PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 2, Fasc. 2 (1982), p. 167-178

LINEAR ESTIMATORS OF THE MEAN VECTOR IN LINEAR MODELS: PROBLEM OF ADMISSIBILITY

BY

WITOLD KLONECKI (WROCŁAW)

Abstract. In this paper we consider linear estimators in linear models with a general covariance structure. Necessary conditions for admissibility of the linear estimators with quadratic loss function are given and they are shown to be sufficient when only positive definite covariance matrices are admitted. In the case where the set of admitted covariance matrices coincides with all nonnegative definite matrices, it is shown that LY is admissible for the expected value $\mathrm{E}Y$ if and only if the eigenvalues of the matrix L are in the closed interval $[0, 1]$.

1. Introduction. Let Y be an n-variate random vector with expectation $\theta = \mathrm{E}Y$ and covariance matrix $V = \mathrm{Cov}\, Y$. The parameters are (θ, V), and the parameter space, denoted by $\mathscr{P}$, is assumed to be of the form $\mathscr{R}^n \times \mathscr{V}$, where $\mathscr{V}$ is a closed convex cone of nonnegative definite (n.n.d., for short) matrices of order $n \times n$. The paper is concerned with the problem of characterization of admissible linear estimators of θ with the squared distance as loss when different restrictions are imposed on $\mathscr{V}$.

By a *linear estimator* of θ we understand a function LY, where L is an $n \times n$ real matrix of constants. Denote by R the expected squared distance, i.e. let

$$R(\theta, V|L) = \mathrm{E}\,[(LY-\theta)'(LY-\theta)], \quad (\theta, V) \in \mathscr{P}.$$

As usual, MY is said to be *as good as* LY if $R(\theta, V|M) \leqslant R(\theta, V|L)$ throughout $\mathscr{P}$, and LY is *better than* MY if, in addition, strict inequality holds for some point in $\mathscr{P}$. The estimator LY is called *admissible* for θ within model with parameter space $\mathscr{P}$ among linear estimators (admissible within model $\mathscr{P}$, for short) if no other linear estimator is better than LY. Finally, LY is called *locally best* at a point in $\mathscr{P}$ if no other linear estimator is better at this point than LY.

A characterization of admissible linear estimators within the model treated here has been given in a particular case by Cohen in [1]. Cohen assumed that $\mathscr{V}$ is generated by the unit matrix and proved that LY is admissible if and only if L is symmetric and the eigenvalues of L are in the closed interval $[0, 1]$. His proof was based on the observation that MY, where $M = I-[(I-L)'(I-L)]^{1/2}$, I being the unit matrix, is better than LY when L is asymmetric. Next Rao [9] has shown in his 1976 Wald lecture paper that Cohen's result may be extended to models in which $\mathscr{V}$ is generated by a single positive definite (p.d., for short) matrix. He deduced this result from Cohen's theorem using a lemma given by Shinozaki in [10].

This paper is concerned with further generalizations of Cohen's result. For the model considered here we succeeded in establishing only necessary conditions, i.e. if LY is admissible within model $\mathscr{R}^n \times \mathscr{V}$, then (i) the eigenvalues of L are in $[0, 1]$ and (ii) the product LV is symmetric for some nonzero matrix $V \in \mathscr{V}$. However, we have also shown that conditions (i) and (ii) are sufficient if it is assumed in addition that $\mathscr{V}$ consists only of p.d. matrices (except for the zero matrix) and, consequently, that LY is admissible within model $\mathscr{R}^n \times \mathscr{V}$ if and only if LY is admissible within model $\mathscr{R}^n \times \mathscr{V}$ for some nonzero matrix $V \in \mathscr{V}$. Here $[V]$ denotes the convex cone generated by the matrix V. This latter result is not necessarily valid if the condition that $\mathscr{V}$ consists of p.d. matrices is removed. It remains to be true that if LY is admissible within model $\mathscr{R}^n \times [V]$, then LY is admissible within any model $\mathscr{R}^n \times \mathscr{V}$ provided $V \in \mathscr{V}$. Finally, we show that condition (i) mentioned above is necessary and sufficient for admissibility in the case where $\mathscr{V}$ coincides with the family of all n.n.d. matrices. Some of the results presented in this paper have already been announced in [3] and [4].

In our considerations we use a method developed first by Olsen et al. [7] and then extended by La Motte in [6]. The essential tools in this approach are Lemma 1.2 which states that, roughly speaking, each admissible estimator is locally best at some nonzero point in an "extended" parameter space and Lemma 1.1 which gives a simple characterization of locally best linear estimators. The details are as follows.

