ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XV, 3 (1976)

R. ZMYŚLONY (Wrocław)

# ON ESTIMATION OF PARAMETERS IN LINEAR MODELS

**1. Introduction.** In the paper we extend some known results on uniformly minimum variance unbiased linear estimation (shortly, UMVULE) and uniformly minimum variance unbiased quadratic estimation (shortly, UMVUQE) in random models under the assumption of normality.

Suppose that $y$ is a random normal $n$-vector with expectation $X\beta$ and covariance matrix $\sum_{i=1}^{m} \sigma_i^2 V_i$, where $\beta = (\beta_1, \ldots, \beta_p)$ and $\sigma = (\sigma_1^2, \ldots, \sigma_m^2)$ are unknown. It is assumed that for at least one $\sigma$ the covariance matrix of $y$ is positive-definite. Let $\mathscr{G}$ be the collection of all parametric functions of $\beta$ which have a linear unbiased estimator $a'y$, and let $\mathscr{H}$ be the collection of all parametric functions of $\sigma$ and $\beta$ which have a quadratic unbiased estimator $y'Ay$.

Theorem 1 gives necessary and sufficient conditions for each function in $\mathscr{G}$ to have a UMVULE. Theorem 2 gives necessary and sufficient conditions for each function in $\mathscr{H}$ to have a UMVUQE. Moreover, Corollary 3 gives necessary and sufficient conditions for each function in $\mathscr{G}$ to have a UMVULE and for each function in $\mathscr{H}$ to have a UMVUQE. Theorem 2 extends a theorem of Seely [3] and Theorem 2 of the author [5].

The proofs of Theorems 1 and 2 are based on the fundamental lemma on UMVUE of Lehmann and Scheffé. Corollary 3 is deduced from Theorems 1 and 2 by applying Lemma 1 which gives a decomposition of the space $\mathscr{B}$ of all $(n\times n)$-symmetric matrices into three subspaces two of which are quadratic subspaces of $\mathscr{B}$.

**2. Preliminaries.** Let $y$ be a normal random vector such that $y = X\beta + e$, where $X$ is a given $(n\times p)$-matrix, $\beta$ is a $p$-vector of unknown parameters, $e$ is a random vector with mean value 0 and covariance matrix

$$V(\sigma) = \sum_{i=1}^{m} \sigma_i V_i.$$

Here $V_i$ $(i = 1, \ldots, m)$ are given $(n\times n)$-symmetric matrices and $\sigma = (\sigma_1, \ldots, \sigma_m)'$, while the $\sigma$'s are unknown parameters. We assume that $\beta \in \Omega_1 = R^p$ and that $\sigma \in \Omega_2 \subset R^m$, where $\Omega_2$ contains a non-void open set in $R^m$. Moreover, we suppose that $V(\sigma)$ is positive-definite for some $\sigma \in \Omega_2$, and that $\beta$ and $\sigma$ are functionally independent so that the whole parameter space is $\Omega = \Omega_1 \times \Omega_2$. Let $\theta = [\beta, \sigma]$.

Throughout the paper, $R^k(\cdot, \cdot)$ denotes the $k$-dimensional Euclidean space with the usual inner product, and $\mathscr{B}\langle\cdot, \cdot\rangle$ the vector space of $(n \times n)$-symmetric matrices with the trace inner product. Finally, $A^-$ stands for the generalized inverse of the matrix $A$, and $R(A)$ for the space generated by columns of $A$.

We recall the following terminology introduced by Seely [1] and Seely and Zyskind [4]. Let $g(\theta)$ be a *parametric function*, i.e. a function from $\Omega$ into $R$.

Definition 1. A parametric function $g(\theta)$ is said to be *$\mathscr{A}$-estimable* if the set

$$\mathscr{A}_g = \{(a, y) \colon a \in R^n, \mathrm{E}_\theta(a, y) = g(\theta)\}$$

is non-empty.

