KYBERNETIKA — VOLUME *50* (2014), NUMBER 3, PAGES 310–321

ADMISSIBLE INVARIANT ESTIMATORS IN A LINEAR MODEL

Czesław Stępniak

Let $\mathbf{y}$ be observation vector in the usual linear model with expectation $\mathbf{A}\beta$ and covariance matrix known up to a multiplicative scalar, possibly singular. A linear statistic $\mathbf{a}^T\mathbf{y}$ is called invariant estimator for a parametric function $\phi = \mathbf{c}^T\beta$ if its MSE depends on β only through ϕ. It is shown that $\mathbf{a}^T\mathbf{y}$ is admissible invariant for ϕ, if and only if, it is a BLUE of ϕ, in the case when ϕ is estimable with zero variance, and it is of the form $k\widehat{\phi}$, where $k \in \langle 0, 1\rangle$ and $\widehat{\phi}$ is an arbitrary BLUE, otherwise. This result is used in the one- and two-way ANOVA models. Our paper is self-contained and accessible, also for non-specialists.

Keywords: linear estimator, invariant estimator, admissibility, one-way/two-way ANOVA

Classification: 62J05, 62C05

1. BACKGROUND

From real point of view, any observation vector $\mathbf{y}$ in a statistical experiment (or model) may be perceived as a deformation of a signal β. This deformation involves a deterministic component, say $\mathbf{f}(\beta)$, and a random one, say $\mathbf{e}$, treated as a noise or error. In a linear model this deterministic component takes the form of linear transformation, say $\mathbf{A}\beta$. Usually a linear form $\mathbf{c}^T\beta$, of the signal, is estimated by a linear form $\mathbf{d}^T\mathbf{y}$, of the observation.

A class of potential estimators is ordered by a preference rule and it may be reduced by some initial conditions. The preference rule is usually induced by the Mean Squared Error while the most popular initial conditions include unbiasedeness, invariance and admissibility.

Earlier results on admissibility in the context of linear model relate to the unbiased estimators. Cohen ([4, 5]), Rao [19] and Stępniak [21] went beyond this classical framework. Further results in this area were given, among others, by LaMotte [14], Klonecki [11], Klonecki and Zontek [12], Baksalary and Markiewicz ([1, 2, 3]), Groß and Markiewicz [8] and Stępniak [24]. They were based on the Loewner order of nonnegative definite matrices (see, e. g., Stępniak [22], or Groß ([6, 7])).

Subsequent works by Stępniak [23], Zontek [29], LaMotte [15] and Synówka-Bejenka and Zontek [28] introduced a new tool in the problem of admissibility. It is based on the

DOI: 10.14736/kyb-2014-3-0310

limits of the unique locally best linear estimators. In this way one can use additional information on the location of the unknown parameters.

With no doubt the fundamental results on admissibility in linear estimation were obtained by Cohen ([4] and [5]). The first work presents a geometric characterization of the set of all admissible estimators for a scalar parametric function in a simple linear model with nonsingular covariance matrix. This result has been modified in Rao [19] and Stępniak [21] for the general linear model. Its application is, however, limited by the fact that not all admissible estimators have good global properties. Thus we restrict ourselves to such estimators whose Mean Squared Error is invariant with respect to the nuisance parameters.

This paper completes results in this area. Auxiliary results are collected in Sections 2, which may serve as a good introduction to linear models. The main result, presented in Section 3, is used to the classical one- and two-way balanced, fixed, ANOVA models with restraints on the parameters. Such restraints lead to singular covariance matrix.

The paper is self-contained and accessible, also for non-specialists..

2. DEFINITIONS, NOTATION AND KNOWN RESULTS

For clarity, the symbols of matrices are emphasized by boldface; capital for many columns and lower-case for a single one, i. e. vector. The set of all n-vectors is denoted by R^n. For deterministic vectors we use letters $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\mathbf{g}$ and $\mathbf{x}$ (with possible indices) and for random ones: $\mathbf{y}$, $\mathbf{u}$, $\mathbf{z}$ and $\mathbf{e}$. Greek letters stand for parameters.

Given a matrix $\mathbf{M}$, of size $n \times p$, its range and kernel is defined, respectively, by

$$\mathcal{R}(\mathbf{M}) = \{\mathbf{a} \in R^n : \mathbf{a} = \mathbf{M}\mathbf{c} \text{ for some } \mathbf{c} \in R^p\}$$

and

$$\mathcal{N}(\mathbf{M}) = \{\mathbf{c} \in R^p : \mathbf{M}\mathbf{c} = \mathbf{0}\}.$$

It is well known from a course in linear algebra (see, for instance, Halmos [9]) that any vector $\mathbf{a} \in R^n$ may be presented in the form

$$\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2, \tag{1}$$

where $\mathbf{a}_1 \in \mathcal{R}(\mathbf{M})$, $\mathbf{a}_2 \in \mathcal{N}(\mathbf{M}^T)$ and $\mathbf{a}_1$ and $\mathbf{a}_2$ are orthogonal in the sense $\mathbf{a}_1^T\mathbf{a}_2 = 0$. Moreover,

$$\mathcal{R}(\mathbf{M}\mathbf{M}^T) = \mathcal{R}(\mathbf{M}). \tag{2}$$

Let $\mathbf{P} = \mathbf{P}_\mathbf{M}$ be a square matrix satisfying the condition

$$\mathbf{P}\mathbf{a} = \begin{cases} \mathbf{a}, & \text{if} \quad \mathbf{a} \in \mathcal{R}(\mathbf{M}) \\ \mathbf{0} & \text{if} \quad \mathbf{a} \in \mathcal{N}(\mathbf{M}^T). \end{cases} \tag{3}$$

By (1) such a matrix is unique. This matrix is called the orthogonal projector from R^n onto $\mathcal{R}(\mathbf{M})$.