Let $\mathscr{S}_{n\times n}$ denote the class of all $n \times n$ real matrices and let, for $A, B \in \mathscr{S}_{n\times n}$, the expression $A \otimes B$ denote a linear operator on $\mathscr{S}_{n\times n}$ defined for each $C \in \mathscr{S}_{n\times n}$ by $(A \otimes B)C = ACB'$. Following La Motte [6] define for the model $\mathscr{P}$ a subset $\mathscr{T}$ of $\mathscr{S}_{n\times n} \times \mathscr{S}_{n\times n}$ as

$$\mathscr{T} = \{(\theta\theta', V): (\theta, V) \in \mathscr{P}\},$$

and let $[\mathscr{T}]$ denote the smallest closed convex cone in $\mathscr{S}_{n\times n} \times \mathscr{S}_{n\times n}$ containing $\mathscr{T}$. Now for each $L \in \mathscr{S}_{n\times n}$ let $R(\cdot|L): [\mathscr{T}] \to \mathscr{R}$ be a function defined for each $(\Phi, V) \in [\mathscr{T}]$ by

$$R(\Phi, V|L) = \operatorname{tr}[(I-L)'(I-L)\Phi + L'LV], \tag{1.1}$$

where tr stands for trace. Clearly, $R(\theta\theta', V|L) = R(\theta, V|L)$ for all $(\theta, V) \in \mathscr{P}$. In the sequel we shall also treat (1.1) for fixed (Φ, V) as a function of L defined on $\mathscr{S}_{n\times n}$.

Now let $\mathscr{K}$ be a nonempty subspace of $\mathscr{R}^n$ and let Π be the orthogonal projection on $\mathscr{K}$ such that the range $\mathscr{R}(\Pi)$ of Π coincides with $\mathscr{K}$. Finally, let

$$\mathscr{L} = \{L: \mathscr{R}(L'-L'_0) \subset \mathscr{K}\}, \quad \text{where } \mathscr{R}(L'_0) \subset \mathscr{K}^{\perp}.$$

The estimator LY is said to be *locally best* at $(\Phi, V) \in [\mathscr{T}]$ in the class $\mathscr{L}$ if $R(\Phi, V|L) \leqslant R(\Phi, V|M)$ for all $M \in \mathscr{L}$.

LEMMA 1.1. *Let $L \in \mathscr{L}$ and let $(\Phi, V) \in [\mathscr{T}]$. Then the estimator LY is locally best at (Φ, V) in the class $\mathscr{L}$ if and only if*

$$L(\Phi+V)\Pi = \Phi\Pi. \tag{1.2}$$

This result has predecessors in [2], [3], [6], and [8]. For the later use we shall now specialize matrices L_0 and Π to

$$L_0 = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \Pi = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}, \tag{1.3}$$

where the subscripts r and $n-r$ denote the orders of the unit matrices. Clearly, $\mathscr{L}$ is now the set of all matrices of the form

$$\begin{pmatrix} I_r & L_{12} \\ 0 & L_{22} \end{pmatrix},$$

where L_{12} and L_{22} may be any matrices of orders $r\times(n-r)$ and $(n-r)\times\times(n-r)$, respectively. Partitioning correspondingly Φ and V, we can easily verify that (1.2) takes the form

$$L_{12}(\Phi_{22}+V_{22}) = -V_{12}, \quad L_{22}(\Phi_{22}+V_{22}) = \Phi_{22}. \tag{1.4}$$

Lemma 1.2 below will be formulated under the additional assumption

$$(L_0-I)(I-\Pi)\Phi\Pi = 0 \tag{1.5}$$

for all $\Phi \geqslant 0$.

This condition is clearly fulfilled for $\Pi = I$. It also holds for the above-specified matrices L_0 and Π. Note that under (1.5) formula (1.2) becomes

$$L(\Pi\Phi\Pi + V\Pi) = \Pi\Phi\Pi.$$

Now let

$$\mathscr{T}_\Pi = \{(\Pi\Phi\Pi, V\Pi): (\Phi, V) \in [\mathscr{T}]\}$$

and

$$\mathscr{E} = \{(\Pi\Phi\Pi, V\Pi): \Phi \geqslant 0, V \geqslant 0, \operatorname{tr}(\Pi\Phi\Pi\Phi + V^2\Pi) = 1\}.$$

Moreover, denote by $\mathscr{W}$ the convex hull of $\mathscr{T}_\Pi \cap \mathscr{E}$. Note that $0 \notin \mathscr{W}$.

