ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIII, 4 (1973)

R. ZMYŚLONY (Wrocław)

ESTIMATION OF VARIANCE COMPONENTS IN RANDOM MODELS

1. Introduction. This paper is concerned with quadratic unbiased estimation of variance components in random models under the assumption of normality. This problem has been considered by others, e.g. by Graybill and Hultquist [1] and by Seely [3].

Graybill and Hultquist stated conditions under which the best unbiased estimates of variance components can be obtained from the analysis of variance. Introducing the notion of quadratic subspaces, Seely extended that result to mixed models. Moreover, Seely obtained necessary and sufficient conditions for the existence of best quadratic unbiased estimates for variance components in random models when the vector of observations has a zero mean.

The aim of this paper is to get similar results for random models without assuming that the vector of observations has a zero mean. The main result is stated in theorem 2. Also, an example is given showing that the assumptions of theorem 2 can be satisfied, and, at the same time, the assumptions of the other mentioned here theorems can be not.

2. Preliminaries and definitions. Throughout the paper $\mathscr{A}$, $\langle -, -\rangle$, denotes the finite-dimensional Hilbert space of symmetric matrices of order $n \times n$ with the trace inner product, and R^p, $(-, -)$, the p-dimensional space of real vectors with the usual inner product. The transpose of matrix X is denoted by X', the trace of X by $\mathrm{tr}(X)$, and the vector space generated by the column vectors of X by $M(X)$. Moreover, $M(X)^{\perp}$ stands for the space orthogonal to $M(X)$, and X^- for a generalized inverse of X. As usual, the notation $y \sim N(\mu, V)$ means that the random vector y has a multivariate normal distriution with vector of means μ and covariance matrix V.

In this paper the model

$$y = \sum_{i=1}^{k} X_i b_i$$

is considered. The X's stand for known matrices of n rows, and the b's except of b_1, for random vectors. In addition, $X_1 = \mathbf{1} = (1, \ldots, 1)'$ and $X_k = I$, I being the unit matrix. For $i = 2, \ldots, k$, it is assumed that $\mathrm{E}(b_i) = 0$ and $\mathrm{E}(b_i b_i') = \sigma_i^2 V_i$, where the matrix V_i is given and σ_i^2 is a non-negative parameter. Moreover, it is assumed that $V_k = I$ and $\mathrm{E}(b_i b_j') = 0$ for $i \neq j$; $i, j = 2, \ldots, k$. For convenience, b_1 is denoted throughout the paper by σ_1.

Models with the described structure are called *random*.

Let Ω be a subset of R^k defined as follows:

$$\Omega = \{\sigma\colon \sigma' = (\sigma_1^2, \ldots, \sigma_k^2),\ \sigma_i^2 \geqslant 0,\ i = 1, \ldots, k\}.$$

Definition 1. A linear function h of $\sigma_1^2, \ldots, \sigma_k^2$ is called a *parametric function.*

Introducing the notation $c' = (c_1, \ldots, c_k)$, it can be written as $h = c'\sigma$.

Definition 2. A parametric function $h = c'\sigma$ is called *estimable* if there exists a matrix $D \in \mathscr{A}$ such that $\mathrm{E}(y'Dy) = h$ for all $\sigma \in \Omega$.

Definition 3. The function $y'Dy$, where $D \in \mathscr{A}$, is called a *quadratic unbiased estimator* of $h = c'\sigma$ if $\mathrm{E}(y'Dy) = h$ for all $\sigma \in \Omega$.

3. The least squares estimators. Similarly as in [1], let the mapping $S\colon \mathscr{A} \to R^m$, where $m = n(n+1)/2$, be defined as

$$S(A) = (a_{11}, \sqrt{2}a_{12}, \ldots, \sqrt{2}a_{1n}, a_{22}, \sqrt{2}a_{23}, \ldots, \sqrt{2}a_{2n}, \ldots, a_{nn});$$

while $A = [a_{ij}] \in \mathscr{A}$.

