Mathematica Slovaca

Lubomír Kubáček
Repeated regression experiment and estimation of variance components

Mathematica Slovaca, Vol. 34 (1984), No. 1, 103--114

Persistent URL: http://dml.cz/dmlcz/128793

Math. Slovaca 34, 1984, No. 1, 103—114

REPEATED REGRESSION EXPERIMENT AND ESTIMATION OF VARIANCE COMPONENTS

LUBOMÍR KUBÁČEK

Introduction

In the regression model $\boldsymbol{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ the covariance matrix of the vector $\boldsymbol{\varepsilon}$ (i.e. the covariance matrix of the random vector $\boldsymbol{Y}$) is considered in the form $\boldsymbol{\Sigma}=v_1\mathbf{V}_1+\ldots+v_m\mathbf{V}_m$; $v_1, \ldots, v_m$ are variance components.

The aim is to estimate the components $v_1, \ldots, v_m$ on the basis of the $(k+1)$-tuple stochastically independent realizations of a normally distributed vector $\boldsymbol{Y}\sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, when the matrix $\mathbf{X}$ and the symmetric matrices $\mathbf{V}_1, \ldots, \mathbf{V}_m$ are known. The vector $\boldsymbol{\beta}$ is a nuisance parameter. (Procedure for estimating the vector $\boldsymbol{\beta}$ from the results of a repeated regression experiment see in [2].)

Repeated realizations of the vector $\boldsymbol{Y}$ or, which is the same, the realization of the $(k+1)$-tuple stochastically independent random vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}$ with the same normal distribution $N_n(\mathbf{X}\boldsymbol{\beta}, v_1\mathbf{V}_1+\ldots+v_m\mathbf{V}_m)$ generate a realization of a random matrix $k\mathbf{S}=\sum_{i=1}^{k+1}(\boldsymbol{Y}_i-\bar{\boldsymbol{Y}})(\boldsymbol{Y}_i-\bar{\boldsymbol{Y}})'\left(\bar{\boldsymbol{Y}}=[1/(k+1)]\sum_{i=1}^{k+1}\boldsymbol{Y}_i\right)$ with the Wishart distribution $k\mathbf{S}\sim W_n(k, \boldsymbol{\Sigma})$. Thus not only the vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}$ but the vector $\bar{\boldsymbol{Y}}$ and the matrix $\boldsymbol{S}$ as well are at our disposal for estimating the components $v, \ldots, v_m$. The last two random quantities are stochastically independent (in detail see [1]).

Procedures for estimating the components $v_1, \ldots, v_m$ based on the realization of the vector $\boldsymbol{Y}$ (i.e. based on the realization of the vector $\bar{\boldsymbol{Y}}$ as well) are described in detail in [6]. A natural question arising in the case of repeated experiments is how the knowledge of the realization of the matrix $\mathbf{S}$ contributes to estimating the components $v_1, \ldots, v_m$.

1. Symbols and auxiliary statements

Let $(\mathscr{A}, \langle ., ..\rangle)$ be a Hilbert space of symmetric $n\times n$ matrices, $\langle ., ..\rangle$ denotes the inner product given by $\langle \mathbf{A}, \mathbf{B}\rangle=\operatorname{Tr}(\mathbf{AB})$, $\mathbf{A}, \mathbf{B}\in\mathscr{A}$ [7]; here $\operatorname{Tr}(\mathbf{C})$ denotes the trace of the matrix $\mathbf{C}$.

A function which is to be estimated from the realizations of the vector $\boldsymbol{Y}$ and from the realization of the matrix $\mathbf{S}$, respectively, is denoted by the symbol $g(.)$ and we assume the linearity of it, i.e. $g(\boldsymbol{v})=\boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v}=(v_1, \dots, v_m)'$, $\boldsymbol{\lambda} \in \mathcal{R}^m$ ($\mathcal{R}^m$ means the m-dimensional Euclidean space). The symbol $\boldsymbol{v}*$ denotes a set of the space $\mathcal{R}^m$ in which the vector $\boldsymbol{v}$ is located; a closed sphere with a positive radius included into the set $\boldsymbol{v}*$ is assumed. The estimator of the function $g(.):\boldsymbol{v}*\to\mathcal{R}^1$ based on the realization of the matrix $\mathbf{S}$ is considered in the form $\tau_g(\mathbf{S})=\langle \mathbf{A}, \mathbf{S}\rangle = \mathrm{Tr}\,(\mathbf{AS})$, $\mathbf{A}\in\mathcal{A}$, $\tau_g(\mathbf{S})\in\bar{\mathcal{A}}=\{\langle \mathbf{A}, \mathbf{S}\rangle : \mathbf{A}\in\mathcal{A}\}$.

$E_v(\langle \mathbf{A}, \mathbf{S}\rangle)$ denotes the mean value of the random quantity $\langle \mathbf{A}, \mathbf{S}\rangle$. The subspace of the space $\mathcal{A}$ generated by symmetric matrices $\mathbf{V}_1, \dots, \mathbf{V}_m \in \mathcal{A}$ is denoted by $\mathscr{E}$.

Definition 1.1. *The function $g(.):\boldsymbol{v}*\to\mathcal{R}^1$ is $\bar{\mathcal{A}}$-estimable if there exists a matrix $\mathbf{A}\in\mathcal{A}$ with the property: $\forall\{\boldsymbol{v}\in\boldsymbol{v}*\}\ E_v[\mathrm{Tr}\,(\mathbf{AS})]=\boldsymbol{\lambda}'\boldsymbol{v}=g(\boldsymbol{v})$.*

Lemma 1. 1. *The class of all $\bar{\mathcal{A}}$-estimable functions is*

$$\mathscr{G}=\left\{g(.): g(\boldsymbol{v})=\sum_{i=1}^{m} v_i\ \mathrm{Tr}\,(\mathbf{AV}_i),\ \boldsymbol{v}\in\boldsymbol{v}*,\ \mathbf{A}\in\mathcal{A}\right\}.$$

Proof is obvious.

