3-1

PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 1, Fasc. 1 (1980), p. 1-13

BEST UNBIASED LINEAR ESTIMATION, A COORDINATE FREE APPROACH

BY

S. GNOT, W. KLONECKI AND R. ZMYŚLONY (WROCŁAW)

Abstract. This paper gives further developments of the theory of uniformly minimum variance unbiased estimation (UMVUE) in Euclidean vector spaces as originated by W. Kruskal, G. Zyskind and J. Seely. It gives necessary and sufficient conditions for the existence of a UMVUE for each estimable function in any subspace of linear estimators with no restrictions posed on the covariance operators. Also construction of UMVUE's in a given subspace of linear estimators, if they exist, is considered. The developed theory is illustrated by two examples: estimation of variance components in a general mixed linear model and estimation of the mean in a multivariate linear model.

1. Introduction. In the paper we consider a random element Z with values in an arbitrary Euclidean vector space $\mathscr{K}$ endowed with an inner product denoted by $[\cdot,\cdot]$. The expectation $\mathrm{E}Z$ and the covariance $\operatorname{Cov} Z$ (for definitions, see [10]) are assumed to exist. $\mathrm{E}Z$ is assumed to be an element of a known subspace $\mathscr{E}$ of $\mathscr{K}$. The symbol Θ stands for the minimal convex cone containing all covariance operators of Z and is assumed to be known. For our considerations we may assume without loss of generality that there is no functional relationship between $\mathrm{E}Z$ and $\operatorname{Cov} Z$. Let $\mathscr{M}(\mathscr{E}, \Theta)$ denote a model as described above.

The problem we consider is the estimation of functions of the form $[A, \mathrm{E}Z]$ defined on the parameter space $\mathscr{K} \times \mathscr{E}$, say. We confine attention to the class of estimators of the form $[B, Z]$, where B may stand for any element in a given subspace $\mathscr{K}_0$ of $\mathscr{K}$. The function $[A, \mathrm{E}Z]$ is said to be $\mathscr{K}_0$-*estimable* if there exists a $B \in \mathscr{K}_0$ such that $[B, \mathrm{E}Z] = [A, \mathrm{E}Z]$. Within this model context the entire collection of $\mathscr{K}_0$-estimable functions can be presented as $\{[A, \mathrm{E}Z]\colon A \in \mathscr{K}_0\}$. The estimator $[A, Z]$, where

$A \in \mathcal{K}_0$, is called *$\mathcal{K}_0$-best for a $\mathcal{K}_0$-estimable function g* if $[A, \mathrm{E}Z] = g$ and if $\mathrm{Var}[A, Z] \leqslant \mathrm{Var}[B, Z]$ for each $B \in \mathcal{K}_0$ such that $[B, \mathrm{E}Z] = g$. The model $\mathcal{M}(\mathcal{E}, \Theta)$ is said to be *$\mathcal{K}_0$-regular* if there exists a $\mathcal{K}_0$-best estimator for each $\mathcal{K}_0$-estimable function. Throughout the paper, $\mathcal{K}$-best estimators are called simply *best estimators.*

The purpose of this paper is to provide conditions which are necessary and sufficient for $\mathcal{M}(\mathcal{E}, \Theta)$ to be $\mathcal{K}_0$-regular. The problem of finding $\mathcal{K}_0$-best estimators, when ones exist, is also considered.

Most of the problems of linear estimation in linear models and of quadratic estimation of variance components in mixed linear models are special cases of the problems considered in this paper. Also some problems of estimation in multivariate linear models are covered by the developed theory.

In Section 2 we summarize for the sake of completeness a number of facts on linear operators which are used throughout the paper. In Section 3, Lemma 3.2 extends a result of Seely [17] on the representation of estimable functions. Theorem 3.1 extends a result due to Seely and Zyskind [20]. It gives a necessary and sufficient condition for a model $\mathcal{M}(\mathcal{E}, \Theta)$ to be $\mathcal{K}_0$-regular. In Section 4 we introduce a concept of a $\mathcal{K}_0$-Gauss-Markov estimator ($\mathcal{K}_0$-GME) which is as related to $\mathcal{K}_0$-estimable functions as the usual GME is related to $\mathcal{K}$-estimable functions. These results extend recent works on Gauss-Markov estimators of Mitra and Moore [14] and of Drygas [1].

In the next section we deal with the problem of constructing $\mathcal{K}_0$-best estimators under various assumptions posed on $\mathcal{K}_0$. In particular, we consider the situation where $\mathcal{K}_0^{\perp} \subset \mathcal{E}$, which is of special interest, because under this assumption each $\mathcal{K}_0$-best estimator is also $\mathcal{K}$-best. The results in Section 5 are closely related to the work on locally best estimation of variance components developed by Rao [15], [16], and Kleffe [8].

In the last section we show how to deduce from the developed theory the results regarding the existence of best unbiased estimators of variance components established under various degrees of generality by Seely [17]-[19], Seely and Zyskind [20], Zmyślony [21], [22], Kleffe and Pincus [9], Gnot et al. [5]-[7] and by Drygas [2].

2. Preliminaries. Throughout the paper, elements of the Euclidean vector space $\mathcal{K}$ are denoted by capital Latin letters $(A, B, \ldots)$. For subspaces $\mathcal{A}$ and $\mathcal{B}$ of $\mathcal{K}$, the expression $\mathcal{A}+\mathcal{B}$ stands for the set of all vectors $A+B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$, and the expression $\mathcal{A} \cap \mathcal{B}$ for the set of all vectors belonging to $\mathcal{A}$ and to $\mathcal{B}$.

Note that

$$(\mathcal{A}+\mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap \mathcal{C} + \mathcal{B}$$

provided $\mathcal{B} \subset \mathcal{C} \subset \mathcal{K}$.

We write $\mathscr{A} \oplus \mathscr{B}$ instead of $\mathscr{A} + \mathscr{B}$ if $\mathscr{A} \cap \mathscr{B} = \{0\}$. The set of all vectors orthogonal to $\mathscr{A} \subset \mathscr{K}$ is denoted by $\mathscr{A}^{\perp}$. Then

$$(\mathscr{K}_1 + \mathscr{K}_2)^{\perp} = \mathscr{K}_1^{\perp} \cap \mathscr{K}_2^{\perp} \quad \text{and} \quad (\mathscr{K}_1 \cap \mathscr{K}_2)^{\perp} = \mathscr{K}_1^{\perp} + \mathscr{K}_2^{\perp},$$

provided $\mathscr{K}_1$ and $\mathscr{K}_2$ are subspaces of $\mathscr{K}$.