Clearly, S is an isomorphism of $\mathscr{A}$, $\langle -, - \rangle$, onto R^m, $(-, -)$, since, for $B_i \in \mathscr{A}$ and $c_i \in R$, $i = 1, \ldots, p$,

$$S\Big(\sum_{i=1}^{p} c_i B_i\Big) = \sum_{i=1}^{p} c_i S(B_i), \tag{1}$$

and, for $A, B \in \mathscr{A}$,

$$\langle A, B \rangle = \mathrm{tr}(AB) = \big(S(A), S(B)\big). \tag{2}$$

Introducing the notation $A_1 = \mathbf{1}\mathbf{1}'$, $A_i = X_i V_i X_i'$, $i = 2, \ldots, k$,

$$Wx = \sum_{i=2}^{k} x_i A_i \quad \text{and} \quad W_1 x = \sum_{i=1}^{k} x_i A_i,$$

where $x = (x_1, \ldots, x_k)' \in R^k$, the covariance matrix V of y and the expected value of yy' can be represented in the forms $V = W\sigma$ and $\mathrm{E}(yy') = W_1\sigma$, $\sigma \in \Omega$, respectively.

Finally, introducing the notation $S(yy') = z$, $S(A_i) = a_i$, $i = 1, \ldots, k$, $A = [a_1, \ldots, a_k]$ and using (1), it is easily seen that, for $\sigma \in \Omega$,

$$\mathrm{E}(z) = \sum_{i=1}^{k} \sigma_i^2 S(A_i) = \sum_{i=1}^{k} \sigma_i^2 a_i$$

or, equivalently,

$$\mathrm{E}(z) = A\sigma. \tag{3}$$

If $y'Dy$ is a quadratic unbiased estimator of $h = c'\sigma$ and if $S(D) = d$, then (2) implies that

$$\begin{aligned}\mathrm{E}(y'Dy) &= \mathrm{E}\big(\mathrm{tr}(Dyy')\big) = \mathrm{E}(\langle D, yy'\rangle) = \langle D, \mathrm{E}(yy')\rangle \\ &= \big(d, \mathrm{E}(z)\big) = \mathrm{E}(d'z).\end{aligned}$$

This establishes the following result:

LEMMA 1. *The parametric function $h = c'\sigma$ is estimable if and only if there exists a vector $d \in R^m$ such that $\mathrm{E}(d'z) = h$ for all $\sigma \in \Omega$.*

Clearly, in view of (3), the parametric function $h = c'\sigma$ is estimable if and only if $c \in M(T)$, where $T = A'A$. Also, $d'z$ is an unbiased estimator of $h = c'\sigma$ if and only if $d = f+g$, where $f' = c'T^{-}A'$, while $g \in M(A)^{\perp}$.

Because of the isomorphism described, the following lemma can be stated:

LEMMA 2. *If $h = c'\sigma$ is an estimable parametric function, then $y'Dy$ is an unbiased estimator of h if and only if $D = F+G$, where $F = W_1^{-}(T^{-}c)$ and $G \in \mathscr{A}_1^{\perp}$, while $\mathscr{A}_1 \subset \mathscr{A}$ is the linear subspace generated by $A_1, \ldots, A_k$.*

The estimator $y'Fy$ is called the *least squares estimator* (LSE – for short) of h.

4. The main results. In this section it is assumed that the random vectors $b_2, \ldots, b_k$ have multivariate normal distributions. Then, in the framework of the assumptions considered, $y \sim N(\sigma_1 \mathbf{1}, W\sigma)$, where $\sigma \in \Omega$. Also, there hold the relations

$$\mathrm{Var}(y'Py) = 2\,\mathrm{tr}(PW\sigma PW\sigma) + 4\sigma_1^2 \mathrm{tr}(PW\sigma PA_1),$$

where $\sigma \in \Omega$ and $P \in \mathscr{A}$, and

$$\mathrm{Cov}(y'Py, y'Ry) = 2\,\mathrm{tr}(PW\sigma RW\sigma) + 4\sigma_1^2 \mathrm{tr}(PW\sigma RA_1), \tag{4}$$

where $\sigma \in \Omega$ and $P, R \in \mathscr{A}$.

