

ON MAXIMUM LIKELIHOOD ESTIMATION IN MIXED NORMAL MODELS WITH TWO VARIANCE COMPONENTS

MARIUSZ GRZĄDZIEL

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

e-mail: mariusz.grzadziel@up.wroc.pl

Abstract

In the paper we deal with the problem of parameter estimation in the linear normal mixed model with two variance components. We present solutions to the problem of finding the global maximizer of the likelihood function and to the problem of finding the global maximizer of the REML likelihood function in this model.

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

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

1. INTRODUCTION

The most popular strategy for computing maximum likelihood estimates of variance components in mixed linear normal models is to use numerical optimization procedures. However, applying these methods may result in "trapping in local maxima" of the likelihood function. The possibility of multimodality of the likelihood function in the linear mixed model with two variance components was demonstrated *e.g.* in [6, Chapter 7].

An alternative approach to the problem of finding the global maximizer of the likelihood function in the linear mixed model is to determine all its stationary points. In some special cases of the linear mixed normal model with two variance components, such as the one-way classification random model, this can be

done via finding all the real roots of a certain rational or polynomial expression, corresponding to the model and the data vector, see [4, 3] and [6].

In this paper we deal with the problem of computing the maximum likelihood estimate in the linear mixed normal model with two variance components in its general form. Using the results obtained by Gnot *et al.* [4] we show how this problem can be reduced to finding all the real roots of an appropriately defined polynomial. More precisely, we show that once the real roots of this polynomial have been computed, a finite set containing all the global maximizers of the likelihood function can be easily constructed (provided that the maximum likelihood estimate in this model, for a given realization of the observation vector, exists). We give an upper bound for the degree the mentioned polynomial. These results are presented in Section 2. In Section 3 we describe similar results for the REML estimation. We close the paper with concluding remarks.

1.1. Notation

For a given $m \times n$ matrix A , we will denote by A' its transpose, by A^+ its Moore-Penrose inverse, by $\text{rank}(A)$ its rank and by $\mathcal{M}(A)$ the space spanned by the columns of A . For a given $m \times n$ matrix A_1 and a given $m \times p$ matrix A_2 , we will denote by $[A_1, A_2]$ the partitioned $m \times (n + p)$ matrix consisting of A_1 and A_2 . We will write $|B|$ for the determinant of a square matrix B , I_n for the identity matrix of order n , $PD(n)$ for the set of positive definite symmetric matrices of order n . The n -dimensional vector having all coordinates equal to 0 we will denote by $\mathbf{0}^{(n)}$. The degree of a polynomial $P(x)$ we will denote by $\text{deg}(P(x))$. We will use the notation $y \sim N(\mu, \Sigma)$ if the random vector y has the multivariate normal distribution with the mean vector μ and the variance-covariance matrix Σ . For a real-valued function f with domain S we define

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

2. COMPUTING MAXIMUM LIKELIHOOD ESTIMATES OF VARIANCE COMPONENTS

2.1. The model and the likelihood function

Let us consider the normal linear mixed model with two variance components $\mathcal{N}(Y, X\beta, \Sigma(s))$, in which Y is an $n \times 1$ normally distributed random vector with

$$(1) \quad E(Y) = X\beta, \quad \text{Cov}(Y) = \Sigma(s) = \sigma_1^2 V + \sigma_2^2 I_n,$$

where X is an $n \times p$ matrix of full rank, $p < n$, β is a $p \times 1$ parameter vector, V is an $n \times n$ non-negative definite symmetric non-zero matrix of rank $k < n$

and $s = (\sigma_1^2, \sigma_2^2)'$ is an unknown vector of variance components belonging to $\mathcal{S} = \{s : \sigma_1^2 \geq 0, \sigma_2^2 > 0\}$.