Let $\mathscr{L}$ be the space of all linear operators transforming $\mathscr{K}$ into $\mathscr{K}$. Elements of $\mathscr{L}$ are denoted by capital Greek letters ($\Gamma, \Sigma, \ldots$). The adjoint of $\Gamma \in \mathscr{L}$ is denoted by Γ'. The operator $\Gamma \in \mathscr{L}$ is said to be *self-adjoint* if $\Gamma = \Gamma'$, *nonnegative-definite* if $[\Gamma A, A] \geqslant 0$ for all $A \in \mathscr{K}$, and *positive-definite* if $[\Gamma A, A] > 0$ for all $A \in \mathscr{K}$ except for $A = 0$. For any $\Gamma \in \mathscr{L}$, $\Gamma(\mathscr{A})$ and $\Gamma^{-1}(\mathscr{A})$ stand for the image and the inverse image of $\mathscr{A}$ by Γ, respectively. In case $\mathscr{A} = \mathscr{K}$, we write $\mathscr{R}(\Gamma)$ instead of $\Gamma(\mathscr{K})$. Furthermore, $\mathscr{N}(\Gamma)$ stands for the null space $\{A \in \mathscr{K}\colon \Gamma(A) = 0\}$. For $\Gamma_1, \Gamma_2 \in \mathscr{L}$ and for each $\Gamma \in \mathscr{L}$ we define a linear operator $\Gamma_1 \otimes \Gamma_2$ from $\mathscr{L}$ into itself by

$$(\Gamma_1 \otimes \Gamma_2)\,\Gamma = \Gamma_1 \Gamma \Gamma_2'.$$

As usual, Γ^- denotes a generalized inverse and Γ^+ the Moore-Penrose inverse of $\Gamma \in \mathscr{L}$.

Let L be a linear operator transforming $\mathscr{K}$ into a linear space $\mathscr{W}$ endowed with an inner product $(\cdot,\cdot)$, The following standard set theory facts are useful in the sequel:

$$L^{-1}\{L(\mathscr{A})\} = \mathscr{A} + \mathscr{N}(L), \tag{2.1}$$

$$LL^{-1}(\mathscr{A}) = \mathscr{R}(L) \cap \mathscr{A}, \tag{2.2}$$

$$\Gamma^{-1}(\mathscr{A}) = \Gamma^{+}\{\mathscr{A} \cap \mathscr{R}(\Gamma)\} + \mathscr{N}(\Gamma). \tag{2.3}$$

For a self-adjoint operator $\Gamma \in \mathscr{L}$,

$$\Lambda = \Gamma^- \quad \text{iff} \quad \Gamma\Gamma^+ \Lambda \Gamma\Gamma^+ = \Gamma^+. \tag{2.4}$$

Now let $\mathscr{K}_1$ and $\mathscr{K}_2$ be disjoint subspaces of $\mathscr{K}$. Any vector $A \in \mathscr{K}_1 \oplus \mathscr{K}_2$ has a unique decomposition $A = A_1 + A_2$, where $A_1 \in \mathscr{K}_1$ and $A_2 \in \mathscr{K}_2$. Then Π is said to be a *projection operator onto* $\mathscr{K}_1$ *along* $\mathscr{K}_2$ if $\Pi A = A_1$ for every $A \in \mathscr{K}_1 \oplus \mathscr{K}_2$ and the corresponding A_1. The projection operator Π is *unique* iff $\mathscr{K}_1 \oplus \mathscr{K}_2$ coincides with the entire space $\mathscr{K}$.

If Π is a projection operator onto $\mathscr{K}_1$ along $\mathscr{K}_2$, then

(i) for any $\mathscr{A} \subset \mathscr{K}_1 \oplus \mathscr{K}_2$,

$$A \in \mathscr{A} + \mathscr{K}_2 \quad \text{iff} \quad \Pi A \in \Pi(\mathscr{A})$$

provided $A \in \mathscr{K}_1 \oplus \mathscr{K}_2$;

(ii) $\Pi' A \in (\mathscr{K}_1 \oplus \mathscr{K}_2)^{\perp}$ for each $A \in \mathscr{K}_1^{\perp}$;

(iii) $\Pi' A - A \in (\mathscr{K}_1 \oplus \mathscr{K}_2)^{\perp}$ for each $A \in \mathscr{K}_2^{\perp}$.

From (ii) and (iii) it follows that $\Pi'(\mathscr{K}) \subset \mathscr{K}_2^{\perp}$. Moreover, Π' is the projection operator onto $\mathscr{K}_2^{\perp}$ along $\mathscr{K}_1^{\perp}$ if $\mathscr{K}_1 \oplus \mathscr{K}_2 = \mathscr{K}$.

3. Main results. In this section we present results which are basic in our discussion of the problem of existence of $\mathcal{K}_0$-best estimators.

Throughout the paper, $\mathcal{K}_0$ stands for a subspace of $\mathcal{K}$ and $\mathcal{F} = \mathcal{E}+\mathcal{K}_0^{\perp}$.

LEMMA 3.1 (Lehmann and Scheffé [13]). *Let* $A \in \mathcal{K}_0$. *Then* $[A, Z]$ *is* $\mathcal{K}_0$*-best for* $[A, EZ]$ *iff* $\Gamma A \in \mathcal{F}$ *for all* $\Gamma \in \Theta$.

LEMMA 3.2. *Let* $\mathcal{A}$ *be a subspace of* $\mathcal{K}_0$ *and let* $\mathcal{G}_0$ *be the class of all* $\mathcal{K}_0$*-estimable functions. Then*

$$\mathcal{G}_0 = \{[A, EZ]\colon A \in \mathcal{A}\} \quad \text{iff} \quad \mathcal{A}+\mathcal{F}^{\perp} = \mathcal{K}_0.$$

The proof of Lemma 3.2 for $\mathcal{K}_0 = \mathcal{K}$ is given by Seely [17]. The above-given extension may be established in the same manner.

