JOURNAL OF MULTIVARIATE ANALYSIS 24, 11–30 (1988)

On the Structure of Admissible Linear Estimators

W. KLONECKI AND S. ZONTEK

Institute of Mathematics,
Polish Academy of Sciences,
Wroclaw, Poland

Communicated by the Editors

Using a technique originated by A. Olsen, J. Seely, and D. Birkes (*Ann. Statist.* **4** (1976), 878–890) and developed by L. R. LaMotte (*Ann. Statist.* **10** (1982), 245–256) we establish necessary conditions of C. R. Rao's type (*Ann. Statist.* **4** (1976), 1023–1037) for a linear estimator to be admissible among the class of linear estimators in a general linear model. They are shown to be sufficient for the regression model with a nonnegative definite covariance matrix and for the model with the mean lying in a subspace and the covariance operators varying through the set of all nonnegative definite symmetric matrices. From these results necessary and/or sufficient conditions for admissibility of nonhomogeneous estimators are also derived.

1. INTRODUCTION

The current interest in the problem of characterizing admissible linear estimators in general linear models derives from two sources. The first one is characterization of linear admissible estimators in regression models. Here the paper of Rao [16] published in 1976 should be mentioned, where a complete characterization of admissible linear estimators has been given for the regression model with covariance $\sigma^2 V$ with a known pd (positive definite) matrix V. Results regarding this topic were also published in [1, 2, 6, 7, 9, 13, 18, 20].

The second source of interest is the work aimed at the characterization of admissible quadratic invariant (unbiased and biased) estimators for variance components in mixed linear models. This field of research was stimulated mainly by a paper of Olsen *et al.* [14], also published in 1976. Generalizations and extensions of these results may be found in [4, 5, 8, 12, 13].

Received April 1983; revised September 12, 1984.

AMS 1970 subject classifications: Primary 62 J 99; Secondary 62 F 10.

Key words and phrases: general linear model, linear estimators, admissibility.

0047-259X/88 $3.00

In particular, in [13], LaMotte extended the method of characterizing admissible estimators which was originated in [14]. It is a step-wise procedure and it applies to any general linear model with no restrictions on the relation between mean vectors and covariance matrices. At each step it has to be examined whether or not the considered estimator is best among an affine set of linear estimators, with the dimension of the affine sets decreasing at each step.

Similarly, as in [16], the purpose of this paper is to characterize the structure of admissible linear estimators. Using the technique developed by LaMotte we establish easy to check (one-step) necessary conditions for admissibility of homogeneous linear estimators in a general linear model. Also we show that these conditions are sufficient in two particular cases: (i) the regression model with covariance matrix $\sigma^2 V$ with a known nnd (non-negative definite) matrix V; (ii) the model with the mean vectors varying through a linear subspace and the covariance matrices varying through the set of all nnd matrices. Making use of the fact that within the general linear model considered in this paper, nonhomogeneous estimators are a particular case of homogeneous estimators, necessary and/or sufficient conditions for admissibility of nonhomogeneous estimators are also given.

For the reason of simplicity of notation we restrict our considerations to random vectors Y taking values in the Euclidean vector space $\mathscr{R}^n$ endowed with the usual inner product. However, the results presented carry over to random vectors, taking values in a finite-dimensional vector space endowed with an arbitrary inner product (see [5]).

2. PRELIMINARIES

Let $Y \in \mathscr{R}^n$ be a random vector with an unknown distribution belonging to $\mathscr{P}$, say. Suppose that the expected value μ and the covariance V of Y exist for all distributions in $\mathscr{P}$. It is desired to estimate $C'\mu$, where C is an $n \times t$ matrix, while C' stands for the transpose matrix of C. The estimators considered are $L'Y$, where L is an $n \times t$ matrix. For simplicity we refer to the estimator $L'Y$ of $C'\mu$ in terms "estimator L of C." To compare estimators we use the risk function $E(L'Y - C'\mu)'\,(L'Y - C'\mu)$. Writing the risk as $[L, VL] + [L - C, \mu\mu'(L - C)]$, we see that it depends on the distribution of Y only through V and $\mu\mu'$. Here $[A, B]$ stands for the trace of matrix $A'B$. Taking $(V, \mu\mu')$ as the parameter, the parameter space, to be denoted by $\mathscr{T}$, is a subset of the Cartesian product $\mathscr{V}_n \times \mathscr{V}_n$, where $\mathscr{V}_n$ is the set of all $n \times n$ *nnd* matrices. The aim of this paper is to study the structure of admissible linear estimators within models with parameter spaces $\mathscr{T} \subset \mathscr{V}_n \times \mathscr{V}_n$.

For a given $n \times s$ matrix F let $\mathscr{R}(F)$ denote the subspace spanned by the columns of F and $\mathscr{N}(F)$ the subspace $\{x \in \mathscr{R}^s: Fx = 0\}$ called the null space of F. Moreover, let $\mathscr{K}(F) = \mathscr{R}(F \otimes I_t)$, I_t being the $t \times t$ unit matrix, be the subspace spanned by $F \otimes I_t$, where $(A \otimes B)\, C = ABC'$. As usual F^+ denotes the Moore–Penrose generalized inverse matrix of F and $\|F\| = [F, F]^{1/2}$. For a given set $\mathscr{A} \subset \zeta_n \times \cdots \times \zeta_{n'}$ where ζ_n is the space of the $n \times n$ symmetric matrices, denote by $[\mathscr{A}]$ the minimal closed convex cone containing $\mathscr{A}$. Element $A \in [\mathscr{A}]$, $A = (A_1, ..., A_m)$ say, is referred to as maximal in $[\mathscr{A}]$ if

$$\mathscr{N}(A_1 + \cdots + A_m) = \bigcap_{(B_1, ..., B_m) \in [\mathscr{A}]} \mathscr{N}(B_1 + \cdots + B_m).$$

There exists a maximal element $(V_{\max}, \phi_{\max})$ in $[\mathscr{T}]$ such that $V_{\max}$ and $\phi_{\max}$ are maximal in $\mathscr{V}$ and Φ, respectively. Here $\mathscr{V} = \{V: \exists \phi, (V, \phi) \in [\mathscr{T}]\}$ and $\Phi = \{\phi: \exists V, \ (V, \phi) \in [\mathscr{T}]\}$. The subspace $\mathscr{N}(V_{\max} + \phi_{\max})$ is referred to as the null space of the maximal elements in $[\mathscr{T}]$. Also throughout the paper

$$C_0 = \phi_{\max} \phi_{\max}^+ C \qquad \text{and} \qquad \mathscr{C} = C_0 + \mathscr{K}(I - \phi_{\max} \phi_{\max}^+).$$

Following LaMotte [13] denote by $\mathscr{W}$ a subspace of $\zeta_n \times \zeta_n$ such that $\mathscr{T} \subset \mathscr{W}$.

Denote by $\rho(S; L)$ the extension of the risk function from $\mathscr{T}$ to $\mathscr{W}$ defined for every $S = (S_1, S_2) \in \mathscr{W}$ by

$$\rho(S; L) = [L, S_1 L] + [L - C, S_2(L - C)]. \tag{2.1}$$

Estimator L in the affine set $\mathscr{L} = L_0 + \mathscr{K}(F)$ is called best among $\mathscr{L}$ at point $S \in \mathscr{W}$ if $\rho(S; L) \leqslant \rho(S; M)$ for all $M \in \mathscr{L}$. For $L \in \mathscr{L}$, say $L = L_0 + FZ$, formula (2.1) becomes $\rho(S; L) = V_S(Z, Z) + \rho(S; L_0)$, where, for $s \times t$ matrices Z_1 and Z_2, $V_S(Z_1, Z_2) = [Z_1, T(S)\, Z_2] + [Z_1 + Z_2, U(S)]$, with $T(S) = F'(S_1 + S_2)\, F$ and $U(S) = F'S_1 L_0 + F'S_2(L_0 - C_0)$. If $T(S)$ is nnd, then $V_S(Z_1, Z_2)$ is a quasi-inner product on the space of $s \times t$ matrices. Consequently, by using the projection theorem for a quasi-inner product (see Drygas [3] or Theorem 3.1 in LaMotte [13]), we get that L is best among $\mathscr{L}$ at $S \in \mathscr{W}$ iff $T(S)$ is nnd and $T(S)\, Z + U(S) = 0$. A point $S \in \mathscr{W}$ is said to be a trivial point for $\mathscr{L}$ if every L is best among $\mathscr{L}$ at S. By Theorem 3.1 in [13], the set of trivial points $\mathscr{S}$ for $\mathscr{L}$ is given by $\mathscr{S} = \{S \in \mathscr{W}: T(S) = U(S) = 0\}$. For each set $\mathscr{X} \subset \mathscr{W}$ the relations "as good as" and "better than" among $\mathscr{L}$ on $\mathscr{X}$ may be defined in the usual way. Estimator L is said to be admissible among $\mathscr{L}$ on $\mathscr{X}$ if $L \in \mathscr{L}$ and if no other estimator in $\mathscr{L}$ is better than L on $\mathscr{X}$. As noted in [13] the relations among $\mathscr{L}$ are equivalent on $\mathscr{T}$, $[\mathscr{T}]$ and on