Definition 2. A parametric function $g(\theta)$ is said to be *$\mathscr{B}$-estimable* if the set

$$\mathscr{B}_g = \{\langle B, yy'\rangle \colon B \in \mathscr{B}, \mathrm{E}_\theta\langle B, yy'\rangle = g(\theta)\}$$

is non-empty.

In the remaining we denote $(a, y)$ by $\bar{a}$, and $\langle B, yy'\rangle$ by $\bar{B}$.

Definition 3. An element $\bar{a} \in \mathscr{A}_g$ is said to be *$\mathscr{A}$-best* for a parametric function $g(\theta)$ if $\mathrm{Var}_\theta \bar{a} \leqslant \mathrm{Var}_\theta \bar{b}$ for every $\theta \in \Omega$ and every $\bar{b} \in \mathscr{A}_g$.

Definition 4. An element $\bar{B} \in \mathscr{B}_g$ is said to be *$\mathscr{B}$-best* for a parametric function $g(\theta)$ if $\mathrm{Var}_\theta \bar{B} \leqslant \mathrm{Var}_\theta \bar{C}$ for every $\theta \in \Omega$ and every $\bar{C} \in \mathscr{B}_g$.

**3. $\mathscr{A}$-best estimators.** As it is well known, $\mathscr{G} = \{(\lambda, \beta) \colon \lambda \in R(X')\}$ is the collection of $\mathscr{A}$-estimable functions. In the sequel we state necessary and sufficient conditions for every function in $\mathscr{G}$ to have an $\mathscr{A}$-best estimator. Moreover, the theorem gives an explicit form of an $\mathscr{A}$-best estimator for every function in $\mathscr{G}$.

THEOREM 1. *Suppose that $V_0 = V(\sigma_0)$ is positive-definite. Then for every function $g = (\lambda, \beta)$ in $\mathscr{G}$ the expression*

$$\big(\lambda, (X'V_0^{-1}X)^- X'V_0^{-1}y\big) \tag{1}$$

*represents an $\mathscr{A}$-best estimator if and only if*

$$V_i V_0^{-1} X(X'V_0^{-1}X)^- X' = X(X'V_0^{-1}X)^- X'V_0^{-1} V_i \quad (i = 1, \ldots, m). \tag{2}$$

*The estimator* (1) *does not depend on the choice of the generalized inverse matrix.*

Proof. Let $I$ be the $(n\times n)$-unit matrix. There exists a matrix $B$ such that $B'B = V_0^{-1}$ or, equivalently, $BV_0B' = I$. Then the expectation and covariance matrices of $z = By$ are

$$\mathrm{E}_\theta z = BX\beta \quad \text{and} \quad \mathrm{Var}_\theta z = W(\sigma) = \sum_{i=1}^{m} \sigma_i W_i,$$

respectively, where $W_i = BV_iB'$. Note that $\mathrm{Var}_\theta z = I$ for $\sigma = \sigma_0$. Now Corollary 5.2 in [1] states that for each parametric function in $\mathscr{G}$ there exists an $\mathscr{A}$-best estimator if and only if

$$W_i[R(BX)] \subset R(BX) \quad (i = 1, \ldots, m). \tag{3}$$

Since $W_i$ is a symmetric operator, formula (3) is equivalent to

$$PW_i = W_iP \quad (i = 1, \ldots, m), \tag{4}$$

where $P$ is the projection on $R(BX)$, i.e.

$$P = BX(X'V_0^{-1}X)^- X'B'.$$

Multiplying (4) from the left by $B^{-1}$ and from the right by $(B')^{-1}$ we get (2).

Formula (1) and the uniqueness of the estimator given by (1) follow from the fact that the minimum of the variance of (1) is attained at $\sigma = \sigma_0$ and from the assumption that $V(\sigma_0)$ is positive-definite.

