Applications of Mathematics

Júlia Volaufová

On variance of the two-stage estimator in variance-covariance components model

Applications of Mathematics, Vol. 38 (1993), No. 1, 1--9

Persistent URL: http://dml.cz/dmlcz/104529

ON VARIANCE OF THE TWO-STAGE ESTIMATOR IN VARIANCE-COVARIANCE COMPONENTS MODEL

JÚLIA VOLAUFOVÁ, Bratislava

(Received October 9, 1989)

Summary. The paper deals with a linear model with linear variance-covariance structure, where the linear function of the parameter of expectation is to be estimated. The two-stage estimator is based on the observation of the vector Y and on the invariant quadratic estimator of the variance-covariance components. Under the assumption of symmetry of the distribution and existence of finite moments up to the tenth order one approach to determining the upper bound for the difference in variances of the estimators is proposed, which uses the estimated covariance matrix instead of the real one.

Keywords: two-stage estimator, symmetrically distributed estimator, unbiased estimator

AMS classification: 62J10

1. INTRODUCTION

In most situations when treating the linear model the main objective in analyzing the data is to make inference about linear combinations of the unknown fixed parameters of mean. If the variance-covariance matrix of the vector of observations is known the problem has been discussed by many authors (see e.g. C.R. Rao (1973)).

If the variance-covariance components that are connected with the variance matrix are unknown, then the traditional procedure in estimating the linear combinations (functions) of the unknown mean parameter is first to estimate the variance components and then to consider these estimators as if they were the true values. In general, the statistical properties of such two-stage estimators are not specified. Several authors studied unbiasedness of such estimators, e.g. Kackar and Harville (1981) showed that the procedure gives unbiased estimators provided the distribution of the data vector is symmetric about its expected value and provided the variance component estimators are translation invariant and are even functions of the data vector. Such kind of symmetry arguments were used by Kakwani (1967) and later e.g. by Don and Magnus (1980), Seely and Hogg (1982), and Khatri and Shah (1981) who investigated the estimation of fixed effects in a mixed model for growth curves, and except unbiasedness they derived an expression for the increase in variance due to substitution of the variances by their estimates.

Here, we attempt to give an expression for the approximation of the variance of the two-stage estimator of the linear function of mean-parameter, provided the distribution of the observation vector is symmetric about its expectation and the variance component estimators are quadratic translation-invariant functions of data vectors. Further we assume that there exist finite moments up to the tenth order.

2. Estimation of linear functions of mean parameter

Most of the linear statistical models can be viewed as a special case of a model given as follows:

$$Y = X\beta + \varepsilon,$$

where Y is an $n \times 1$ observable random vector; X is an $n \times k$ matrix of known elements, β is a $k \times 1$ fixed vector of unknown and unobservable parameters; ε is an $n \times 1$ random error vector.

It is assumed that ε is symmetrically distributed around zero and that the variance-covariance matrix of ε is of the form $\sum_{i=1}^{p} \vartheta_i V_i$ with $\vartheta = (\vartheta_1, \ldots, \vartheta_p)'$ a $p \times 1$ vector of unknown parameters, $\vartheta \in \Theta \subset \mathbb{R}^p$, provided $V(\vartheta) = \sum_{i=1}^{p} \vartheta_i V_i$ is positive definite for all $\vartheta \in \Theta$.