Lemma 1.2. *The projection of the matrix $\mathbf{A}\in\mathcal{A}$ on the subspace $\mathscr{E}$ is $P(\mathbf{A}) = \sum_{i=1}^{m} p_i\mathbf{V}_i$; the vector $\mathbf{p}=(p_1, \dots, p_m)'$ is a solution of the consistent system of linear equations $\mathbf{Kp}=(\mathrm{Tr}\,(\mathbf{AV}_1, \dots, \mathrm{Tr}\,(\mathbf{AV}_m))'$; $\{\mathbf{K}\}_{i,j}$ — the (i, j)-th element of the matrix $\mathbf{K}$ is $\{\mathbf{K}\}_{i,j}=\mathrm{Tr}\,(\mathbf{V}_i\mathbf{V}_j)$, $i, j=1, \dots, m$. The matrix $P(\mathbf{A})$ does not depend on the choice of the solution $\mathbf{p}$.*

Proof: The properties of a projection operator imply $\forall\{i = 1, \dots, m\}$ $\forall\{\mathbf{A}\in\mathcal{A}\}\langle\mathbf{A}, \mathbf{V}_i\rangle = \langle\mathbf{A}, P(\mathbf{V}_i)\rangle = \langle P(\mathbf{A}), \mathbf{V}_i\rangle = \sum_{j=1}^{m}\varkappa_j\langle\mathbf{V}_j, \mathbf{V}_i\rangle \Rightarrow \sum_{j=1}^{m}\{\mathbf{K}\}_{.j}\varkappa_j\in\mathcal{M}(\mathbf{K})$; $\{\mathbf{K}\}_{.j}$ is the j-th column of the matrix $\mathbf{K}$ and $\mathcal{M}(\mathbf{K})$ is the column space of it. Thus the system $\mathbf{Kp}=(\mathrm{Tr}\,(\mathbf{AV}_1), \dots, \mathrm{Tr}\,(\mathbf{AV}_m))'$ is consistent for each matrix $\mathbf{A}\in\mathcal{A}$. The following $m+1$ relations have to be valid simultaneously for the matrix $P(\mathbf{A})$:$\langle\mathbf{A}-P(\mathbf{A}), \mathbf{V}_i\rangle=0$, $i=1, \dots, m$ and $P(\mathbf{A}) = \sum_{i=1}^{m}p_i\mathbf{V}_i\in\mathscr{E}$, which immediately implies the second part of the statement. Let $\mathbf{p}_1$ and $\mathbf{p}_2$ be different solutions of the system $\mathbf{Kp} = (\mathrm{Tr}(\mathbf{AV}_1), \dots, \mathrm{Tr}\,(\mathbf{AV}_m))'$. Then $\left\langle\sum_{j=1}^{m}\{\mathbf{p}_1\}_j\mathbf{V}_j - \sum_{j=1}^{m}\{\mathbf{p}_2\}_j\mathbf{V}_j,\ \mathbf{V}_i\right\rangle = \{\mathbf{K}\}_{i.}\mathbf{p}_1-\{\mathbf{K}\}_{i.}\mathbf{p}_2=0$, $i=1, \dots, m$ and thus $\sum_{i=1}^{m}\{\mathbf{p}_1\}_j\mathbf{V}_j-\sum_{j=1}^{m}\{\mathbf{p}_2\}_j\mathbf{V}_j\in\mathscr{E}^{\perp}$ (orthogonal complement of the subspace $\mathscr{E}$). At the same time $\sum_{j=1}^{m}\{\mathbf{p}_1\}_j\mathbf{V}_j-\sum_{j=1}^{m}\{\mathbf{p}_2\}_j\mathbf{V}_j\in\mathscr{E}$ and thus $\sum_{j=1}^{m}\{\mathbf{p}_1\}_j\mathbf{V}_j=\sum_{j=1}^{m}\{\mathbf{p}_2\}_j\mathbf{V}_j$.

Lemma 1.3. *Let* $\boldsymbol{Z} \sim N_n(\boldsymbol{0}, \Sigma)$, $R(\Sigma)$ (*the rank of the matrix* $\Sigma) = r \leq n$, $r > 0$ *and* $\mathbf{J}$ *be an* $n \times r$ *matrix with the property* $\Sigma = \mathbf{J}\mathbf{J}'$. *Then there exists a random vector* $\boldsymbol{U} \sim N_r(\boldsymbol{0}, \mathbf{I})$ ($\mathbf{I}$ *denotes the identity matrix*) *for which* $P\{\boldsymbol{Z} = \mathbf{J}\boldsymbol{U}\} = 1$.

Proof. Let us consider a random vector $\boldsymbol{U} = \mathbf{J}^{-}\boldsymbol{Z}$ ($\mathbf{J}^{-}$ denotes the g-inversion of the matrix $\mathbf{J}$ (see [5])), $\mathbf{J}^{-}\mathbf{J} = \mathbf{I}$. For the covariance matrix of the vector $\boldsymbol{Z} - \mathbf{J}\boldsymbol{U}$ we have $E[(\boldsymbol{Z} - \mathbf{J}\boldsymbol{U})(\boldsymbol{Z} - \mathbf{J}\boldsymbol{U})'] = \mathbf{0}$. This and $E(\boldsymbol{Z} - \mathbf{J}\boldsymbol{U}) = \boldsymbol{0}$ imply the validity of the statement.

Lemma 1.4. *Let* $\mathbf{A}, \mathbf{B} \in \mathscr{A}$; *then* $\operatorname{cov}[\operatorname{Tr}(\mathbf{AS}), \operatorname{Tr}(\mathbf{BS})] = (2/k) \operatorname{Tr}(\mathbf{A}\Sigma\mathbf{B}\Sigma) = (2/k)\sum_{i=1}^{m}\sum_{j=1}^{m} v_i v_j \operatorname{Tr}(\mathbf{AV}_i\mathbf{BV}_j)$.

Proof. From the definition of the Wishart matrix $k\mathbf{S}$ (see [1]) it follows that $k\mathbf{S} = \sum_{\alpha=1}^{k} \boldsymbol{Z}_\alpha \boldsymbol{Z}'_\alpha$, where $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k$ are stochastically independent, normal and equally distributed random vectors, $\boldsymbol{Z}_\alpha \sim N_n\left(\boldsymbol{0}, \Sigma = \sum_{i=1}^{m} v_i \mathbf{V}_i\right)$, $\alpha = 1, \ldots, k$.