LEMMA 1.2. *Let $\mathscr{W}$ be a compact set. If LY is admissible within model $\mathscr{P}$ and if $L \in \mathscr{L}$, then there exists a point (Φ, V) with $(\Pi\Phi\Pi, V\Pi)$ in $\mathscr{W}$ such that the matrix L minimizes $R(\Phi, V|\cdot)$ in the class $\mathscr{L}$.*

In the case where $\Pi = I$, the set $\mathscr{W}$ is compact, and then Lemma 1.2 reduces to La Motte's result in [6]. For Π specified in (1.3), the set $\mathscr{W}$ is compact when $\mathscr{V} = [V]$, but it is not compact when $\mathscr{V}$ coincides with the set of all n.n.d. matrices.

For convenience of the reader the proofs of Lemmas 1.1 and 1.2 are given in the Appendix.

We end this section with two corollaries which will be useful in the sequel.

COROLLARY 1.1. *If $L \in \mathscr{L}$ has an r-fold degeneracy for the eigenvalue $\lambda = 1$ and if L satisfies (1.4) at a nonzero point $(\Phi_{22}, V_{12}, V_{22})$, then V_{22} cannot be the zero matrix.*

Proof. Suppose to the contrary that $V_{22} = 0$. Then $V_{12} = 0$ and (1.4) reduces to $L_{12}\Phi_{22} = 0, L_{22}\Phi_{22} = \Phi_{22}$. But in this case $\Phi_{22} \neq 0$ and, therefore, there exists a nonzero vector $Q \in \mathscr{R}^{n-r}$ such that $L_{22}Q = Q$ and $L_{12}Q = 0$, which contradicts the assumption that L has an r-fold degeneracy for $\lambda = 1$. The desired result that $V_{22} \neq 0$ is hence established.

Under the assumption that $\Pi = I$ we deduce next the following result:

COROLLARY 1.2. *Let L and M be $(n \times n)$-matrices. If $LQ = \lambda Q$ for $0 \leqslant \lambda \leqslant 1$, if $R(\Phi, V|M) \leqslant R(\Phi, V|L)$ for $V = (1-\lambda)QQ'$, and $\Phi = \lambda QQ'$, then $MQ = \lambda Q$.*

Proof. Since $\Phi + V = QQ'$, we have $L(\Phi+V) = LQQ' = \Phi$. Hence, by Lemma 1.1, L minimizes $R(\Phi, V|\cdot)$. From the second assumption we then obtain $R(\Phi, V|M) = R(\Phi, V|L)$ so that $M(\Phi+V) = \Phi$. Now $\Phi+V = QQ'$ applies once more to show that $MQ = \lambda Q$.

2. Necessary and sufficient conditions for admissibility. Rao has shown in [9] that LY is admissible within model $\mathscr{R}^n \times [V]$, where V is any p.d. matrix, if and only if the eigenvalues of L are in $[0, 1]$ and LV is symmetric. For the more general model $\mathscr{R}^n \times \mathscr{V}$ treated in this paper, Theorem 2.1 below gives necessary conditions for admissibility which, for the model considered by Rao, are equivalent to Rao's conditions. As will be demonstrated later, they are not sufficient for the general model.

THEOREM 2.1. *If LY is admissible within model $\mathscr{R}^n \times \mathscr{V}$, then*

(i) *the eigenvalues of L are in $[0, 1]$,*

(ii) *there exists a nonzero matrix $V \in \mathscr{V}$ such that LV is symmetric.*

Proof. If LY is admissible, then Lemma 1.2 with $\mathscr{K} = \mathscr{R}^n$ guarantees the existence of a nonzero point $(\Phi, V) \in [\mathscr{T}]$ such that

$$L(\Phi+V) = \Phi. \tag{2.1}$$

Since the left-hand side of (2.1) is symmetric, the matrix L has $r = \operatorname{rank}(\Phi+V)$ independent real eigenvectors, say $P_1, \ldots, P_r$, such that

$$\Phi+V = \sum_{i=1}^{r} \tau_i P_i P_i',$$

where $\tau_1, \ldots, \tau_r$ are some positive constants. Suppose that $P_1, \ldots, P_r$ correspond to the eigenvalues $\lambda_1, \ldots, \lambda_r$ of L, respectively. Then

$$\Phi = L(\Phi+V) = \sum_{i=1}^{r} \tau_i \lambda_i P_i P_i' \tag{2.2}$$

so that

$$V = (\Phi+V)-\Phi = \sum_{i=1}^{r} \tau_i (1-\lambda_i) P_i P_i'. \tag{2.3}$$

Since Φ and V are n.n.d., we infer from (2.2) that $\lambda_i \geqslant 0$ and from (2.3) that $1-\lambda_i \geqslant 0$. Consequently, $0 \leqslant \lambda_i \leqslant 1$ for $i = 1, \ldots, r$.

