

*Discussiones Mathematicae*  
*Probability and Statistics 34 (2014) 159–167*  
doi:10.7151/dmps.1164

**A NOTE ON THE EXISTENCE OF THE MAXIMUM  
LIKELIHOOD ESTIMATE IN VARIANCE  
COMPONENTS MODELS**

MARIUSZ GRZĄDZIEL

*Department of Mathematics*  
*Wrocław University of Environmental and Life Sciences*  
*Grunwaldzka 53, 50–357 Wrocław, Poland*

**e-mail:** mariusz.grzadzziel@up.wroc.pl

AND

ANDRZEJ MICHALSKI

*Department of Mathematics*  
*Wrocław University of Environmental and Life Sciences*  
*Grunwaldzka 53, 50–357 Wrocław, Poland*

**e-mail:** andrzej.michalski@up.wroc.pl

**Abstract**

In the paper, the problem of the existence of the maximum likelihood estimate and the REML estimate in the variance components model is considered. Errors in the proof of Theorem 3.1 in the article of Demidenko and Massam (*Sankhyā* 61, 1999), giving a necessary and sufficient condition for the existence of the maximum likelihood estimate in this model, are pointed out and corrected. A new proof of Theorem 3.4 in the Demidenko and Massam's article, concerning the existence of the REML estimate of variance components, is presented.

**Keywords:** variance component, linear mixed model, maximum likelihood.

**2010 Mathematics Subject Classification:** Primary: 62J05; Secondary: 62F10.

## 1. INTRODUCTION

The paper of Demidenko and Massam [1] gives a definitive answer to the problem of the existence of the maximum likelihood estimate in the variance components model. Checking the necessary and sufficient condition for its existence given in this paper should be done before using numerical procedures for computing it. Such a check should be also performed if we are going to use the algebraic methods for computing this estimate: In this approach all critical points of the log-likelihood function are computed via solving a system of polynomial equations, see the paper of Gross *et al.* [2].

The main purpose of this paper is to point out errors in the proof of Theorem 3.1 in [1], giving a necessary and sufficient condition for the existence of the maximum likelihood estimate in the variance components model, and to correct them. We also show that the result of this theorem can be extended to the case when the design matrix is rank deficient. In consequence, we obtain a new proof of Theorem 3.4 in [1] which gives a necessary and sufficient condition for the existence of the restricted maximum likelihood (REML) estimate in the variance components model.

The paper is organized as follows. In Section 2, after recalling the Demidenko and Massam's theorem [1, Theorem 3.1], we point out two errors in its proof given by the authors and we correct these errors. The problem of the existence of the REML estimate is considered in Section 3.

## 1.1. Notation

For a vector  $\mathbf{x} \in \mathbb{R}^n$  we will denote its Euclidean norm by  $\|\mathbf{x}\|$ . For a given  $m \times n$  matrix  $\mathbf{A}$ , we will denote by  $\mathbf{A}'$  its transpose, by  $\text{rank}(\mathbf{A})$  its rank and by  $\mathcal{M}(\mathbf{A})$  the space spanned by the columns of  $\mathbf{A}$ . For given matrices  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p$  of dimension  $m \times n_1, m \times n_2, \dots, m \times n_p$ , respectively, we will denote by  $[\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p]$  the partitioned  $m \times (n_1 + n_2 + \dots + n_p)$  matrix consisting of  $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p$ . For brevity we will write  $\mathcal{M}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p)$  instead of  $\mathcal{M}([\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p])$ . We will write  $|\mathbf{B}|$  for the determinant of a square matrix  $\mathbf{B}$ ,  $\mathbf{I}_n$  for the identity matrix of order  $n$ . The matrix of order  $m \times n$  in which every entry is equal to 0 we will denote by  $\mathbf{0}_{m \times n}$ . If the order of the matrix is clear from the context we will denote the zero matrix simply by  $\mathbf{0}$ . The  $n$ -dimensional vector having all coordinates equal to 1 we will denote by  $\mathbf{1}^{(n)}$ , the  $n$ -dimensional vector having all coordinates equal to 0 we will denote by  $\mathbf{0}^{(n)}$ . For an  $m \times n$  matrix  $\mathbf{G}$  and a subspace  $\mathcal{V}$  of  $\mathbb{R}^n$ , we will denote by  $\mathbf{G}\mathcal{V}$  the image of the transformation corresponding to  $\mathbf{G}$  of the subspace  $\mathcal{V}$ . We will use the notation  $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  if the random vector  $\mathbf{Y}$  has the multivariate normal distribution with the mean vector  $\boldsymbol{\mu}$  and the variance-covariance matrix  $\boldsymbol{\Sigma}$ . For a real-valued function  $f$  with domain  $S$  we