Taking E to be the expectation operator the following is assumed:

$$E(\varepsilon) = 0; \quad E(\varepsilon\varepsilon') = V(\vartheta) = \sum_{i=1}^{p} \vartheta_i V_i.$$

From the symmetry of the distribution all matrices of higher odd moments are identically zero matrices, e.g. the matrix of the third moments is $E(\varepsilon \otimes \varepsilon\varepsilon') = 0$, etc. Analogously to the variance matrix we can denote the matrices of higher even orders as

$$E(\varepsilon\varepsilon'^{2\otimes}) = \psi; \quad E(\varepsilon\varepsilon'^{3\otimes}) = \xi;$$

$$E(\varepsilon\varepsilon'^{4\otimes}) = \chi; \quad \text{and} \quad E(\varepsilon\varepsilon'^{5\otimes}) = \omega.$$

Here $A \otimes B$ stands for the Kronecker product of A and B defined as $A \otimes B = (a_{ij}B)$ if $A = (a_{ij})$. The matrices ψ, ξ, χ and ω in general depend on the second order moments of the vector ε which in particular means that they are dependent on the vector parameter ϑ.

The following notation partly due to Kleffe (1978) and Volaufová, Witkovský and Bognárová (1988), is used.

Let A be an $n \times n$ matrix and B an $n^2 \times n^2$ matrix decomposed into n^2 $n \times n$ matrices (B_{ij}). Then the operation $A \mathbin{\text{⮡}} B$ denotes the $n \times n$ matrix with i,j-th element $\operatorname{tr} AB_{ij}$, where "tr" stands for the trace of a matrix.

This matrix operation can be expressed in terms of the Hadamard product of matrices defined as $A * B = (A_{ij} \otimes B_{ij})$, with partitioned matrices A and B. Let 1 denote the vector of n units $1 = (1, \ldots, 1)'$. Let the matrix B be defined as above. Then the matrix $\operatorname{tr}_n B$ is defined as the matrix with i,j-th element

$$(\operatorname{tr}_n B)_{ij} = (\operatorname{tr} B_{ij}).$$

Now we can write

$$A \mathbin{\text{⮡}} B = \operatorname{tr}_n[(1\,1' \otimes A) * B].$$

Let $f(\beta) = p'\beta$ be a linear function of β, $p \in \mathbb{R}^k$ a known vector. $f(\beta)$ is to be estimated. It is well known that if ϑ is a known vector parameter, then $f(\beta)$ is estimated by a linear function of Y. $f(\beta)$ is said to be linearly estimable if there exists a linear unbiased estimate for $f(\beta)$ and this occurs if and only if $p \in \mathcal{R}(X')$, i.e. p is a linear combination of the rows of the matrix X. Then

$$\widetilde{p'\beta} = (X'V(\vartheta)^{-1}X)^{-1}X'V(\vartheta)^{-1}Y$$

is the best linear unbiased estimator (BLUE) of $p'\beta$ (see C.R. Rao (1973)).

Let the variance components ϑ_j be estimated by the quadratic invariant unbiased estimators, say $\hat{\vartheta}_j = Y'A_jY$, where $A_jX = 0$, A_j are symmetric $n \times n$ matrices, and $trA_jV_i = \delta_{ij}$, where δ_{ij} means the Kronecker δ. Then for $p \in \mathcal{R}(X')$ the following is valid:

$$\widehat{p'\beta} = p'(X'V(\hat{\vartheta})^{-1}X)^{-1}X'V(\hat{\vartheta})^{-1}Y$$

is the two-stage estimator of the linear function $f(\beta) = p'\beta$. It is known (e.g. Seely and Hogg (1982)) that $\widehat{p'\beta}$ is an unbiased estimator of $p'\beta$.

We shall further assume that the rank of the matrix X is full, i.e. $r(X) = k$, which implies that for all $p \in \mathbb{R}^k$, $p'\beta$ is an unbiasedly estimable function.