To prove that the remaining eigenvalues of L are in $[0, 1]$ we may assume without loss of generality that

$$L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix}$$

and that $\lambda_1, \ldots, \lambda_r$ are the eigenvalues of the $(r\times r)$-matrix L_{11}. By Shinozaki's lemma mentioned in Section 1, the estimator $L_{22}Z$ is then admissible for $\mathrm{E}Z$, where Z is a random variable with the parameter space $\mathscr{R}^{n-r}\times \times\{UVU' : V \in \mathscr{V}\}$, $U = (0 \quad I_{n-r})$. As before, we can show that L_{22} has at least one eigenvalue in $[0, 1]$ which evidently is also an eigenvalue of L. If necessary, we may continue in this way to conclude finally that all the eigenvalues of L are in $[0, 1]$, which completes the proof of (i).

Assertion (ii) is evident in view of (2.1) when $V \neq 0$. If (2.1) reduces to $L\Phi = \Phi$, then the matrix L has 1 as its eigenvalue. Suppose that it has an r-fold degeneracy for $\lambda = 1$. We may then assume without loss of generality that L is as above but $L_{11} = I_r$.

Let $\mathscr{W}$ be defined as in Section 1 with Π specified in (1.3). We distinguish now two cases.

(1) If $\mathscr{W}$ is compact, then Lemma 1.2 with L_0 as in (1.3) and Corollary 1.1 guarantee the existence (see (1.4)) of a point (Φ, V) in $[\mathscr{T}]$ such that

$$L_{12}(\Phi_{22}+V_{22}) = -V_{12}, \tag{2.4}$$

$$L_{22}(\Phi_{22}+V_{22}) = \Phi_{22}, \tag{2.5}$$

where $V_{22} \neq 0$.

We shall now show that LV is symmetric. For this purpose we write

$$LV = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where

$$A_{11} = V_{11}+L_{12}V'_{12}, \quad A_{12} = V_{12}+L_{12}V_{22},$$
$$A_{21} = L_{22}V'_{12}, \quad A_{22} = L_{22}V_{22}.$$

Now it follows from (2.4) and (2.5) that A_{11} and A_{22}, respectively, are symmetric. Moreover, by (2.4) and (2.5), we have

$$A_{12} = -L_{12}\Phi_{22} = -L_{12}(\Phi_{22}+V_{22})L'_{22} = V_{12}L'_{22} = A'_{21}.$$

This shows that LV is symmetric.

(2) If $\mathscr{W}$ is not compact, then there exists a sequence $(\Phi^{(n)}, V^{(n)}) \in [\mathscr{T}]$, $n = 1, 2, \ldots$, such that $A_n = (\Pi\Phi^{(n)}\Pi, V^{(n)}\Pi) \in \mathscr{W}$ and $A_n \to A_0$ as $n \to \infty$, but $A_0 \notin \mathscr{W}$. Since no subsequence of $\{PV^{(n)}P'\}$ with $P = (I_r \;\; 0)$ may converge to an n.n.d. matrix, the elements of $\{PV^{(n)}P'\}$ are not bounded. As a consequence, there exists a subsequence $\{n_i\}$ such that

$$\left\{\frac{1}{\operatorname{tr} V^{(n_i)}} V^{(n_i)}\right\} \to V \in \mathscr{R}(\Gamma\otimes\Gamma) \text{ as } i \to \infty, \quad V \notin 0,$$

where $\Gamma = P'P$. It is now obvious that $V \in \mathscr{V}$ and that LV is symmetric since $\mathscr{V}$ is closed and $L\Gamma = \Gamma$. This completes the proof of Theorem 2.1.

The next theorem may be considered as an extension of Cohen's result to n.n.d. matrices.

As before, let Y denote a random vector with the parameter space $\mathscr{P} = \mathscr{R}^n \times \mathscr{V}$, let $P = (I_r \;\; 0)$, and let $\Gamma = P'P$. Assume that $\mathscr{V} \subset \mathscr{R}(\Gamma\otimes\Gamma)$. Moreover, let Z be a random vector with the parameter space $\tilde{\mathscr{P}} = \mathscr{R}^r \times \times \{PVP' : V \in \mathscr{V}\}$.

THEOREM 2.2. *The estimator LY is admissible within model $\mathscr{P}$ if and only if*

(i) $\mathscr{R}(L-I) \subset \mathscr{R}(\Gamma(L-I)\Gamma)$,

(ii) *$PLP'Z$ is admissible within model $\tilde{\mathscr{P}}$.*

Proof. Partitioning L as

$$L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \quad L_{11} \in \mathscr{S}_{r\times r}$$

we note easily that L fulfills (i) if and only if

(2.6) $$L_{21} = 0, \quad L_{22} = I_{n-r}, \quad L_{12} = (I_r - L_{11})H,$$

where H may be any $[r\times(n-r)]$-matrix.

