a(u)\}]du$$

$$= \Big\{2^{k/2-1}\Gamma\Big[\frac{1}{2}(|a|+k-2)\Big]/\Gamma\Big[\frac{1}{2}|a|\Big]\Big\}$$

$$E\{t_a(X^2{}_{|a|+k-2})[2\chi^2{}_{|a|+k-2}-2k-t_a(\chi^2_{|a|+k-2})]\}$$

$$\geqslant \Big\{2^{k/2-1}\Gamma\Big[\frac{1}{2}(|a|+k-2\Big]/\Gamma\Big[\frac{1}{2}|a|\Big]\Big\}$$

$$E(t_a(\chi^2{}_{|a|+k-2})[2\chi^2{}_{|a|+k-2}-2(|a|+k-2)]\} \quad (A.17)$$

as $t_a(u) \leqslant 2(|a|-2)$. Now, $t_a(y)$ and $2y-2(|a|+k-2)$ are both monotone $\nearrow$ in y and $E\chi_q^2 = q$, $\forall\, q \geqslant 0$. Therefore, the right hand side of (A.17) greater than or equal to

$$\{2^{k/2-1}\Gamma((|a|+k-2)/2)/\Gamma((|a|)/2)\}.\ E_0\{t_a(\chi^2{}_{|a|+k-2})\}$$

$$E_0\{2\chi^2{}_{|a|k-2}-2(|a|+k-2)]\} = 0,\ \forall\, k \geqslant 0 \text{ and } a \subsetneq N_p^*. \quad \ldots \quad (A.18)$$

Finally, by (4.4), the $\psi_k(\theta_{a:a'},\Sigma_{a:a'})$ are all nonnegative, and hence, by (A.13) through (A.18), we conclude that (A.12) is nonnegative. Q.E.D.

Finally, we consider the proof of Theorem 5.1. Note that Lemma A.1 extends directly to the case where the positive orthant is replaced by any positively homogeneous cone, and hence, we may virtually repeat the proof of Theorem 4.1 where we need to replace (3.2) by (5.3). For intended brevity, the details are therefore omitted.

Acknowlegements. The authors are grateful to the Co-editor and the referees for their useful comments on the manuscript.

References

BARANCHIK, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution. *Ann. Math. Statist.* **41**, 642-645.

BARLOW, R. E., BARTHOLOMEW, D. J., BERMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions*, John Wiley, New York.

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd Ed. Springer-Verlag, New York.

CHANG, Y.-T (1981). Stein-type estimators for parameters restricted by linear inequalities. *Keio Sci. Tech. Rep.* **34**, 83-95.

——— (1982). Stein-type estimators of parameters in truncated spaces. *Keio Sci. Tech. Rep.* **35**, 185-193.

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. 4th. Berkeley Symp. Math. Statist. Probability*, **1**, 361-379.

JUDGE, G. G. and YANCEY, T. A. (1986). *Improved Methods of Inference in Econometrics*, North Holland, Amsterdam.

KUDO, A. (1963). A multivariate analogue of the sided test. *Biometrika*, **50**, 403-418.

NUESCH, P. E. (1966). On the problem of testing location in multivariate problems for restricted alternatives. *Ann. Math. Statist.* **37**, 113-119.

PERLMAN, M. D. (1969). One sided testing problems in multivariate analysis. *Ann. Math. Statist.* **40**, 549-567.

ROBERTSON, T., WRIGHT, F. T. and DYKSTRA, R. L. (1988). *Order Restricted Statistical Inference*, John Wiley, New York.

SEN, P. K. (1984). A James-Stein detour of U-statistics. *Commun. Statist. Theor. Meth.* **A13**, 2725-2747.

SENGUPTA, D. and SEN, P. K. (1987). Shrinkage estimation in a restricted parameter space. *Inst. Statist., Univ. N. Carolina Mimeo Rep.* 1731.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. 3rd Berkeley Symp. Math. Statist. Probability* **1**, 197-206.

STRAWDEDMAN, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.* **42**, 385-388.

DIVISION OF THEORETICAL
STATISTICS AND MATHEMATICS
INDIAN STATISTICAL INSTITUTE
203 B. T. ROAD
CALCUTTA 700 035
INDIA

DEPARTMENT OF BIOSTATISTICS
SCHOOL OF PUBLIC HEALTH
THE UNIVERSITY OF NORTH CAROLINA
CBH 7400 RESNNAU HALL
CHAPEL HILL NC. 27599-7400.
U.S.A.