From Theorem 1 the following conclusions can be easily deduced:

COROLLARY 1. *Suppose that $V(\sigma) = I$ for some $\sigma \in \Omega_2$. Then for every function in $\mathscr{G}$ there exists an $\mathscr{A}$-best estimator if and only if $PV_i = V_iP$ $(i = 1, \ldots, m)$, where $P$ is the projection on $R(X)$.*

COROLLARY 2. *Suppose that $V(\sigma) = I$ for some $\sigma \in \Omega_2$ and that $P = X(X'X)^-X'$ commutes with each $V_i$ $(i = 1, \ldots, m)$. Then the $\mathscr{A}$-best estimator for $(\lambda, \beta) \in \mathscr{G}$ is $(\lambda, \hat{\beta})$, where $\hat{\beta}$ is a solution of $X'X\beta = X'y$.*

**4. $\mathscr{B}$-best estimators.** In this section we state necessary and sufficient conditions for every $\mathscr{B}$-estimable function to have a $\mathscr{B}$-best estimator.

First we introduce some additional notation. Let $U = zz'$ and let $H = BX$. Moreover, let $H_\beta = H\beta\beta'H'$. Under the assumption of normality, the expectation and covariance operators of $U$ are

$$\mathrm{E}_\theta U = H_\beta + W(\sigma)$$

and

$$\begin{aligned} \mathrm{Cov}_\theta(\langle A, U\rangle, \langle B, U\rangle) &= \langle \Sigma_\theta A, B\rangle \\ &= 2\langle W(\sigma)AW(\sigma) + W(\sigma)AH_\beta + H_\beta AW(\sigma), B\rangle, \end{aligned} \tag{5}$$

respectively.

Remark 1. It may be worth-while to point out that the assumption of normality of $y$ is used only to obtain the formula for the covariance matrix of $U$. In other words, the stated results hold provided the covariance matrices of $U$ are of the form as above.

Defining $\mathscr{E} = \mathrm{sp}\{\mathrm{E}_\theta U\colon\ \theta \in \Omega\}$, we note that

$$\mathscr{E} = \mathrm{sp}\{H_\beta, W_1, \ldots, W_m, \ \beta \in \Omega_1\}.$$

Now, let $h_1, \ldots, h_p$ denote the columns of $H$, and let

$$H_{ii} = h_i h_i' \ (i = 1, \ldots, p) \quad \text{and} \quad H_{ij} = h_i h_j' + h_j h_i' \ (1 \leqslant i < j \leqslant p).$$

Then, by the lemma of Seely (see [2], Lemma 1), we have

$$\mathscr{E} = \mathrm{sp}\{H_{11}, H_{12}, \ldots, H_{pp}, W_1, \ldots, W_m\}. \tag{6}$$

THEOREM 2. *For each $\mathscr{B}$-estimable function there exists a $\mathscr{B}$-best estimator if and only if $\mathscr{E}$ is a quadratic subspace of $\mathscr{B}$.*

Proof. Let $\theta_0 = [0, \sigma_0]$. It follows from (5) that $\frac{1}{2}\Sigma_{\theta_0}$ is the identity operator. Consequently, in view of a well-known result of Seely (see Corollary 5.2 in [4]), for each $\mathscr{B}$-estimable function there exists a $\mathscr{B}$-best estimator if and only if $\mathscr{E}$ is an invariant subspace of $\Sigma_\theta$ for all $\theta \in \Omega$.

Now assume that $\mathscr{E}$ is a quadratic subspace of $\mathscr{B}$. Then by (5) we have

$$\Sigma_\theta(A) = 2\big(W(\sigma)AW(\sigma) + W(\sigma)AH_\beta + H_\beta AW(\sigma)\big). \tag{7}$$

Since $W(\sigma)$, $A$ and $H_\beta$ are elements of $\mathscr{E}$ and since $\mathscr{E}$ is a quadratic subspace, it follows from Lemma 4 in [5] that $\Sigma_\theta(A) \in \mathscr{E}$ for all $\theta \in \Omega$.