By virtue of Lemmas 3.1 and 3.2 we obtain immediately the following result:

LEMMA 3.3. *The model* $\mathcal{M}(\mathcal{E}, \Theta)$ *is* $\mathcal{K}_0$*-regular iff there exists a subspace* $\mathcal{H} \subset \mathcal{K}$ *such that*

$$\mathcal{H}+\mathcal{F}^{\perp} = \mathcal{K}_0 \tag{3.1}$$

and

$$\Gamma\mathcal{H} \subset \mathcal{F} \quad \text{for all } \Gamma \in \Theta. \tag{3.2}$$

A subspace $\mathcal{H} \subset \mathcal{K}_0$ which satisfies conditions (3.1) and (3.2) is said to be $\mathcal{K}_0$*-best*. A $\mathcal{K}_0$-best subspace is said to be *minimal* if none of its proper subspaces is $\mathcal{K}_0$-best. A $\mathcal{K}_0$-best subspace $\mathcal{H}$ is minimal iff $\mathcal{H} \cap \mathcal{F}^{\perp} = \{0\}$.

If $\mathcal{H}$ is $\mathcal{K}_0$-best, then for each $\mathcal{K}_0$-estimable function there exists an element $A \in \mathcal{H}$ such that $[A, Z]$ is its $\mathcal{K}_0$-best estimator, and for each $A \in \mathcal{H}$ the estimator $[A, Z]$ is $\mathcal{K}_0$-best for $[A, EZ]$. Moreover, if $\mathcal{H}$ is a minimal $\mathcal{K}_0$-best subspace, then for each $\mathcal{K}_0$-estimable function g there exists exactly one element $A \in \mathcal{H}$ such that $[A, Z]$ is a $\mathcal{K}_0$-best estimator of g. Finally, if $\mathcal{H}$ is a $\mathcal{K}_0$-best subspace, then

$$\mathcal{H} \subset \bigcap_{\Gamma\in\Theta} \Gamma^{-1}(\mathcal{F})$$

by Lemma 3.3.

To include in our considerations the case where there is no invertible operator in Θ we introduce the concept of a maximal element. The operator $\Sigma \in \Theta$ is said to be *maximal* in Θ if

$$\mathcal{N}(\Sigma) = \bigcap_{\Gamma\in\Theta} \mathcal{N}(\Gamma).$$

Clearly, if Σ is maximal in Θ, then for all $\Gamma \in \Theta$ we have $\mathcal{R}(\Gamma) \subset \mathcal{R}(\Sigma)$. One can easily show that there exists a maximal element in Θ (see Lemma 1 in [12]). Throughout the paper, Σ will stand for a maximal element in Θ.

Now we are in a position to prove the main result of the paper.

THEOREM 3.1. *The model* $\mathcal{M}(\mathcal{E}, \Theta)$ *is* $\mathcal{K}_0$*-regular iff*

$$\Gamma(\mathcal{K}_0 \cap \Sigma^{-1}(\mathcal{F})) \subset \mathcal{F} \quad \textit{for all } \Gamma \in \Theta. \tag{3.3}$$

Proof. If $\mathcal{M}(\mathcal{E}, \Theta)$ is $\mathcal{K}_0$-regular, then the subspace

$$\mathcal{H}_0 = \mathcal{K}_0 \cap \Big[\bigcap_{\Gamma \in \Theta} \Gamma^{-1}(\mathcal{F})\Big]$$

satisfies conditions (3.1) and (3.2). Thus to prove the sufficiency one needs only to show that

$$\mathcal{H}_0 = \mathcal{K}_0 \cap \Sigma^{-1}(\mathcal{F}). \tag{3.4}$$

Using an easily proved set theory fact

$$\mathcal{R}(\Sigma) \subset \Sigma(\mathcal{F}^\perp) + \mathcal{K}_0^\perp + \mathcal{F} \cap \mathcal{K}_0$$

and the assumption

$$\mathcal{R}(\Gamma) \subset \mathcal{R}(\Sigma) \quad \text{for all } \Gamma \in \Theta,$$

we obtain

$$\Gamma(\mathcal{F}^\perp) \subset \Sigma(\mathcal{F}^\perp) + \mathcal{K}_0^\perp + \mathcal{F} \cap \mathcal{K}_0 \quad \text{for all } \Gamma \in \Theta.$$

Hence

$$\mathcal{H}_0^\perp \subset \Sigma(\mathcal{F}^\perp) + \mathcal{K}_0^\perp + \mathcal{F} \cap \mathcal{K}_0. \tag{3.5}$$

Using (3.1) we get

$$\mathcal{H}_0^\perp \cap \mathcal{F} \cap \mathcal{K}_0 = \{0\}.$$

On the other hand, since $\Sigma \in \Theta$, we obtain

$$\mathcal{H}_0 \subset \mathcal{K}_0 \cap \Sigma^{-1}(\mathcal{F}). \tag{3.6}$$

Consequently, by (3.5),

$$\mathcal{H}_0^\perp \subset \Sigma(\mathcal{F}^\perp) + \mathcal{K}_0^\perp. \tag{3.7}$$

Now, combining (3.6) and (3.7), we get (3.4), and the sufficiency is established.

Now suppose that (3.3) holds. Using elementary set theory methods one can easily show that

$$\mathcal{K}_0 \cap \Sigma^{-1}(\mathcal{F}) + \mathcal{F}^\perp = \mathcal{K}_0.$$

Thus $\mathcal{K}_0 \cap \Sigma^{-1}(\mathcal{F})$ satisfies conditions (3.1) and (3.2), and the necessity is established.

Theorem 3.1 shows that, once it is established that $\mathcal{M}(\mathcal{E}, \Theta)$ is $\mathcal{K}_0$-regular, calculating $\mathcal{K}_0$-best estimators within $\mathcal{M}(\mathcal{E}, \Theta)$ reduces to calculating $\mathcal{K}_0$-best estimators within $\mathscr{M}(\mathscr{E}, \Theta_0)$, where $\Theta_0 = \{\varkappa\Sigma \colon \varkappa \in \mathscr{R}^+\}$. Obviously, for $\mathscr{K}_0 = \mathscr{K}$ condition (3.3) reduces to

$$\Gamma\Sigma^{-1}(\mathscr{E}) \subset \mathscr{E} \quad \text{for all } \Gamma \in \Theta.$$

Moreover, the $\mathscr{K}_0$-estimation within $\mathscr{M}(\mathscr{E}, \Theta)$ is in a sense equivalent to the estimation in the class of all linear estimators within an appropriately induced linear model. To show this let us introduce some additional notation.