$[\mathcal{T}+\mathcal{S}]$. The term "among $\mathcal{L}$" will be omitted if $\mathcal{L}$ coincides with the space of all $n \times t$ matrices.

The following result of LaMotte (see Corollary 3.10 in [13]) plays an important role in characterizing admissible linear estimators.

THEOREM 2.1. *In order that L be admissible among $\mathcal{L}$ on $\mathcal{T}$ it is necessary and sufficient that either $\mathcal{T} \subset \mathcal{S}$ or there exists a point S in $[\mathcal{T}+\mathcal{S}]$ which is not trivial for $\mathcal{L}$, such that L is best among $\mathcal{L}$ at S and is admissible among $\mathcal{L}_1 = \{M \in \mathcal{L} : M$ is best among $\mathcal{L}$ at $S\}$ on $\mathcal{T}$.*

The following results will be required later. As above, let $\mathcal{L} = L_0 + \mathcal{K}(F)$ and let P be the orthogonal projection on the null space of the maximal elements in $[\mathcal{T}]$.

PROPOSITION 2.2. *If L is an admissible estimator among $\mathcal{L}$ on $\mathcal{T}$ and if the risk function of an estimator M such that $(I-P)M \in (I-P)(\mathcal{L})$ is on $\mathcal{T}$ identical with the risk function of L, then $(I-P)M = (I-P)L$.*

Proposition 2.2 is a consequence of the fact that the risk function is strictly convex with respect to L at each point $(V, \phi) \in [\mathcal{T}]$ *provided* $V+\phi$ *is* pd.

PROPOSITION 2.3. *Let W be any nonsingular $n \times n$ matrix. If L is an admissible estimator of C among $\mathcal{L}$ on $\mathcal{T}$, then WL is an admissible estimator of WC among $W(\mathcal{L})$ on $\mathcal{T}_* = \{((W')^{-1} VW^{-1}, (W')^{-1} \mu\mu' W^{-1}) : (v, \mu\mu') \in \mathcal{T}\}$. If $W^2 = A + I - AA^+$, where A is nnd, then $W^{-1}AW^{-1} = AA^+$.*

To establish the first part of Proposition 2.3 it is sufficient to notice that $L'Y = (WL)'(W')^{-1} Y$ and that $\mathcal{T}_*$ is the parameter space of $(W')^{-1} Y$.

In view of this proposition we may assume without loss of generality that there exists an idempotent maximal element in $\mathcal{V}$ or that there exists an idempotent maximal element in Φ.

PROPOSITION 2.4. *If L is an admissible estimator of C among $\mathcal{L}$ on $\mathcal{T}$, then for any $t \times s$ matrix T, the estimator LT is admissible for CT among $L_0 + \mathcal{R}(F \otimes I_s)$ on $\mathcal{T}$.*

Proposition 2.4 is a slight generalization of an important lemma due to Shinozaki [18] (see also Rao [16]) which says that if L is an admissible estimator of C under the risk function $E(L'Y - C'\mu)'(L'Y - C'\mu)$ introduced in Section 2, then it is also admissible under any risk function of the form $E(L'Y - C'\mu) M(L'Y - C'\mu)$, where M is nnd.

To prove this proposition one needs a more general formulation of Theorem 2.1 for affine sets $L_0 + \mathcal{R}(\pi)$, where π may be an arbitrary linear operator with its range in the space of $n \times t$ matrices. The proof of this extension of LaMotte's theorem proceeds along the same lines as the proof of Theorem 2.1 and will be omitted.

PROPOSITION 2.5. *If L is an admissible estimator among $\mathscr{L}$ on $\mathscr{T}$, then L is an admissible estimator among $\mathscr{L}$ within any model $\mathscr{T}_*$ provided $[\mathscr{T}] \subset [\mathscr{T}_*]$ and provided the null spaces of the maximal elements in $[\mathscr{T}]$ and $[\mathscr{T}_*]$ coincide.*

This proposition is an immediate consequence of Proposition 2.2. It shows an interesting fact that the property of admissibility within linear models is not lost by augmentation of the parameter space when the null space of the maximal elements in the new parameter space coincides with the null space of the maximal elements in the original space.

3. NECESSARY CONDITIONS FOR ADMISSIBILITY

The next theorem gives necessary conditions for the admissibility of a linear estimator within a general linear model. We assume only that the dimension of the null space of the maximal elements in $[\mathscr{T}]$ is equal to zero. This is no loss of generality. In case it is not fulfilled, the matrix L appearing in the conditions of the theorem should be replaced (see [13, p. 251]) by $(I-P)L$, where P is the orthogonal projection on the null space of the maximal elements in $[\mathscr{T}]$.

THEOREM 3.1. *If L is an admissible estimator of C, then under the specified assumptions and notation*

(i) $\mathscr{R}(L') = L'\mathscr{R}(\phi_{\max}) \subset \mathscr{R}(C_0')$

(ii) $\mathscr{R}(L'-C') = \mathscr{R}((L-C)' V_{\max}) \subset C'\mathscr{R}(V_{\max})$, $\forall C \in \mathscr{C}$

(iii) *if $Lx = \lambda Cx$ for some $C \in \mathscr{C}$ and $x = a + ib$, where $a, b \in \mathscr{R}^t$, and if $Cx \neq 0$, then λ is a real number and is in the closed interval* $[0, 1]$

(iv) *there exists a nonzero matrix $V \in [\mathscr{V}]$ such that for every $C \in \mathscr{C}$*

(a) *$L'VC$ is symmetric*

(b) *$L'VL \leqslant L'VC$.*

Proof. In order to show that $\mathscr{R}(L') = \mathscr{R}(L'\phi_{\max})$ we may assume by Propositions 2.3 and 2.5 that $\phi_{\max}$ is idempotent and that $\mathscr{V} = \mathscr{V}_n$. Let G be any matrix such that $\phi_{\max} LL' = \phi_{\max} LL' \phi_{\max} G'$. Put $A = \phi_{\max} L$, $B = (I - \phi_{\max}) L$, $D = (I - \phi_{\max}) G$, and $M = (I + D) A$. Then

$$\mathscr{R}(M') \subset \mathscr{R}(A') = \mathscr{R}(M'\phi_{\max}). \tag{3.1}$$

Because $L = A + B$ and because $AB' = AA'D'$, we easily verify that $LL' - MM' = BB' - DAA'D'$ which is nnd. On the other hand, $(L - C)' \phi(L - C) = (M - C)' \phi(M - C)$ for each nnd matrix ϕ such that $\mathscr{R}(\phi) \subset \mathscr{R}(\phi_{\max})$ since $\phi_{\max} B = \phi_{\max} D = 0$. From this and the assumption that L is admissible we conclude that $M = L$ by Proposition 2.2. The first part of (i) follows now from (3.1).

Since L is admissible for C, it follows from Proposition 2.4 that $L[I - (\phi_{\max} C)^+ (\phi_{\max} C)]$ is admissible for 0. Hence $[I - C'\phi_{\max}(C'\phi_{\max})^+]\ L' = 0$ by Proposition 2.2. This proves the second part of (i).