define

$$\operatorname{argmin}_{x \in S} f(x) := \{z \in S : f(z) \leq f(x) \text{ for all } x \in S\}.$$

## 2. THE EXISTENCE OF THE MAXIMUM LIKELIHOOD ESTIMATE

The variance components model considered in the Demidenko and Massam's paper [1] can be expressed in the form

$$(1) \quad \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=1}^r \mathbf{Z}_i \mathbf{u}_i + \boldsymbol{\epsilon},$$

where  $\mathbf{Y}$  is an  $n \times 1$  random vector,  $\mathbf{X}$  is an  $n \times m$  known design matrix of full column rank,  $m < n$ ,  $\boldsymbol{\beta}$  is an  $m \times 1$  vector of fixed parameters,  $\mathbf{Z}_i$  are known  $n \times k_i$  matrices such that  $\sum_{i=1}^r k_i < n$  and  $\mathbf{u}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 \mathbf{I}_{k_i})$ ,  $i = 1, \dots, r$ , while  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_0^2 \mathbf{I}_n)$ . We assume that all the random terms are stochastically independent. The parameter space is equal to

$$(2) \quad \Theta := \{\boldsymbol{\theta} := (\boldsymbol{\beta}, \sigma_0^2, \sigma_1^2, \dots, \sigma_r^2) = (\boldsymbol{\beta}, \boldsymbol{\sigma}^2) \in \mathbb{R}^m \times (0, \infty) \times [0, \infty)^r\},$$

where  $\boldsymbol{\sigma}^2 := (\sigma_0^2, \sigma_1^2, \dots, \sigma_r^2)$  is the vector of variance components. It can be seen that the covariance matrix of the vector  $\mathbf{Y}$  can be expressed as

$$\mathbf{V}(\boldsymbol{\sigma}^2) := \sigma_0^2 \mathbf{I}_n + \sum_{i=1}^r \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i'.$$

Minus twice the log-likelihood function is given, up to an additive constant, by

$$(3) \quad l(\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{Y}) := \log |\mathbf{V}(\boldsymbol{\sigma}^2)| + (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{V}^{-1}(\boldsymbol{\sigma}^2) (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).$$

For a given realization  $\mathbf{y}$  of the observation vector  $\mathbf{Y}$  the maximum likelihood estimate of  $\boldsymbol{\theta} \in \Theta$  is defined as

$$(4) \quad \operatorname{argmin}_{\boldsymbol{\tau} \in \Theta} l(\boldsymbol{\tau}, \mathbf{y}).$$

Put

$$(5) \quad \kappa_0 := \sigma_0^2, \quad \kappa_i := \frac{\sigma_i^2}{\sigma_0^2}, \quad i = 1, 2, \dots, r,$$

$$(6) \quad \boldsymbol{\kappa} := (\kappa_1, \dots, \kappa_r), \quad \tilde{\mathbf{V}}(\boldsymbol{\kappa}) := \mathbf{I}_n + \sum_{i=1}^r \kappa_i \mathbf{Z}_i \mathbf{Z}_i'.$$