Proposition 1. *Under the given assumptions we have for all $L \in \mathbb{R}^n$*

$$\operatorname{cov}(L'Y, Y'AY) = 0, \quad \textit{whenever} \quad AX = 0.$$

P r o o f. For the proof it is enough to consider the relation

$$\mathrm{cov}(L'Y, Y'AY) = 2L'V(\vartheta)AX\beta.$$

□

The main idea of the approach used in this paper to give the upper bound of the variance of $\widehat{p'\beta}$ is to express $\widehat{p'\beta} - p'\beta$ according to Taylor's expression. Obviously the following holds:

$$\begin{aligned}\widehat{p'\beta} - p'\beta &= p'(X'V(\hat\vartheta)^{-1}X)^{-1}X'V(\hat\vartheta)^{-1}(Y - X\beta)\\ &= p'(X'V(\hat\vartheta)^{-1}X)^{-1}X'V(\hat\vartheta)^{-1}\varepsilon \equiv U(\varepsilon, \hat\vartheta).\end{aligned}$$

We express Taylor's expression at the point $(\varepsilon, \vartheta_0)$, where ϑ_0 is the actual value of the parameter ϑ. Then we can write

$$(1)\quad U(\varepsilon,\hat\vartheta) = U(\varepsilon,\vartheta_0) + \left[\frac{\partial U(\varepsilon,\vartheta)}{\partial\vartheta'}\right]_{(\vartheta_0)}(\hat\vartheta - \vartheta_0) + \frac{1}{2}\left[\frac{\partial^2 U(\varepsilon,\vartheta)}{\partial\vartheta'\partial\vartheta'}\right]_{(\tau)}(\hat\vartheta - \vartheta_0)^{2\otimes},$$

where τ is defined from Taylor's theorem as $\tau = \theta\vartheta_0 + (1-\theta)\hat\vartheta$, $0 < \theta < 1$.

Obviously

$$E(U(\varepsilon,\vartheta_0)) = 0, \quad \mathrm{var}_{\vartheta_0}(U(\varepsilon,\vartheta_0)) = p'(X'V(\vartheta_0)^{-1}X)^{-1}p.$$

In our following considerations we shall use the notation

$$(2)\qquad A(\vartheta,\varepsilon) = \frac{\partial U(\varepsilon,\vartheta)}{\partial\vartheta'}, \quad B(\vartheta,\varepsilon) = \frac{1}{2}\left[\frac{\partial^2 U(\varepsilon,\vartheta)}{\partial\vartheta'\partial\vartheta'}\right],$$

and analogously when dealing with the derivative at a given point we shall use

$$(3)\qquad A(\vartheta_0,\varepsilon) = \left[\frac{\partial U(\varepsilon,\vartheta)}{\partial\vartheta'}\right]_{\vartheta=\vartheta_0}, \quad B(\tau,\varepsilon) = \frac{1}{2}\left[\frac{\partial^2 U(\varepsilon,\vartheta)}{\partial\vartheta'\partial\vartheta'}\right]_{\vartheta=\tau}.$$

Consequently, (1) can be presented as

$$U(\varepsilon,\hat\vartheta) = U(\varepsilon,\vartheta_0) + A(\vartheta_0,\varepsilon)(\hat\vartheta - \vartheta_0) + B(\tau,\varepsilon)(\hat\vartheta - \vartheta_0)^{2\otimes}.$$

For the next proposition see also Volaufová, Witkovský and Bognárová (1988).

Proposition 2. *For the derivatives $A(\vartheta,\varepsilon)$ and $B(\vartheta,\varepsilon)$ the following statements are valid:*

(a)

$$A(\vartheta,\varepsilon) = p'(X'V(\vartheta)^{-1}X)^{-1}X'V(\vartheta)^{-1}\left[-V_1(MV(\vartheta)M)^{+}\varepsilon,\dots,-V_p(MV(\vartheta)M)^{+}\varepsilon\right],$$

where $(MV(\vartheta)M)^{+} = V(\vartheta)^{-1}M_{V(\vartheta)}$, *and*

$$M_{V(\vartheta)} = I - X(X'V(\vartheta)^{-1}X)^{-1}X'V(\vartheta)^{-1} = I - P_{V(\vartheta)},$$

(b)

$$B(\vartheta,\varepsilon) = p'(X'V(\vartheta)^{-1}X)^{-1}X'V(\vartheta)^{-1}\left[V_1(MV(\vartheta)M)^{+}Z,\dots,V_p(MV(\vartheta)M)^{+}Z\right],$$

where $Z = [V_1(MV(\vartheta)M)^{+}\varepsilon,\dots,V_p(MV(\vartheta)M)^{+}\varepsilon]$.