The assertion (ii) follows from (i) by noting that the risk defined by (2.1) is symmetric in the sense that $\rho((S_1, S_2); L) \equiv \rho((S_2, S_1); C - L)$.

To prove (iii) suppose to the contrary that λ is not in $[0, 1]$. Let $\lambda = \alpha + i\beta$, where $\alpha, \beta \in \mathscr{R}$. Since L is admissible for $C \in \mathscr{C}$, it follows from Proposition 2.4 that $L(a, b)$ is admissible for $C(a, b)$. Now $Lx = \lambda Cx$ implies

$$La = \alpha Ca - \beta Cb$$
$$Lb = \alpha Cb + \beta Ca,$$

so that the risk of $L(a, b)$ at point $(V, \phi) \in [\mathscr{T}]$ becomes

$$\begin{aligned}\rho((V, \phi); L(a, b)) = {} & \alpha^2([Ca, VCa] + [Cb, VCb]) \\ & + (\alpha - 1)^2\ ([Ca, \phi Ca] + [Cb, \phi Cb]) \\ & + \beta^2([Ca, (V + \phi)\ Ca] + [Cb, (V + \phi)\ Cb]).\end{aligned}$$

Since $Cx \neq 0$, the expression in the last parentheses is positive for $(V, \phi) = (V_{\max}, \phi_{\max})$. This leads to a contradiction in case $\alpha \notin [0, 1]$ or $\beta \neq 0$, since suitable multiples of $L(a, b)$ would improve $L(a, b)$ at $(V, \phi) = (V_{\max}, \phi_{\max})$.

(α) To prove (iv) we may assume without loss of generality that $\mathscr{W} = \operatorname{span} \mathscr{T}$. In view of Theorem 2.1 there exists a nonzero point (V_0, ϕ_0) in $[\mathscr{T}]$ such that for every C in $\mathscr{C}$

$$(V_0 + \phi_0)\ L = \phi_0 C. \tag{3.2}$$

Transposing and postmultiplying with C yields $L'V_0C = C'V_0C - L'(V_0 + \phi_0)\ L$. Also $L'V_0C - L'V_0L = (L - C)'\ \phi_0(L - C)$. This obviously establishes (iv) with $V = V_0$ when $V_0 \neq 0$.

(β) In the case when $V_0 = 0$, Eq. (3.2) reduces to $\phi_0(L - C) = 0$ with $\phi_0 \neq 0$. This implies that $L \in \mathscr{L}_1 = C + \mathscr{N}(N_1)$. Here C may be any element in $\mathscr{C}$ and N_1 is the orthogonal projection on $\mathcal{N}(\phi_0)$. Since $\phi_{\max}$ is a pd matrix when $N_1=0$, we need to consider only the case when $N_1 \neq 0$.

To show that there exists a nonzero matrix V in $[\mathscr{V}]$ which fulfills (iv) we apply Theorem 2.1 to $\mathscr{L}_1$. By this theorem there exists a nontrivial point (W_1, W_2) in $[\mathscr{T}+\mathscr{S}_1]$, such that $N_1(W_1+W_2)L=N_1W_2C$ for every $C\in\mathscr{C}$. Here $\mathscr{S}_1$ is the set of trivial points for $\mathscr{L}_1$ and the given C. Since for every C in $\mathscr{C}$ there exists a matrix Z such that $L=C+N_1Z$, this becomes

$$N_1(W_1+W_2)N_1Z=-N_1W_1C. \tag{3.3}$$

Because $(W_1, W_2)\in[\mathscr{T}+\mathscr{S}_1]=[[\mathscr{T}]+\mathscr{S}_1]$, there exist two sequences $\{(V^{(m)}, \phi^{(m)})\}\subset[\mathscr{T}]$ and $\{(S_1^{(m)}, S_2^{(m)})\}\subset\mathscr{S}_1$ such that $(V^{(m)}+S_1^{(m)}, \phi^{(m)}+S_2^{(m)})\to(W_1, W_2)$ as $n\to\infty$.

First let us consider the case when there exists a subsequence $\{m_i\}$ such that $V^{(m_i)}\to V_1\neq 0$ as $i\to\infty$. Then, necessarily, $S^{(m_i)}\to S_1$, say, and $V_1\in[\mathscr{V}]$. Moreover, (3.3) implies that

$$N_1(V_1+W_2)N_1Z=-N_1V_1C, \tag{3.4}$$

provided that $\mathscr{W}=\operatorname{span}\mathscr{T}$. Hence $Z'N_1V_1C=-Z'N_1(V_1+W_2)N_1Z$, which in turn implies that

$$L'V_1C=C'V_1C-Z'N_1(V_1+W_2)N_1Z$$

and

$$L'V_1C-L'V_1L=Z'N_1W_2N_1Z.$$

Thus conditions (iv) are met for $V=V_1$.

Next let us consider the case when no subsequence of $\{V^{(m)}\}$ is convergent. Define $U^{(m)}=V^{(m)}/\|V^{(m)}\|$ and let $U^{(m_i)}\to U_1$ as $i\to\infty$. Clearly, $U_1\neq 0$ and $U_1\in(\mathscr{V}]$. Partitioning each $U^{(m_i)}$ as

$$\begin{aligned}U^{(m_i)}&=N_1U^{(m_i)}N_1+N_1U^{(m_i)}(I-N_1)+(I-N_1)U^{(m_i)}N_1\\&\quad+(I-N_1)U^{(m_i)}(I-N_1)\end{aligned}$$

and noting that $\{N_1U^{(m_i)}N_1\}$ converges to the zero matrix, we obtain that $U_1=(I-N_1)U_1(I-N_1)$. Consequently, $L'U_1C=C'U_1C$ and $L'U_1L=L'U_1C$. Thus, in this case, conditions (iv) hold with $V=U_1$.

In the remaining case when $V^{(m_i)}\to 0$ as $i\to\infty$, relation (3.4) reduces to $N_1W_2N_1Z=0$, i.e., $N_1W_2(L-C)=0$, where $N_1W_2N_1\neq 0$. Hence $L\in\mathscr{L}_2=C+\mathscr{K}(N_2)$, where, as above, C may be any element in $\mathscr{C}$, while N_2 is the orthogonal projection on $\mathscr{R}(N_1)\cap\mathcal{N}(N_1W_2N_1)$.

It should be noted that $\phi_{\max}$ must be pd when $N_2 = 0$. In fact, since $\mathscr{R}(W_2) \subset \mathscr{R}(\phi_{\max})$ and since $\mathscr{R}(I - N_1) \subset \mathscr{R}(\phi_{\max})$, we have that $\mathscr{R}(W_2 + I - N_1) \subset \mathscr{R}(\phi_{\max})$. But $W_2 + I - N_1$ is pd when $N_2 = 0$. Thus, if $N_2 = 0$, then $\mathscr{C} = \{C_0\}$ so that $L = C_0$ and in this case (iv) is trivially fulfilled.

On the other hand, when $N_2 \neq 0$ we still need to consider a further step: applying Theorem 2.1 to $\mathscr{L}_2$ we return to part (β). Since at each time, the dimension of the resulting affine set decreases, we ultimately establish the desired property (iv). The proof of Theorem 3.1 is hence completed.

The following example illustrates that the conditions appearing in Theorem 3.1 are not sufficient.

Example 3.2. Consider a model with parameter space $\mathscr{T} = \mathscr{V} \times \mathscr{V}_2$, where $\mathscr{V} = \{I, N\}$, $N = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, while $\mathscr{V}_2$ is the set of all 2×2 nnd matrices. Then $L = \begin{pmatrix} 1 & 0 \\ 2 & 1/2 \end{pmatrix}$ and $C = I$ meet conditions (i)–(iv) of Theorem 3.1. However, L is inadmissible on $\mathscr{T}$. In fact $L \in \mathscr{L} = N + \mathscr{K}(I - N)$, but L is best at no nontrivial point in $[\mathscr{T} + \mathscr{S}]$, where $\mathscr{S}$ is the set of trivial points for $\mathscr{L}$, i.e., $\mathscr{S} = \{(\alpha N, \Psi - (I - N)\ \Psi(I - N)) : \alpha \in \mathscr{R}, \Psi \text{ symmetric}\}$.