In the first step $\Sigma = \mathbf{I}$ is assumed. Then $\operatorname{cov}[\operatorname{Tr}(\mathbf{A}\boldsymbol{Z}_\alpha\boldsymbol{Z}'_\alpha), \operatorname{Tr}(\mathbf{B}\boldsymbol{Z}_\alpha\boldsymbol{Z}'_\alpha)] = \mathrm{E}\{[\boldsymbol{Z}'_\alpha\mathbf{A}\boldsymbol{Z}_\alpha - \operatorname{Tr}\mathbf{A})][\boldsymbol{Z}'_\beta\mathbf{B}\boldsymbol{Z}_\beta - \operatorname{Tr}(\mathbf{B})\} = E(\boldsymbol{Z}'_\alpha\mathbf{A}\boldsymbol{Z}_\alpha\boldsymbol{Z}'_\beta\mathbf{B}\boldsymbol{Z}_\beta) - \operatorname{Tr}(\mathbf{A})\operatorname{Tr}(\mathbf{B})$. For $\alpha \neq \beta$ we obtain zero. Let $\mathbf{Q}$ be an orthogonal $n \times n$ matrix with the property $\mathbf{QBQ}' = \operatorname{diag}(d_{11}, \ldots, d_{nn})$ (i.e. a diagonal matrix with an indicated diagonal) $= \mathbf{D}$. If $\boldsymbol{U}_\alpha = \mathbf{Q}\boldsymbol{Z}_\alpha$, then obviously $\boldsymbol{U}_\alpha \sim N_n(\boldsymbol{0}, \mathbf{I})$ and for the quantity $E(\boldsymbol{Z}'_\alpha\mathbf{A}\boldsymbol{Z}_\alpha\boldsymbol{Z}'_\alpha\mathbf{B}\boldsymbol{Z}_\alpha)$ we obtain $E(\boldsymbol{Z}'_\alpha\mathbf{A}\boldsymbol{Z}_\alpha\boldsymbol{Z}'_\alpha\mathbf{B}\boldsymbol{Z}_\alpha) = E(\boldsymbol{U}'_\alpha\mathbf{QAQ}'\boldsymbol{U}_\alpha\boldsymbol{U}'_\alpha\mathbf{D}\boldsymbol{U}_\alpha) = E\left(\boldsymbol{U}'_\alpha\mathbf{QAQ}'\boldsymbol{U}_\alpha\sum_{j=1}^{n}\{\boldsymbol{U}_\alpha\}_j^2 d_{jj}\right)$.

As $E(\{\boldsymbol{U}_\alpha\}_i\{\boldsymbol{U}_\alpha\}_j) = \begin{cases} 0 \text{ for } i \neq j \\ 1 \text{ for } i = j \end{cases}$ and $E(\{\boldsymbol{U}_\alpha\}_j^4) = 3$, we have $E(\boldsymbol{U}'_\alpha\mathbf{QAQ}'\boldsymbol{U}_\alpha\sum_{\alpha=1}^{n}\{\boldsymbol{U}_\alpha\}_j^2 d_{jj}) = 2\operatorname{Tr}(\mathbf{QAQ}'\mathbf{D}) + \operatorname{Tr}(\mathbf{QAQ}')\operatorname{Tr}(\mathbf{D}) = 2\operatorname{Tr}(\mathbf{AB}) + \operatorname{Tr}(\mathbf{A})\operatorname{Tr}(\mathbf{B})$.

In the second step $\Sigma \neq \mathbf{I}$ is assumed and the matrix Σ is expressed in the form $\Sigma = \mathbf{J}\mathbf{J}'$, where $\mathbf{J}$ is an $n \times r$ matrix, $r = R(\Sigma)$. With respect to Lemma 1.3 ($\boldsymbol{U}_\alpha = \mathbf{J}\ \boldsymbol{Z}_\alpha$, $\boldsymbol{Z}_\alpha = \mathbf{J}\boldsymbol{U}_\alpha$) and to the result of the firs step we obtain: $E(\boldsymbol{Z}'_\alpha\mathbf{A}\boldsymbol{Z}_\alpha\boldsymbol{Z}'_\alpha\mathbf{B}\boldsymbol{Z}_\alpha) = E(\boldsymbol{U}'_\alpha\mathbf{J}'\mathbf{AJ}\boldsymbol{U}_\alpha\boldsymbol{U}'_\alpha\mathbf{J}'\mathbf{BJ}\boldsymbol{U}_\alpha) = 2\operatorname{Tr}(\mathbf{J}'\mathbf{AJJ}'\mathbf{BJ}) + \operatorname{Tr}(\mathbf{J}'\mathbf{AJ})\operatorname{Tr}(\mathbf{J}'\mathbf{BJ}) = 2\operatorname{Tr}(\mathbf{A}\Sigma\mathbf{A}\Sigma) + \operatorname{Tr}(\mathbf{A}\Sigma)\operatorname{Tr}(\mathbf{B}\Sigma)$. The completion of the proof is now elementary.

Lemma 1.5. *The statistic* $\operatorname{Tr}(\mathbf{AS})$, $\mathbf{A} \in \mathscr{A}$ *estimates its mean value with a minimal dispersion in the class of estimators* $\bar{\mathscr{A}}$ *iff* $\operatorname{cov}[\operatorname{Tr}(\mathbf{AS}), \operatorname{Tr}(\mathbf{BS})] = 0$ *for all* $\mathbf{B} \in \mathscr{A}$ *with the property* $E_v[\operatorname{Tr}(\mathbf{BS})] = 0$, $\boldsymbol{v} \in \mathbf{v}*$.

Proof. It is a consequence of Theorem 5.3 in [3].

Lemma 1.6. *The class of all unbiased estimators in the class* $\bar{\mathscr{A}}$ *which estimate the function* $g(\boldsymbol{v}) = 0$, $\boldsymbol{v} \in \mathbf{v}*$, *is* $\{\operatorname{Tr}(\mathbf{BS}) : \mathbf{B} \in \mathscr{E}^{\perp}\}$.

Proof. Let $\mathbf{B} \in \mathscr{E}^{\perp}$. Then $E_v[\mathrm{Tr}\,(\mathbf{BS})] = \sum_{i=1}^{m} v_i \,\mathrm{Tr}\,(\mathbf{BV}_i) = 0$ because of the assumption $\mathrm{Tr}\,(\mathbf{BV}_i) = 0$, $i = 1, \ldots, m$. Let vice versa $E_v[\mathrm{Tr}\,(\mathbf{BS})] = 0$, $\boldsymbol{v} \in \mathbf{v}*$. Then $\boldsymbol{v}'(\mathrm{Tr}\,(\mathbf{BV}_1), \ldots, \mathrm{Tr}\,(\mathbf{BV}_m))' = 0$, $\boldsymbol{v} \in \mathbf{v}*$ and since the closed sphere with a positive radius is included into the set $\mathbf{v}*$, we have $\mathrm{Tr}\,(\mathbf{BV}_i) = 0$, $i = 1, \ldots, m$. Thus $\mathbf{B} \in \mathscr{E}^{\perp}$.