Making use of $\mathrm{tr}(AB) = \mathrm{tr}(BA) = \langle A, B\rangle$ and of $\mathrm{tr}(aA+bB) = a\,\mathrm{tr}(A) + b\,\mathrm{tr}(B)$ as well as of the fact that $\langle A, B\rangle = \langle A+A', B\rangle/2$ holds for each $(n \times n)$-matrix A and each $B \in \mathscr{A}$, formula (4) can be written in the form

(5) $$\mathrm{Cov}(y'Py, y'Ry) = 2\langle W\sigma PW\sigma + \sigma_1^2(A_1PW\sigma + W\sigma PA_1), R\rangle,$$

where $\sigma \in \Omega$.

In the remaining parts of the paper the notation of quadratic subspaces introduced by Seely [3] is essential. In the following definition of a quadratic subspace as well as some properties associated with these subspaces are given:

Definition 4 (Seely [3]). A subspace $\mathscr{B}$ of $\mathscr{A}$ with the property that $B \in \mathscr{B}$ implies $B^2 \in \mathscr{B}$ is called a *quadratic subspace* of $\mathscr{A}$.

LEMMA 3 (Seely [3]). *Let $\mathscr{B}$ be a subspace of $\mathscr{A}$ and let $\mathscr{B}_1$ be an arbitrary spanning set for $\mathscr{B}$. Then $\mathscr{B}$ is a quadratic subspace of $\mathscr{A}$ if and only if*

(6) $$A, B \in \mathscr{B}_1 \Rightarrow AB + BA \in \mathscr{B}.$$

LEMMA 4. *Let $\mathscr{B}$ be a quadratic subspace of $\mathscr{A}$. Then*

(7) $$A, B \in \mathscr{B} \Rightarrow ABA \in \mathscr{B}$$

and

$$A, B, C \in \mathscr{B} \Rightarrow ABC + CBA \in \mathscr{B}.$$

Proof. The first implication is due to Seely and can be found in [3]. The proof of the second implication is as follows:

Let $A, B, C \in \mathscr{B}$. Then $A + C \in \mathscr{B}$. From (7) it follows that $ABA + CBC \in \mathscr{B}$ and that $(A+C)B(A+C) \in \mathscr{B}$. Since $\mathscr{B}$ is a subspace, it follows from the above-mentioned that $ABC + CBA$ is a member of $\mathscr{B}$.

Let h be an estimable parametric function.

Definition 5. A quadratic unbiased estimator $y'D_0y$ of h is called *uniformly best* if $\mathrm{Var}(y'D_0y) \leqslant \mathrm{Var}(y'Dy)$ for all D such that $\mathrm{E}(y'Dy) = h$ and all $\sigma \in \Omega$.

Theorems 1 and 2 given in the sequel provide some properties of the LSE.

THEOREM 1. *Let $y \sim N(\sigma_1 \mathbf{1}, W\sigma)$, where $\sigma \in \Omega$, and let $h = c'\sigma$ be an estimable parametric function. If there exists a uniformly best quadratic unbiased estimator of h, then this estimator is the LSE.*

Proof. Let $y'Dy$ be a uniformly best quadratic unbiased estimator of h and let $y'Ry$ be any estimator of zero.