Now, let $\mathscr{E}$ be an invariant subspace of $\Sigma_\theta$ for all $\theta \in \Omega$. Putting $\beta = 0$ into (7) we obtain

$$W(\sigma)AW(\sigma) \in \mathscr{E} \quad \text{for } \sigma \in \Omega_2, \ A \in \mathscr{E}, \tag{8}$$

and

$$W(\sigma)AH_\beta + H_\beta AW(\sigma) \in \mathscr{E} \quad \text{for } \sigma \in \Omega_2, \ \beta \in \Omega_1, \ A \in \mathscr{E}. \tag{9}$$

On the other hand, substituting $I$ in place of $A$ into (8) and in place of $W(\sigma)$ into (9) we have

$$W(\sigma)W(\sigma) \in \mathscr{E} \quad \text{for } \sigma \in \Omega_2 \tag{10}$$

and

$$AH_\beta + H_\beta A \in \mathscr{E} \quad \text{for } \beta \in \Omega_1, \ A \in \mathscr{E}. \tag{11}$$

Since, by assumption, $\Omega_2$ contains a non-empty open set in $R^p$, by (10) we have $W_i W_j + W_j W_i \in \mathscr{E}$ for $i, j = 1, \ldots, m$. Now, in view of (11) and Lemma 1 of Seely in [3], we can conclude that $\mathscr{E}$ is a quadratic subspace of $\mathscr{B}$. This completes the proof of Theorem 2.

Remark 2. In the case $X = 0$ Theorem 2 has been proved previously by Seely [4], and in the case $X = (1, \ldots, 1)'$ by the author [5].

COROLLARY 3. *Suppose that* $y \sim N(X\beta, V(\sigma))$ *and* $\theta \in \Omega$. *Then for each* $\mathscr{A}$*-estimable function there exists an* $\mathscr{A}$*-best estimator and for each* $\mathscr{B}$*-estimable function there exists a* $\mathscr{B}$*-best estimator if and only if*

$$V_i P_0 = P_0' V_i \quad (i = 1, \ldots, m),$$

*and*

$$V_i V_0^{-1} M_0 V_j + V_j V_0^{-1} M_0 V_i \in \operatorname{sp}\{V_1 M_0, \ldots, V_m M_0\} \quad (i, j = 1, \ldots, m),$$

*where* $P_0' = X(X'V_0^{-1}X)^{-}X'V_0^{-1}$, *while* $M_0 = I - P_0$.

In order to prove Corollary 3 we need Lemma 1.

Let

$$\mathscr{B}_1 = \{PAP\colon A \in \mathscr{B}\}, \quad \mathscr{B}_2 = \{MAM\colon A \in \mathscr{B}\},$$
$$\mathscr{B}_3 = \{PAM + MAP\colon A \in \mathscr{B}\},$$

where $P = H(H'H)^{-}H'$, while $M = I - P$.

LEMMA 1. *The subspaces* $\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3$ *of the space* $\mathscr{B}$ *have the following properties*:

(a) $\mathscr{B}_1$ *and* $\mathscr{B}_2$ *are quadratic subspaces of* $\mathscr{B}$;

(b) $AB = BA = 0$ *if* $A \in \mathscr{B}_1$ *and* $B \in \mathscr{B}_2$;

(c) $\langle B_i, B_j \rangle = 0$ *if* $B_i \in \mathscr{B}_i$ *and* $B_j \in \mathscr{B}_j$ *for* $i = 1, 2, 3$, *and* $i \neq j$;

(d) $\mathscr{B}$ *is a direct sum of* $\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3$, *i.e.* $\mathscr{B} = \mathscr{B}_1 \oplus \mathscr{B}_2 \oplus \mathscr{B}_3$.

The verification of Lemma 1 is straightforward, and is omitted.