Note that $\mathbf{P}$ depends on $\mathbf{M}$ only through $\mathcal{R}(\mathbf{M})$. Let $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r$ be a basis in $\mathcal{R}(\mathbf{M})$. It is easy to verify that

$$\mathbf{P}_\mathbf{B} = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T, \text{ where } \mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r]. \tag{4}$$

Let $\mathbf{y}$ be a random vector in R^n with expectation $\mathbf{A}\beta$ and covariance matrix $\sigma^2\mathbf{V}$, where $\mathbf{A}$ is a known matrix of size $n \times p$, $\mathbf{V}$ is a known symmetric nonnegative definite matrix of order n, while $\beta = (\beta_1, \ldots, \beta_p)^T$ is an unknown parameter and σ^2 is a positive scalar (known or not). Formally, we assume that (β,σ^2) is running over the Cartesian product $R^p \times S$, where S is a not empty subset of $(0, \infty)$. Such a system of conditions will be called a linear model and will be symbolically written as $\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$ (Notation $[\cdot]$ is adapted from [17].) The fact that $\mathbf{y}$ is observation vector in the model will be shortly expressed in the form $\mathbf{y} \sim\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$.

In this paper we are interested in estimation of a parametric function $\varphi(\beta) = \mathbf{c}^T\beta$, for a given $\mathbf{c} \in R^p$, by linear statistic of the form $\mathbf{d}^T\mathbf{y}$. An estimator $\mathbf{d}^T\mathbf{y}$ is said to be unbiased for φ if $E(\mathbf{d}^T\mathbf{y}) = \varphi(\beta)$ for all $\beta \in R^p$. If such estimator exists then the parametric function φ is said to be estimable. In the context of the model $\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$ we have $E(\mathbf{d}^T\mathbf{y}) = \mathbf{d}^T\mathbf{A}\beta$. Thus, by definition of range, $\varphi = \mathbf{c}^T\beta$ is estimable, if and only if, $\mathbf{c} \in \mathcal{R}(\mathbf{A}^T)$. In consequence, any unbiased estimator of $\mathbf{c}^T\beta$ may be presented in the form $\mathbf{d}^T\mathbf{y}$, where $\mathbf{d}$ is any solution of the equation $\mathbf{A}^T\mathbf{x} = \mathbf{c}$.

It is well known that for any estimable φ in $\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$ there exists a linear unbiased estimator with minimal variance, called a Best Linear Unbiased Estimator (BLUE), say $\mathbf{b}^T\mathbf{y}$. This estimator is characterized by the following theorem.

Theorem 1. (Lehmann-Scheffé [16], Theorem 5.3; cf. also [27], Theorem 1)
Let $\mathfrak{D} = \{f(\mathbf{y})\}$ be a class of potential estimators. If $\mathfrak{D}$ constitutes a linear space then a member $f_1(\mathbf{y})$ of this class is a minimum variance unbiased estimator of its expectation in $\mathfrak{D}$, if and only if, $Cov(f_1(\mathbf{y}), f_0(\mathbf{y})) = 0$ for any $f_0 \in \mathfrak{D}$ such that $E(f_0(\mathbf{x})) = 0$ for all β.

Thus, as a corollary, we get

Theorem 2. (Zyskind [30], Theorem 3)
In the context of the model $\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$ a linear functional $\mathbf{d}'\mathbf{y}$ is a *BLUE* of its expectation, if and only if, $\mathbf{V}\mathbf{d} \in \mathcal{R}(\mathbf{A})$.

In particular, if $\mathbf{V} = \mathbf{I}_n$, then $\mathbf{d}^T\mathbf{y}$ is a BLUE (of its expectation $\varphi = \mathbf{d}^T\mathbf{A}\beta$), if and only if, $\mathbf{d} \in R(\mathbf{A})$. Such estimator is called the the Least Squares Estimator (LSE) for φ.

The question "When does the BLUE coincide with the LSE for all estimable parametric functions?" is very important and it is still undertaken by many authors. The first simple answer to this question is due to Kruskal ([13], Theorems 1 and 3). Caution: There are two theorems with the number 3 in [13].

Theorem 3. (Kruskal [13], Theorem 1 and 3 (p. 74))
Model $\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$ satisfies the desired condition if an only if

$$\mathbf{V}\mathcal{R}(\mathbf{A}) \subseteq \mathcal{R}(\mathbf{A}), \tag{5}$$

i.e. when $\mathcal{R}(\mathbf{A})$ is an invariant subspace of $\mathbf{V}$ (treated as an operator on R^n).