From lemma 2 it follows that $y'Ry$ is an estimator of zero if and only if $R \in \mathscr{A}_1^{\perp}$. Now, the Lehmann and Scheffé theorem 5.3 in [2] can be applied to infer that $\mathrm{Cov}(y'Dy, y'Ry) = 0$ for $R \in \mathscr{A}_1^{\perp}$ and $\sigma \in \Omega$. Thus, by (5),

$$\langle W\sigma DW\sigma + \sigma_1^2(A_1DW\sigma + W\sigma DA_1), R\rangle = 0$$

$$\text{for all } R \in \mathscr{A}_2^{\perp} \text{ and } \sigma \in \Omega.$$

Note that $W\sigma DW\sigma + \sigma_1^2(A_1DW\sigma + W\sigma DA_1)$ is a symmetric matrix and, therefore,

$$W\sigma DW\sigma + \sigma_1^2(A_1DW\sigma + W\sigma DA_1) \in \mathscr{A}_1 \quad \text{for all } \sigma \in \Omega.$$

In particular, putting $\sigma_1^2 = 0$ and selecting σ so that $W\sigma = I$ it is easily seen that $D \in \mathscr{A}_1$. Now, lemma 2 implies that $D = F$, i.e., $y'Dy$ is the LSE of h.

THEOREM 2. *Suppose that $y \sim N(\sigma_1 \mathbf{1}, W\sigma)$, where $\sigma \in \Omega$. For each estimable parametric function there exists a uniformly best quadratic unbiased estimator if and only if $\mathscr{A}_1$ is a quadratic subspace of $\mathscr{A}$.*

Proof. Necessity. Let $\mathscr{H}$ be the set of all estimable parametric functions. From lemma 2 it is clear that the mapping $\varphi\colon \mathscr{H} \to \mathscr{A}_1$ defined by $\varphi(h) = W_1(T^- c)$ is one to one.

Thus, from theorem 1 it follows that

$$W\sigma DW\sigma + \sigma_1^2(A_1DW\sigma + W\sigma DA_1) \in \mathscr{A}_1 \quad \text{for all } D \in \mathscr{A}_1 \text{ and } \sigma \in \Omega. \tag{8}$$

Substituting $\sigma_1^2 = 0$ in (8) leads to

$$W\sigma DW\sigma \in \mathscr{A}_1 \quad \text{for all } \sigma \in \Omega \text{ and } D \in \mathscr{A}_1. \tag{9}$$

But $\mathscr{A}_1$ is a subspace of $\mathscr{A}$ and, therefore, for all $\sigma \in \Omega$ and $D \in \mathscr{A}_1$,

$$\sigma_1^2(A_1DW\sigma + W\sigma DA_1) \in \mathscr{A}_1. \tag{10}$$

In particular, putting $\sigma_1^2 = 1$, $D = I$ and $W\sigma = A_i + I$ for $i = 2, \ldots, k-1$ in (9) and (10), it is easily seen that, for $i = 2, \ldots, k-1$,

$$A_i^2 \in \mathscr{A}_1 \tag{11}$$

and

$$A_iA_1 + A_1A_i \in \mathscr{A}_1. \tag{12}$$

On the other hand, taking $D = I$ and $W\sigma = A_i + A_j + I$, $i, j = 2, \ldots, k-1$, formula (9) leads to

$$A_iA_j + A_jA_i \in \mathscr{A}_1, \quad i, j = 2, \ldots, k-1;\ i \neq j. \tag{13}$$

Moreover, from the definitions of A_1 and A_k it is easily seen that $A_1^2 = nA_1 \in \mathscr{A}_1$ and $A_k^2 = I \in \mathscr{A}_1$. Now, (11), (12), (13) and (6) imply that $\mathscr{A}_1$ is a quadratic subspace of $\mathscr{A}$.