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

$$(2) \quad l_0(\beta, s, Y) := -\log |\Sigma(s)| - (Y - X\beta)' \Sigma^{-1}(s)(Y - X\beta).$$

Put

$$\begin{aligned} M &:= I_n - X(X'X)^{-1}X', \\ R(s) &:= (M\Sigma(s)M)^+ = \Sigma^{-1}(s) - \Sigma^{-1}(s)X(X'\Sigma^{-1}(s)X)^{-1}X'\Sigma^{-1}(s), \\ G_0(s) &:= X'\Sigma(s)X, \quad \tilde{\beta}(s) := G_0^{-1}(s)X'\Sigma^{-1}(s)Y. \end{aligned}$$

It can be verified that

$$l_0(\beta, s, Y) = -\log |\Sigma(s)| - Y'R(s)Y - (\tilde{\beta}(s) - \beta)'G_0(\tilde{\beta}(s) - \beta),$$

so the problem of finding the maximizers of l_0 in the set $\mathbb{R}^p \times \mathcal{S}$ can be reduced to finding the maximizers of the function l given by

$$l(s, Y) := -\log |\Sigma(s)| - Y'R(s)Y$$

in the set \mathcal{S} , compare [10, p. 230].

For a given realization y of the observation vector Y we thus define the maximum likelihood estimate of s as

$$(3) \quad \operatorname{argmax}_{s \in \mathcal{S}} l(s, y).$$

It can be seen that the set

$$\operatorname{argmax}_{(\beta, s) \in \mathbb{R}^p \times \mathcal{S}} l(\beta, s, y)$$

is empty if and only if the set (3) is empty. In such a case we will say that the maximum likelihood estimate, for a given realization y of the observation vector Y , does not exist.

The model (1) with the parameter space \mathcal{S} can be regarded as the matrix form of the following variance components model

$$(4) \quad Y = X\beta + Zu + \epsilon,$$

where Z is an $n \times k$ matrix such that $ZZ' = V$, $u \sim N(0, \sigma_1^2 I_k)$ and $\epsilon \sim N(0, \sigma_2^2 I_n)$. A necessary and sufficient condition for the existence of the maximum likelihood estimate in the model (4) gives the following

Theorem 2.1 (Demidenko and Massam [2], Theorem 3.1). *Let y be a given realization of the vector Y in the model (4). The maximum likelihood estimate of $s = (\sigma_1^2, \sigma_2^2)'$ in this model exists if and only if*

$$(5) \quad y \notin \mathcal{M}([X, Z]).$$

Remark 2.2. Corrections to the proof of this theorem can be found in [7].

Since $\mathcal{M}(Z) = \mathcal{M}(ZZ')$, we immediately obtain

Proposition 2.3. *Let y be a given realization of the vector Y in the model (1). The maximum likelihood estimate in this model exists if and only if*

$$(6) \quad y \notin \mathcal{M}([X, V]).$$

2.2. The results obtained by Gnot et al. (2002)

Gnot *et al.* [4] considered the problem of computing the maximum likelihood estimate of the vector of variance components in the model that differs from our model in that the parameter space in their model is equal to $\mathcal{S}^* = \{s : \Sigma(s) \in PD(n), \sigma_2^2 > 0\}$. We will now recall some results from this paper (they are valid also in the case when the parameter space is equal to \mathcal{S}).

Let B be an $(n - p) \times n$ matrix satisfying the conditions

$$(7) \quad BB' = I_{n-p}, \quad B'B = M.$$

Let

$$(8) \quad BVB' = \sum_{i=1}^{d-1} m_i E_i$$

be the spectral decomposition of BVB' , where $m_1 > \dots > m_{d-1} > m_d = 0$ stand for the ordered sequence of different eigenvalues of BVB' . Let E_d be such that $\sum_{i=1}^d E_i = I_{n-p}$. Let us define

$$(9) \quad T_i := z'E_i z / \nu_i, \quad z := BY, \quad i = 1, \dots, d,$$

where ν_i is the multiplicity of the eigenvalue m_i , $i = 1, \dots, d$. We assume that

$$(10) \quad \nu_d > 0.$$

It can be checked that m_i , ν_i and E_i , $i = 1, \dots, d$, don't depend on the choice of B in (7) [9, Remark 2.1]. Let $\alpha_1 > \alpha_2 > \dots > \alpha_{d_0} = 0$ stand for the ordered sequence of the eigenvalues of V and let s_i stand for the multiplicity of

the eigenvalue α_i , $i = 1, \dots, d_0$. Let y be a realization of the observation vector Y . It can be shown that $s = (\sigma_1^2, \sigma_2^2)'$ is a solution to the system