Remark 3.3. It should be observed that by using the symmetry of the risk function (see proof of (ii)) additional necessary conditions from the results of Theorem 3.1 may be derived. However, it may be worthwhile to add that the problem of estimation is not symmetric with respect to Φ and $\mathscr{V}$, because Φ is spanned by one-dimensional matrices, whereas $\mathscr{V}$ does not need to be.

4. Necessary and Sufficient Conditions for Admissibility for Two Special Models

First we show that the necessary conditions established in Theorem 3.1 are sufficient for models with the mean vectors varying through a subspace $\mathscr{R}_0^n$ of $\mathscr{R}^n$ and covariance matrix $\sigma^2 V$, where V is a known nnd matrix. In this case $\mathscr{T} = \{(\sigma^2 V, \mu\mu') : \sigma^2 \geqslant 0, \mu \in \mathscr{R}_0^n\}$. Note that $V_{\max} = V$ and suppose that $V + \phi_{\max}$ is pd. First we consider the estimation of the mean vector μ.

Theorem 4.1. *The estimator L is an admissible estimator of $C = I$ iff*

(i) $\mathscr{R}(L') \subset \mathscr{R}(\phi_{\max})$

(ii) $\mathscr{R}(L' - I) = \mathscr{R}((L' - I)\ V)$

(iii) (a) *$L'V$ is symmetric*

(b) *$L'VL \leqslant L'V$.*

Proof. The necessity follows immediately from Theorem 3.1.

To prove sufficiency we first establish that every matrix L satisfying conditions (i)–(iii) is semisimple. In fact, by (ii) there exists a matrix A such that $L'(I-VA)(I-VA)'$ is symmetric. Hence part (a) of (iii) implies that $L'[V+(I-VA)(I-VA)']$ is symmetric. This entails the assertion, since the matrix in the big parentheses is pd.

Now let N denote the orthogonal projection on $\mathscr{R}(I-L)$. Then

$$(I-N)L = I-N, \qquad I-N \in \Phi, \tag{4.1}$$

because $\mathscr{R}(I-N) = \mathscr{N}(L'-I) \subset \mathscr{R}(\phi_{\max})$ by (i). Formula (4.1) asserts that L is best at point $(0, I-N) \in [\mathscr{T}]$. If $N=0$, then $L=I$ is admissible.

For $N \neq 0$ define $A = NVLN$ and $B = N(I-L)N$. By (iii) we have

$$AB = NL'VN(I-L)N = NL'V(I-L)N = N(L'V - L'VL)N$$

which is nnd. Since $I-L$ is semisimple, we have $\mathscr{R}(N) \subset \mathscr{R}(B)$. Therefore $\mathscr{R}(A) \subset \mathscr{R}(B)$, so that $\phi = AB^+$ is nnd. Combining (iii) (a) and (4.1) we find

$$(L'-I)\phi = (L'VL - L'V)B^+ = -L'VBB^+ = -L'VN.$$

This shows that

$$N(V+\phi)L = N\phi, \qquad (V, \phi) \in [\mathscr{T}], \tag{4.2}$$

which means that L is best among $\mathscr{L}_1 = I - N + \mathscr{K}(N)$ at point $(V, \phi) \in [\mathscr{T}]$. To verify that L is admissible it is now enough to show that L is the unique best estimator among $\mathscr{L}_1$ at (V, ϕ).

According to (ii) and part (a) of (iii) we have the inclusion $\mathscr{R}(I-L') \subset \mathscr{R}(V)$. This and $\mathscr{R}(N) = \mathscr{R}(N(I-L'))$ ensures that $\mathscr{R}(N) \subset \mathscr{R}(NV)$. Hence $\mathscr{R}(N) = \mathscr{R}(N(V+\phi)N)$, which is a sufficient condition for (4.2) to have a unique solution in $\mathscr{L}_1$. Theorem 4.1 is proven.

It may be worthwhile to note that the sufficiency part of Theorem 4.1 becomes incorrect if the assumption that the mean vector varies through a subspace is omitted. For example, the matrix $L = \left(\begin{smallmatrix} 1/2 & 0 \\ 1/2 & 1 \end{smallmatrix}\right)$ fulfills all conditions of Theorem 4.1 within model $[\mathscr{T}] = \{(\sigma^2 \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)) : \sigma^2, \alpha, \beta \geqslant 0\}$, but L is not admissible for $C = I$ on $\mathscr{T}$. In fact L is best at no nonzero point in $[\mathscr{T}]$.

COROLLARY 4.2. *In order that L be an admissible estimator within model $\mathscr{T} = \{(V, \mu\mu') : \mu \in \mathscr{R}_0^n\}$ it is necessary that L be semisimple.*

The preceding theorem extends to the case where the estimated function is $C'\mu$, while C may be any $n \times t$ matrix.

THEOREM 4.3. *The estimator L is an admissible estimator of C iff*

(i) $\mathscr{R}(L')\subset\mathscr{R}(C_0')$

(ii) $\mathscr{R}(L'-C')=\mathscr{R}((L-C)'V), \forall C\in\mathscr{C}$

(iii) (a) $\mathscr{R}(VL)\subset\mathscr{R}(\phi_{\max})$

(b) $L'VC_0$ *is symmetric*

(c) $L'VL\leqslant L'VC_0$.

Proof. All these conditions except (iii)(a) appear in Theorem 3.1. The latter one follows from part (a) of (iv) which for the underlying model states that $L'VC$ is symmetric for all $C\in\mathscr{C}$.

To show sufficiency we need only consider the case where $\mathscr{R}(C_0)$ is a proper subspace of $\mathscr{R}(\phi_{\max})$. Otherwise there exists in $\mathscr{C}$ a nonsingular matrix C, and with such a selection of C, conditions (i)–(iii) imply that L is admissible by Proposition 2.4 and Theorem 4.1. Also by virtue of Proposition 2.3 we may assume that V is idempotent.

(α) We first consider the case when $\mathscr{R}(\phi_{\max})\subset\mathscr{R}(V)$. Then in view of the adopted assumptions $V=I$, conditions (i)–(iii) reduce to

(a) $\mathscr{R}(L)\subset\mathscr{R}(\phi_{\max})$

(b) $L'C_0$ is symmetric

(c) $L'L\leqslant L'C_0$.

Let $H=L(L'C_0)^+L'$. This is a symmetric matrix and $\mathscr{R}(H)\subset\mathscr{R}(\phi_{\max})$ by (b) and (a), respectively. Because $\mathscr{R}(L')=\mathscr{R}(L'L)\subset\mathscr{R}(L'C_0)$ by (c), we have $L=HC$ for all $C\in\mathscr{C}$. Also $H^2\leqslant H$, by observing that

$$L(L'C_0)^+L'L(L'C_0)^+L'\leqslant L(L'C_0)^+L'C_0(L'C_0)^+L'=L(L'C_0)^+L'.$$

Thus H meets the conditions of Theorem 4.1 with $V=I$, and consequently H is an admissible estimator of I. Proposition 2.4 implies then that L is an admissible estimator of C.

(β) In the case when $\mathscr{R}(\phi_{\max})$ is not a subspace of $\mathscr{R}(V)$ we proceed essentially as in part (α). However, to construct the corresponding matrix H we need to define two matrices A and B, say.

First note that under the adopted assumptions VL is an admissible estimator of VC_0 within model $\mathscr{T}_*=\{(\sigma^2V,\mu\mu'):\sigma^2>0,\mu\in\mathscr{R}^n_*]$, where $\mathscr{R}^n_*=\mathscr{R}(\phi_{\max})\cap\mathscr{R}(V)$. This follows at once by noting that if L and C_0 fulfill conditions (i)–(iii) of Theorem 4.3, within model $\mathscr{T}$, then, respectively, VL and VC_0 fulfill these conditions within model $\mathscr{T}_*$. Therefore, similarly as in part (α), there exists a matrix A which has the following properties: $VL=AVC$ for all $C\in\mathscr{C}$, $\mathscr{R}(A)\subset\mathscr{R}^n_*$, $A'V$ is symmetric, and $A'VA\leqslant A'V$.