Since  $\mathbf{V}(\boldsymbol{\sigma}^2) = \kappa_0 \tilde{\mathbf{V}}(\boldsymbol{\kappa})$ , we will refer to  $\tilde{\mathbf{V}}(\boldsymbol{\kappa})$  as the scaled covariance matrix of the vector  $\mathbf{Y}$ . The parameter space and the counterpart of the function  $l$  corresponding to this new parametrization are given by

$$(7) \quad \tilde{\Theta} := \{\tilde{\boldsymbol{\theta}} := (\boldsymbol{\beta}, \kappa_0, \kappa_1, \dots, \kappa_r) \in \mathbb{R}^m \times (0, \infty) \times [0, \infty)^r\},$$

$$(8) \quad \tilde{l}(\boldsymbol{\beta}, \boldsymbol{\kappa}, \mathbf{Y}) := n \log \kappa_0 + \log |\tilde{\mathbf{V}}(\boldsymbol{\kappa})| + \kappa_0^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa}) (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).$$

The function  $\tilde{l}$  is equal, up to an additive constant, to minus twice the log-likelihood function expressed in terms of  $\boldsymbol{\beta}, \boldsymbol{\kappa}$  and  $\mathbf{Y}$ . If  $\mathbf{y}$  is a given realization of the observation vector  $\mathbf{Y}$ , then the maximum likelihood estimate of  $\tilde{\boldsymbol{\theta}} \in \tilde{\Theta}$  can be expressed as

$$(9) \quad \underset{\boldsymbol{\tau} \in \tilde{\Theta}}{\operatorname{argmin}} \tilde{l}(\boldsymbol{\tau}, \mathbf{y}).$$

We will say that the maximum likelihood estimate of the parameters exists for the given realization  $\mathbf{y}$  of the observation vector  $\mathbf{Y}$  if and only if the set (9) (or the set (4)) is not empty.

A definitive answer to the existence problem in the variance components model is given by Demidenko and Massam in [1, Theorem 3.1]. Before stating it, let us define the following quantities:

$$(10) \quad \mathbf{Z} := [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_r], \quad M := \mathcal{M}(\mathbf{X}), \quad H := \mathcal{M}(\mathbf{Z}),$$

$$(11) \quad s_{X,Z} := \mathbf{y}' \mathbf{P}_{(H+M)^\perp} \mathbf{y},$$

where  $\mathbf{P}_{(H+M)^\perp}$  stands for the orthogonal projection from  $\mathbb{R}^n$  to the orthogonal subspace to  $H + M$ .

**Theorem 2.1** [1, Theorem 3.1]. *The maximum likelihood estimate in the variance components model (1) exists if and only if*

$$(12) \quad \mathbf{y} \notin \mathcal{M}(\mathbf{X}, \mathbf{Z}).$$

The proof of this theorem can be sketched as follows: Theorem 2.1 is an immediate consequence of the following facts mentioned in [1, p. 436–437] presented below in the form of

**Proposition 2.2.** *If for a fixed  $\mathbf{y} \in \mathbb{R}^n$  the condition (12) is satisfied, then:*

$$(a) \quad \tilde{l}(\tilde{\boldsymbol{\theta}}, \mathbf{y}) \geq n \log s_{X,Z} - n \log n + n.$$

(b) *For a given  $A > n \log s_{X,Z} - n \log n + n$  there exists a compact set  $C_A \subset \tilde{\Theta}$  such that*

$$(13) \quad \tilde{l}(\tilde{\boldsymbol{\theta}}, \mathbf{y}) \geq A \quad \text{for each } \tilde{\boldsymbol{\theta}} \in \tilde{\Theta} \setminus C_A.$$

Otherwise:

(c) The infimum of  $\tilde{l}(\tilde{\boldsymbol{\theta}}, \mathbf{y})$  in  $\tilde{\Theta}$  is  $-\infty$ .