Proof. The proof is obvious but needs tedious calculations. □

For the variance of the statistics $U(\varepsilon,\hat{\vartheta})$ the following relation is valid:

$$\begin{aligned}\operatorname{var}_{\vartheta_0}\left(U(\varepsilon,\hat{\vartheta})\right) &= \operatorname{var}_{\vartheta_0}(U(\varepsilon,\vartheta_0)) + \operatorname{var}_{\vartheta_0}\left(A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\right)\\ &+ \operatorname{var}_{\vartheta_0}\left(B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right) + 2\operatorname{cov}_{\vartheta_0}\left(U(\varepsilon,\vartheta_0), A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\right)\\ &+ 2\operatorname{cov}_{\vartheta_0}\left(U(\varepsilon,\vartheta_0), B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right)\\ &+ 2\operatorname{cov}_{\vartheta_0}\left(A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0), B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right).\end{aligned}$$

Using Schwarz' inequality we get

$$\begin{aligned}(4)\quad \left|\operatorname{var}_{\vartheta_0}\left(U(\varepsilon,\hat{\vartheta})\right) - \operatorname{var}_{\vartheta_0}(U(\varepsilon,\vartheta_0))\right| &\leqslant \operatorname{var}_{\vartheta_0}\left(A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\right)\\ &+ \operatorname{var}_{\vartheta_0}\left(B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right) + 2\left|\operatorname{cov}_{\vartheta_0}\left(U(\varepsilon,\vartheta_0), A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\right)\right|\\ &+ 2\left[\operatorname{var}_{\vartheta_0}\left(A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\right)\operatorname{var}_{\vartheta_0}\left(B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right)\right]^{1/2}\\ &+ 2\left[\operatorname{var}_{\vartheta_0}(U(\varepsilon,\vartheta_0))\operatorname{var}_{\vartheta_0}\left(B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right)\right]^{1/2}.\end{aligned}$$

The last inequality offers immediately an upper bound of the error in variance caused by the use of the estimated covariance matrix in the form $V(\hat{\vartheta})$ instead of the actual one $V(\vartheta_0)$ which will be denoted as V_0.

However, the choice of the vector τ needs some discussion. It is clear that the vector $\tau = \theta\vartheta_0 + (1-\theta)\hat{\vartheta}$ is a random vector. For a very rough approximation in the expressions for variances and covariances we shall fix the vector τ_0. It is the subject of further investigation to study the behaviour of the upper bound which regard to the change of the vector τ_0.

Utilizing some technical approach the following theorems give the exact forms of the terms appearing in the expression for the upper bound of the difference of variances.

We recall that the invariant and unbiased estimators of the elements of the vector ϑ are given by $\hat{\vartheta}_i = Y'A_iY$ for $i = 1, 2, \ldots, p$. The matrices A_i satisfy the assumptions $A_iX = 0$ and $\operatorname{tr} A_iV_j = \delta_{ij}$, which ensures the invariance and the unbiasedness of the estimators.

Theorem 1. *The variance of the variable* $A(\vartheta_0, \varepsilon)(\hat{\vartheta} - \vartheta_0)$ *at* ϑ_0 *is given by*

$$
\begin{aligned}
&\operatorname{var}_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta} - \vartheta_0)\big)\\
(5)\quad &= p'(X'V_0^{-1}X)^{-1}X'V_0^{-1}\sum_{i,j} V_i(MV_0M)^+J_{ij}(MV_0M)^+V_jV_0^{-1}X(X'V_0^{-1}X)^{-1}p,
\end{aligned}
$$

where the matrices J_{ij} *are given by*

$$
J_{ij} = A_i \otimes A_j \llcorner\!\!\to \xi(\vartheta_0) - \vartheta_{0i}A_j \llcorner\!\!\to \psi(\vartheta_0) - \vartheta_{0j}A_i \llcorner\!\!\to \psi(\vartheta_0) + \vartheta_{0i}\vartheta_{0j}V_0 \quad i,j = 1,\ldots,p.
$$