If LY is admissible within $\mathscr{P}$, then L must necessarily fulfill (2.6). Otherwise, an estimator with a matrix obtained by replacing in L, respectively, L_{21}, L_{22} and L_{12} by 0, I_{n-r} and $(I_r - L_{11})H$ with H selected to meet the condition

$$(I_r - L_{11})'L_{12} = (I_r - L_{11})'L_{12}(I_r - L_{11})H$$

would be better than LY. Thus (i) must hold.

Now suppose to the contrary that $M_{11}Z$ is better than $L_{11}Z = PLP'Z$ so that

$$R(P\theta, PVP'|M_{11}) \leqslant R(P\theta, PVP'|L_{11}) \tag{2.7}$$

for all $(\theta, V) \in \mathscr{P}$ with strict inequality for at least one point in $\mathscr{P}$. Putting

$$M = \begin{pmatrix} M_{11} & (I-M_{11})H \\ 0 & I \end{pmatrix}$$

and applying (2.7) yields that MY is as good as LY. Since LY is admissible within model $\mathscr{P}$, we may hence conclude that MY and LY are equivalent. But this leads to equality in (2.7), which is a contradiction.

To prove the sufficiency suppose to the contrary that MY is better than LY so that

$$R(\theta, V|M) \leqslant R(\theta, V|L) \tag{2.8}$$

for all $(\theta, V) \in \mathscr{P}$ with strict inequality for at least one point in $\mathscr{P}$. Partitioning M similarly as L, we must have $M_{21} = 0, M_{22} = I_{n-r}$ and $M_{12} = (I_r - M_{11})K$, where K is an $[r \times (n-r)]$-matrix. Applying assumption (ii) and (2.8) it may easily be checked that the risk functions of $M_{11}Z$ and $L_{11}Z$ are identical within model $\tilde{\mathscr{P}}$. Then (2.8) implies that for all $(\theta, V) \in \mathscr{P}$

$$R(\theta, V|L) - R(\theta, V|M) = \theta' \begin{pmatrix} C_{11} & C_{12} \\ C'_{12} & C_{22} \end{pmatrix} \theta \geqslant 0,$$

where

$$C_{11} = 0, \quad C_{12} = (I-L_{11})'(I-L_{11})H - (I-M_{11})'(I-M_{11})K = 0, \tag{2.9}$$

while

$$C_{22} = H'(I-L_{11})'(I-L_{11})H - K'(I-M_{11})'(I-M_{11})K.$$

Applying (2.9) we can show that $C_{22} = 0$ so that $\leqslant$ may be replaced by $=$ in (2.8). But this is a contradiction and the proof is complete.

If $\mathscr{P} = \mathscr{R}^n \times [V]$, where V is any n.n.d. matrix, then using some notation introduced by Rao [9] and Zmyślony [11] we may formulate Theorem 2.2 as follows:

COROLLARY 2.1. *The estimator LY is admissible within model $\mathscr{R}^n \times [V]$ if and only if L satisfies the following conditions*:

(i) *LV is symmetric*,

(ii) $\mathscr{R}(L-I) \subset \mathscr{R}((L-I)V)$,

(iii) $LVL' \leqslant LV$.

It may be worth-while adding that this corollary is derived directly from Lemma 1.2 without making any reference to Cohen's theorem mentioned in Section 1.

Next we establish a uniqueness result for admissible estimators, which will be used in the proofs of the last three theorems.

LEMMA 2.1. *If LY is admissible within model with any parameter space $\mathscr{P}$ and if MY is another estimator with the same risk function, i.e. $R(\theta, V|M) = R(\theta, V|L)$ for all $(\theta, V) \in \mathscr{P}$, then $M = L$.*

Proof. As known, the risk function $R(\theta, V|L)$ as defined in Section 1 is a convex function of L for all $(\theta, V) \in \mathscr{P}$, whereas it is a strict convex function of L when V is p.d. Thus, if there exists a p.d. matrix in $\mathscr{V}$, the assertion is obvious. Otherwise, we may assume without loss of generality that

$$\mathscr{V} \subset \mathscr{R}(\Gamma \otimes \Gamma), \quad \text{where } \Gamma = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \; 0 < r < n.$$

Then, by part (i) of Theorem 2.2, $M_{21} = L_{21} = 0, M_{22} = L_{22} = I$ and there exist matrices K and H such that $M_{12} = (I - M_{11})K$ and $L_{12} = (I - L_{11})H$.