Lemma 1.7. *If $\boldsymbol{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{A}, \mathbf{B} \in \mathscr{A}$, then the random variables $\boldsymbol{Z}'\mathbf{A}\boldsymbol{Z}$, $\boldsymbol{Z}'\mathbf{B}\boldsymbol{Z}$ are stochastically independent iff $\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\Sigma} = \mathbf{0}$.*

Proof. See [5] Theorem 9.4.1.

Lemma 1.8. *Let $\boldsymbol{Z} \sim N_n(\boldsymbol{0}, \boldsymbol{\Sigma})$ and $\mathbf{P} \in \mathscr{A}$. A necessary and sufficient condition for $\boldsymbol{Z}'\mathbf{P}\boldsymbol{Z}$ to be chi-square distributed with r degrees of freedom is $\boldsymbol{\Sigma}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{P}\boldsymbol{\Sigma}$ and $r = R(\boldsymbol{\Sigma}\mathbf{P})$.*

Proof. See [5] Theorem 9.2.1.

2. Unbiased estimability of a linear function of the variance components

Theorem 2.1. *The function $g(\boldsymbol{v}) = \boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v} \in \mathbf{v}*$, is $\mathscr{A}$-estimable if $\boldsymbol{\lambda} \in \mathscr{M}(\mathbf{K})$, where $\{\mathbf{K}\}_{i,j} = \mathrm{Tr}\,(\mathbf{V}_i\mathbf{V}_j)$, $i, j = 1, \ldots, m$.*

Proof. Let $g(.)$ be $\mathscr{A}$-estimable, i.e. there exists a matrix $\mathbf{A} \in \mathscr{A}$ with the property $\forall\{\boldsymbol{v} \in \mathbf{v}*\} E_v[\mathrm{Tr}\,(\mathbf{AS})] = \sum_{i=1}^{m} v_i \,\mathrm{Tr}\,(\mathbf{AV}_i) = \sum_{i=1}^{m} v_i\lambda_i$. Since the set $\mathbf{v}*$ includes the closed sphere with a positive radius, $\boldsymbol{\lambda} = (\mathrm{Tr}\,(\mathbf{AV}_1), \ldots, \mathrm{Tr}\,(\mathbf{AV}_m))'$. With respect to Lemma 1.2 $\boldsymbol{\lambda} \in \mathscr{M}(\mathbf{K})$.

Let $g(\boldsymbol{v}) = \boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v} \in \mathbf{v}*$ and $\boldsymbol{\lambda} \in \mathscr{M}(\mathbf{K})$. An arbitrary solution of the system $\mathbf{K}\boldsymbol{p} = \boldsymbol{\lambda}$ is considered. The matrix $P(\mathbf{A}) = \sum_{j=1}^{m} \{\boldsymbol{p}\}_j \mathbf{V}_j$ from Lemma 1.2 is a projection of some matrix $\mathbf{A} \in \mathscr{A}$ for which $\boldsymbol{\lambda} = (\mathrm{Tr}\,(\mathbf{AV}_1), \ldots, \mathrm{Tr}\,(\mathbf{AV}_m))'$. It implies $E_v[\mathrm{Tr}\,(\mathbf{AS})] = \sum_{i=1}^{m} v_i \,\mathrm{Tr}\,(\mathbf{AV}_i) = \sum_{i=1}^{m} v_i\lambda_i = g(\boldsymbol{v})$, $\boldsymbol{v} \in \boldsymbol{v}*$. Thus the function $g(.)$ is $\mathscr{A}$-estimable.

Corollary. *Every linear function $g(\boldsymbol{v}) = \boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v} \in \mathbf{v}*$, unbiasedly estimable on the base of the realization of the vector $\boldsymbol{Y}_j$, $j = 1, \ldots, k+1$ (i.e. on the base of the vector $\bar{\boldsymbol{Y}}$) is unbiasedly estimable on the base of the realization of the matrix $\mathbf{S}$ as well.*

Proof. The function $g(\boldsymbol{v}) = \boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v} \in \mathbf{v}*$, is unbiasedly estimable on the basis of the realization of the vector $\boldsymbol{Y}$ iff $\boldsymbol{\lambda} \in \mathscr{M}(\mathbf{K}_0)$, where $\{\mathbf{K}_0\}_{i,j} = \mathrm{Tr}\,(\mathbf{V}_i\mathbf{N}\mathbf{V}_j)$, $\mathbf{N} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$ (see [8] and [6], respectively). Thus it is sufficient to show the inclusion $\mathscr{M}(\mathbf{K}_0) \subset \mathscr{M}(\mathbf{K})$. For verification we substitute for the matrix $\mathbf{A}$ from Lemma 1.2 $\mathbf{A} = \frac{1}{2}(\mathbf{V}_j\mathbf{N} + \mathbf{N}\mathbf{V}_j) \in \mathscr{A}$, $j = 1, \ldots, m$; the j-th column of the matrix $\mathbf{K}_0$ is $(\mathrm{Tr}\,(\mathbf{AV}_1), \ldots, \mathrm{Tr}\,(\mathbf{AV}_m))'$, which is obviously an element of the space $\mathscr{M}(\mathbf{K})$.

Example 2.1. Let $\boldsymbol{Y}_j \sim N_2\left(\begin{pmatrix}1\\1\end{pmatrix}\beta,\ \boldsymbol{\Sigma} = \sigma_1^2\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} + c_{12}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\right)$, $|c_{12}| \leq \sigma_1^2$, $j =$

1, ..., $k+1$ (≥ 3). From the realization of the vector $\mathbf{Y}_j$ neither the component σ_1^2 nor the component c_{12} can be estimated because of $(1, 0)' \notin \mathcal{M}(\mathbf{K}_0) = \mathcal{M}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $(0, 1)' \notin \mathcal{M}(\mathbf{K}_0)$. From the matrix $k\mathbf{S} = \sum_{i=1}^{k+1} (\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Y}_i - \bar{\mathbf{Y}})'$ it is nevertheless possible to estimate the arbitrary linear function $g(\sigma_1^2, c_{12}) = \lambda_1\sigma_1^2 + \lambda_2 c_{12}$ because of $\mathcal{M}(\mathbf{K}) = \mathcal{M}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \mathcal{R}^2$.

The example shows how important the repetition of experiments can be, e.g., in the case of estimating the variance components of a stationary random process.