Sufficiency. Let $\mathscr{A}_1$ be a quadratic subspace of $\mathscr{A}$, $h \in \mathscr{H}$ and let $y'Dy$ be the LSE of h. Then (7) implies that $W\sigma DW\sigma \in \mathscr{A}_1$. On the other hand, from lemma 4 it follows that $A_1DW\sigma + W\sigma DA_1 \in \mathscr{A}_1$. Thus, it is clear that

$$2\langle W\sigma DW\sigma + \sigma_1^2(A_1DW\sigma + W\sigma DA_1), R\rangle = \mathrm{Cov}(y'Dy, y'Ry) = 0$$

holds for all $R \in \mathscr{A}_1^{\perp}$ and $\sigma \in \Omega$. Consequently, from theorem 5.3 in [2], it follows that $y'Dy$ is the uniformly best quadratic unbiased estimator of h. This completes the proof of theorem 2.

Example. Let $y = (y_1, \ldots, y_n)' \sim N(\sigma_1 \mathbf{1}, \sigma_2^2 A_2 + \sigma_3^2 A_3 + \sigma_4^2 I)$, where

$$A_2 = \begin{bmatrix} \underset{k\times k}{\mathbf{11}'} & \underset{k\times l}{0} \\ \underset{l\times k}{0} & \underset{l\times l}{0} \end{bmatrix}, \quad A_3 = \begin{bmatrix} \underset{k\times k}{0} & \underset{k\times l}{0} \\ \underset{l\times k}{0} & \underset{l\times l}{\mathbf{11}'} \end{bmatrix},$$

while $k+l = n$.

In this case it is easily seen that

$$A_2 A_3 = A_3 A_2 = 0 \in \mathscr{A}_1,$$

$$A_1 A_2 + A_2 A_1 = k(A_1 + A_2 - A_3) \in \mathscr{A}_1,$$

$$A_1 A_3 + A_3 A_1 = l(A_1 - A_2 + A_3) \in \mathscr{A}_1,$$

$$A_2^2 = k A_2 \in \mathscr{A}_1, \quad A_3^2 = l A_3^2 \in \mathscr{A}_1.$$

Lemma 3 implies that $\mathscr{A}_1$ is a quadratic subspace of $\mathscr{A}$.

Moreover, since A_1, A_2, A_3, I are lineary independent, it follows from theorem 2 that, for each parametric function $h = c_1\sigma_1^2 + c_2\sigma_2^2 + c_3\sigma_3^2 + c_4\sigma_4^2$, where $c_1, c_2, c_3, c_4 \in R$, there exists a uniformly best quadratic unbiased estimator.

Acknowledgement. I wish to thank Dr. W. Klonecki for his helpful discussions.

References

[1] F. A. Graybill and R. A. Hultquist, *Theorems concerning Eisenhart's model II*, Ann. Math. Statist. 32 (1961), p. 261-269.

[2] E. L. Lehmann and H. Scheffé, *Completeness, similar regions and unbiased estimation*, Sankhya 10 (1950), p. 305-314.

[3] J. Seely, *Quadratic subspaces and completeness*, Ann. Math. Statist. 42 (1971), p. 710-721.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WROCŁAW

Received on 11. 1. 1973

R. ZMYŚLONY (Wrocław)

ESTYMACJA KOMPONENTÓW WARIANCYJNYCH W MODELACH LOSOWYCH

STRESZCZENIE

W pracy rozważa się problem estymacji funkcji parametrycznych o postaci $c_1 b_1^2 + c_2 \sigma_2^2 + \ldots + c_k \sigma_k^2$, gdzie b_1 jest wartością oczekiwaną, wspólną dla wszystkich składowych wektora losowego y, natomiast $\sigma_2^2, \ldots, \sigma_k^2$ są komponentami wariancyjnymi. Funkcja parametryczna nazywa się *estymowalną*, jeśli istnieje dla niej nieobciążony estymator o postaci $y'Ay$, zwany *nieobciążonym kwadratowym estymatorem*.

Przy założeniu, że wektor losowy ma rozkład normalny, podano warunek konieczny i dostateczny na to, aby dla każdej estymowalnej funkcji parametrycznej istniał estymator o jednostajnie najmniejszej wariancji w klasie kwadratowych nieobciążonych estymatorów.