As previously, let L be a linear operator transforming $\mathscr{K}$ into $\mathscr{W}$ such that $\mathscr{R}(L') = \mathscr{K}_0$. Now define a random vector $Y = LZ$. Then

$$\mathrm{E}\,Y = L\theta \in \mathscr{E}_L = L(\mathscr{E}) \quad \text{and} \quad \mathrm{Cov}\,Y = L\Gamma L' \in \Theta_L = L \otimes L(\Theta),$$

where $\theta = \mathrm{E}Z$ and $\Gamma = \mathrm{Cov}\,Z$. With this notation we have

$$\{(W, Y) \colon W \in \mathscr{W}\} = \{[L'W, Z] \colon W \in \mathscr{W}\} = \{[A, Z] \colon A \in \mathscr{K}_0\}.$$

If (W, Y) is a uniformly best estimator of $(W, L\theta)$ within $\{(W, Y) \colon W \in \mathscr{W}\}$, then, by Lemma 3.1,

$$L\Gamma L'W \in \mathscr{E}_L \quad \text{for all } \Gamma \in \Theta$$

or, equivalently,

$$\Gamma L'W \in L^{-1}L(\mathscr{E}) \quad \text{for all } \Gamma \in \Theta.$$

Now, by (2.1),

$$L^{-1}\{L(\mathscr{E})\} = \mathscr{E} + \mathscr{K}_0^{\perp}.$$

Applying Lemma 3.1 shows that (W, Y) is uniformly best iff

$$\Gamma L'W \in \mathscr{E} + \mathscr{K}_0^{\perp} \quad \text{for all } \Gamma \in \Theta.$$

Thus the following result is established:

LEMMA 3.4. *The estimator* (W, Y) *is best for* $(W, L\theta)$ *within* $\{(W, Y) \colon W \in \mathscr{W}\}$ *iff* $[L'W, Z]$ *is* $\mathscr{K}_0$*-best for* $[L'W, \theta] = (W, L\theta)$ *within* $\{[A, Z] \colon A \in \mathscr{K}_0\}$.

This lemma yields immediately the following important conclusion:

THEOREM 3.2. *The model* $\mathscr{M}(\mathscr{E}, \Theta)$ *is* $\mathscr{K}_0$*-regular iff the model* $\mathscr{M}(\mathscr{E}_L, \Theta_L)$ *is regular.*

4. $\mathscr{K}_0$-Gauss-Markov estimators. As previously, let $\mathscr{M}(\mathscr{E}, \Theta)$ be a model consisting of a $\mathscr{K}$-valued random vector Z such that $\mathrm{E}Z \in \mathscr{E}$ and $\mathrm{Cov}\,Z \in \Theta$. Let L_0 be the orthogonal projection onto a subspace $\mathscr{K}_0 \subset \mathscr{K}$ and let $Z_0 = L_0 Z$. Finally, let

$$\Omega = \{\Pi \colon \mathscr{K} \to \mathscr{K}, \mathscr{R}(\Pi') \subset \mathscr{K}_0, \mathrm{E}\Pi Z = \mathrm{E}Z_0\}.$$

The estimator $\Pi_0 Z$ is said to be a $\mathscr{K}_0$-*Gauss-Markov estimator* ($\mathscr{K}_0$-GME, for short) if

(i) $\Pi_0 \in \Omega$,

(ii) $\operatorname{Var}[A, \Pi_0 Z] \leqslant \operatorname{Var}[A, \Pi Z]$ for all $A \in \mathscr{K}$ and all $\Pi \in \Omega$.

Clearly, a $\mathscr{K}_0$-GME coincides with the usual GME as defined in [1], say, if $\mathscr{K}_0 = \mathscr{K}$.

Now put $\mathscr{W} = \mathscr{K}$ and let us specialize L to L_0. Let $\mathscr{M}(\mathscr{E}_L, \Theta_L)$ be the induced linear model as defined in Section 3. We shall translate certain known facts about GME in the induced model $\mathscr{M}(\mathscr{E}_L, \Theta_L)$ to equivalent statements on $\mathscr{K}_0$-GME in the original setting of the model $\mathscr{M}(\mathscr{E}, \Theta)$.

THEOREM 4.1. *There exists a $\mathscr{K}_0$-GME iff $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_0$-regular.*

Proof. If $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_0$-regular, then $\mathscr{M}(\mathscr{E}_L, \Theta_L)$ is regular by Theorem 3.2. Thus there exists a GME in $\mathscr{M}(\mathscr{E}, \Theta)$ to be denoted by $\Pi_L Z_0$, where Π_L projects on $\mathscr{E}_L$ along $\Sigma_L(\mathscr{E}_L^{\perp})$. Now let $\Pi = \Pi_L L$. Since $\Pi \in \Omega$ and $[A, \Pi_L Z_0] = [A, \Pi Z]$, $A \in \mathscr{K}$, we infer that ΠZ is a $\mathscr{K}_0$-GME within the original model. To prove the reverse statement note that there exists a linear operator Π_L: $\mathscr{K}_0 \to \mathscr{K}_0$ such that $\Pi = \Pi_L L$. Since $\Pi_L Z_0$ turns out to be then a GME in $\mathscr{M}(\mathscr{E}_L, \Theta_L)$, the original model is $\mathscr{K}_0$-regular.

The following theorem gives a method of constructing a $\mathscr{K}_0$-GME if one exists.