The proofs of the above facts can be found in the course of the proof of Theorem 3.1 in [1]. In the next subsection we will discuss their correctness.

## 2.1. The errors in the proof and their corrections

### 2.1.1. Representation of the scaled covariance matrix

The parts of the proof of Theorem 3.1 in [1] that correspond to the parts (a) and (c) of Proposition 2.2 rely on a certain algebraic fact concerning the representation of the scaled covariance matrix  $\tilde{\mathbf{V}}(\boldsymbol{\kappa})$  of the observation vector  $\mathbf{Y}$ . Unfortunately, the proof of this fact, stated as Proposition 3.2 in [1], is not quite correct.

**Comments on Demidenko and Massam’s proof.** We will restate Proposition 3.2 in [1] as the following

**Proposition 2.3** (Demidenko and Massam). *Let us denote the dimension of the space  $H$  defined in (10) by  $q$ . There exists an  $n \times n$  orthogonal matrix  $\mathbf{U}$  and  $q \times q$  matrices  $\mathbf{A}_1, \dots, \mathbf{A}_r$  satisfying the conditions*

$$(14) \quad \mathbf{A}_i \text{ is positive definite, } i = 1, \dots, r,$$

such that

$$(15) \quad \tilde{\mathbf{V}}(\boldsymbol{\kappa}) = \mathbf{U} \begin{bmatrix} \mathbf{I}_q + \sum_{i=1}^r \kappa_i \mathbf{A}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-q} \end{bmatrix} \mathbf{U}'.$$

Moreover, if  $c = \kappa_1 = \kappa_2 = \dots = \kappa_r$ , there exists an  $n \times n$  orthogonal matrix  $\mathbf{U}$  and a  $q \times q$  diagonal matrix  $\mathbf{D}$  with positive diagonal elements satisfying

$$(16) \quad \tilde{\mathbf{V}}(\boldsymbol{\kappa}) = \mathbf{U} \begin{bmatrix} \mathbf{I}_q + c\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-q} \end{bmatrix} \mathbf{U}'.$$

The proof of this proposition given in [1, p. 442] is not correct: The  $q \times q$  matrices  $\mathbf{A}_1, \dots, \mathbf{A}_r$  constructed during the proof satisfy the condition  $\text{rank}(\mathbf{A}_i) = \text{rank}(\mathbf{Z}_i)$ ,  $i = 1, 2, \dots, r$ , and if there exists  $i$  such that  $\text{rank}(\mathbf{Z}_i) < q$  (the ranks of the matrices  $\mathbf{Z}_i$  are not equal in the general case), the matrix  $\mathbf{A}_i$  will have rank less than  $q$ , which will contradict its positive definiteness.

**The correction.** Let us consider the weakened version of Proposition 2.3 in which the condition (14) is replaced by the following condition

$\mathbf{A}_i$  is non-negative definite,  $i = 1, \dots, r$ , and  $\sum_{i=1}^r \mathbf{A}_i$  is positive definite.

It can be verified that using the arguments from the proof of Proposition 3.2 in [1, p. 442] we can obtain the mentioned above weakened version of Proposition 2.3, which is sufficient for the purposes of the proof of [1, Theorem 3.1].

### 2.1.2. Attaining the supremum of the log-likelihood function

**Comments on Demidenko and Massam's proof.** Let us now discuss the part of the proof of Theorem 3.1 in [1] that corresponds to the part (b) of Proposition 2.2 and concerns attaining the infimum of the function  $\tilde{l}$  — or attaining the supremum of the likelihood function. It contains the following statement: It is possible to construct a set  $C_A$  satisfying the condition (13) by:

- choosing  $\epsilon \in (0, 1)$  such that

$$(17) \quad \text{for } \kappa_0 \in (0, \epsilon), \quad n \log \kappa_0 + \kappa_0^{-1} s_{X,Z} \geq A;$$

$$(18) \quad \text{for } \kappa_0 > \epsilon^{-1}, \quad n \log \kappa_0 \geq A;$$

$$(19) \quad \text{for } \epsilon \leq \kappa_0 \leq \epsilon^{-1}, \quad s_{X,Z} \geq (A - n \log \epsilon) \epsilon^{-1};$$

- choosing any bounded set  $B \subset M$ ;
- choosing  $b > 0$  such that

$$(20) \quad \text{if } \max_{1 \leq i \leq r} \kappa_i \geq b, \quad \log |\tilde{V}(\boldsymbol{\kappa})| \geq A - n \log \epsilon;$$

- putting  $C_A := \{\boldsymbol{\beta} : \mathbf{X}\boldsymbol{\beta} \in B\} \times (\epsilon, \epsilon^{-1}) \times [0, b]^r$ ,

see Demidenko and Massam [1, p. 436].

Unfortunately, this claim is not true. To see this, let us consider a set  $C^0$  constructed by:

- choosing  $A$  such that  $A > m + \epsilon_0$ , where  $m$  is the infimum of  $\tilde{l}$  in  $\tilde{\Theta}$  while  $\epsilon_0$  is a given positive number;
- choosing  $\tilde{\boldsymbol{\theta}}_0 = (\boldsymbol{\beta}_0, \kappa_0^0, \kappa_1^0, \dots, \kappa_r^0)$  such that  $\tilde{l}(\tilde{\boldsymbol{\theta}}_0) \leq m + \epsilon_0$ ;
- choosing a bounded set  $B_0 \subset M$  such that  $\mathbf{X}\boldsymbol{\beta}_0 \notin B_0$ ;

- choosing  $\epsilon_1 \in (0, 1)$  and  $b_1 > 0$  such that for  $\epsilon = \epsilon_1$ ,  $b = b_1$  and  $B = B_0$  the conditions (17)–(20) are satisfied;
- putting

$$C^0 := \{\boldsymbol{\beta} : \mathbf{X}\boldsymbol{\beta} \in B_0\} \times (\epsilon_1, \epsilon_1^{-1}) \times [0, b_1]^r.$$

According to Demidenko and Massam [1, p. 436] the condition (13) is satisfied for  $C_A = C^0$  (for the chosen  $A$ ). This implies that  $\tilde{l}(\tilde{\boldsymbol{\theta}}_0) \geq A > m + \epsilon_0$ , and we have obtained a contradiction.

**The correction.** We will now present the corrected version of the proof of the part (b) of Proposition 2.2.

*Proof of the part (b) of Proposition 2.2.* Let us choose

- (i)  $\epsilon \in (0, 1)$  such that the conditions (17) and (18) are satisfied;
- (ii)  $b$  such that the condition (20) is satisfied;
- (iii)  $t$  such that for  $\boldsymbol{\beta} \in \mathbb{R}^m$  satisfying  $\|\boldsymbol{\beta}\| \geq t$ :

$$(21) \quad \text{If } \kappa_0 \in [\epsilon, \epsilon^{-1}] \text{ and } \max_{1 \leq i \leq r} \kappa_i < b, \text{ then}$$

$$(22) \quad \kappa_0^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \tilde{\mathbf{V}}^{-1}(\boldsymbol{\kappa})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \geq A - n \log \epsilon.$$

The existence of  $b$  mentioned in (ii) stems from the fact that

$$\lim_{\lambda \rightarrow \infty} \log |\tilde{\mathbf{V}}(\boldsymbol{\kappa} + \lambda \mathbf{v})| = \infty,$$

where  $\mathbf{v} \in [0, \infty)^r \setminus \{\mathbf{0}^{(r)}\}$ . To prove the existence of  $t$  in (iii) we can use the fact that if the condition (21) is satisfied, then the matrix  $\mathbf{V}^{-1}(\boldsymbol{\kappa}) - \mathbf{V}^{-1}(b\mathbf{1}^{(r)})$  is non-negative definite, see [4, p. 70]. It can be seen that  $C_A := [-t, t]^m \times [\epsilon, \epsilon^{-1}] \times [0, b]^r$  satisfies the condition (13). ■