Proof. The variance $\operatorname{var}_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta} - \vartheta_0)\big)$ from the definition is expressed as

$$
(6)\qquad E_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta} - \vartheta_0)(\hat{\vartheta} - \vartheta_0)'A(\vartheta_0,\varepsilon)'\big) - \big[E_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta} - \vartheta_0)\big)\big]^2.
$$

First we investigate the second term. According to Proposition 1 (a)

$$
\begin{aligned}
&E_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta} - \vartheta_0)\big)\\
&\quad = E_{\vartheta_0}\big(-p'(X'V_0^{-1}X)^{-1}X'V_0^{-1}\big[V_1(MV_0M)^+\varepsilon,\ldots,V_p(MV_0M)^+\varepsilon\big](\hat{\vartheta} - \vartheta_0)\big)\\
&\quad = -p'(X'V_0^{-1}X)^{-1}X'V_0^{-1}\sum_{i=1}^{p} V_i(MV_0M)^+\big[E_{\vartheta_0}(\varepsilon\varepsilon'A_i\varepsilon) - \vartheta_{0i}E_{\vartheta_0}(\varepsilon)\big] = 0.
\end{aligned}
$$

The first term in (6) can be expressed with regard to Proposition 1 (a) as

$$
\begin{aligned}
&E_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta} - \vartheta_0)(\hat{\vartheta} - \vartheta_0)'A(\vartheta_0,\varepsilon)'\big)\\
&\quad = p'(X'V_0^{-1}X)^{-1}X'V_0^{-1}\\
&\qquad \times \sum_{i,j} V_i(MV_0M)^+E_{\vartheta_0}\big[(\hat{\vartheta}_i - \vartheta_{0i})(\hat{\vartheta}_j - \vartheta_{0j})\varepsilon\varepsilon'\big](MV_0M)^+V_j\\
&\qquad \times V_0^{-1}X(X'V_0^{-1}X)^{-1}p.
\end{aligned}
$$

The expectation in the middle of the last term is evaluated as follows:

$$
\begin{aligned}
&E_{\vartheta_0}\big[(\hat{\vartheta}_i - \vartheta_{0i})(\hat{\vartheta}_j - \vartheta_{0j})\varepsilon\varepsilon'\big]\\
&\quad = E_{\vartheta_0}\big(\hat{\vartheta}_i\hat{\vartheta}_j\varepsilon\varepsilon' - \vartheta_{0i}\hat{\vartheta}_j\varepsilon\varepsilon' - \vartheta_{0j}\hat{\vartheta}_i\varepsilon\varepsilon' + \vartheta_{0i}\vartheta_{0j}\varepsilon\varepsilon'\big)\\
&\quad = E_{\vartheta_0}(\hat{\vartheta}_i\hat{\vartheta}_j\varepsilon\varepsilon') - \vartheta_{0i}E_{\vartheta_0}(\hat{\vartheta}_j\varepsilon\varepsilon') - \vartheta_{0j}E_{\vartheta_0}(\hat{\vartheta}_i\varepsilon\varepsilon') + \vartheta_{0i}\vartheta_{0j}E_{\vartheta_0}(\varepsilon\varepsilon')\\
&\quad = E_{\vartheta_0}\big[(\operatorname{tr}(A_i \otimes A_j)(\varepsilon\varepsilon' \otimes \varepsilon\varepsilon'))\varepsilon\varepsilon'\big] - \vartheta_{0i}E_{\vartheta_0}\big[(\operatorname{tr} A_j\varepsilon\varepsilon')\varepsilon\varepsilon'\big]\\
&\qquad - \vartheta_{0j}E_{\vartheta_0}\big[(\operatorname{tr} A_i\varepsilon\varepsilon')\varepsilon\varepsilon'\big] + \vartheta_{0i}\vartheta_{0j}V_0\\
&\quad = A_i \otimes A_j \llcorner\!\!\to \xi(\vartheta_0) - \vartheta_{0i}A_j \llcorner\!\!\to \psi(\vartheta_0) - \vartheta_{0j}A_i \llcorner\!\!\to \psi(\vartheta_0) + \vartheta_{0i}\vartheta_{0j}V_0 = J_{ij},
\end{aligned}
$$