This theorem is a direct consequence of Theorem 2 and the fact that the set of all LSE's in the model $\mathcal{L}(\mathbf{A}\beta,[\mathbf{V}])$ is of the form $\{\mathbf{d}^T\mathbf{y} : \mathbf{d} \in \mathcal{R}(\mathbf{A})\}$.

We will prove

Lemma 4. The condition (5) may be expressed in the matrix form

$$\mathbf{V}\mathbf{P}_\mathbf{A} = \mathbf{P}_\mathbf{A}\mathbf{V}. \tag{6}$$

Proof. Let us rewrite the left side of (5) in the form

$$\mathbf{V}\mathcal{R}(\mathbf{A}) = \{\mathbf{a} \in R^n : \mathbf{a} = \mathbf{V}\mathbf{A}\mathbf{c} \ \text{ for some } \mathbf{c} \in R^p\},$$

and next, by (3), as $\{\mathbf{V}\mathbf{P}_\mathbf{A}\mathbf{b} : \mathbf{b} \in R^n\}$. Thus (5) is equivalent to $(\mathbf{I} - \mathbf{P}_\mathbf{A})\mathbf{V}\mathbf{P}_\mathbf{A}\mathbf{b} = \mathbf{0}$ for all $\mathbf{b} \in R^n$. The last one is satisfied, if and only if, $\mathbf{V}\mathbf{P}_\mathbf{A}$ is symmetric, i. e. when (6) holds. □

The problem posed by Kruskal may be restricted to a given estimable φ, and namely: when the LSE of ϕ coincides with its BLUE? Some necessary and sufficient conditions may be found in ([25], Theorem 2).

By Gauss–Markov theorem (cf. [20], Theorem 2, p.14, or [26], Theorem 4.1 with its proof in the spirit of this section), the LSE for estimable $\varphi = \mathbf{c}^T\beta$ is unique and is given by $\mathbf{c}^T\widehat{\beta}$, where $\widehat{\beta}$ is any solution of the normal equation

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{y}. \tag{7}$$

(By (2) this equation is consistent.)

However, the BLUE for ϕ may not be unique. It follows from Theorem 1 that any BLUE for φ may be presented in the form $\mathbf{b}^T\mathbf{y}$, where $\mathbf{b}$ is a member of the class

$$\mathcal{B} = \{\mathbf{b}_0 + \mathbf{x} : \mathbf{b}_0^T\mathbf{y} \text{ is an arbitrary BLUE for } \varphi,\ \mathbf{A}^T\mathbf{x} = \mathbf{0} \text{ and } \mathbf{V}\mathbf{x} = \mathbf{0}\}. \tag{8}$$

In particular we get the following corollary.

Corollary 5. The BLUE for an estimable parametric function $\varphi = \mathbf{c}^T\beta$ in the model $\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ is unique, if and only if, the equation

$$\begin{bmatrix}\mathbf{A}^T \\ \mathbf{V}\end{bmatrix}\mathbf{x} = \mathbf{0} \tag{9}$$

has a unique solution $\mathbf{x} = \mathbf{0}$.

Now let us go to estimation of a parametric function $\varphi = \mathbf{c}^T\beta$ by a linear (not necessarily unbiased) statistic $\mathbf{d}^T\mathbf{y}$ with respect to the Mean Squared Error (MSE), defined by the formula

$$MSE(\mathbf{d}^T\mathbf{y}, \mathbf{c}^T\beta) = E(\mathbf{d}^T\mathbf{y} - \mathbf{c}^T\beta)^2.$$

An estimator $\mathbf{d}^T\mathbf{y}$ is said to be admissible if there is no other estimator, say $\mathbf{d}_0^T\mathbf{y}$, such that $MSE(\mathbf{d}_0^T\mathbf{y}, \mathbf{c}^T\beta) \le MSE(\mathbf{d}^T\mathbf{y}, \mathbf{c}^T\beta)$ for all β and σ with the strict inequality for some β_0 and σ_0.

Cohen [4] considered a linear model $\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ with $\mathbf{A} = \mathbf{I}$ and nonsingular $\mathbf{V}$. For simplicity such model will be denoted by $\mathfrak{L}(\beta,[\mathbf{V}])$. In this case every linear function $\varphi = \mathbf{c}^T\beta$ is estimable. Cohen proved the following theorem.

Theorem 6. Let $\mathbf{y}$ be observation vector in the model $\mathfrak{L}(\beta,[\mathbf{V}])$ with nonsingular $\mathbf{V}$. Then a linear statistic $\mathbf{d}^T\mathbf{y}$ is admissible estimator of a parametric function $\varphi = \mathbf{c}^T\beta$, if and only if, $\mathbf{d}$ belongs to the ellipsoid

$$\left(\mathbf{d}-\frac{\mathbf{c}}{2}\right)^T \mathbf{V}\left(\mathbf{d}-\frac{\mathbf{c}}{2}\right) \leq \frac{\mathbf{c}^T\mathbf{V}\mathbf{c}}{4}. \tag{10}$$

The first characterization of admissible estimators for an estimable parametric function $\varphi = \mathbf{c}^T\beta$ in a linear model $\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ with possibly singular matrix $\mathbf{V}$, presented in ([19], Section 4, and its correction), seems not be clear. A more precise result of this kind (see [21]) may be presented as follows.