Next we show that there exists a matrix B such that $\mathscr{R}(I-VB) \subset \mathscr{R}(\phi_{\max})$. Since $\mathscr{R}(C_0) \subset \mathscr{R}(\phi_{\max})$ there exist matrices G and Z such that

$$GC_0' = 0 \qquad \text{and} \qquad GZ' = I - \phi_{\max}\phi_{\max}^{+}. \tag{4.3}$$

Put $C_* = C_0 + (I - \phi_{\max}\phi_{\max}^{+}) Z$. Since $C_* \in \mathscr{C}$, there exists by assumption (iii) of Theorem 4.3 a matrix B such that

$$L' - C_*' = (L - C_*)' VB,$$

and we may choose B so that $VB = B$ and $BV = V$. Multiplying from the left-hand side by G and then using (4.3) and the inclusion $\mathscr{R}(L') \subset \mathscr{R}(C_0')$ we obtain $I - \phi_{\max}\phi_{\max}^{+} = (I - \phi_{\max}\phi_{\max}^{+}) VB$. Thus B has the required property.

Now let $H = I + B'(A - V)$. Because, obviously, $L = HC$ for all $C \in \mathscr{C}$ it remains to show that H is admissible for I on $\mathscr{T}$. This will be accomplished by showing that H fulfills the conditions of Theorem 4.1.

Using the stated properties of A and B it is easy to verify that $(H-I)' V = (A - V)' V$. Hence (ii) is valid. Furthermore, noting that $H'V = [I + (A - V)' B] V = A'V$ and that $H'VH = A'V[I + B'(A - V)] = A'VA$ we obtain (iii). Now this, together with Proposition 2.4, implies that L is admissible. This terminates the proof of Theorem 4.3.

Theorem 4.3 includes as special cases the results established by Karlin [7], Cohen [1, 2], Shinozaki [18], Rao [16], and Klonecki [8]. The following result is slightly more general than Theorem 3.4 in [16].

COROLLARY 4.4. *Within model* $\mathscr{T} = \{(\sigma^2 V, \mu\mu') : \sigma^2 \geqslant 0, \mu \in \mathscr{R}_0^n\}$, *where V is nnd, the following two statements are equivalent*:

(i) *$L'V$ is symmetric and Lp is an admissible estimator of p for every $p \in \mathscr{R}^n$*

(ii) *L is an admissible estimator of I.*

Proof. That (ii) implies (i) follows from Theorem 4.1 and Proposition 2.4. The reversed implication is a direct consequence of Theorems 4.1 and 4.3 and the fact that if $p'\mathscr{R}(A) \subset p'\mathscr{R}(B)$ for every $p \in \mathscr{R}^n$, then $\mathscr{R}(A) \subset \mathscr{R}(B)$.

In proving Theorem 4.3 we have established that within the regression model L is an admissible estimator of C iff there exists an $n \times n$ matrix H such that $L = HC$ and such that H is admissible for I. Zontek [20] has defined a class of linear models for which both these implications are still valid.

In the conclusion of our investigation of the regression model $\mathscr{T} = \{(\sigma^2 V, \mu\mu') : \sigma^2 \geqslant 0, \mu \in \mathscr{R}_0^n\}$ with $\mathscr{R}(V) + \mathscr{R}_0^n = \mathscr{R}^n$, we formulate without proof alternative conditions which characterize admissible linear estimators.

The estimator L is an admissible estimator of $C=I$ iff there exist two matrices ϕ_0 and ϕ_1 in Φ such that

(i) $\phi_0 L=\phi_0$

(ii) $N(V+\phi_1)L=N\phi_1$

(iii) $\mathscr{R}(N(V+\phi_1)N)=\mathscr{R}(N)$,

where $N=I-\phi_0\phi_0^+$.

Since (iii) guarantees that $V+\phi_0+\phi_1$ is pd, a standard calculation leads to the following result. The estimator L is an admissible estimator of C iff there exist two matrices ϕ_0 and ϕ_1 in Φ such that

$$L=\lim_{\lambda\to 0}(\lambda V+\phi_\lambda)^{-1}\phi_\lambda C, \tag{4.4}$$

where $\phi_\lambda=(1-\lambda)\phi_0+\lambda\phi_1$, while $\lambda\in(0,1)$. Using the terminology as in [16] and referring to [14], we can therefore say that for the regression model under consideration the Bayes homogeneous linear estimators of C with respect to all matrices ϕ in Φ which fulfill the condition $\mathscr{N}(V+\phi)=0$ and their limits form a complete class of estimators.

Finally, if V is pd, then L is an admissible estimator of C iff it may be represented as

$$L=V^{-1}X(X'V^{-1}X)^{-1/2}A(X'V^{-1}X)^{-1/2}X'C \tag{4.5}$$

for some symmetric matrix A with all its eigenvalues in $[0,1]$. Here X stands for any matrix of full rank whose columns span $\mathscr{R}_0^n$. For $V=I$ formula (4.4) has been established by LaMotte in [13] and formula (4.5) is essentially due to Cohen [2] and Rao [16].

Next we establish necessary and sufficient conditions for the admissibility of a linear estimator within model $\mathscr{T}$ with the mean vectors varying through a subspace of $\mathscr{R}^n$ and the covariances varying through $\mathscr{V}_n$. Such models are of special interest, because by Proposition 2.5 any estimator admissible within some model $\mathscr{T}_*$ such that $[\mathscr{T}_*]\subset[\mathscr{T}]$ and with a positive definite maximal element must be admissible also within model $\mathscr{T}$.

THEOREM 4.6. *In order that L be an admissible estimator of C within model* $\mathscr{T}=\mathscr{V}_n\times\{\mu\mu':\mu\in\mathscr{R}_0^n\}$ *it is necessary and sufficient that*

(i) $\mathscr{R}(L')=\mathscr{R}(L'\phi_{\max})\subset\mathscr{R}(C_0')$

(ii) *for every* $C\in\mathscr{C}$ *if* $Lx=\lambda Cx$, $Cx\neq 0$, *then* $\lambda\in[0,1]$.

Proof. Because of Theorem 3.1 we need only prove the sufficiency part.

(α) To begin with suppose that $\phi_{\max}$ is pd and that $C_0 = I$. Then (i) and (ii) assert only that the eigenvalues of L are in $[0, 1]$.

Step 1. Denote by $P_1, ..., P_r$ the eigenvectors of L' and suppose they correspond to eigenvalues $\lambda_1, ..., \lambda_r$, respectively. Let

$$V = \sum_{i=1}^{r} (1 - \lambda_i) P_i P_i' \qquad \text{and} \qquad \phi = \sum_{i=1}^{r} \lambda_i P_i P_i'.$$

Then $(V + \phi) L = \phi$, where $(V, \phi) \in [\mathscr{T}]$. The class of estimators best at (V, ϕ) is $\mathscr{L}_1 = L + \mathscr{K}(N_1)$, where N_1 is the orthogonal projection on the null space of $\sum_{i=1}^{r} P_i P_i'$. For $N_1 = 0$ the assertion follows. Otherwise we proceed to the second step.