which, after substitution, implies the statement of the theorem. □

Theorem 2. *The variance of the variable* $B(\tau_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}$ *at* ϑ_0 *is given by*

$$
\begin{aligned}
(7)\quad & \operatorname{var}_{\vartheta_0}\left(B(\tau_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right)\\
&= p'(X'V_{\tau_0}^{-1}X)^{-1}X'V_{\tau_0}^{-1}\\
&\times \sum_{i,j,k,l} V_i(MV_{\tau_0}M)^+V_j(MV_{\tau_0}M)^+K_{ijkl}(MV_{\tau_0}M)^+V_k(MV_{\tau_0}M)^+V_l\\
&\times V_{\tau_0}^{-1}X(X'V_{\tau_0}^{-1}X)^{-1}p,
\end{aligned}
$$

where K_{ijkl} *is given as a sum of products of the type "*$\llcorner\!\!\to$*" involving all matrices of even moments up to the tenth order, and the explicit formula is*

$$
\begin{aligned}
K_{ijkl} = {} & A_i\otimes A_j\otimes A_k\otimes A_l \llcorner\!\!\to \omega(\vartheta_0)\\
& - \vartheta_{0i}A_j\otimes A_k\otimes A_l \llcorner\!\!\to \chi(\vartheta_0) - \vartheta_{0j}A_i\otimes A_k\otimes A_l \llcorner\!\!\to \chi(\vartheta_0)\\
& - \vartheta_{0k}A_j\otimes A_i\otimes A_l \llcorner\!\!\to \chi(\vartheta_0) - \vartheta_{0l}A_j\otimes A_k\otimes A_i \llcorner\!\!\to \chi(\vartheta_0)\\
& + \vartheta_{0i}\vartheta_{0j}A_k\otimes A_l \llcorner\!\!\to \xi(\vartheta_0) + \vartheta_{0i}\vartheta_{0k}A_j\otimes A_l \llcorner\!\!\to \xi(\vartheta_0) + \vartheta_{0i}\vartheta_{0l}A_j\otimes A_k \llcorner\!\!\to \xi(\vartheta_0)\\
& + \vartheta_{0j}\vartheta_{0k}A_i\otimes A_l \llcorner\!\!\to \xi(\vartheta_0) + \vartheta_{0j}\vartheta_{0l}A_k\otimes A_i \llcorner\!\!\to \xi(\vartheta_0) + \vartheta_{0k}\vartheta_{0l}A_i\otimes A_j \llcorner\!\!\to \xi(\vartheta_0)\\
& - \vartheta_{0i}\vartheta_{0j}\vartheta_{0k}A_l \llcorner\!\!\to \psi(\vartheta_0) - \vartheta_{0i}\vartheta_{0j}\vartheta_{0l}A_k \llcorner\!\!\to \psi(\vartheta_0) - \vartheta_{0i}\vartheta_{0k}\vartheta_{0l}A_j \llcorner\!\!\to \psi(\vartheta_0)\\
& - \vartheta_{0j}\vartheta_{0k}\vartheta_{0l}A_i \llcorner\!\!\to \psi(\vartheta_0) + \vartheta_{0i}\vartheta_{0j}\vartheta_{0k}\vartheta_{0l}V_0.
\end{aligned}
$$