From the course of the above proof follows immediately

**Proposition 2.4.** *Let us assume that the sequence  $(\tilde{\boldsymbol{\theta}}_n)$  of elements of the parameter space  $\tilde{\Theta}$ ,  $\tilde{\boldsymbol{\theta}}_n = (\boldsymbol{\beta}^{(n)}, \kappa_0^{(n)}, \boldsymbol{\kappa}^{(n)})$ , satisfies at least one of the following conditions:  $\|\boldsymbol{\beta}^{(n)}\| \rightarrow \infty$ ,  $\kappa_0^{(n)} \rightarrow 0$ ,  $\kappa_0^{(n)} \rightarrow \infty$  or  $\|\boldsymbol{\kappa}^{(n)}\| \rightarrow \infty$ . If  $\mathbf{y}$  satisfies the condition (12), then*

$$\lim_{n \rightarrow \infty} \tilde{l}(\tilde{\boldsymbol{\theta}}_n, \mathbf{y}) = \infty.$$

## 2.2. The case of rank deficient design matrix

Throughout the paper, except this subsection, we assume that the design matrix  $\mathbf{X}$  has a full column rank. It can be shown that

**Proposition 2.5.** *Theorem 2.1 remains valid if the assumption " $\mathbf{X}$  is a matrix of full column rank" is dropped.*

*Proof.* Let us consider the case  $\mathbf{X} \neq \mathbf{0}_{n \times m}$  first. Choose a matrix of full column rank  $\mathbf{X}_1$  such that  $\mathcal{M}(\mathbf{X}_1) = \mathcal{M}(\mathbf{X})$ . Observe that for a given realization  $\mathbf{y}$  of the observation vector  $\mathbf{Y}$  the maximum likelihood estimate of  $\boldsymbol{\sigma}^2$  in the new model (with the design matrix  $\mathbf{X}_1$ ) exists if and only if it exists in the model with the design matrix  $\mathbf{X}$ . It follows from Theorem 2.1 that the maximum likelihood estimate exists in the former model if and only if  $\mathbf{y} \notin \mathcal{M}(\mathbf{X}_1, \mathbf{Z}) = \mathcal{M}(\mathbf{X}, \mathbf{Z})$ , which completes the proof of the part of the proposition concerning the case of non-zero design matrix.

Let us now turn our attention to the case  $\mathbf{X} = \mathbf{0}_{n \times m}$ , when we can assume, without loss of generality, that  $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ , where  $\boldsymbol{\beta}_0$  is an arbitrarily chosen element of  $\mathbb{R}^m$ . We can thus replace the sample space  $\tilde{\Theta}$  by  $\tilde{\Theta}_1 := \{\{\boldsymbol{\beta}_0\} \times (0, \infty) \times [0, \infty)^r\}$ . It can be seen that the analogue of Proposition 2.2 corresponding to the "modified sample space"  $\tilde{\Theta}_1$  holds true. This completes the proof. ■

## 3. THE EXISTENCE OF THE REML ESTIMATE

Let  $\mathbf{K}$  be an  $n \times (n - m)$  matrix of rank  $n - m$  satisfying the condition

$$(23) \quad \mathbf{K}'\mathbf{X} = \mathbf{0}_{(n-m) \times m};$$

let us recall that throughout the paper, except Subsection 2.2, we assume that the design matrix  $\mathbf{X}$  has a full column rank. The vector  $\mathbf{K}'\mathbf{Y}$  follows the multivariate normal distribution  $\mathcal{N}(\mathbf{0}, \mathbf{K}'\mathbf{V}(\boldsymbol{\sigma}^2)\mathbf{K})$ . The function