Step 2. Denote by $Q_1, ..., Q_s$ the eigenvectors of $N_1 L'$ in $\mathscr{R}(N_1)$. Suppose they correspond to the eigenvalues $\beta_1, ..., \beta_s$. Now let

$$\begin{aligned} S_1 &= A_{11} + A_{12} + A_{12}' \\ S_2 &= B_{11} + B_{12} + B_{12}', \end{aligned} \tag{4.6}$$

where

$$A_{11} = \sum_{i=1}^{s} (1 - \beta_i) Q_i Q_i'$$

$$A_{12} = \sum_{i=1}^{s} Q_i Q_i' L K_1 K_1^+, \qquad K_1 = (I - N_1)(I - L)(I - N_1)$$

$$B_{11} = \sum_{i=1}^{s} \beta_i Q_i Q_i'$$

$$B_{12} = -\sum_{i=1}^{s} Q_i Q_i' L M_1 M_1^+, \qquad M_1 = (I - N_1) L (I - N_1).$$

Since $\beta_i \in [0, 1]$, $i = 1, ..., s$, by Lemma A1, therefore $(S_1, S_2) \in [\mathscr{T} + \mathscr{S}_1]$, where $\mathscr{S}_1$ is the space of trivial points for $\mathscr{L}_1$. From (4.6) it follows that $N_1(S_1 + S_2) N = \sum_{i=1}^{s} Q_i Q_i'$ is a nonzero matrix. Moreover, $N_1(S_1 + S_2) L = N_1 S_2$, since $N_1 L N_1 = L N_1$. Now the set of all best estimators of C among $\mathscr{L}_1$ is $\mathscr{L}_2 = L + \mathscr{K}(N_1 N_2)$, where N_2 is the orthogonal projection on the null space of $\sum_{i=1}^{s} Q_i Q_i'$. Obviously, $\dim(\mathscr{L}_2) < \dim(\mathscr{L}_1)$. If $N_2 N_1 = 0$, the admissibility of L follows. Otherwise we need to consider a further step.

Step 3. Denote by $R_1, ..., R_t$ the eigenvectors of $N_2 N_1 L'$ in $\mathscr{R}(N_1 N_2)$, and by $\gamma_1, ..., \gamma_t$ the corresponding eigenvalues. Let $A_{21}, A_{22}, B_{21}, B_{22}$ be defined as $A_{11}, A_{12}, B_{11}, B_{12}$ in Step 2 with the β's, Q's, and N_1 replaced by the γ's, R's, and $N_1 N_2$, respectively, and let $T_1 = A_{21} + A_{22} + A'_{22}$ and $T_2 = B_{21} + B_{22} + B'_{22}$. Since the eigenvalues of $N_1 N_2 L'$ are in $[0, 1]$, it follows that $(T_1, T_2) \in [\mathcal{T} + \mathcal{S}_2]$, where $\mathcal{S}_2$ is the space of trivial points for $\mathcal{L}_2$. Utilizing $N_1 N_2 L N_1 N_2 = L N_1 N_2$ we arrive at $N_1 N_2 (T_1 + T_2) L = N_1 N_2 T_2$. The dimension of $\mathcal{L}_3 = L + \mathcal{K}(N_1 N_2 N_3)$, where N_3 is the orthogonal projection on the null space of $\sum_{i=1}^{r} R_i R'_i$, is less than the dimension of $\mathcal{L}_2$. If $N_1 N_2 N_3 = 0$, the desired result follows. Otherwise, we still need to consider a further step, and so on. Continuing in this manner altogether at most n times we ultimately establish that, in the considered case, L is an admissible estimator of C.

(β) We now proceed to the case when $\phi_{\max}$ is pd and C is singular. As seen from part (α) of the proof of Theorem 4.6 and from Proposition 2.4 it is enough to show that there exists a matrix F such that the eigenvalues of the matrix $H = LC^+ + F(I - CC^+)$ are in $[0, 1]$. By Corollary A3 there exists such a matrix F iff rank $\{(L - \lambda I) CC^+\} < \text{rank}\{CC^+\}$ is valid for λ in $[0, 1]$ only. And this is precisely assured by (i) and (ii).

(γ) Now using the results established in parts (α) and (β) we prove the assertion in the remaining case when $\phi_{\max}$ is not pd. In view of Proposition 2.4 we may assume without loss of generality that C is idempotent and that $C = C_0$.

For brevity let $N = \phi_{\max}\phi_{\max}^+$. Condition (i) implies then that $L = LN$. Moreover, since (i) and (ii) remain valid with L replace by NLN, it follows that NLN is an admissible estimator of C on $\mathcal{T}_* = \{(NVN, \phi) : (V, \phi) \in \mathcal{T}\}$. Thus by Theorem 2.1 (here we assume that $\mathcal{W} = \{(S_1, S_2) : S_1$ and S_2 symmetric$\}$), there exist $r \leqslant \text{rank}(N)$ points $S_1, ..., S_r$ in $\mathcal{W}$ combined with $r+1$ affine sets $\mathcal{L}_1 \supset \cdots \supset \mathcal{L}_{r+1}$, where $\mathcal{L}_{r+1} = NLN + \mathcal{K}(I - N)$, which meet the necessary conditions of Theorem 2.1 with $\mathcal{T} \subset \mathcal{S}_{r+1}$. In particular, $S_i \in [\mathcal{T} + \mathcal{S}_i]$, $S_i \notin \mathcal{S}_i$, for $i = 1, ..., r$.

Using the necessary conditions for admissibility of NLN within model $\mathcal{T}_*$ we now establish sufficient conditions for the admissibility of L on $\mathcal{T}$.

Since $L = NLN + (I - N) LN$, it should be obvious that for $i = 1, ..., r$, estimator L is best among $\mathcal{L}_i$ at S_i and admissible among $\mathcal{L}_{i+1}$ on $\mathcal{T}$. Since $\mathcal{T} \not\subset \mathcal{S}_{r+1}$, the above conditions are not sufficient on $\mathcal{T}$. However, we shall show that, by adding one more point $S_{r+1} \in [\mathcal{T} + \mathcal{S}_{r+1}]$, the resulting conditions become sufficient. In fact, we show that there exists a point $S_{r+1} = (V, 0) \in [\mathcal{T}]$ such that V is pd, $(I - N) V(I - N) = I - N$, and $(I - N) VL = 0$. This latter equation becomes $[I - N + (I - N) VN] L = 0$, which in turn may be written as $L'NV(I - N) = -L'(I = N)$. Since $\mathcal{R}(L'(I - N)) \subset \mathcal{R}(L'N)$ by (ii), it has a solution with respect to $NV(I - N)$. Thus there exists a pd matrix V with the stated properties and these imply that L is the only estimator in $\mathcal{L}_{r+1}$ best among $\mathcal{L}_{r+1}$ at S_{r+1}. Theorem 4.6 is fully proved.

5. Necessary and Sufficient Conditions for Admissibility of Nonhomogeneous Linear Estimators

Comparing nonhomogeneous linear estimators $\delta + L'Y$, where $\delta \in \mathscr{R}^t$, while L is a $n \times t$ matrix, under the risk function $E(\delta + L'Y - C'\mu)'(\delta + L'Y - C'\mu)$, the problem of determining the admissibility of a nonhomogeneous linear estimator transforms to a problem of determining the admissibility of a homogeneous linear estimator. In fact, the estimator $\delta + L'Y$ of $C'\mu$ on $\mathscr{T}$ becomes a homogeneous estimator $\binom{L}{\delta'}'\binom{Y}{1}$ of $\binom{C}{0}'\binom{\mu}{1}$ on $\mathscr{T}_* = \{(\begin{pmatrix} V & 0 \\ 0 & 0 \end{pmatrix}, \binom{\mu}{1}\binom{\mu}{1}') : (V, \mu\mu') \in \mathscr{T}\}$. In consequence necessary and/or sufficient conditions for admissibility of a nonhomogeneous linear estimator may be derived directly from those of the corresponding homogeneous linear estimator.

Let $(V_{\max}, \phi_{\max})$ be a maximal element in $\mathscr{T}$ and suppose that $V_{\max} + \phi_{\max}$ is pd. Also, as before, let $\mathscr{C} = C_0 + \mathscr{K}(I - \phi_{\max}\phi_{\max}^+)$.

Theorem 5.1. *In order that $\delta + L'Y$ be an admissible estimator of $C'\mu$ among the class of all nonhomogeneous linear estimators on $\mathscr{T}$ it is necessary that*

(i) $\delta \in \mathscr{R}((L - C)'\phi_{\max})$

(ii) $\mathscr{R}(L', \delta) = \mathscr{R}(L'\phi_{\max}, \delta) \subset \mathscr{R}(C_0')$

(iii) $\mathscr{R}(L' - C', \delta) = \mathscr{R}((L - C)' V_{\max}) \subset C'\mathscr{R}(V_{\max})$, $\forall C \in \mathscr{C}$

(iv) *if* $Lx = \lambda Cx$ *for some* $C \in \mathscr{C}$ *and if* $Cx \neq 0$, $\delta'x = 0$, *then* $\lambda \in [0, 1]$

(v) *there exists a nonzero* $V \in \mathscr{V}$ *such that for each* $C \in \mathscr{C}$

   (a) $L'VC$ *is symmetric*

   (b) $L'VL \leqslant L'VC$.