Proof. The proof goes along the same lines as the previous one:

$$
\begin{aligned}
(8)\qquad & \operatorname{var}_{\vartheta_0}\left(B(\tau_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right)\\
&= E_{\vartheta_0}\left[B(\tau_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right]^2 - \left[E_{\vartheta_0}\left(B(\tau_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right)\right]^2.
\end{aligned}
$$

As before, it can be shown that the second term in (8) vanishes. For the first term we have

$$
\begin{aligned}
& E_{\vartheta_0}\left[B(\tau_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\right]^2\\
&= p'(X'V_{\tau_0}^{-1}X)^{-1}X'V_{\tau_0}^{-1}\Bigg\{\sum_{i,j,k,l} V_i(MV_{\tau_0}M)^+V_j(MV_{\tau_0}M)^+\\
&\times E_{\vartheta_0}\left((\hat{\vartheta}_i-\vartheta_{0i})(\hat{\vartheta}_j-\vartheta_{0j})(\hat{\vartheta}_k-\vartheta_{0k})(\hat{\vartheta}_l-\vartheta_{0l})\varepsilon\varepsilon'\right)\\
&\times (MV_{\tau_0}M)^+V_k(MV_{\tau_0}M)^+V_l\Bigg\}V_{\tau_0}^{-1}X(X'V_{\tau_0}^{-1}X)^{-1}p.
\end{aligned}
$$

The expression with the expectation can be denoted as the matrix K_{ijkl} from the statement of the theorem if we use the symbol "$\llcorner\!\!\to$" and the matrices up to the 10th order moments of the vector ε. □

Theorem 3. *The covariance of the variables $U(\varepsilon,\vartheta_0)$ and $A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)$ at ϑ_0 is given by*

$$\operatorname{cov}_{\vartheta_0}\big(U(\varepsilon,\vartheta_0),A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\big)$$
$$=p'(X'V_0^{-1}X)^{-1}X'V_0^{-1}\Big(\sum_i V_i(MV_0M)^+A_i\,\llcorner\!\!\rightarrow\psi(\vartheta_0)\Big)V_0^{-1}X(X'V_0^{-1}X)^{-1}p.$$

Proof. For the proof it is enough to express $E_{\vartheta_0}\big(\hat{\vartheta}_i\varepsilon\varepsilon'-\vartheta_{0i}\varepsilon\varepsilon'\big)$, which leads to $A_i\,\llcorner\!\!\rightarrow\psi(\vartheta_0)$. □

Summarizing the previous considerations we get the following theorem.

Theorem 4. *Under the conditions given above the upper bound for the increase of the variance of the unbiased estimator of $p'\beta$ using the estimated value of ϑ instead of the actual value ϑ_0 is given by (4).*

In that case $\operatorname{var}_{\vartheta_0}\big(A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\big)$ *is given by (5),* $\operatorname{var}_{\vartheta_0}\big(B(\tau,\varepsilon)(\hat{\vartheta}-\vartheta_0)^{2\otimes}\big)$ *is given by (7), and* $\operatorname{cov}_{\vartheta_0}\big(U(\varepsilon,\vartheta_0),A(\vartheta_0,\varepsilon)(\hat{\vartheta}-\vartheta_0)\big)$ *is given by (9).*

Note. The crucial point is to express the matrices of even higher order. A very frequent situation is that the vector ε is normally distributed. Then the elements $\{\omega(\vartheta_0)\}_{ij,kl,mn,op,rs}$ of the matrix $\omega(\vartheta_0)$ can be derived using the characteristic function of the vector ε in the form

$$\frac{\partial^{10}\varphi(\boldsymbol{t})}{\partial t_i\partial t_j\partial t_k\partial t_l\partial t_m\partial t_n\partial t_o\partial t_p\partial t_r\partial t_s}\bigg|_{\boldsymbol{t}=0},$$

where the characteristic function $\varphi(\boldsymbol{t})=\exp\left(-\frac{1}{2}\boldsymbol{t}'V_0\boldsymbol{t}\right)$, $\boldsymbol{t}\in\mathbb{R}^n$.