$$(11) \quad \frac{\partial l(s, y)}{\partial \sigma_1^2} = 0, \quad \frac{\partial l(s, y)}{\partial \sigma_2^2} = 0$$

if and only if

$$(12) \quad \begin{aligned} \sum_{i=1}^{d-1} \frac{\nu_i m_i}{(m_i \sigma_1^2 + \sigma_2^2)^2} t_i &= \sum_{j=1}^{d_0-1} \frac{s_j \alpha_j}{\alpha_j \sigma_1^2 + \sigma_2^2}, \\ \sum_{i=1}^d \frac{\nu_i}{(m_i \sigma_1^2 + \sigma_2^2)^2} t_i &= \sum_{j=1}^{d_0} \frac{s_j}{\alpha_j \sigma_1^2 + \sigma_2^2}, \end{aligned}$$

where t_i , $i = 1, \dots, d$, are the quantities obtained as the result of the substitution $Y = y$ in (9), see [4, p. 286].

Let us observe that we may reparametrize the model (1) by defining

$$\sigma^2 := \sigma_1^2 + \sigma_2^2, \quad \rho := \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}.$$

Note that

$$(13) \quad \sigma_1^2 = \sigma^2 \rho, \quad \sigma_2^2 = \sigma^2 (1 - \rho).$$

Let us define the algebraic expression $\phi_\mu(\rho)$ by

$$\phi_\mu(\rho) := (\mu - 1)\rho + 1$$

and the rational algebraic expressions $H_1(\rho)$, $H_2(\rho)$ and $h(\rho)$ by

$$H_1(\rho) := \sum_{i=1}^{d-1} \frac{\nu_i m_i}{\phi_{m_i}^2(\rho)} t_i, \quad H_2(\rho) := \sum_{j=1}^{d_0-1} \frac{\alpha_j s_j}{\phi_{\alpha_j}(\rho)}, \quad h(\rho) := \frac{H_1(\rho)}{H_2(\rho)}.$$

The conditions (12), assuming that $\sigma_1^2 + \sigma_2^2 \neq 0$, are equivalent to:

$$(14) \quad \sigma^2 = h(\rho),$$

$$(15) \quad \sum_{i=1}^d \frac{\nu_i}{\phi_{m_i}^2(\rho)} t_i = h(\rho) \sum_{j=1}^{d_0} \frac{s_j}{\phi_{\alpha_j}(\rho)},$$

compare [4, p. 287]. It can be seen that if we find a solution of (15) with respect to ρ , we will be able to compute a solution to the system (12) using (14) and (13).

Remark 2.4. Let us observe that $\sigma^2 = \sigma_1^2 + \sigma_2^2 > 0$ if $s = (\sigma_1^2, \sigma_2^2)' \in \mathcal{S} = \{s : \sigma_1 \geq 0, \sigma_2^2 > 0\}$, so we will not "lose any solution" of (12) belonging to the parameter set \mathcal{S} if we use this new parametrization of the model (1).

2.3. Computing the maximum likelihood estimate — the polynomial approach

After reviewing the facts from [4] we are ready to present our main results concerning computing the maximum likelihood estimate of variance components in the model (1) via finding the real roots of an appropriately defined polynomial.

Let us define the polynomials

$$Q_1(\rho) := \prod_{i=1}^{d-1} \phi_{m_i}^2(\rho), \quad Q_2(\rho) := \prod_{j=1}^{d_0-1} \phi_{\alpha_j}(\rho).$$

Let $P_1(\rho)$, $P_2(\rho)$, $P_3(\rho)$ and $P_4(\rho)$ be the polynomials obtained as the result of simplifying the rational algebraic expressions

$$\begin{aligned} R_1(\rho) &:= \sum_{i=1}^d \frac{\nu_i}{\phi_{m_i}^2(\rho)} t_i Q_1(\rho) (1 - \rho)^2, & R_2(\rho) &:= \sum_{j=1}^{d_0-1} \frac{\alpha_j s_j}{\phi_{\alpha_j}(\rho)} Q_2(\rho), \\ R_3(\rho) &:= \sum_{i=1}^{d-1} \frac{\nu_i m_i}{\phi_{m_i}^2(\rho)} Q_1(\rho) t_i, & R_4(\rho) &:= \sum_{j=1}^{d_0} \frac{s_j}{\phi_{\alpha_j}(\rho)} Q_2(\rho) (1 - \rho)^2, \end{aligned}$$

respectively. Put $P(\rho) := P_1(\rho)P_2(\rho) - P_3(\rho)P_4(\rho)$. Let us note that if ρ_0 is a root of $P(\rho)$ and $\rho_0 \notin \{m_1, \dots, m_d\} \cup \{\alpha_1, \dots, \alpha_{d_0}\}$, then ρ_0 is also a solution to the equation (15). It can be shown that