Proof. Let $L_* = (L', \delta)'$ and let $C_* = (C', 0)'$. Moreover, let $N = (0, ..., 0, 1)'(0, ..., 0, 1)$. Since L_* is admissible for C_* among $(I - N)L_* + \mathscr{K}(N)$, there exists by Theorem 2.1 a point $(0, \phi_*) \in [\mathscr{T}_* + \mathscr{S}_*]$ such that $N\phi_*N \neq 0$ and $N\phi_*(L_* - C_*) = 0$. This entails (i). The remaining conditions are immediate consequences of Theorem 3.1.

From Theorems 4.3 and 5.1 it follows that if $\delta + L'Y$ is an admissible estimator of $C'\mu$ within the regression model considered in Section 4, then δ fulfills condition (i) of Theorem 5.1 and $L'Y$ is an admissible estimator of $C'\mu$ among the class of all homogeneous linear estimators. Now we show that this implication may be reversed. We first set down a result for homogeneous linear estimation which will be required in the proof of the next theorem.

To the end of this section we assume that

$$\mathscr{T} = \mathscr{V} \times \{X\beta\beta'X' : \beta \in \mathscr{R}^p\}, \tag{5.1}$$

where $\mathscr{V} \subset \mathscr{V}_n$.

PROPOSITION 5.2. *If $L'Y$ is an admissible estimator of $C'X$ on $\mathscr{T}$, then, for every $\gamma \in \mathscr{R}^p$, $L'(Y+X\gamma)$ is an admissible homogeneous linear estimator of $C'X(\beta+\gamma)$.*

Proof. For $\mathscr{T}$ given by (5.1) the risk (2.1) may be rewritten as $\rho((V, X\beta); L) = \rho((V, X\beta\beta'X'); L) = [L, VL] + [L-C, \quad X\beta\beta'X'(L-C)]$, where $(V, \beta) \in \mathscr{V} \times \mathscr{R}^p$.

If there would exist an estimator $M'(Y+X\gamma)$ better than $L'(Y+X\gamma)$ for $C'X(\beta+\gamma)$, that means if $\rho((V, X(\beta+\gamma); M) \leqslant \rho((V, X(\beta+\gamma)); L)$ for all $(V, \beta) \in \mathscr{V} \times \mathscr{R}^p$, then also $\rho((V, X\beta); M) \leqslant \rho((V, X\beta); L)$ for all $(V, \beta) \in \mathscr{V} \times \mathscr{R}^p$. From this, it would then follow that the risk functions of L and M are identical on $\mathscr{T}$. This contradiction concludes the proof.

THEOREM 5.3. *In order that $\delta + L'Y$ be an admissible estimator of $C'X\beta$ among nonhomogeneous linear estimators on $\mathscr{T}$ it is necessary and sufficient that*

(i) $\delta \in \mathscr{R}((L-C)'X)$

(ii) *$L'Y$ be an admissible estimator of $C'X\beta$ among the linear estimators on $\mathscr{T}$.*

Proof. First note that if $\delta + L'Y$ is admissible for $C'X\beta$, then (i) is fulfilled by Theorem 5.1, i.e., there exists a vector $\gamma \in \mathscr{R}^p$ such that $\delta = (L-C)'X\gamma$.

Now the assertion of Theorem 5.3 is established by noting that the following statements are equivalent for an arbitrary $\gamma \in \mathscr{R}^p$:

(a) $(L-C)'X\gamma + L'Y$ is admissible for $C'X\beta$ among nonhomogeneous linear estimators

(b) $L'(Y+X\gamma)$ is admissible for $C'X(\beta+\gamma)$

(c) $L'Y$ is admissible for $C'X\beta$.

The equivalence of (a) and (b) is a consequence of the considered risk functions; the equivalence of (b) and (c) follows straightforwardly from Proposition 5.2.

Conditions for admissibility of a nonhomogeneous linear estimator within the regression model with a pd covariance have been established by Rao in [16] and LaMotte in [11].

APPENDIX

LEMMA A1. *Let $P_1, ..., P_r$ be linearly independent and real eigenvectors of matrix L and let N be the orthogonal projection on the space spanned by these r vectors. Then*

(i) *each real nonzero eigenvalue of $(I-N)L$ is an eigenvalue of L,*

(ii) *each eigenvalue of L which corresponds to none of the eigenvectors $P_1, ..., P_r$ is an eigenvalue of $(I-N)L$.*

Proof. For $N=I$, conditions (i) and (ii) are obviously fulfilled. Thus suppose that $N \neq I$. Let β be a real, nonzero eigenvalue of $(I-N)L$ corresponding to eigenvector Q, say. Moreover, suppose that $\beta \neq \lambda_i$, where $\lambda_1, ..., \lambda_r$ are the eigenvalues of L corresponding to $P_1, ..., P_r$, respectively.

To prove (i) it is sufficient to show that there exists a vector X such that $(L-\beta I)(Q+NX)=0$, where $Q+NX \neq 0$. Writing this equation in the form $(L-\beta I)NX=-NLQ$, we note that it has a solution in X when $\mathscr{R}(N) \subset \mathscr{R}((L-\beta I)N)$. However, such an inclusion holds, because $\mathscr{R}((L-\beta I)N)=\mathscr{R}((\lambda-\beta)P_1, ..., (\lambda_r-\beta)P_r)$. Moreover, since $Q \in \mathscr{R}(I-N)$ when $\beta \neq 0$ and since $\mathscr{R}(N) \cap \mathscr{R}(I-N)=\{0\}$, it follows that $Q+NX \neq 0$. Thus β is an eigenvalue of L as asserted.

Now let $LQ=\beta Q$, where $\beta \neq \lambda_i$, $i=1, ..., r$, while $Q \neq 0$. Then $(I-N)LQ=\beta(I-N)Q$. Since $(I-N)L(I-N)=(I-N)L$, we obtain $(I-N)LT=\beta T$, where $T=(I-N)Q$. Now we show that $T \neq 0$. Suppose to the contrary that $T=0$. Then $Q=\sum_{j=1}^{s} \rho_j P_{i_j}$ for some $\rho_j \neq 0$ and $1 \leqslant s \leqslant r$ so that $LQ=\sum_{j=1}^{s} \rho_j \lambda_{i_j} P_{i_j}=\sum_{j=1}^{s} \beta\rho_j P_{i_j}$. Consequently, $\sum_{j=1}^{s} \rho_j(\beta-\lambda_{i_j}) P_{i_j}=0$. Since $P_1, ..., P_r$ are linearly independent, this yields that $\beta=\lambda_{i_0}$ for some $1 \leqslant i_0 \leqslant r$. This contradiction proves (ii).

In the remainder of this section let A be a complex $(n-1)\times n$ matrix. Denote by $\mathscr{Q}$ the set of all characteristic polynomials $W_B(\lambda)$ of the $n \times n$ matrices $\binom{A}{B}$, where $B=(b_1, ..., b_n)$ varies over all $1 \times n$ complex vectors. Denote by $W_1=W_1(\lambda), ..., W_n=W_n(\lambda)$ the cofactors of the elements $b_1, ..., b_n-\lambda$ which appear in the last row of the matrix $\binom{A}{B}-\lambda I$, respectively. With this notation we have

$$W_B(\lambda)=\sum_{i=1}^{n} b_i W_i(\lambda)-\lambda W_n(\lambda).$$

A polynomial $P(\lambda)$ is called the greatest divisor for $\mathscr{Q}$ if it is a divisor of each polynomial in $\mathscr{Q}$ and if there exists no other polynomial of degree greater than the degree of $P(\lambda)$, which is a divisor of all the polynomials in $\mathscr{Q}$.