In turn, part (ii) of Theorem 2.2 together with the assumption that MY and LY have the same risk function imply then that $M_{11} = L_{11}$. Using this fact and once more the assumption that LY and MY have the same risk function, we obtain $(I - L_{11})K - (I - L_{11})H = 0$, whence $M = L$, which completes the proof.

Theorem 2.3 below gives a condition under which the necessary conditions appearing in Theorem 2.1 are also sufficient. It is more general than Rao's theorem mentioned at the beginning of this section.

THEOREM 2.3. *If $\mathscr{V}$ consists only of p.d. matrices (except for the zero matrix), then LY is admissible within model $\mathscr{P} = \mathscr{R}^n \times \mathscr{V}$ if and only if all eigenvalues of L are in $[0, 1]$ and LV is symmetric for some nonzero matrix V in $\mathscr{V}$.*

Proof. The necessity has already been established in Theorem 2.1. To show the sufficiency suppose that MY is as good as LY and that LV_0, where V_0 is a p.d. matrix in $\mathscr{V}$, is symmetric. Then, by Corollary 2.1, the estimator LY is admissible within model $\mathscr{R}^n \times [V_0]$. Consequently, the risk functions of MY and LY are identical for all $(\theta, V_0) \in \mathscr{P}$ so that $M = L$ by Lemma 2.1. This proves that LY is admissible.

The assertion of Theorem 2.3 may be rephrased as follows:

The estimator LY is admissible within model $\mathscr{R}^n \times \mathscr{V}$ if and only if there exists a nonzero matrix $V \in \mathscr{V}$ such that LY is admissible within model $\mathscr{R}^n \times [V]$.

The following example shows that this corollary may not hold when $\mathscr{V}$ contains singular matrices.

Example. Let Y be a random vector with the parameter space $\mathscr{R}^2 \times \mathscr{V}$, where $\mathscr{V}$ is generated by the unit matrix and by

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

From Lemma 2.1 it then follows that the estimator LY, where

$$L = \begin{pmatrix} 1/2 & 1 \\ 0 & 1/2 \end{pmatrix},$$

is admissible. In fact, if MY is as good as LY, then M must be of the form

$$\begin{pmatrix} 1/2 & m \\ 0 & n \end{pmatrix}$$

by Corollary 1.2, and, moreover, M must satisfy (1.2) with

$$\Pi = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi = \begin{pmatrix} 16 & 4 \\ 4 & 1 \end{pmatrix}, \quad \text{and} \quad V = I,$$

since L satisfies (1.2) with the above-specified matrices Π, Φ, and V. This leads to $m = 1$ and $n = 1/2$ so that $M = L$. Hence LY is admissible.

If there existed a nonzero matrix $V \in \mathscr{V}$ such that LY were admissible within model $\mathscr{R}^2 \times [V]$, the matrix LV would be symmetric by Theorem 2.1. Now, since

$$L\begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} 1/2 & \alpha \\ 0 & \alpha/2 \end{pmatrix},$$

LV is symmetric if and only if

$$V = \begin{pmatrix} \tau^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

The matrix L, however, does not meet the conditions of Corollary 2.1 with V as above and $\tau^2 > 0$. Hence LY is inadmissible within each model $\mathscr{R}^2 \times [V]$ when V ranges over all nonzero matrices in $\mathscr{V}$.

A further immediate consequence of Lemma 2.1 is the following result:

THEOREM 2.4. *Each estimator admissible within model* $\mathscr{P}_1 = \mathscr{R}^n \times \mathscr{V}_1$ *is also admissible within any model* $\mathscr{P}_2 = \mathscr{R}^n \times \mathscr{V}_2$ *if* $\mathscr{V}_1 \subset \mathscr{V}_2$.

Proof. Let LY be admissible within model $\mathscr{P}_1$ and suppose to the contrary that it is not admissible in model $\mathscr{P}_2$. Now, if MY is better than LY within model $\mathscr{P}_2$, it must be therefore as good as LY in $\mathscr{P}_1$. But LY is admissible in $\mathscr{P}_1$ so that $M = L$ by Lemma 2.1, which is a contradiction.

We shall conclude the paper with a theorem referring to the case where $\mathscr{V}$ coincides with the family $\mathscr{V}_n$ of all n.n.d. matrices. To this end let Y_n, $n = 1, 2, \ldots$, be a random vector with the parameter space $\mathscr{P}_n = \mathscr{R}^n \times \mathscr{V}_n$.

THEOREM 2.5. *A necessary and sufficient condition for the estimator LY_n to be admissible for* $\mathrm{E}Y_n$ *within model* $\mathscr{P}_n$ *is that the eigenvalues of* L *are in* $[0, 1]$.