Proof of Corollary 3. First note that any subspace $\mathscr{D}$ of the space $\mathscr{B}$ can be decomposed into subspaces $\mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3$ such that

$$\mathscr{D} \subset \mathscr{D}_1 \oplus \mathscr{D}_2 \oplus \mathscr{D}_3, \quad \text{where } \mathscr{D}_1 \subset \mathscr{B}_1,\ \mathscr{D}_2 \subset \mathscr{B}_2,\ \mathscr{D}_3 \subset \mathscr{B}_3.$$

In particular, let $\mathscr{E}_1, \mathscr{E}_2, \mathscr{E}_3$ and $\mathscr{W}_1, \mathscr{W}_2, \mathscr{W}_3$ be such decompositions of $\mathscr{E}$ and $\mathscr{W} = \operatorname{sp}\{W_1, \ldots, W_m\}$, respectively. Using the fact that $\mathscr{B}_1 = \operatorname{sp}\{H_{11}, \ldots, H_{pp}\}$ and using (d) we can represent $\mathscr{E}$ in the form of $\mathscr{E} \subset \mathscr{B}_1 \oplus \mathscr{W}_2 \oplus \mathscr{W}_3$. Note that if $PW = WP$ for each $W \in \mathscr{W}$, then $\mathscr{W}_3 = \{0\}$ and $\mathscr{E} = \mathscr{B}_1 \oplus \mathscr{W}_2$. Now, in view of assertion (b) of Lemma 1, the relations $A \in \mathscr{B}_1$ and $W \in \mathscr{W}_2$ imply that $AW = WA = 0$. However, this shows that if one of the subspaces of $\mathscr{E}$ or $\mathscr{W}_2$ is quadratic, then so are both of them. In view of Theorems 1 and 2, this completes the proof of Corollary 3.

## References


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

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

[3] — *Quadratic subspaces and completeness*, ibidem 42 (1971), p. 710-721.
[4] — and G. Zyskind, *Linear spaces and minimum variance unbiased estimation*, ibidem 42 (1971), p. 691-703.
[5] R. Zmyślony, *Estimation of variance components in random models*, Zastosow. Matem. 13 (1973), p. 521-527.

MATHEMATICAL INSTITUTE
POLISH ACADEMY OF SCIENCES
51-617 WROCŁAW

*Received on 22. 8. 1975*

---

R. ZMYŚLONY (Wrocław)

## ESTYMACJA PARAMETRÓW W MODELACH LINIOWYCH

### STRESZCZENIE

Zakładamy, że $y$ jest wektorem losowym o rozkładzie normalnym o wartości oczekiwanej $X\beta$ i macierzy kowariancji $\sum_{i=1}^{m} \sigma_i V_i$. Wielkościami nie znanymi są

$$\beta = (\beta_1, \ldots, \beta_p) \in \Omega_1 = R^p \quad \text{oraz} \quad \sigma = (\sigma_1, \ldots, \sigma_m) \in \Omega_2 \subset R^m.$$

Zakładamy, że dla co najmniej jednego $\sigma \in \Omega_2$ macierz kowariancji wektora $y$ jest nieosobliwa. Niech $\mathcal{G}$ będzie klasą wszystkich funkcji parametrycznych $\beta$, które są liniowo estymowalne, $\mathcal{H}$ zaś klasą wszystkich funkcji parametrycznych $\sigma$ i $\beta$, które są kwadratowo estymowalne. Twierdzenie 1 podaje warunki konieczne i dostateczne na to, aby dla każdej funkcji w klasie $\mathcal{G}$ istniał jednostajnie najlepszy nieobciążony estymator liniowy. Twierdzenie 2 podaje warunki konieczne i dostateczne na to, aby dla każdej funkcji w klasie $\mathcal{H}$ istniał jednostajnie najlepszy nieobciążony estymator kwadratowy. Ponadto wniosek 3 podaje warunki konieczne i dostateczne na to, aby dla każdej funkcji w klasie $\mathcal{G}$ istniał jednostajnie najlepszy nieobciążony estymator liniowy oraz aby dla każdej funkcji w klasie $\mathcal{H}$ istniał jednostajnie najlepszy nieobciążony estymator kwadratowy.

---