3. Natural estimation and γ-estimation

Let the matrices $\mathbf{V}_i$, $i = 1, ..., m$ be positive semidefinite and $v_i > 0$, $i = 1, ..., m$.Then for each matrix $\mathbf{V}_i$ there exists an $n \times R(\mathbf{V}_i)$ matrix $\mathbf{J}_i$ which satisfies the condition $\mathbf{V}_i = \mathbf{J}_i\mathbf{J}_i'$. With respect to Lemma 1.3 the vector $\mathbf{Z}_\alpha$, $\alpha = 1, ..., k$ can be expressed in the form $\mathbf{Z}_\alpha = \mathbf{J}_1\mathbf{U}_{\alpha 1} + ... + \mathbf{J}_m\mathbf{U}_{\alpha m}$, where $\mathbf{U}_{\alpha j} \sim N_{r_j}(\mathbf{0}, v_j\mathbf{I})(r_j = R(\mathbf{V}_j))$, $\alpha = 1, ..., k$; $j = 1, ..., m$ and the random vectors $\mathbf{U}_{\alpha j}$, $\alpha = 1, ..., k$; $j = 1, ..., m$ are stochastically independent.

When the realizations of the vectors $\mathbf{U}_{\alpha j}$, $\alpha = 1, ..., k$; $j = 1, ..., m$ are known, then the natural estimator of the component v_i, $i = 1, ..., m$, is the statistic $\hat{v}_i = (1/k)\sum_{\alpha=1}^{k} \mathbf{U}_{\alpha i}'\mathbf{U}_{\alpha i}/r_i$, $i = 1, ..., m$. The estimators v_i are unbiased with a minimal dispersion. Therefore their linear combination is again an unbiased estimator of its own mean value with a minimal dispersion.

The natural estimator of the function $g(\mathbf{v}) = \boldsymbol{\lambda}'\mathbf{v}$, $\mathbf{v} \in \mathbf{v}_*$, is thus the statistic $\widehat{\boldsymbol{\lambda}'\mathbf{v}} = (1/k)\sum_{\alpha=1}^{k} \mathbf{U}_\alpha'\Delta\mathbf{U}_\alpha$, where $\mathbf{U}_\alpha' = (\mathbf{U}_{\alpha 1}', ..., \mathbf{U}_{\alpha m}')$,

$$\Delta = \begin{bmatrix} \frac{\lambda_1}{r_1}\mathbf{I}_{r_1}, & \mathbf{0}, & \dots, & \mathbf{0} \\ \mathbf{0}, & \frac{\lambda_2}{r_2}\mathbf{I}_{r_2}, & \dots, & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0}, & \mathbf{0}, & \dots, & \frac{\lambda_m}{r_m}\mathbf{I}_{r_m} \end{bmatrix},$$

and $\mathbf{I}_{r_i}$ is the $r_i \times r_i$ identity matrix.

Let $\mathbf{T}$ denote the matrix defined by $k\mathbf{T} = \sum_{\alpha=1}^{k} \mathbf{U}_\alpha\mathbf{U}_\alpha'$; then $k\mathbf{T}$ has the Wishart distribution and the natural estimator of the function $g(.)$ can be expressed in the form $\widehat{\boldsymbol{\lambda}'\mathbf{v}} = \mathrm{Tr}\,(\Delta\mathbf{T})$. The difference between the estimators Tr ($\mathbf{AS}$) and Tr ($\Delta\mathbf{T}$), respectively, can be expressed in the form $\mathrm{Tr}\,(\mathbf{AS})-\mathrm{Tr}\,(\mathbf{\Delta T}) = \mathrm{Tr}\,[(\mathbf{J'AJ}-\mathbf{\Delta})\mathbf{T}]$, where $\mathbf{J}$ is an $n\times\sum_{i=1}^{m} r_i$ matrix for which $\mathbf{J}=(\mathbf{J}_1, \ldots, \mathbf{J}_m)$, $\mathbf{JJ'}=\mathbf{V}_1+\ldots+\mathbf{V}_m=\mathbf{V}$ (it is sufficient to take into account that $\mathbf{JTJ'}=\mathbf{S}$).

Definition 3.1. *The minimum norm unbiased estimator (MINUE) of the function* $g(v)=\boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v}\in\boldsymbol{v}*$, *is a statistic* $\tau_g(\mathbf{S})=\mathrm{Tr}\,(\mathbf{AS})$, $\mathbf{A}\in\mathscr{A}$, *where the matrix* $\mathbf{A}$ *minimizes the Euclidean norm of the quantity* $\mathbf{J'AJ}-\mathbf{\Delta}$ *and satisfies the conditions* $\mathrm{Tr}\,(\mathbf{AV}_i)=\lambda_i$, $i=1, \ldots, m$.

Theorem 3.1. *Let the matrix* $\mathbf{V}=\mathbf{V}_1+\ldots+\mathbf{V}_m$ *be regular. The MINUE of function* $g(\boldsymbol{v})=\boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v}\in\boldsymbol{v}*$, $\boldsymbol{\lambda}\in\mathscr{M}(\mathbf{K})$ *is* $\mathrm{Tr}\left(\sum_{i=1}^{m}\varkappa_i\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}\mathbf{S}\right)$, *where the vector* $\boldsymbol{\varkappa}=(\varkappa_1, \ldots, \varkappa_m)$ *is a solution of the linear system* $\mathbf{M}\boldsymbol{\varkappa}=\boldsymbol{\lambda}$. *The* (i, j)*-th element of the matrix* $\mathbf{M}$ *is* $\{\mathbf{M}\}_{i,j}=\mathrm{Tr}\,(\mathbf{V}_i\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1})$, $i, j=1, \ldots, m$. *Further* $\mathscr{M}(\mathbf{K})=\mathscr{M}(\mathbf{M})$.