THEOREM 4.2. *Let $\mathscr{M}(\mathscr{E}, \Theta)$ be a $\mathscr{K}_0$-regular model. The estimator ΠZ is a $\mathscr{K}_0$-GME iff Π projects on $\mathscr{E}_0$ along $\Sigma(\mathscr{F}^{\perp}) + \mathscr{K}_0^{\perp}$, where $\mathscr{E}_0$ is the orthogonal projection of $\mathscr{E}$ on $\mathscr{K}_0$.*

Proof. To begin with note that, by (2.2),

$$LL^{-1}(\mathscr{E}^{\perp}) = \mathscr{R}(L) \cap \mathscr{E}^{\perp} = \mathscr{F}^{\perp}. \tag{4.1}$$

If ΠZ is a $\mathscr{K}_0$-GME in $\mathscr{M}(\mathscr{E}, \Theta)$, then there exists a projection Π_L on $\mathscr{E}_L = \mathscr{E}_0$ along $\Sigma_L(\mathscr{E}_L^{\perp})$ such that $\Pi = \Pi_L L$. Since, by (4.1),

$$\Sigma_L(\mathscr{E}_L^{\perp}) = L\Sigma L\big(L(\mathscr{E})^{\perp}\big) = L\Sigma(\mathscr{F}^{\perp}),$$

$\Pi = \Pi_L L$ projects on $\mathscr{E}_L = \mathscr{E}_0$ along $\Sigma(\mathscr{F}^{\perp}) + \mathscr{K}_0^{\perp}$.

Now suppose that Π projects as indicated in the theorem. Then there exists an operator Π_L: $\mathscr{K}_0 \to \mathscr{K}_0$ such that $\Pi = \Pi_L L$. To prove that ΠZ is a $\mathscr{K}_0$-GME within $\mathscr{M}(\mathscr{E}, \Theta)$ it is sufficient to show that $\Pi_L Z_0$ is a GME within $\mathscr{M}(\mathscr{E}_L, \Theta_L)$. However, this follows from the assumption and from (4.1).

5. Representation of $\mathscr{K}_0$-best classes. It is often difficult to find an explicit formula of a projection operator as specified in Theorem 4.2. There is another way of calculating $\mathscr{K}_0$-best estimators, when a set of vectors which spans a $\mathscr{K}_0$-best class is given (each best estimator may be then represented as a certain linear combination of vectors from this set). For this reason we shall be concerned in this section with representations of $\mathscr{K}_0$-best classes in terms of some easily constructed operators. We shall establish them under the following assumptions:

$$\mathscr{R}(\Sigma)+\mathscr{K}_0^{\perp} = \mathscr{K}, \quad \mathscr{E} \subset \mathscr{R}(\Sigma) \quad \text{and} \quad \mathscr{K}_0^{\perp} \subset \mathscr{E} \subset \mathscr{R}(\Sigma).$$

Note that the first one is satisfied if $\mathscr{K}_0 \subset \mathscr{R}(\Sigma)$.

Now let L be defined as in Section 3, i.e. let $L\colon \mathscr{K} \to \mathscr{W}$ and $\mathscr{R}(L') = \mathscr{K}_0$.

THEOREM 5.1. *Let $\mathscr{M}(\mathscr{E}, \Theta)$ be $\mathscr{K}_0$-regular and let $\mathscr{R}(\Sigma)+\mathscr{K}_0^{\perp} = \mathscr{K}$. Then*

$$L'(L\Sigma L')^-(\mathscr{E}_L) \tag{5.1}$$

is a minimal $\mathscr{K}_0$-best class.

Proof. Since $L\Sigma L'$ is maximal in Θ_L, it follows from Theorem 3.1 that

$$\mathscr{H}_L = (L\Sigma L')^{-1}(\mathscr{E}_L)$$

is best for the induced model $\mathscr{M}(\mathscr{E}_L, \Theta_L)$. The corresponding $\mathscr{K}_0$-best class in the original model is then easily seen to be $L'(\mathscr{H}_L)$. To show that $L'(\mathscr{E}_L)$ coincides with (5.1) note first that $\mathscr{R}(\Sigma)+\mathscr{K}_0^{\perp} = \mathscr{K}$ yields $\mathscr{R}(L\Sigma L') = \mathscr{R}(L)$. In fact, since $L^{-1}\{L\mathscr{R}(\Sigma)\} = \mathscr{K}$, we obtain $\mathscr{R}(L) = L(\mathscr{R}(\Sigma))$, which is obviously equivalent to $\mathscr{R}(L\Sigma L') = \mathscr{R}(L)$. Next note that by applying (2.3) we get

$$\mathscr{H}_L = (L\Sigma L')^+(\mathscr{E}_L)+\mathscr{N}(L\Sigma L').$$

Using these results yields

$$L'(\mathscr{H}_L) = L'(L\Sigma L')^+(\mathscr{E}_L).$$

To conclude the proof note that, in view of (2.4), the Moore-Penrose inverse may be here replaced by any generalized inverse.

Another important theorem is the following one:

THEOREM 5.2. *Let $\mathscr{M}(\mathscr{E}, \Theta)$ be $\mathscr{K}_0$-regular and let $\mathscr{E} \subset \mathscr{R}(\Sigma)$. Let $\mathscr{E}_1$ be the projection of $\mathscr{E}$ on $\Sigma(\mathscr{K}_0)$ along $\mathscr{K}_0^{\perp}$. Then $\Sigma^+(\mathscr{E}_1)$ presents the orthogonal projection of the $\mathscr{K}_0$-best class $\mathscr{K}_0 \cap \Sigma^{-1}(\mathscr{F})$ on $\mathscr{R}(\Sigma)$.*

Proof. The assertion follows from the formula

$$\Sigma^+(\mathscr{E}_1) = \Sigma\Sigma^+(\mathscr{K}_0) \cap \Sigma^+\{\mathscr{F} \cap \mathscr{R}(\Sigma)\}, \tag{5.2}$$

which may be derived under the adopted assumption $\mathscr{E} \subset \mathscr{R}(\Sigma)$. Since the algebra is standard, we omit the details.

To construct a $\mathscr{K}_0$-best class in case $\mathscr{K}_0^{\perp} \subset \mathscr{E} \subset \mathscr{R}(\Sigma)$ we introduce some additional notation. Let Λ stand for an invertible and self-adjoint operator in $\mathscr{L}$ such that $\Lambda^{-1}A = \Sigma^+A$ for any $A \in \mathscr{R}(\Sigma)$. One of the possible choices of Λ is $\Sigma+I-\Sigma\Sigma^+$, where, as usual, I stands for the identity operator in $\mathscr{L}$. Dealing with an invertible Λ instead of with Σ simplifies the problem of constructing a $\mathscr{K}_0$-best class.