Theorem 7. Let $\mathbf{y}$ be observation vector in a linear model $\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ and let $\varphi = \mathbf{c}^T\beta$ be estimable in this model. Then a linear statistic $\mathbf{d}^T\mathbf{y}$ is admissible for φ, if and only if,

(i) $\mathbf{V}\mathbf{d} \in \mathcal{R}(\mathbf{A})$,
(ii) $\mathbf{d}^T\mathbf{V}\mathbf{b} \geq \mathbf{d}^T\mathbf{V}\mathbf{d}$, and,
(iii) either $\mathbf{A}^T(\mathbf{d}-\mathbf{b}) = \mathbf{0}$, or $\mathbf{V}(\mathbf{d}-\mathbf{b}) \neq \mathbf{0}$,
where $\mathbf{b}^T\mathbf{y}$ is a BLUE of φ.

The following remarks throw some light on the conditions (i)-(iii).

Remark 8. Under assumption of Theorem 6 we get $\mathbf{b} = \mathbf{c}$ and the condition (ii) is equivalent to (10).

Remark 9. The condition (i) does not depend on φ, while (ii) and (iii) do.

Remark 10. The class of admissible linear estimators for φ, presented by Theorem 7, does not depend on the choice of its BLUE. To verify this we only need to use the representation (8).

One can ask whether the set of the conditions (i)-(iii) is minimal in the sense that neither of them is implied by the others. We will show by examples that the answer is YES.

Example 11. Condition (i) is not implied by (ii) and (iii).

Consider model $\mathbf{y} \sim \mathcal{L}(\mu\mathbf{1}_n,[\mathbf{I}_n])$, where $\mathbf{1}_n$ means the column of n ones and μ is a scalar. If $\varphi = \mu$ then $\mathbf{b}^T\mathbf{y}$ with $\mathbf{b} = \frac{1}{n}\mathbf{1}_n$ is a BLUE of ϕ and $var(\mathbf{b}^T\mathbf{y}) = \frac{1}{n}$. Let us set $\mathbf{d} = \frac{\mathbf{b}}{2} + \mathbf{a}$, where $\mathbf{a}^T\mathbf{1}_n = 0$ and $0 < \mathbf{a}^T\mathbf{a} < \frac{1}{4n}$. Then (ii) and (iii) are met but (i) is not.

Example 12. Condition (ii) is not implied by (i) and (iii).

Consider model $\mathbf{y} \sim \mathcal{L}(\beta,[\mathbf{I}_n])$. In this case $\mathbf{b} = \mathbf{c}$ and the condition (i) is met for arbitrary $\mathbf{d}$. By setting $\mathbf{d} = 2\mathbf{c} \neq \mathbf{0}$, we get $\mathbf{V}(\mathbf{d}-\mathbf{b}) \neq \mathbf{0}$ and $\mathbf{d}^T\mathbf{V}\mathbf{b} < \mathbf{d}^T\mathbf{V}\mathbf{d}$. Thus (i) and (iii) are satisfied but (ii) is not.

Example 13. Condition (iii) is not implied by (i) and (ii).

Consider model $\mathbf{y} \sim\mathcal{L}\left(\beta, \left[\left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right)\right]\right)$. In this case the condition (i) is always met and $\mathbf{b} = \mathbf{c}$. By setting $\mathbf{c} = [1, 0]^T$ and $\mathbf{d} = [1, 1]^T$, we get $\mathbf{d}^T\mathbf{V}\mathbf{b} = \mathbf{d}^T\mathbf{V}\mathbf{d}$, $\mathbf{A}^T(\mathbf{d}-\mathbf{b}) \neq \mathbf{0}$ and $\mathbf{V}(\mathbf{d} - \mathbf{b}) = \mathbf{0}$. Thus (i) and (ii) are met but (iii) is not.

Let us end this section by the following remark.

Remark 14. Ip, Wong and Liu [10] claim that any linear form defined by their formula (4.2) is both linearly sufficient and admissible for KB. Holding the notation used in [10] let us put $X = \Sigma = K = 1$ and $V = 0$. Then the trivial statistic $0Y$ satisfy (4.2) but is neither linearly sufficient nor admissible for KB because is dominated by Y.

3. MAIN RESULT AND ITS APPLICATIONS

We mention that the class of admissible linear estimators presented in Theorems 6 and 7 is very large, usually represented by a multidimensional set in R^n. This set may be reasonably restricted by an invariance condition imposed on the Mean Squared Error of potential estimators.

Definition 15. We shall say that a linear functional $\mathbf{d}^T\mathbf{y}$ is an invariant estimator for a parametric function $\varphi(\beta) = \mathbf{c}^T\beta$ in a linear model $\mathbf{y} \sim\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ if its MSE depends on β only through $\varphi(\beta)$, and totally invariant, if the MSE does not depend on β.