Theorem 2.5.

- (a) *The degree of $P(\rho)$ is less than or equal to $2d + d_0 - 4$.*
- (b) *If the condition (6) is satisfied, then $P(\rho)$ is a non-zero polynomial.*

In order to prove this theorem we need the following

Lemma 2.6. *Let us assume that for y , the given realization of the observation vector Y , the condition (6) is satisfied. If the sequence of pairs $(\theta_1^j, \theta_2^j)'$, $\theta_1^j \geq 0$, $\theta_2^j > 0$, $j \in \mathbb{N}$, satisfies the condition*

$$(16) \quad \lim_{n \rightarrow \infty} \frac{\theta_1^n}{\theta_2^n} = \infty,$$

then $l((\theta_1^j, \theta_2^j)', y) \rightarrow \infty$ for $j \rightarrow \infty$.

Proof. The proof follows from the fact that $l_0(\beta, s, y) \leq l(s, y)$ for $\beta \in \mathbb{R}^p$, $s \in \mathcal{S}$ and $y \in \mathbb{R}^n$ and from [7, Proposition 2.4]. ■

Proof of Theorem 2.5. (a) This part of the theorem follows from the fact that $\text{deg}(P_1(\rho)) \leq 2d-2$, $\text{deg}(P_2(\rho)) \leq d_0-2$, $\text{deg}(P_3(\rho)) \leq 2d-4$ and $\text{deg}(P_4(\rho)) \leq d_0$.

(b) It can be seen that the function L defined by $L(\rho) := l(\rho h(\rho), (1-\rho)h(\rho))$ is differentiable on $(0, 1)$. Let us assume that, for the given y , $P(\rho)$ is the zero polynomial. This implies that $(\rho h(\rho), (1-\rho)h(\rho))'$ is a solution to the likelihood equations if $\rho \in (0, 1)$ and L is constant on $(0, 1)$. Now let (ρ_n) be a sequence of numbers belonging to $(0, 1)$ converging to 1. From Lemma 2.6 follows that $L(\rho_n) \rightarrow \infty$, and we have obtained a contradiction. ■

We are now ready to state the following

Theorem 2.7. *Let us assume that the model (1) satisfies the condition (10) and y , the given realization of the vector Y , satisfies (6). Then:*

- (a) *The set of all solutions to the maximum likelihood equation system (12) that belong to the parameter space \mathcal{S} is a subset of the finite set Ψ_1 constructed by:*
 - (i) *finding all the real roots of the polynomial $P(\rho)$ that lie in the set $[0, 1) \setminus (\{m_1, \dots, m_d\} \cup \{\alpha_1, \dots, \alpha_{d_0}\})$;*
 - (ii) *computing solutions to (12) that correspond to the elements obtained in (i) according to the formula (13).*

- (b) *The maximum likelihood estimate of s , denoted by s_{ML} , is given by*

$$s_{ML} := \operatorname{argmax}_{s \in \Psi} l(s, y),$$

where $\Psi := (\Psi_1 \cap \mathcal{S}) \cup \{(0, s_*^2)'\}$,

$$s_*^2 := \frac{1}{n}(y - Xb)'(y - Xb),$$

and b stands for the ordinary least squares estimate of β .