THEOREM 5.3. *Let $\mathscr{M}(\mathscr{E}, \Theta)$ be a $\mathscr{K}_0$-regular model and let $\mathscr{E}_2$ be the projection of $\mathscr{E}$ on $\Lambda(\mathscr{K}_0)$ along $\mathscr{K}_0^{\perp}$. If $\mathscr{K}_0^{\perp} \subset \mathscr{E} \subset \mathscr{R}(\Sigma)$, then $\Lambda^{-1}(\mathscr{E}_2)$ is a minimal $\mathscr{K}_0$-best class within the model $\mathscr{M}(\mathscr{E}, \Theta)$.*

Proof. It is sufficient to show that $\Lambda^{-1}(\mathscr{E}_2)$ satisfies (3.1), (3.2) and that $\Lambda^{-1}(\mathscr{E}_2) \cap \mathscr{E}^{\perp} = \{0\}$.

Since Λ is invertible, we have

$$\Lambda(\mathscr{K}_0) \oplus \mathscr{K}_0^{\perp} = \mathscr{K}.$$

Using $\mathscr{K}_0^{\perp} \subset \mathscr{E}$ yields

$$\mathscr{E} = \mathscr{E} \cap \Lambda(\mathscr{K}_0) \oplus \mathscr{K}_0^{\perp}.$$

Hence

$$\mathscr{E}_2 = \mathscr{E} \cap \Lambda(\mathscr{K}_0).$$

Consequently, using $\mathscr{E} \subset \mathscr{R}(\Sigma)$ leads to

$$\Lambda^{-1}(\mathscr{E}_2) = \mathscr{K}_0 \cap \Lambda^{-1}(\mathscr{E}) = \mathscr{K}_0 \cap \Sigma^{+}(\mathscr{E}) \subset \mathscr{K}_0 \cap \Sigma^{-1}(\mathscr{E}). \tag{5.3}$$

The set $\Lambda^{-1}(\mathscr{E}_2)$ satisfies condition (3.2) because the considered model is $\mathscr{K}_0$-regular. Now, since $\Lambda^{-1}(\mathscr{E}_2) \oplus \mathscr{E}_2^{\perp} = \mathscr{K}$ and since by (5.3) the relation $\Lambda^{-1}(\mathscr{E}_2) \subset \mathscr{K}_0$ holds, we note that

$$\Lambda^{-1}(\mathscr{E}_2) \oplus \mathscr{E}_2^{\perp} \cap \mathscr{K}_0 = \mathscr{K}_0.$$

This is equivalent to

$$\Lambda^{-1}(\mathscr{E}_2) \oplus \mathscr{E}^{\perp} = \mathscr{K}_0,$$

since $\mathscr{E} = \mathscr{E}_2 + \mathscr{K}_0^{\perp}$. Hence (3.1) is satisfied. Finally, (5.3) implies that $\Lambda^{-1}(\mathscr{E}_2) \cap \mathscr{E}^{\perp} = \{0\}$. Thus the proof of the theorem is completed.

To end this section we present three theorems which may be derived immediately from Theorems 5.1, 5.2 and 5.3.

THEOREM 5.4. *Let $\mathscr{R}(\Sigma) + \mathscr{K}_0^{\perp} = \mathscr{K}$. Then $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_0$-regular iff*

$$\Gamma L'(L\Sigma L')^{-}(\mathscr{E}_L) \subset \mathscr{E} \quad \textit{for all } \Gamma \in \Theta.$$

THEOREM 5.5. *Let $\mathscr{E} \subset \mathscr{R}(\Sigma)$. Then $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_0$-regular iff*

$$\Gamma \Sigma^{-}(\mathscr{E}_1) \subset \mathscr{F} \quad \textit{for all } \Gamma \in \Theta.$$

THEOREM 5.6. *Let $\mathscr{K}_0^{\perp} \subset \mathscr{E} \subset \mathscr{R}(\Sigma')$. Then $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_0$-regular iff*

$$\Gamma \Lambda^{-1}(\mathscr{E}_2) \subset \mathscr{E} \quad \textit{for all } \Gamma \in \Theta.$$

6. Two special cases.

1. Let y be a normal random vector of order $n \times 1$. Let $\mathrm{E}y = X\beta$ and let

$$\operatorname{Cov} y = \sum_{i=1}^{k} \sigma_i V_i.$$

The $(n \times p)$-matrix X and the $n \times n$ nonnegative-definite matrices $V_1, \ldots, V_k$ are assumed to be known. The $(p \times 1)$-vector β and the $(k \times 1)$-vector $\sigma = (\sigma_1, \ldots, \sigma_k)'$ are unknown. When the problem consists in the best estimation of $c'\sigma$, where $c \in \mathscr{R}^k$, attention is usually restricted either to all quadratic unbiased estimators or to quadratic unbiased estimators which are invariant with respect to all translations moving y into $y + X\beta$. Let $\mathscr{G}_1$ and $\mathscr{G}_2$ stand for the parametric functions $c'\sigma$ which have a quadratic unbiased estimator and an invariant quadratic unbiased estimator, respectively. Note that $\mathscr{G}_1 = \mathscr{G}_2$ if $PV_i = V_i P$ $(i = 1, \ldots, k)$, where P is the orthogonal projection on $\mathscr{R}(X)$.

Suppose that we want to characterize those models in which there exists a uniformly minimum variance quadratic unbiased estimator (UMVQUE) for each function in $\mathscr{G}_1$ and those ones in which there exists a uniformly minimum variance invariant quadratic unbiased estimator (UMVIQUE) for each function in $\mathscr{G}_2$. To apply the developed theory we specialize $\mathscr{K}$ to the space of all $n \times n$ symmetric matrices endowed with the usual trace inner product. Putting $Z = yy'$ we note that

$$\mathrm{E}Z = \operatorname{Cov} y + \mathrm{E}y(\mathrm{E}y)' \tag{6.1}$$

and

$$\operatorname{Cov} Z = 2[\mathrm{E}Z \otimes \mathrm{E}Z - \mathrm{E}y(\mathrm{E}y)' \otimes \mathrm{E}y(\mathrm{E}y)']. \tag{6.2}$$

Thus we are led to a model $\mathscr{M}(\mathscr{E}, \Theta)$, where $\mathscr{E}$ is the subspace spanned by (6.1) and Θ is the minimal convex cone containing all covariance matrices (6.2).