From representation

$$MSE(\mathbf{d}^T\mathbf{y}, \mathbf{c}^T\beta) = \sigma\mathbf{d}^T\mathbf{V}\mathbf{d} + (\mathbf{d}^T\mathbf{A}\beta - \mathbf{c}^T\beta)^2 \tag{11}$$

we derive the following corollaries.

Corollary 16. If φ is estimable then $\widehat{\varphi}$ is an totally invariant for φ if and only if it is unbiased.

Corollary 17. If φ is estimable and $\widehat{\varphi}$ is an unbiased estimator for φ, then $k\widehat{\varphi}$ is invariant for arbitrary $k \in R$.

We shall prove

Theorem 18. If $\varphi = \mathbf{c}^T\beta$ is estimable in the model $\mathbf{y} \sim\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ then $\mathbf{d}^T\mathbf{y}$ is invariant for φ, if and only if, it may be presented in the form $k\widehat{\varphi}$, where $\widehat{\varphi}$ is unbiased for φ and $k \in R$.

Proof. By Corollary 17 it remains to show the necessity of the condition.

Let $\mathbf{c}^T\beta$ be estimable and let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ be an orthonormal basis in $\mathcal{R}(\mathbf{A}^T)$ such that $\mathbf{a}_1 = \frac{\mathbf{c}}{||\mathbf{c}||}$. We shall use the reparametrization $\theta = \mathbf{T}\beta$, where

$$\mathbf{T} = \begin{bmatrix} \mathbf{a}_1^T \\ \ldots \\ \mathbf{a}_r^T \end{bmatrix}.$$

In this reparametrization model $\mathfrak{L}(\beta,[\mathbf{V}])$ may be written in the form $\mathfrak{L}(\mathbf{G}\theta,[\mathbf{V}])$, where $\mathbf{G} = \mathbf{AT}^T$, $\theta_1 = \frac{\varphi}{||\mathbf{c}||}$, and the parameter $\theta = (\ \theta_1, \theta_2, \ldots, \theta_r)'$ is running over R^r. In consequence, the second component in (11) may be presented in the form

$$\left[\sum_{i=1}^{r}(\mathbf{d}^T\mathbf{g}_i)\theta_i - \|\mathbf{c}\|\,\theta_1\right]^2,$$

where $\mathbf{g}_i$ denotes the $i-$th column of the matrix $\mathbf{G}$. Since the first component in (11) is independent of β, $\mathbf{d}^T\mathbf{y}$ is invariant for φ, if and only if $\mathbf{d}^T\mathbf{g}_i = 0$ for $i = 2, \ldots, r$, i. e. when desired condition is met. □

Now we are ready to state the main result in this section.

Theorem 19. If $\varphi = \mathbf{c}^T\beta$ is estimable in the model $\mathbf{y} \sim\mathfrak{L}(\mathbf{A}\beta,[\mathbf{V}])$ then a linear form $\mathbf{d}^T\mathbf{y}$ is admissible among invariant estimators for φ, if and only if, it coincides with its BLUE, in the case when ϕ is estimable with zero variance, and $\mathbf{d}^T\mathbf{y}$ is of the form $k\widehat{\varphi}$, where $\widehat{\varphi}$ is a BLUE and $k \in \langle 0, 1\rangle$, otherwise.

P r o o f. If $var(\widehat{\varphi}) = 0$ then the MSE of $\widehat{\varphi}$ is zero and hence it is minimal among all estimators. Now suppose that $var(\widehat{\varphi}) = v > 0$. By representation (11), Theorem 18 and definition of BLUE it follows that any admissible estimator $\mathbf{d}^T\mathbf{y}$ may be presented in the form kz, where z is the BLUE of φ. In consequence, kz is admissible, if and only if, it is admissible in the scalar model $z{\sim}\mathfrak{L}(\varphi,[v])$. Now, by condition (ii) in Theorem 7 we get the desired result. □

Let us recall that the class of all BLUEs for φ is presented by the formula (8).

Now we shall use Theorem 19 in the Analysis of Variance (ANOVA) with one- and two-way balanced classification and so called fixed (i. e. deterministic) effects.

3.1. One-way ANOVA with restraint

Let us consider the model with t treatments, each with q replications. Such a model may be presented in the form

$$y_{ij} = \mu + \alpha_i + e_{ij},\ \ i = 1, \ldots, t;\ j = 1, \ldots, q, \tag{12}$$

where μ is the general mean (or intercept), α_i is the effect of the ith treatment, while e_{ij}, $i = 1, \ldots, t$; $j = 1, \ldots, q$, are not correlated experimental errors with mean zero and (known or unknown) common variance σ. The observation vector $\mathbf{y} = (y_{11}, \ldots, y_{1q}; \ldots; y_{t1}, \ldots y_{tq})^T$ may be concisely written in the form

$$\mathbf{y} = \mathbf{A}\beta + \mathbf{e}, \tag{13}$$

where

$$\mathbf{A} = \begin{bmatrix}\mathbf{1}_n & \mathbf{I}_t \otimes \mathbf{1}_q\end{bmatrix}$$

$$\beta = (\mu, \alpha_1, \ldots, \alpha_t)^T$$

and

$$\mathbf{e} = (e_{11}, \ldots, e_{1q}; \ldots; e_{t1}, \ldots e_{tq})^T,$$

while $n = tq$ and $\otimes$ means the Kronecker product.