Proof. The part (a) follows from the fact that under the assumptions of the theorem $\sigma^2 = \sigma_1^2 + \sigma_2^2$ is positive. To prove the part (b) it suffices to observe that $(0, s_*^2)'$ is the maximum likelihood estimate of s in the model (1) with σ_1^2 fixed to 0 if the assumptions of the theorem are satisfied, see [11, p. 37]. ■

3. RESTRICTED MAXIMUM LIKELIHOOD (REML) ESTIMATION

The REML estimator of s in the model (1) is defined as the maximizer of the likelihood function based on $z := BY$, where B is any matrix satisfying the conditions (7), see [8, p. 13]. Since $z \sim N(\mathbf{0}^{(n-p)}, B\Sigma(s)B')$, the REML equation system has the form

$$(17) \quad \begin{aligned} \sum_{i=1}^{d-1} \frac{\nu_i m_i}{(m_i \sigma_1^2 + \sigma_2^2)^2} t_i &= \sum_{j=1}^{d-1} \frac{\nu_j m_j}{m_j \sigma_1^2 + \sigma_2^2}, \\ \sum_{i=1}^d \frac{\nu_i}{(m_i \sigma_1^2 + \sigma_2^2)^2} t_i &= \sum_{j=1}^d \frac{\nu_j}{m_j \sigma_1^2 + \sigma_2^2}, \end{aligned}$$

where t_i , $i = 1, \dots, d$, are (as in Section 2) the quantities obtained as the result of the substitution $Y = y$ in (9), see [4, p. 291], and y stands for a given realization of the observation vector Y .

A necessary and sufficient condition for the existence of the REML estimate of s in the model (1) is

$$(18) \quad y \notin \mathcal{M}(MZ) = \mathcal{M}(MV),$$

where M and Z are as in Subsection 2.1, see [2, Theorem 3.4] and [7, Chapter 3].

Computing the REML estimate of s can be reduced to finding all the real roots of the appropriately defined polynomial in a similar way as it was the case with computing the maximum likelihood estimate of s . In order to construct such a polynomial let us define the following rational algebraic expressions:

$$(19) \quad R_1^*(\rho) := \sum_{i=1}^d \frac{\nu_i}{\phi_{m_i}^2(\rho)} t_i \sum_{j=1}^{d-1} \frac{m_j \nu_j}{\phi_{m_j}(\rho)}, \quad R_2^*(\rho) := \sum_{i=1}^{d-1} \frac{\nu_i m_i}{\phi_{m_i}^2(\rho)} t_i \sum_{j=1}^d \frac{\nu_j}{\phi_{m_j}(\rho)},$$

$$(20) \quad Q_1^*(\rho) := \prod_{i=1}^d \phi_{m_i}^2(\rho), \quad \text{and} \quad P_0^*(\rho) := (R_1^*(\rho) - R_2^*(\rho))Q_1^*(\rho).$$

Let us note that $P^*(\rho)$ can be rewritten as follows:

$$(21) \quad P_0^*(\rho) = \sum_{i=1}^d \sum_{\substack{j=1 \\ i \neq j}}^d \frac{\nu_i \nu_j (m_j - m_i)}{\phi_{m_i}^2(\rho) \phi_{m_j}(\rho)} t_i Q_1^*(\rho).$$

It can be seen that $P_0^*(\rho)$ simplifies to a polynomial which we will denote by $P^*(\rho)$.

Theorem 3.1.

- (a) *The degree of the polynomial $P^*(\rho)$ does not exceed $2d - 3$.*
- (b) *If the condition (18) is satisfied, then $P^*(\rho)$ is a non-zero polynomial.*

In order to prove this theorem we need the following

Lemma 3.2. *Let us assume that for a given observation vector y the condition (18) is satisfied. If the sequence of pairs $(\theta_1^j, \theta_2^j)'$, $\theta_1^j \geq 0$, $\theta_2^j > 0$, $j \in \mathbb{N}$, satisfies the condition*

$$(22) \quad \lim_{n \rightarrow \infty} \frac{\theta_1^n}{\theta_2^n} = \infty,$$

then $l_0(\mathbf{0}^{(n-p)}, (\theta_1^j, \theta_2^j)', By) \rightarrow \infty$ for $j \rightarrow \infty$.

Proof. See [7, Proposition 3.2]. ■

Proof of Theorem 3.1. The part (a) follows immediately from the fact that $P^*(\rho)$ can be presented in the form (21). The part (b) can be proved by analogy with the proof of part (b) of Theorem 2.5 (using Lemma 3.2 instead of Lemma 2.6). ■

Let b denote the ordinary least squares estimate of β and let

$$s_0^2 := \frac{1}{n - p} (y - Xb)'(y - Xb).$$

If the condition (18) is satisfied, then s_0^2 , the mean squared estimate of σ_2^2 , is also its REML estimate in the model (1) with σ_1^2 fixed to 0 [1, p. 307].