$$l_K(\boldsymbol{\sigma}^2, \mathbf{Y}) := \log |\mathbf{K}'\mathbf{V}(\boldsymbol{\sigma}^2)\mathbf{K}| + \mathbf{Y}'\mathbf{K}(\mathbf{K}'\mathbf{V}(\boldsymbol{\sigma}^2)\mathbf{K})^{-1}\mathbf{K}'\mathbf{Y}$$

is equal, up to an additive constant, to minus twice the log-likelihood function based on the vector  $\mathbf{K}'\mathbf{Y}$  [3, p. 13]. Let  $\mathbf{y}$  be a realization of the observation vector  $\mathbf{Y}$ . The REML estimate of the vector of variance components  $\boldsymbol{\sigma}^2$  is defined as

$$(24) \quad \underset{\boldsymbol{\sigma}^2 \in \Xi}{\operatorname{argmin}} l_K(\boldsymbol{\sigma}^2, \mathbf{y}),$$

where  $\Xi := (0, \infty) \times [0, \infty)^r$ . It can be shown that the set (24) does not depend on the choice of the matrix  $\mathbf{K}$  satisfying the condition (23), see [3, p. 48]. It follows

immediately from Proposition 2.5 that the necessary and sufficient condition for the existence of the REML estimate of the vector  $\sigma^2$  has the form  $\mathbf{K}'\mathbf{y} \notin \mathbf{K}'H$ , which is equivalent to  $\mathbf{K}\mathbf{K}'\mathbf{y} \notin \mathbf{K}\mathbf{K}'H$ . The last condition is in turn equivalent to

$$(25) \quad \mathbf{y} \notin \mathbf{N}H, \quad \text{where} \quad \mathbf{N} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

We have thus proved Theorem 3.4 from [1], which we restate as

**Theorem 3.1.** *The REML estimate of the vector of the variance components  $\sigma^2$  in the variance components model (1) exists if and only if the condition (25) is satisfied.*

It can be also verified that the following analogue of Proposition 2.4 holds true.

**Proposition 3.2.** *Let us assume that the sequence  $(\boldsymbol{\xi}_n)$  of elements of the parameter space  $\Xi$ ,  $\boldsymbol{\xi}_n = (\xi_0^{(n)}, \xi_1^{(n)}, \dots, \xi_r^{(n)})$ , satisfies at least one of the following conditions:*

$$(26) \quad \lim_{n \rightarrow \infty} \frac{\xi_k^{(n)}}{\xi_0^{(n)}} = \infty \quad \text{for some} \quad k \in \{1, \dots, r\},$$

$$(27) \quad \lim_{n \rightarrow \infty} \xi_0^{(n)} = 0 \quad \text{or} \quad \lim_{n \rightarrow \infty} \xi_0^{(n)} = \infty.$$

If  $\mathbf{y}$  satisfies the condition (25), then

$$\lim_{n \rightarrow \infty} l_K(\boldsymbol{\xi}_n, \mathbf{y}) = \infty.$$

It might be expected that this proposition, as well as Proposition 2.4, will prove to be useful for future research concerning maximum likelihood estimation in variance components models.

#### REFERENCES

- [1] E. Demidenko and H. Massam, *On the existence of the maximum likelihood estimate in variance components models*, Sankhyā A. Methods and Techniques **61** (1999) 431–443.
- [2] E. Gross, M. Drton and S. Petrović, *Maximum likelihood degree of variance components models*, Electronic Journal of Statistics **6** (2012) 993–1016. doi:10.1214/12-EJS702
- [3] J. Jiang, *Linear and Generalized Linear Mixed Models and Their Applications* (Springer, Dordrecht, 2007). doi:10.1007/978-0-387-47946-0
- [4] C.R. Rao, *Linear statistical inference and its applications* (Wiley, New York, 1973). doi:10.1002/9780470316436

Received 15 October 2014  
Revised 25 November 2014