We note that $\mathbf{A}$ is a matrix of size $n \times (t+1)$ with $rank(\mathbf{A}) = t$ and, therefore, of not full column rank. In consequence, the parameters $\mu, \alpha_1, \ldots, \alpha_t$ are not identifiable by the model (12) in the sense that $E\mathbf{y}$ is not uniquely represented by them. The identifiability may be satisfied by an additional condition, for instance of the form $\sum_{i=1}^{t} \alpha_i = 0$, called restraint. The model (13) together with this restraint may be presented in the form

$$\mathbf{u} \sim \mathfrak{L}(\mathbf{B}\beta, [\mathbf{V}]), \tag{14}$$

where

$$\mathbf{u} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} \mathbf{1}_n & \mathbf{I}_t \otimes \mathbf{1}_q \\ 0 & \mathbf{1}_t^T \end{bmatrix}$$

with singular

$$\mathbf{V} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_n \\ \mathbf{0}_n^T & 0 \end{bmatrix}.$$

Since $rank(\mathbf{B}) = t+1$, any parametric function $\varphi = c_0\mu + \sum_{i=1}^{t} c_i\alpha_i$ in the model (14) is estimable. We are interested in admissible invariant estimators for φ. By Theorem 19 this problem reduces to the BLUE's for the parameters $\mu, \alpha_1, \ldots, \alpha_t$.

We shall start from the Least Squares Estimator for the vector $\beta = (\mu, \alpha_1, \ldots, \alpha_t)^T$. The classical way leads by solving the normal equation

$$\mathbf{B}^T\mathbf{B}\beta = \mathbf{B}^T\mathbf{u}$$

and, in consequence, we get

$$\widehat{\beta} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{u}. \tag{15}$$

To this aim we only need to invert the patterned matrix

$$\mathbf{B}^T\mathbf{B} = \begin{bmatrix} n & q\mathbf{1}_t^T \\ q\mathbf{1}_t & q\mathbf{I}_t + \mathbf{1}_t\mathbf{1}_t^T \end{bmatrix}.$$

One can easy verify that

$$(\mathbf{B}^T\mathbf{B})^{-1} = \begin{bmatrix} \frac{1}{t} + \frac{1}{n} & -\frac{1}{t^2}\mathbf{1}_t^T \\ -\frac{1}{t^2}\mathbf{1}_t & \frac{1}{q}\mathbf{I}_t + (\frac{1}{t^2} - \frac{1}{n})\mathbf{1}_t\mathbf{1}_t^T \end{bmatrix}.$$

In consequence we get

$$(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \begin{bmatrix} \frac{1}{n}\mathbf{1}_n^T & -\frac{1}{t} \\ \frac{1}{q}\mathbf{I}_t \otimes \mathbf{1}_q^T - \frac{1}{n}\mathbf{1}_t\mathbf{1}_n^T & \frac{1}{t}\mathbf{1}_t \end{bmatrix}$$

and the LSE of the parameter vector $\beta = (\mu, \alpha_1, \ldots, \alpha_t)^T$ takes the form

$$\widehat{\beta} = \begin{bmatrix} \widehat{\mu} \\ \widehat{\alpha}_1 \\ \cdots \\ \widehat{\alpha}_t \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i,j} y_{ij} \\ \frac{1}{q}\sum_j y_{1j} - \frac{1}{n}\sum_{i,j} y_{ij} \\ \cdots \\ \frac{1}{q}\sum_j y_{tj} - \frac{1}{n}\sum_{i,j} y_{ij} \end{bmatrix}. \tag{16}$$

Remark 20. A similar result has been obtained in Scheffé ([20], chapter 3) for one-way classification without intercept.

In order to ensure that the BLUE of β is unique and it coincides with its LSE we only need to verify the condition (6), which is actually in the form $\mathbf{P_B V} = \mathbf{V P_B}$, and that the equation

$$\begin{bmatrix} \mathbf{B}^T \\ \mathbf{V} \end{bmatrix} \mathbf{x} = \mathbf{0}$$

has the unique solution of $\mathbf{x} = \mathbf{0}$. By formula (4)

$$\mathbf{P_B} = \begin{bmatrix} \frac{1}{q} diag(\mathbf{1}_q \mathbf{1}_q^T, \ldots, \mathbf{1}_q \mathbf{1}_q^T) & \mathbf{0}_n \\ \mathbf{0}_n^T & 1 \end{bmatrix}$$

and the desired conditions are met.