Proof. The square of the Euclidean norm of $\mathbf{J'AJ}-\mathbf{\Delta}$ is $\|\mathbf{J'AJ}-\mathbf{\Delta}\|^2 = \mathrm{Tr}\,[(\mathbf{J'AJ}-\mathbf{\Delta})(\mathbf{J'AJ}-\mathbf{\Delta})] = \mathrm{Tr}\,(\mathbf{AVAV}) - 2\,\mathrm{Tr}\,(\mathbf{\Delta J'AJ}) + \mathrm{Tr}\,(\mathbf{\Delta}^2)$. As the matrix $\mathbf{A}$ has to satisfy m conditions $\mathrm{Tr}\,(\mathbf{AV}_i) = \lambda_i$, $i=1, \ldots, m$ there holds $\mathrm{Tr}\,(\mathbf{\Delta J'AJ}) = \sum_{i=1}^{m}(\lambda_i/r_i)\,\mathrm{Tr}\,(\mathbf{I}_{r_i}\mathbf{J}_i'\mathbf{AJ}) = \sum_{i=1}^{m}\lambda_i^{\prime}r_i = \mathrm{Tr}\,(\mathbf{\Delta}^2)$. Thus $\|\mathbf{J'AJ}-\mathbf{\Delta}\|^2 = \mathrm{Tr}\,(\mathbf{AVAV}) - \mathrm{Tr}\,(\mathbf{\Delta}^2)$. The matrix $\mathbf{A}$ minimizing the quantity $\mathrm{Tr}\,(\mathbf{AVAV})$ and satisfying the given conditions can be determined by the method of the Lagrange undetermined multipliers. The Lagrange auxiliary function is $\Phi(\mathbf{A})=\mathrm{Tr}\,(\mathbf{AVAV}) - 2\sum_{i=1}^{m}\varkappa_i[\mathrm{Tr}\,(\mathbf{AV}_i)-\lambda_i]$, where $\varkappa_i$, $i=1, \ldots, m$, are the Lagrange multipliers

$$(\partial\Phi(\mathbf{A})/\partial\mathbf{A}=)4\mathbf{VAV}-4\sum_{i=1}^{m}\varkappa_i\mathbf{V}_i-\left[2\,\mathrm{diag}\,(\mathbf{VAV})-2\sum_{i=1}^{m}\varkappa_i\,\mathrm{diag}\,(\mathbf{V}_i)\right]=\mathbf{0}$$

$$\Leftrightarrow \mathbf{VAV}=\sum_{i=1}^{m}\varkappa_i\mathbf{V}_i\,.$$

(The symbol $\mathrm{diag}\,(\mathbf{C})$ denotes a diagonal matrix, which diagonal is identical with the diagonal of the matrix $\mathbf{C}$.) For each symmetric matrix $\mathbf{D}$ satisfying the conditions $\mathrm{Tr}\,(\mathbf{DV}_i)=0$, $i=1, \ldots, m$ there is $\mathrm{Tr}\,[(\mathbf{A}+\mathbf{D})\mathbf{V}(\mathbf{A}+\mathbf{D})\mathbf{V}] = \mathrm{Tr}\,(\mathbf{AVAV})+\mathrm{Tr}\,(\mathbf{DVDV})$ because of $\mathrm{Tr}\,(\mathbf{DVAV}) = \mathrm{Tr}\left(\mathbf{D}\sum_{i=1}^{m}\varkappa_i\mathbf{V}_i\right) = \sum_{i=1}^{m}\varkappa_i\,\mathrm{Tr}\,(\mathbf{DV}_i)=0$. Further $\mathrm{Tr}\,(\mathbf{DVDV}) = \mathrm{Tr}\,(\mathbf{J'DJJ'DJ})\geq 0$ and thus the matrix $\mathbf{A}=\sum_{i=1}^{m}\varkappa_i\mathbf{V}^{-1}\ \mathbf{V}_i\mathbf{V}^{-1}$, where $\varkappa_i$, $i=1, \ldots, m$ are solutions of the equations $\mathrm{Tr}\,(\mathbf{AV}_i)=\lambda_i$, $i=1, \ldots, m$, minimizes the quantity $\mathrm{Tr}\,(\mathbf{AVAV})$ and satisfies the given conditions. The system of the conditions $\mathrm{Tr}\,(\mathbf{AV}_i)=\lambda_i$, $i=1, \ldots, m$ can be obviously written in the form $\mathbf{M}\boldsymbol{\varkappa}=\boldsymbol{\lambda}$, which proves the first part of the statement.

The validity of $\mathcal{M}(\mathbf{K})=\mathcal{M}(\mathbf{M})$ can be proved by means of Lemma 1.2. For the i-th column of the matrix $\mathbf{M}$, $\{\mathbf{M}\}_{.i} = (\mathrm{Tr}\,(\mathbf{V}_1\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}), \ldots, \mathrm{Tr}\,(\mathbf{V}_m\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}))'$ the matrix $\mathbf{A}$ is chosen from Lemma 1.2 in the form $\mathbf{A}=\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}$, which immediately implies $\{\mathbf{M}\}_{.i} = (\mathrm{Tr}\,(\mathbf{V}_1\mathbf{V}_i), \ldots, \mathrm{Tr}\,(\mathbf{V}_m\mathbf{V}_i))'$ the matrix $\mathbf{A}=\mathbf{J}'\mathbf{V}_i\mathbf{J}$ is chosen and the fact is taken into account that the (i, j)-th element of the matrix $\mathbf{M}$ is $\{\mathbf{M}\}_{i,j} = \mathrm{Tr}\,(\mathbf{V}_i\mathbf{V}^{-1}\mathbf{V}_j\mathbf{V}^{-1}) = \mathrm{Tr}\,(\mathbf{J}^{-1}\mathbf{V}_i\mathbf{J}'^{-1}\mathbf{J}^{-1}\mathbf{V}_j\mathbf{J}'^{-1})$. The vector $(\mathrm{Tr}\,(\mathbf{A}\mathbf{J}^{-1}\mathbf{V}_1\mathbf{J}'^{-1}), \ldots, \mathrm{Tr}\,(\mathbf{A}\mathbf{J}^{-1}\mathbf{V}_m\mathbf{J}'^{-1}))'$ is then the i-th column of the matrix $\mathbf{K}$ and obviously an element of $\mathcal{M}(\mathbf{M})$. Thus $\mathcal{M}(\mathbf{M})=\mathcal{M}(\mathbf{K})$.

Corollary. *For each unbiasedly estimable function $g(v)=\lambda' v$, $v \in v_*$ there exists the MINUE.*

Remark 3.1. The MINUE is an analogy of the MINQUE [4], which is based on the realization of the vector $\boldsymbol{Y}$. MINQUE, however, does not exist for each unbiasedly estimable function $g(v)=\lambda' v$, $v \in v_*$, $\lambda \in \mathcal{M}(\mathbf{K}_0)$.