Proof. Clearly, the necessity of this theorem is ensured by Theorem 2.1. Now suppose that the eigenvalues of the matrix L are in $[0, 1]$. Since we may assume without loss of generality that L is an upper triangle matrix, the sufficiency can be derived straightforward by induction over m from the following result:

The estimator MY_m, *where*

$$M = \begin{pmatrix} \lambda & M_{12} \\ 0 & M_{22} \end{pmatrix},$$

is admissible for $\mathrm{E}Y_m$ *within model* $\mathscr{P}_m$ *if* $\lambda \in [0, 1]$ *and if* $M_{22} Y_{m-1}$ *is admissible for* $\mathrm{E}Y_{m-1}$ *within model* $\mathscr{P}_{m-1}$, $m = 2, 3, \ldots$

To prove this latter result suppose that NY_m is as good as MY_m. Then necessarily

$$N = \begin{pmatrix} \lambda & N_{12} \\ 0 & N_{22} \end{pmatrix}$$

by Corollary 1.1. Partitioning correspondingly

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} V_{11} & V_{12} \\ V'_{12} & V_{22} \end{pmatrix},$$

we find after some computations that for all $(\theta, V) \in \mathscr{P}_m$

$$\begin{aligned} 0 &\leqslant R(\theta, V|M) - R(\theta, V|N) \\ &= R(\theta_2, V_{22}|M_{22}) - R(\theta_2, V_{22}|N_{22}) + \operatorname{tr}(M'_{12}M_{12} - N'_{12}N_{12})(\theta_2\theta'_2 + V_{22}) + \\ &\quad + 2(M_{12} - N_{12})[\lambda V'_{12} - (1-\lambda)\theta_1\theta'_2]. \end{aligned}$$

Since θ_1 and θ_2 may be arbitrary and since V_{12} is subject only to the condition $\mathscr{R}(V'_{12}) \subset \mathscr{R}(V_{22})$, we have $N_{12} = M_{12}$ and, therefore, the last two terms on the right-hand side cancel. Consequently, $N_{22} Y_{m-1}$ is as good as $M_{22} Y_{m-1}$, which implies that $N_{22} = M_{22}$, since $M_{22} Y_{m-1}$ is admissible within $\mathscr{P}_{m-1}$. Hence MY_m is admissible for $\mathrm{E}Y_m$ as asserted. This completes the proof of Theorem 2.5.

Appendix.

Proof of Lemma 1.1. Lemma 1.1 states that $L_0 + Z_0\Pi$, where Z_0 is an $(n \times n)$-matrix, minimizes $R(\Phi, V|\cdot)$ in $\mathscr{L}$, i.e. that

$$R(\Phi, V|L_0 + Z_0\Pi) = \min_{Z \in \mathscr{S}_{n \times n}} R(\Phi, V|L_0 + Z\Pi)$$

if and only if

$$(L_0 + Z_0\Pi)(\Phi + V)\Pi = \Phi\Pi. \tag{A.1}$$

To show this it will be convenient to introduce the following notation:

$$S = \Pi(\Phi+V)\Pi \quad \text{and} \quad T = \Phi\Pi - L_0(\Phi+V)\Pi.$$

With this notation we obtain $\mathscr{R}(T') \subset \mathscr{R}(S^+)$, so that $SS^+T' = T'$. Using this fact, by straightforward computations we see that

$$R(\Phi, V|L_0+Z\Pi) = \operatorname{tr}(ZS-T)S^+(ZS-T)' - \operatorname{tr} TS^+T' + R(\Phi, V|L_0).$$

Since $\mathscr{R}(SZ'-T') \subset \mathscr{R}(S^+)$ and S^+ is n.n.d., it is now evident that the right-hand side reaches its minimum value with respect to Z if and only if $ZS = T$, which in terms of the original notation reduces to (A.1).

Proof of Lemma 1.2. This lemma follows by arguments similar to those used by Olsen et al. [7] to show that an admissible unbiased estimator is locally best at some nonzero point and by La Motte [6] to show that an admissible estimator is locally best at a nonzero point in $[\mathscr{T}]$.

Under assumption (1.5) it is sufficient to show that

$$0 \in W = \{L(\Pi\Phi\Pi - V\Pi) - \Pi\Phi\Pi \colon (\Pi\Phi\Pi, V\Pi) \in \mathscr{W}\}$$

because $0 \notin \mathscr{W}$. Suppose to the contrary that $0 \notin W$. Since by assumption $\mathscr{W}$ is a compact convex set, the separating hyperplane theorem assures the existence of a matrix H such that

$$\text{(A.2)} \qquad \operatorname{tr}\{[L(\Pi\Phi\Pi + V\Pi) - \Pi\Phi\Pi]H'\} < 0$$

for all $(\Pi\Phi\Pi, V\Pi)$ in $\mathscr{W}$.