Now, by Theorem 3, Lemma 4 and Theorem 19, we get the following

Conclusion 21. A linear form $\mathbf{a}^T\mathbf{y}$ is admissible invariant estimator for a parametric function $\varphi = c_0\mu + \sum_{i=1}^t c_i\alpha_i$ in the model (14), if and only if, $\mathbf{a}^T\mathbf{y}$ coincides with the LSE $\widehat{\varphi} = c_0\widehat{\mu} + \sum_{i=1}^t c_i\widehat{\alpha}_i$ of ϕ, in the case when $c_1 = c_2 = \ldots = c_t$, and $\mathbf{a}^T\mathbf{y} = k\widehat{\varphi}$ with $k \in \langle 0, 1\rangle$, otherwise.

3.2. Additive two-way ANOVA with restraints

Suppose $n = tq$ experimental units are subject to two independent classifications with t and q subclasses, respectively. Then the observation y_{ij}, corresponding to the ith subclass in the first classification and the jth subclass in the second one, may be presented in the form

$$y_{ij} = \mu + \alpha_i + \theta_j + e_{ij}, \tag{17}$$

where α_i and θ_j, $i = 1, \ldots, t; j = 1, \ldots, q$ are the effects of the ith and jth subclass in the respective classification. As in Section 5 we shall assume that e_{ij}, $i = 1, \ldots, t; j = 1, \ldots, q$ are not correlated experimental errors with mean zero and the common variance σ. The model (17) may be presented in the concise form

$$\mathbf{y} = \mathbf{A}\beta + \mathbf{e}, \tag{18}$$

where

$$\mathbf{A} = [\mathbf{1}_n; \mathbf{I}_t \otimes \mathbf{1}_q; \mathbf{1}_t \otimes \mathbf{I}_q]$$

and

$$\beta = (\mu; \alpha_1, \ldots, \alpha_t; \theta_1, \ldots \theta_q)^T.$$

For the same reason as in one-way ANOVA we shall use additional restraints $\sum_{i=1}^{t}\alpha_i = \sum_{j=1}^{q}\beta_j = 0$ on the parameters. Now the model (18) with the restraints may be presented in the form

$$\mathbf{z} \sim \mathfrak{L}(\mathbf{B}\beta, [\mathbf{V}]), \tag{19}$$

where

$$\mathbf{z} = \begin{bmatrix} \mathbf{y} \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{1}_n & \mathbf{I}_t \otimes \mathbf{1}_q & \mathbf{1}_t \otimes \mathbf{I}_q \\ 0 & \mathbf{1}_t^T & \mathbf{0}_q^T \\ 0 & \mathbf{0}_t^T & \mathbf{1}_q^T \end{bmatrix}$$

and

$$\mathbf{V} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times 2} \\ \mathbf{0}_{2\times n} & \mathbf{0}_{2\times 2} \end{bmatrix}.$$

We are interested in admissible invariant estimators for $\varphi = c_0\mu + \sum_{i=1}^{t} c_i\alpha_i + \sum_{j=1}^{q} d_j\theta_j$. By Theorem 19 this problem reduces to the BLUE's for the parameters $\mu, \alpha_1, \ldots, \alpha_t; \theta_1, \ldots, \theta_q$.

It is easy to verify that

$$(\mathbf{B}^T\mathbf{B})^{-1} = \begin{bmatrix} \frac{1}{t^2} + \frac{1}{q^2} + \frac{1}{n} & -\frac{1}{t^2}\mathbf{1}_t^T & -\frac{1}{q^2}\mathbf{1}_q^T \\ -\frac{1}{t^2}\mathbf{1}_t & \frac{1}{q}\mathbf{I}_t + (\frac{1}{t^2} - \frac{1}{n})\mathbf{1}_t\mathbf{1}_t^T & \mathbf{0} \\ -\frac{1}{q^2}\mathbf{1}_q & \mathbf{0}^T & \frac{1}{t}\mathbf{I}_q + (\frac{1}{q^2} - \frac{1}{n})\mathbf{1}_q\mathbf{1}_q^T \end{bmatrix}.$$

In consequence the LSE of the parametric vector β takes the form

$$\widehat{\beta} = \begin{bmatrix} \widehat{\mu} \\ \widehat{\alpha}_1 \\ \cdots \\ \widehat{\alpha}_t \\ \widehat{\theta}_1 \\ \cdots \\ \widehat{\theta}_q \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i,j} y_{ij} \\ \frac{1}{q}\sum_j y_{1j} - \frac{1}{n}\sum_{i,j} y_{ij} \\ \cdots \\ \frac{1}{q}\sum_j y_{pj} - \frac{1}{n}\sum_{i,j} y_{ij} \\ \frac{1}{t}\sum_i y_{i1} - \frac{1}{n}\sum_{i,j} y_{ij} \\ \cdots \\ \frac{1}{t}\sum_i y_{iq} - \frac{1}{n}\sum_{i,j} y_{ij} \end{bmatrix}. \tag{20}$$

In the same way as in one-way ANOVA we get the following

Conclusion 22. A linear form $\mathbf{a}^T\mathbf{y}$ is admissible invariant estimator for a parametric function $\varphi = c_0\mu + \sum_{i=1}^{t} c_i\alpha_i + \sum_{j=1}^{q} d_j\theta_j$ in the model (19), if and only if, $\mathbf{a}^T\mathbf{y}$ coincides with the LSE $\widehat{\varphi} = c_0\widehat{\mu} + \sum_{i=1}^{t} c_i\widehat{\alpha}_i + \sum_{j=1}^{q} d_j\widehat{\theta}_j$ of ϕ, in the case when $c_0 = 0$, $c_1 = c_2 = \ldots = c_t$ and $d_1 = d_2 = \ldots = d_q$, and $\mathbf{a}^T\mathbf{y} = k\widehat{\varphi}$ with $k \in \langle 0, 1\rangle$, otherwise.