We are now ready to state the following

Theorem 3.3. *Let us assume that the model (1) satisfies the condition (10) and y , a given realization of the vector Y , satisfies the condition (18). Then:*

- (a) *The set of all solutions to the REML equation system (17) that belong to the parameter space \mathcal{S} is a subset of the finite set Ξ_1 constructed by:*
 - (i) *finding all the real roots of the polynomial $P^*(\rho)$ that lie in the set $[0, 1) \setminus \{m_1, \dots, m_d\}$;*
 - (ii) *computing solutions to (17) that correspond to the elements obtained in (i) according to the formula (13).*

(b) *The REML estimate of s , denoted by s_{REML} , is given by*

$$s_{REML} = \operatorname{argmax}_{s \in \Xi} l_0(\mathbf{0}^{(n-p)}, s, By),$$

where $\Xi := (\Xi_1 \cap \mathcal{S}) \cup \{(0, s_0^2)'\}$ and B is any matrix satisfying the conditions (7).

Proof. The theorem can be proved by analogy with Theorem 2.7. ■

Remark 3.4. A matrix B satisfying the conditions (7) can be obtained by finding an orthonormal basis of $\mathcal{M}(M)$: A matrix whose columns are the vectors from this basis satisfies the conditions (7).

4. CONCLUDING REMARKS

The proposed approach to the problem of computing the maximum likelihood estimate and the REML estimate of the vector of variance components in the considered class of mixed linear normal models involves calculating the real roots of a polynomial and diagonalization of real symmetric matrices. Computing the coefficients of this polynomial within a given tolerance may pose a challenge. To tackle this task, one can use the results from [5, Chapter 8] concerning the error bounds for the approximate solutions to the symmetric eigenvalue problem. Thus, it can be expected that the implementation of the methods proposed in this work will result in obtaining reliable procedures for computing the maximum likelihood estimate and the REML estimate of the vector of variance components in the considered class of mixed models.

REFERENCES

- [1] R. Christensen, *Plane answers to complex questions: the theory of linear models*, 4th Edition (Springer, New York, 2011). doi:10.1007/978-1-4419-9816-3
- [2] 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.
- [3] S. Gnot, A. Michalski and A. Urbańska-Motyka, *On some properties of ML and REML estimators in mixed normal models with two variance components*, *Discuss. Math. Prob. and Stat.* **24** (2004) 109–126.
- [4] S. Gnot, D. Stemann, G. Trenkler and A. Urbańska-Motyka, *Maximum likelihood estimation in mixed normal models with two variance components*, *Statistics* **36** (2002) 283–302. doi:10.1080/02331880213197

- [5] G. Golub and C. Van Loan, *Matrix Computations*, 4th Edition (The John Hopkins University Press, Baltimore, 2013).
- [6] E. Gross, M. Drton and S. Petrović, *Maximum likelihood degree of variance components models*, *Electronic J. Stat.* **6** (2012) 993–1016. doi:10.1214/12-EJS702
- [7] M. Grządziel and A. Michalski, *A note on the existence of maximum likelihood estimates in variance components models*, accepted for publication in *Discuss. Math. Prob. and Stat.* (2014).
- [8] J. Jiang, *Linear and Generalized Linear Mixed Models and Their Applications* (Springer, New York, 2007). doi:10.1007/978-0-387-47946-0
- [9] A. Olsen, J. Seely and D. Birkes, *Invariant quadratic estimation for two variance components*, *Ann. Stat.* **4** (1976) 878–890. doi:10.1214/aos/1176343586
- [10] C. R. Rao and J. Kleffe, *Estimation of Variance Components and Applications* (North Holland, New York, 1988). doi:10.1002/bimj.4710320518
- [11] C. R. Rao and H. Toutenburg, *Linear Models: Least Squares and Alternatives*, 2nd Edition (Springer, New York, 1999).

Received 18 October 2014
Revised 26 November 2014