For example, the matrix $\psi(\vartheta_0)=E_{\vartheta_0}(\varepsilon\varepsilon'\otimes\varepsilon\varepsilon')$, $\psi_{ik}=E_{\vartheta_0}(\varepsilon_i\varepsilon_k\varepsilon\varepsilon')$, $\psi_{ij,kl}=E_{\vartheta_0}(\varepsilon_i\varepsilon_k\varepsilon_j\varepsilon_l)$ is given by its entries as

$$\psi_{ij,kl}=v_{ij}v_{kl}+v_{ik}v_{jl}+v_{il}v_{kj},\quad i,j,k,l=1,\dots,n,$$

where v_{ij}, $i,j=1,\dots,n$ are the elements of the matrix V_0. For more details see Kubáček (1988).

References

[1] *F.J.H. Don and J.R. Magnus*: On the unbiasedness of iterated GLS estimators, Commun. Statist.-Theory Meth. *A9(5)* (1980), 519–527.
[2] *R.N. Kackar and D.A. Harville*: Unbiasedness of two-stage estimation and prediction procedures for mixed linear models, Commun. Statist.-Theory Meth. *A10(3)* (1981), 1249–1261.
[3] *N.C. Kakwani*: The unbiasedness of Zellner's seemingly unrelated regression equations estimators, Journal of the American Statistical Association *67* (1967), 141–142.
[4] *C.G. Khatri and K.R. Shah*: On the unbiased estimation of fixed effects in a mixed model for growth curves, Commun. Statist.-Theory Meth. *A10(4)* (1980), 401–406.
[5] *J. Kleffe*: Simultaneous estimation of expectation and covariance matrix in linear models, Math. Operationsforsch. Statist., Series Statistics *9(3)* (1978), 443–478.
[6] *L. Kubáček*: Foundations of Estimation Theory, Volume 9 of *Fundamental Studies in Engineering*, first edition, Elsevier, Amsterdam, Oxford, New York, Tokyo, 1988.
[7] *C.R. Rao*: Linear Statistical Inference and Its Applications, John Wiley, New York, first edition, 1973.
[8] *C.R. Rao and S.K. Mitra*: Generalized Inverse of Matrices and Its Applications, first edition, John Wiley & Sons, New York, London, Sidney, Toronto, 1971.
[9] *J. Seely and R.V. Hogg*: Symmetrically distributed and unbiased estimators in linear models, Commun. Statist.-Theory Meth. *A11(7)* (1982), 721–729.
[10] *J. Volaufová, V. Witkovský, and M. Bognárová*: On the confidence region of parameter of mean in mixed linear models, Poster at European meeting of Statisticians, Berlin, August 1988.

Súhrn

DISPERZIA DVOJETAPOVÉHO ODHADU V MODELI
S VARIANČNO-KOVARIANČNÝMI KOMPONENTAMI

JÚLIA VOLAUFOVÁ

V práci je odvodená horná hranica pre odhad chyby, ktorej sa dopúšťame, keď namiesto skutočnej kovariančnej matice v lineárnom modeli s variančno-kovariančnými komponentami použijeme jej invariantný kvadratický odhad, ktorý dosadíme do optimálneho odhadu lineárnej funkcie parametrov strednej hodnoty. Výsledok je odvodený za predpokladu symetrie rozdelenia a existencie konečných momentov desiateho rádu.

Author's address: *Júlia Volaufová*, Institute of Measurement Science of the SAS, Dúbravská 9, 842 19 Bratislava.