In the following the values $v_1, \ldots, v_m$ of components are assumed to be known at such a level of accuracy that for a vector of a priori values $\gamma=(\gamma_1, \ldots, \gamma_m)'$ there is $\gamma \in \{\boldsymbol{x}: \boldsymbol{x} \in \mathcal{R}^m, (\boldsymbol{x}-v)'(\boldsymbol{x}-v)<\varrho^2\} = O(v, \varrho)$, $\varrho>0$. The value ϱ is so small that the matrix $\mathbf{V}^{(\gamma)}=\sum_{i=1}^{m}\gamma_i\mathbf{V}_i$ is regular in the neighbourhood $O(v, \varrho)$ (obviously $\mathbf{V}^{(v)}=\boldsymbol{\Sigma}$).

Definition 3.2. *The MINUγE of a function $g(v)=\lambda' v$, $v \in v_*$, $\lambda \in \mathcal{M}(\mathbf{K})$ is a statistic $\tau_q(\mathbf{S})=\mathrm{Tr}\,(\mathbf{AS})$, $\mathbf{A} \in \mathcal{A}$, where the matrix $\mathbf{A}$ minimizes the Euclidean norm of $\mathbf{J}^{(\gamma)\prime}\mathbf{A}\mathbf{J}^{(\gamma)}-\Delta^{(\gamma)}$ and satisfies the conditions $\mathrm{Tr}\,(\mathbf{A}\mathbf{V}_i)=\lambda_i$, $i=1, \ldots, m$; $\mathbf{J}^{(\gamma)} = (\mathbf{J}_1^{(\gamma)}, \ldots, \mathbf{J}_m^{(\gamma)})$, $\mathbf{J}_i^{(\gamma)}\mathbf{J}_i^{(\gamma)\prime} = \gamma_i\mathbf{V}_i$, $i=1, \ldots, m$ and*

$$\Delta^{(\gamma)}= \begin{bmatrix} \frac{\lambda_1}{r_1}\gamma_1\mathbf{I}_{r_1}, & \boldsymbol{0}, & \ldots, & \boldsymbol{0} \\ \boldsymbol{0}, & \frac{\lambda_2}{r_2}\gamma_2\mathbf{I}_{r_2}, & \ldots, & \boldsymbol{0} \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{0}, & \boldsymbol{0}, & \ldots, & \frac{\lambda_m}{r_m}\gamma_m\mathbf{I}_{r_m} \end{bmatrix}.$$

Remark 3.2. When in the consideration preceding the definition 3.1 the matrix $\mathbf{J}_i^{(\gamma)}=\sqrt{\gamma_i}\mathbf{J}_i$ is substituted for $\mathbf{J}_i$ and the vector $\boldsymbol{U}_{\alpha i}^{(\gamma)}=(1/\sqrt{\gamma_i})\boldsymbol{U}_{\alpha i}$ for the vector $\boldsymbol{U}_{\alpha i}$, the natural estimator of the function $g(.)$ can be written in the form $\widehat{\lambda' v} = (1/k)\sum_{\alpha=1}^{m}\boldsymbol{U}_\alpha^{(\gamma)\prime}\Delta(\gamma)\boldsymbol{U}_\alpha^{(\gamma)}=\mathrm{Tr}\,(\Delta^{(\gamma)}\mathbf{T}^{(\gamma)})$; $k\mathbf{T}^{(\gamma)} = \sum_{\alpha=1}^{k}\boldsymbol{U}_\alpha^{(\gamma)}\boldsymbol{U}_\alpha^{(\gamma)\prime}$. Analogously the difference between the estimators Ts $(\mathbf{AS})$ and $\mathrm{Tr}\,(\Delta^{(\gamma)}\mathbf{T}^{(\gamma)})$ can be expressed as $\mathrm{Tr}\,(\mathbf{AS}) - \mathrm{Tr}\,(\Delta^{(\gamma)}\mathbf{T}^{(\gamma)}) = \mathrm{Tr}\,[(\mathbf{J}^{(\gamma)\prime}\mathbf{A}\mathbf{J}^{(\gamma)} - \Delta^{(\gamma)})\mathbf{T}^{\gamma}]$.

Theorem 3.2. *The MINUγE of the function $g(v)=\lambda' v$, $v \in v_*$, is the statistic*

$\tau_g(\mathbf{S}) = \mathrm{Tr}\left(\sum_{i=1}^{m} \varkappa_i \mathbf{V}^{(\gamma)^{-1}}\mathbf{V}_i\mathbf{V}^{(\gamma)^{-1}}\mathbf{S}\right)$, *where* $\mathbf{V}^{(\gamma)} = \gamma_1\mathbf{V}_1 + \ldots + \gamma_m\mathbf{V}_m$ *and the vector* $\varkappa = (\varkappa_1, \ldots, \varkappa_m)'$ *is a solution of the equation* $\mathbf{M}^{(\gamma)}\varkappa = \lambda$; *the* (i, j)*-th element of the matrix* $\mathbf{M}^{(\gamma)}$ *is* $\{\mathbf{M}^{(\gamma)}\}_{i,j} = \mathrm{Tr}\,(\mathbf{V}_i\mathbf{V}^{(\gamma)^{-1}}\mathbf{V}_j\mathbf{V}^{(\gamma)^{-1}})$, $i, j = 1, \ldots, m$. *The matrix* $\mathbf{M}^{(\gamma)}$ *has the property* $\mathcal{M}(\mathbf{M}^{(\gamma)}) = \mathcal{M}(\mathbf{K})$.

The proof is analogous to the proof of Theorem 3.1.

Corollary. *If* $\gamma = cv^{(0)}$, $c \in (0, \infty)$, $v^{(0)} \in \mathbf{v}*$, *then the MINUγE is a locally best estimator of the function* $g(v) = \lambda'v$, $v \in \mathbf{v}*$, *at the point* $v = v^{(0)}$.

Proof. Regarding Lemma 1.4 we have $\mathscr{D}[\mathrm{Tr}\,(\mathbf{AS})] = (2/k)\,\mathrm{Tr}\,(\mathbf{A\Sigma A\Sigma})$. If $\mathbf{\Sigma} = \mathbf{V}^{(v^{(0)})}$ and $\gamma = cv^{(0)}$, then the minimization of the quantity $\mathrm{Tr}\,(\mathbf{AV}^{(\gamma)}\mathbf{AV}^{(\gamma)}) = c^2\mathscr{D}[\mathrm{Tr}\,(\mathbf{AS})]k/2$ by a suitable choice of the matrix $\mathbf{A}$ satisfying the conditions $\mathrm{Tr}\,(\mathbf{AV}_i) = \lambda_i$, $i = 1, \ldots, m$ is equivalent to a determination of a locally best estimator.