Now define for each $\gamma \in \mathscr{R}$ a matrix $M = L+\gamma H\Pi$. Clearly, $M \in \mathscr{L}$. Taking into account (1.5) it can easily be verified that

$$\pi(\Phi, V, \gamma) = R(\Phi, V|M) - R(\Phi, V|L) = a\gamma^2 + 2b\gamma,$$

where $a = \operatorname{tr} H\Pi(\Phi+V)\Pi H'$ and $b = \operatorname{tr}\{[L(\Pi\Phi\Pi + V\Pi) - \Pi\Phi\Pi]H'\}$. Now (A.2) gives $a = a(\Phi, V) > 0$ and $b = b(\Phi, V) < 0$ for all $(\Pi\Phi\Pi, V\Pi)$ in $\mathscr{W}$. Then for an arbitrary but fixed pair of matrices $(\Pi\Phi\Pi, V\Pi) \in \mathscr{W}$ the quadratic polynomial $\pi(\Phi, V, \gamma)$ in γ achieves its minimum value $-b^2/a$ when $\gamma = g = -b/a$. Since g, considered as a mapping from $\mathscr{W}$ to $\mathscr{R}$, is continuous and strictly positive on the compact set $\mathscr{W}$, there exists an $\varepsilon > 0$ such that $g(\Pi\Phi\Pi, V\Pi) \geqslant \varepsilon$ for all $(\Pi\Phi\Pi, V\Pi) \in \mathscr{W}$. Therefore

$$\varepsilon^{-1}\pi(\Phi, V, \varepsilon) \leqslant g^{-1}(\Pi\Phi\Pi, V\Pi)\pi(\Phi, V, g(\Pi\Phi\Pi, V\Pi)) < 0$$

for all $(\Pi\Phi\Pi, V\Pi)$ in $\mathscr{W}$. This proves that

$$R(\Phi, V|L+\gamma H\Pi) \leqslant R(\Phi, V|L) \quad \text{for all } (\Phi, V) \in [\mathscr{T}]$$

with strict inequality if $(\Pi\Phi\Pi, V\Pi)$ is in $\mathscr{W}$, contradicting the admissibility of LY. Hence the proof is complete.

Acknowledgement. This work was completed while the author was engaged in the Statistical Laboratory of the University of California in Berkeley. A preliminary version of this paper appeared as Technical Report No. ONR 80-05 at Berkeley. The author wants to express his gratitude to Professor J. Neyman for the kind hospitality offered to him during his stay in Berkeley.

REFERENCES

[1] A. Cohen, *All admissible linear estimates of the mean vector*, Ann. Math. Statist. 37 (1966), p. 458-463.

[2] S. Gnot, *Locally best linear estimation in Euclidean vector spaces*, p. 144-151 in: Mathematical Statistics and Probability Theory (W. Klonecki, A. Kozek and J. Rosiński, eds.), Lecture Notes in Statistics 2, Springer Verlag, 1980.

[3] W. Klonecki, *Admissible linear estimation in linear models*, Presented at the 12th European Meeting of Statisticians in Varna, Bulgaria, 1979.

[4] — *Linear estimators of the mean vector in linear models, problem of admissibility*, Technical Report, Statistical Laboratory, University of California, Berkeley 1980.

[5] L. R. La Motte, *Bayes linear estimators*, Technometrics 20 (1978), p. 281-289.

[6] — *Some results on biased linear estimation applied to variance component estimation*, p. 266-274 in: Mathematical Statistics and Probability Theory (W. Klonecki, A. Kozek and J. Rosiński, eds.), Lecture Notes in Statistics 2, Springer Verlag, 1980.

[7] A. Olsen, J. Seely and D. Birkes, *Invariant quadratic unbiased estimation for two variance components*, Ann. Statist. 4 (1976), p. 878-890.

[8] C. R. Rao, *Unified theory of linear estimation*, Sankhyā, Ser. A, 33 (1971), p. 371-394.

[9] — *Estimation of parameters in a linear model*, Ann. Statist. 4 (1976), p. 1023-1037.

[10] N. Shinozaki, *A study of generalized inverse of matrix and estimation with quadratic loss*, Ph. D. Thesis submitted to the Keio University, Japan, 1975.

[11] R. Zmyślony, *On admissible estimation for parametric functions in linear models*, Preprint No. 216, Institute of Mathematics, Polish Academy of Sciences, 1980.

Institute of Mathematics
Polish Academy of Sciences
ul. Kopernika 18, 51-617 Wrocław, Poland

Received on 26. 7. 1979