LEMMA A2. *The degree of the greatest divisor $P(\lambda)$ for $\mathcal{Q}$ is equal to r if and only if $W_1, ..., W_n$ span an $(n-r)$-dimensional set.*

Proof. First we show that if $W_1, ..., W_n$ span a set of dimension $n-r$, then the degree of the greatest divisor for $\mathcal{Q}$ is at least equal to r. Without loss of generality we may assume that A has the Jordan normal form, i.e.,

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 & C_1 \\ 0 & A_2 & \cdots & 0 & C_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_m & C_m \end{pmatrix},$$

where, for $j = 1, ..., m$,

$$A_j = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \lambda_j & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_j \end{pmatrix}$$

is $i_j \times i_j$, $C_j = (c_{j1}, ..., c_{ji_j})'$ and $\sum_{j=1}^m i_j = n$.

Also without loss of generality we may assume that $\lambda_1 = \cdots = \lambda_s$ and $\lambda_i \neq \lambda_1$ for $i > s$. Now let

$$g_j = \begin{cases} 0 & \text{when } c_{jk} = 0,\ 1 \leqslant k \leqslant i_j \\ \max\{k : c_{jk} \neq 0,\ 1 \leqslant k \leqslant i_j\} & \text{otherwise,} \end{cases}$$

and let $g = \max\{g_i : i = 1, ..., s\}$. Clearly, we may assume without loss of generality that $g = g_1$. To simplify the notation let $r_0 = 0$ and let $r_k = \sum_{j=1}^k i_j$, $k = 1, ..., m$.

Since the greatest divisor of $W_{r_{i-1}+1}, ..., W_{r_i}$, $i \leqslant s$, is $(\lambda_1 - \lambda)^{r_s - g_i} h(\lambda)$, where

$$h(\lambda) = \prod_{j=s+1}^{m} (\lambda_j - \lambda)^{i_j},$$

therefore,

$$H(\lambda) = (\lambda_1 - \lambda)^{r_s - g} h(\lambda)$$

is the greatest divisor of $W_1, ..., W_{r_s}$.

Note that $W_1 = W_2 = \cdots = W_{r_s} = 0$ when $g = 0$. In case $g \neq 0$, the polynomials $W_1, ..., W_g$ are linearly independent, while $W_{g+1}, ..., W_{r_s}$ may be expressed as linear combinations of $W_1, ..., W_g$. In fact, let $V_i = W_i/H$ for $i = 1, ..., g$. Then the coordinates of $V_1, ..., V_g$ with respect to the basis 1,

$\lambda_1-\lambda, ..., (\lambda_1-\lambda)^{g-1}$ are $((-1)^{n-g+1}c_{1g}, (-1)^{n-g+2}c_{1,g-1}, ..., (-1)^n c_{11})$, $(0, (-1)^{n-g+1}c_{1g}, ..., (-1)^{n-1} c_{12}), ..., (0, ..., 0, (-1)^{n-g+1}c_{1g})$, respectively. Since $c_{1g}\neq 0$, it follows that $V_1, ..., V_g$ are linearly independent and that $V_{g+1}, ..., V_{r_s}$ may be expressed as linear combinations of $V_1, ..., V_g$. This shows that $(\lambda_1-\lambda)^{r_s-g}$, where r_s-g is equal to the number of linearly independent polynomials among $W_1, ..., W_{r_s}$, is a divisor of $W_1, ..., W_n$.

To prove the necessity part of Lemma A2 suppose that $W_1, ..., W_n$ span a set of dimension $n-s$ and that the degree of the greatest divisor for $\mathscr{Q}$ equals r. With this notation $n-s<n-r$. On the other hand, from the above established result it follows that $s\leqslant r$. Thus the necessity is established.

Sufficiency also follows straightforwardly from the established result by noting that the intersection of the set spanned by $W_1, ..., W_{r_s}$ and the set spanned by $W_{r_s+1}, ..., W_n$ has dimension zero.

Because the roots of the maximal divisor $P(\lambda)$ are roots of $W_B(\lambda)$, whatever be B, there exists a matrix B such that the eigenvalues of $\binom{A}{B}$ are in $[0, 1]$ iff the roots of $P(\lambda)$ are real and in $[0, 1]$. For our considerations we formulate this as follows without need of further proof.

COROLLARY A3. *Let C be a $n\times n$ real idempotent symmetric matrix and let E be a $n\times n$ real matrix. There exists a $n\times n$ real matrix F such that the eigenvalue of $EC+F(I-C)$ are in $[0, 1]$ iff* $\operatorname{rank}\{(E-\lambda I)C\}<\operatorname{rank}\{C\}$ *is valid for $\lambda\in[0, 1]$ only.*

ACKNOWLEDGMENTS

The authors are thankful to two referees for their constructive comments and suggestions which led to improvement in the presentation of the paper.

REFERENCES

[1] COHEN, A. (1965). Estimates of linear combinations of the parameters in the mean vector of a multivariate distribution, *Ann. Math. Statist.* **36** 78–87.

[2] COHEN, A. (1966). All admissible linear estimates of the mean vector, *Ann. Math. Statist.* **37** 458–463.

[3] DRYGAS, H. (1972). The estimation of residual variance in regression analysis, *Math. Operationsforsch. Statist.* **3** 373–388.

[4] DRYGAS, H. (1977). Best quadratic unbiased estimation in variance-covariance component models, *Math. Operationsforsch. Statist. Ser. Statist.* **8** 211–231.

[5] GNOT, S., KLONECKI, W., AND ZMYŚLONY, R. (1980). Best unbiased linear estimation, a coordinate free approach, *Probab. Math. Statist.* **1** 1–13.

[6] HOFFMAN, K. (1977) Admissibility of linear estimators with respect to restricted parameter sets, *Math Operationsforsch. Statist. Ser. Statist.* **4** 425–438.

[7] KARLIN, S. (1958). Admissibility for estimation with quadratic loss, *Ann. Math. Statist.* **29** 406–436.

[8] KLONECKI, W. (1980). Invariant quadratic unbiased estimation for variance components, *in* "Proceedings, Sixth International Conference, Wisła, Poland, 1978, "Lecture Notes in Statistics Vol. 2, pp. 201–223, Springer-Verlag, New York/Berlin.

[9] KLONECKI, W. (1982). Linear estimators of the mean vector in linear models, problem of admissibility, *Probab. Math. Statist.* **2** 167–178.

[10] LAMOTTE, L. R. (1977). A canonical form for the general linear model, *Ann. Statist.* **5** 787–789.

[11] LAMOTTE, L. R. (1978). Bayes linear estimators, *Technometrics* **20** 281–290.

[12] LAMOTTE, L. R. (1980). Some results on biased linear estimation applied to variance component estimation, *in* "Proceedings, Sixth International Conference, Wisła, Poland 1978," Lecture Notes in Statistics Vol. 2, pp. 266–274, Springer-Verlag, New York/Berlin.

[13] LAMOTTE, L. R. (1982). Admissibility in linear estimation, *Ann. Statist.* **10** 245–256.

[14] OLSEN, A., SEELY, J., AND BIRKES, D. (1976). Invariant quadratic unbiased estimation for two variance components, *Ann. Statist.* **4** 878–890.

[15] RAO, C. R. (1971). Unified theory of linear estimation, *Sankhyá Ser. A* **33** 371–394.

[16] RAO, C. R. (1976). Estimation of parameters in a linear model, *Ann. Statist.* **4** 1023–1037.

[17] SEELY, J., AND ZYSKIND, G. (1971). Linear spaces and minimum variance unbiased estimation, *Ann. Math. Statist.* **42** 691–703.

[18] SHINOZAKI, N. (1975). "A Study of Genealized Inverse of Matrix and Estimation with Quadratic Loss," Ph. D. thesis, Keino University, Japan, 1975.

[19] ZONTEK, S. (1983). "Linear Estimation of Parameters in Linear Models," Ph. D. thesis. Mathematical Institute of the Polish Academy of Sciences, Poland. [Polish]

[20] ZONTEK, S. (in press). On characterization of linear admissible estimators: An extension of a result due to C. R. Rao, *J. Multivariate Anal.*