Now let S be maximal in the convex cone spanned by (6.1). Using routine algebra one may show that $\Sigma = S \otimes S$ is maximal in Θ and that $\mathscr{E} \subset \mathscr{R}(\Sigma)$.

To deduce the desired necessary and sufficient conditions we specify $\mathscr{K}_0$ as $\mathscr{K}_{01} = \mathscr{N}(P \otimes P)$ for $\mathscr{G}_1$ and as $\mathscr{K}_{02} = \mathscr{R}(M \otimes M)$ for $\mathscr{G}_2$. Here $P = XX^+$ and $M = I - P$, I being the unit matrix. Clearly, $\mathscr{K}_{02} \subset \mathscr{K}_{01}$. The operator

$$\Pi_1 = I \otimes I - G \otimes G, \quad \text{where } G = P(PS^- P)^- PS^+,$$

projects on $\Lambda(\mathscr{K}_{01})$ along $\mathscr{K}_{01}^{\perp}$, where $\Lambda = S_1 \otimes S_1$, $S_1 = S + I - SS^+$, while the operator

$$\Pi_2 = (I - G) \otimes (I - G)$$

projects on $\Sigma(\mathscr{K}_{02})$ along $\mathscr{K}_{02}^{\perp}$. Finally, let

$$\mathscr{V} = \operatorname{span}\{V_1, \ldots, V_k\}.$$

Since $\mathscr{K}_{01}^{\perp} \subset \mathscr{E} \subset \mathscr{R}(\Sigma)$, Theorem 5.6 implies that $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_{01}$-

regular iff

$$(AS_1^{-1} \otimes AS_1^{-1})\Pi_1(\mathscr{E}) \subset \mathscr{E} \quad \text{for all } A \in \mathscr{E},$$

where $\Pi_1(\mathscr{E}) = (I \otimes I - G \otimes G)(\mathscr{V})$.

The model $\mathscr{M}(\mathscr{E}, \Theta)$ being $\mathscr{K}_{01}$-regular, $y'Ay$ is a UMVQUE of its expected value if A can be represented in the form

$$A = (S^{-1} \otimes S^{-1})B \quad \text{for } B \in \Pi_1(\mathscr{E}).$$

In case $S = I$, this expression reduces to the well-known formula

$$A = (I \otimes I - P \otimes P)B \quad \text{for } B \in \mathscr{V}.$$

Moreover, since $\mathscr{E} \subset \mathscr{R}(\Sigma)$, it follows from Theorem 5.5 that the considered model is $\mathscr{K}_{02}$-regular iff

$$(AS_2^- \otimes AS_2^-)\Pi_2(\mathscr{E}) \subset \Pi_2(\mathscr{E})$$

for all $A \in \Pi_2(\mathscr{E})$, where $\Pi_2(\mathscr{E}) = \operatorname{span}\{(I-G)V_i(I-G')\colon i = 1, \ldots, k\}$ and $S_2 = (I-G)S(I-G')$.

As known [2] the above-given condition is satisfied iff $\Pi_2(\mathscr{E})$ is an S^+-quadratic subspace of $\mathscr{K}$, i.e. iff $AS^+A \in \Pi_2(\mathscr{E})$ for all $A \in \Pi_2(\mathscr{E})$. The model $\mathscr{M}(\mathscr{E}, \Theta)$ being $\mathscr{K}_{02}$-regular, it follows from Theorem 5.2 that $y'Ay$ is a UMVIQUE of its expected value iff

$$A \in \Sigma^+\Pi_2(\mathscr{E}) + \mathscr{N}(\Sigma) \cap \mathscr{K}_{02}. \tag{6.3}$$

In fact, Theorem 5.2 insures that

$$\mathscr{H} \subset \Sigma^+\Pi_2(\mathscr{E}) + \mathscr{N}(\Sigma), \tag{6.4}$$

where, as usual, $\mathscr{H} = \mathscr{K}_{02} \cap \Sigma^{-1}(\mathscr{F})$. Since $S^+(I-G) = MS^+(I-G)$, we easily see that

$$\Sigma^+\Pi_2(\mathscr{E}) \subset \mathscr{K}_{02}. \tag{6.5}$$

Hence

$$\mathscr{H} \subset \Sigma^+\Pi_2(\mathscr{E}) + \mathscr{N}(\Sigma) \cap \mathscr{K}_{02}. \tag{6.6}$$

Next, using $\mathscr{E} \subset \mathscr{R}(\Sigma)$, we have $\Pi_2(\mathscr{E}) \subset \mathscr{F}$. Hence

$$\Sigma^+\Pi_2(\mathscr{E}) \subset \Sigma^{-1}(\mathscr{F}). \tag{6.7}$$

Combining (6.5) and (6.7) leads to

$$\Sigma^+\Pi_2(\mathscr{E}) \subset \mathscr{H}. \tag{6.8}$$

The inclusion reverse to (6.6) follows from (6.8) because $\mathscr{N}(\Sigma) \subset \Sigma^{-1}(\mathscr{F})$. Thus the assertion is established.

If $\mathscr{R}(M) \subset \mathscr{R}(S)$, which is equivalent to $\mathscr{K}_{02} \subset \mathscr{R}(\Sigma)$, formula (6.3) reduces to

$$A \in \Sigma^{+} \Pi_2(\mathscr{E}). \tag{6.9}$$

Finally, note that under the assumption $\mathscr{R}(M) \subset \mathscr{R}(S)$ Theorem 5.1 gives the following formula alternative to (6.8):

$$A \in [(MV_0M) \otimes (MV_0M)]^{+}(\mathscr{V}), \tag{6.10}$$

where V_0 is maximal in the convex cone spanned by $\mathscr{V}$.

Formulas (6.9) and (6.10) have been derived first by Rao [15], [16] under slightly stronger assumptions.