ACKNOWLEDGEMENTS.

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge. The author would like to express his sincere thanks to one of the reviewers for valuable comments and suggestions. They helped him to improve presentation of this work.

(Received October 23, 2012)

REFERENCES

[1] J. K. Baksalary and A. Markiewicz: Admissible linear estimators in the general Gauss–Markov model. J. Statist. Plann. Inference *19* (1988), 349–359.

[2] J. K. Baksalary and A. Markiewicz: A matrix inequality and admissibility of linear estimators with respect to the mean squared error criterion. Linear Algebra Appl. *112* (1989), 9–18.

[3] J. K. Baksalary and A. Markiewicz: Admissible linear estimators of an arbitrary vector parametric function in the general Gauss–Markov model. J. Statist. Plann. Inference *26* (1990), 161–173.

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

[5] A. Cohen: All admissible estimates of the mean vector. Ann. Math. Statist. *37* (1966), 458–463.

[6] J. Groß: More on concavity of a matrix function. SIAM J. Matrix Anal. Appl. *19* (1998), 365–368.

[7] J. Groß: Löwner partial ordering and space preordering of Hermitian non-negative definite matrices. Linear Algebra Appl. *326* (2001), 215–223.

[8] J. Groß and A. Markiewicz: Characterization of admissible linear estimators in the linear model. Linear Algebra Appl. *388* (2004), 239–248.

[9] P. R. Halmos: Finite-Dimensional Vector Spaces. Second edition. Springer-Verlag, New York 1993.

[10] W. Ip, H. Wong, and J. Liu: Sufficient and admissible estimators in general multivariate linear model. J. Statist. Plann. Inference *135* (2005), 371–383.

[11] W. Klonecki: Linear estimators of mean vector in linear models: Problem of admissibility. Probab. Math. Statist. *2* (1982), 167–178.

[12] W. Klonecki and S. Zontek: On the structure of admissible linear estimators. J. Multivariate Anal. *24* (1988), 11–30.

[13] W. Kruskal: When are Gauss–Markov and least squares estimators identical? A coordinate-free approach. Ann. Math. Statist. *39* (1968), 70–75.

[14] L. R. LaMotte: Admissibility in linear models. Ann. Statist. *10* (1982), 245–255.

[15] L. R. LaMotte: On limits of uniquely best linear estimators. Metrika *45* (1997), 197–211.

[16] E. L. Lehmann and H. Scheffé: Completeness, similar regions, and unbiased estimation – Part 1. Sankhyā A, *10* (1950), 305–340.

[17] A. Olsen, J. Seely, and D. Birkes: Invariant quadratic unbiased estimation for two variance components. Ann. Statist. *4* (1976), 823–1051.

[18] C. R. Rao: Linear Statistical Inference. Wiley, New York 1973.

[19] C. R. Rao: Estimation of parameters in a linear model. Ann. Statist. *4* (1976), 1023–1037. Correction Ann. Statist. *7* (1979), 696–696.

[20] H. Scheffé: The Analysis of Variance. Wiley, New York 1959.

[21] C. Stępniak: On admissible estimators in a linear model. Biom. J. *26* (1984), 815–816.

[22] C. Stępniak: Ordering of nonnegative definite matrices with application to comparison of linear models. Linear Algebra Appl. *70* (1985), 67–71.

[23] C. Stępniak: A complete class for linear estimation in a general linear model. Ann. Inst. Statist. Math. A *39* (1987), 563–573.

[24] C. Stępniak: Admissible linear estimators in mixed linear models. J. Multivariate Anal. *31* (1989), 90–106.

[25] C. Stępniak: Perfect linear models and perfect parametric functions. J. Statist. Plann. Inference *139* (2009), 151–163.

[26] C. Stępniak: From equivalent linear equations to Gauss–Markov theorem. J. Inequal. Appl. (2010), ID 259672, 5 pages.

[27] C. Stępniak: On a problem with singularity in comparison of linear experiments. J. Statist. Plann. Inference *141* (2011), 2489–2493.

[28] E. Synówka-Bejenka and S. Zontek: A characterization of admissible linear estimators of fixed and random effects in linear models. Metrika *68* (2008), 157–172.

[29] S. Zontek: Admissibility of limits of the unique locally best estimators with application to variance components models. Probab. Math. Statist. *9* (1988), 29–44.

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

Czesław Stępniak, Department of Differential Equations and Statistics, Faculty of Mathematics and Natural Sciences, University of Rzeszów, Pigonia 1, 35-310 Rzeszów. Poland.
e-mail: stepniak@umcs.lublin.pl