2. Now consider a multivariate random variable $Z = \{y_{ij}\}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, p$. Assume that $\mathrm{E}Z = \{\alpha_i + \beta_j\}$, where α's and β's are unknown, and that $\operatorname{Cov} Z = I \otimes V$, where V is unknown and ranges over the set of all $p \times p$ covariance matrices. Note that

$$\mathscr{E} = \mathscr{R}(I \otimes \mathbb{1}) + \mathscr{R}(\mathbb{1} \otimes I), \quad \text{where } \mathbb{1} = (1, \ldots, 1)',$$

and that

$$\Theta = \{I \otimes V \colon V \geqslant 0\}.$$

Now we specify $\mathscr{K}$ as the space of all $(n \times p)$-matrices with the usual trace inner product, and restrict attention to estimators of the form $[A, Z]$, where $A \in \mathscr{K}$. As shown by Eaton [3], the model $\mathscr{M}(\mathscr{E}, \Theta)$ is not regular, i.e. not all parametric functions of the form $[A, \mathrm{E}Z]$ admit a best estimator. However, there exists a best estimator for each parametric function of the form $(\alpha'\mathbb{1})(b'\mathbb{1}) + nb'\beta$, where $\alpha = (\alpha_1, \ldots, \alpha_n)'$, $\beta = (\beta_1, \ldots, \beta_p)'$ and $b \in \mathscr{R}^p$, i.e. for each $\mathscr{K}_0$-estimable function, where $\mathscr{K}_0 = \mathscr{R}(\mathbb{1} \otimes I) + \mathscr{E}^{\perp}$.

Assuming $I \otimes I$ to be maximal in Θ and noting that

$$\mathscr{K}_0^{\perp} \subset \mathscr{E} \quad \text{and} \quad \mathscr{K}_0 \cap \Sigma^{-1}(\mathscr{F}) = \mathscr{R}(\mathbb{1} \otimes I),$$

we can write condition (3.3) in the form

$$(I \otimes V)\big(\mathscr{R}(\mathbb{1} \otimes I)\big) \subset \mathscr{R}(\mathbb{1} \otimes I) + \mathscr{R}(I \otimes \mathbb{1}).$$

Since this condition is satisfied for each $p \times p$ covariance matrix V, we conclude that $\mathscr{M}(\mathscr{E}, \Theta)$ is $\mathscr{K}_0$-regular, as asserted.

Acknowledgment. The authors are grateful to the referees for their comments which helped very much to improve this paper.

REFERENCES

[1] H. Drygas, *Gauss-Markov estimation for multivariate linear models with missing observations*, Ann. Statist. 4 (1976), p. 779-787.

[2] — *Best quadratic unbiased estimation in variance-covariance component models*, Math. Operationsforsch. Statist., Ser. Statist., 8 (1977), p. 211-231.

[3] M. L. Eaton, *Gauss-Markov estimation for multivariate linear models: a coordinate free approach*, Ann. Math. Statist. 41 (1970), p. 528-538.

[4] S. Gnot, W. Klonecki and R. Zmyślony, *Uniformly minimum variance unbiased estimation in Euclidean vector spaces*, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astronom. Phys., 24 (1976), p. 281-286.

[5] — *Uniformly minimum variance unbiased estimation in various classes of estimators*, Math. Operationsforsch. Statist., Ser. Statist., 8 (1977), p. 199-210.

[6] — *Best linear plus quadratic unbiased estimation for parameters in mixed linear models*, Zastos. Mat. 15 (1977), p. 455-462.

[7] — *Linear spaces and the theory of best linear unbiased estimation*, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astronom. Phys., 26 (1978), p. 69-72.

[8] J. Kleffe, *A note on MINQUE for normal models*, Math. Operationsforsch. Statist. 7 (1976), p. 707-714.

[9] — and R. Pincus, *Bayes and best quadratic unbiased estimators for parameters of the covariance matrix in a normal linear model*, ibidem 5 (1974), p. 43-67.

[10] W. Kruskal, *The coordinate-free approach to Gauss-Markov estimation, and its application to missing and extra observations*, Fourth Berkeley Symp. Math. Statist. Prob. 1 (1961), p. 435-451.

[11] — *When are Gauss-Markov and least squares estimators identical? A coordinate-free approach*, Ann. Math. Statist. 39 (1968), p. 70-75.

[12] L. R. LaMotte, *A canonical form for the general linear model*, Ann. Statist. 4 (1977), p. 787-789.

[13] E. L. Lehmann and H. Scheffé, *Completeness, similar regions and unbiased estimation*, Part 1, Sankhyā 10 (1950), p. 305-340.

[14] S. K. Mitra and B. J. Moore, *Gauss-Markov estimation with an incorrect dispersion matrix*, Part 2, Sankhyā, Ser. A, 35 (1973), p. 139-152.

[15] C. R. Rao, *Estimation of variance and covariance components — MINQUE theory*, J. Multivariate Anal. 1 (1971), p. 257-275.

[16] — *Minimum variance quadratic unbiased estimation of variance components*, ibidem 1 (1971), p. 445-456.

[17] J. Seely, *Linear spaces and unbiased estimation*, Ann. Math. Statist. 41 (1970), p. 1725-1734.

[18] — *Linear spaces and unbiased estimation — application to the mixed linear model*, ibidem 41 (1970), p. 1735-1748.

[19] — *Quadratic subspaces and completeness*, ibidem 42 (1971), p. 710-721.

[20] — and G. Zyskind, *Linear spaces and minimum variance unbiased estimation*, ibidem 42 (1971), p. 691-703.

[21] R. Zmyślony, *Estimation of variance components in random models*, Zastos. Mat. 13 (1973), p. 521-527.

[22] — *On estimation of parameters in linear models*, ibidem 15 (1976), p. 271-276.

[23] G. Zyskind, *On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models*, Ann. Math. Statist. 38 (1967), p. 1092-1109.

W. Klonecki and R. Zmyślony
Institute of Mathematics
Polish Academy of Sciences
ul. Kopernika 18
51-617 Wrocław, Poland

S. Gnot
Institute of Immunology
and Experimental Therapy
Polish Academy of Sciences
ul. Czerska 12
53-114 Wrocław, Poland

Received on 8. 5. 1979