Remark 3.3. For an arbitrary but fixed realization of the matrix $\mathbf{S}$ the function $f(\gamma_1, \ldots, \gamma_m) = \mathrm{Tr}\left(\sum_{j=1}^{m} \varkappa_j \mathbf{V}^{(\gamma)-1}\mathbf{V}_j\mathbf{V}^{(\gamma)-1}\mathbf{S}\right)$, $\gamma \in O(v^{(0)}, \varrho)$, is continuous at the point $v^{(0)}$. That is why the MINUγE in a sufficient small neighbourhood of the point $v^{(0)}$ is unsubstantially deviated from the locally best estimator.

Remark 3.4. The matrix $\mathbf{M}^{(\gamma)}$ from Theorem 3.2 is related to the Fisher information matrix $\mathbf{F}(v) = E(-\partial^2 \ln f(\mathbf{S}, \sum_{i=1}^{m} v_i\mathbf{V}_i)/\partial v\, \partial v')$, where $f(\mathbf{S}, \mathbf{\Sigma}) = k^{\frac{kn}{2}}\, 2^{-\frac{kn}{2}}\, \pi^{-\frac{n(n-1)}{2}} \left\{\prod_{j=1}^{n} \Gamma\left[\frac{1}{2}(k+1-j)\right]\right\}^{-1} \det(\mathbf{S}) \exp\left[-\frac{k}{2}\mathrm{Tr}\,(\mathbf{\Sigma}^{-1}\mathbf{S})\right] [\det(\mathbf{\Sigma})]^{-k/2}$.

By means of $\partial\mathbf{\Sigma}^{-1}(t)/\partial t = -\mathbf{\Sigma}^{-1}(t)[\partial\mathbf{\Sigma}(t)/\partial t]\mathbf{\Sigma}^{-1}(t)$ and $\partial \ln[\det(\mathbf{\Sigma}(t))]/\partial t = \mathrm{Tr}\,[\mathbf{\Sigma}^{-1}(t)\, \partial\mathbf{\Sigma}(t)/\partial t]$, respectively, we can easily obtain

$$\partial \ln f(\mathbf{S}, \mathbf{\Sigma})/\partial v_i = (k/2)\,\mathrm{Tr}\,(\mathbf{\Sigma}^{-1}\mathbf{V}_i\mathbf{\Sigma}^{-1}\mathbf{S}) - (k/2)\,\mathrm{Tr}\,(\mathbf{V}_i\mathbf{\Sigma}^{-1}),$$

$$\partial^2 \ln f(\mathbf{S}, \mathbf{\Sigma})/\partial v_i\, \partial v_j = -(k/2)\,\mathrm{Tr}\,(\mathbf{\Sigma}^{-1}\mathbf{V}_j\mathbf{\Sigma}^{-1}\mathbf{V}_i\mathbf{\Sigma}^{-1}\mathbf{S} + \mathbf{\Sigma}^{-1}\mathbf{V}_i\mathbf{\Sigma}^{-1}\mathbf{V}_j\mathbf{\Sigma}^{-1}\mathbf{S}) +$$
$$+ (k/2)\,\mathrm{Tr}\,(\mathbf{V}_i\mathbf{\Sigma}^{-1}\mathbf{V}_j\mathbf{\Sigma}^{-1}).$$

Thus $\{\mathbf{F}(v)\}_{i,j} = (k/2)\,\mathrm{Tr}\,(\mathbf{\Sigma}^{-1}\mathbf{V}_i\mathbf{\Sigma}^{-1}\mathbf{V}_j) = (k/2)\{\mathbf{M}^{(v)}\}_{i,j}$. If $\gamma = v$, then the vector of the Lagrange multipliers $\varkappa$ from Theorem 3.2 is $\varkappa = (k/2)\mathbf{F}^{-1}(v)\lambda$ and for the dispersion of the estimator $\tau_g(\mathbf{S}) = \mathrm{Tr}\left(\sum_{i=1}^{m} \varkappa_i \mathbf{V}^{(v)^{-1}}\mathbf{V}_i\mathbf{V}^{(v)^{-1}}\mathbf{S}\right)$ with respect to Lemma 1.4 there holds $\mathscr{D}[\tau_g(\mathbf{s})] = (2/k)\,\mathrm{Tr}\left(\sum_{i=1}^{m} \varkappa_i\mathbf{\Sigma}^{-1}\mathbf{V}_i\mathbf{\Sigma}^{-1}\mathbf{\Sigma}\sum_{j=1}^{m} \varkappa_j\mathbf{\Sigma}^{-1}\mathbf{V}_j\mathbf{\Sigma}^{-1}\mathbf{\Sigma}\right)$

$= (2/k)^2 \varkappa' \mathbf{F}(\nu)\varkappa = \lambda' \mathbf{F}^{-1}(\nu)\lambda$. Thus the dispersion of the MINUνE is equal to the Rao—Cramér lower bound at the point ν.

Theorem 3.3. *If the variance components are eigenvalues of the covariance matrix, then each of the components is $\bar{\mathscr{A}}$-estimable and for each component ν_i, $i=1, \ldots, m$ there exists in the class $\bar{\mathscr{A}}$ a uniformly best estimator $\hat{\nu}_i$ with the same distribution as that of the random variable $\nu_i\chi^2(kr_i)/(kr_i)$, where $\chi^2(kr_i)$ is a random variable with a chi-square distribution with kr_i degrees of freedom and for $i \neq j$ these estimators are stochastically independent. The dispersions of the estimators $\hat{\nu}_i$ are equal to the Rao-Cramér lower bound.*

Proof. With respect to our assumption $\mathbf{\Sigma} = \nu_1\mathbf{V}_1 + \ldots + \nu_m\mathbf{V}_m$, where $\mathbf{V}_i$, $i = 1, \ldots, m$ are projection matrices and for $i \neq j$ there holds $\mathbf{V}_i\mathbf{V}_j = \mathbf{0}$. The matrix $\mathbf{K}$ from Theorem 2.1 is $\mathbf{K} = \text{diag}\,[\text{Tr}\,(\mathbf{V}_1), \ldots, \text{Tr}\,(\mathbf{V}_m)]$, where $\text{Tr}\,(\mathbf{V}_i) = R(\mathbf{V}_i) > 0$ and therefore all components are $\bar{\mathscr{A}}$-estimable.

Let $\hat{\nu}_i = \text{Tr}\,(\mathbf{V}_i\mathbf{S})/\text{Tr}\,(\mathbf{V}_i)$. Then $E_\nu(\hat{\nu}_i) = \nu_i$, $\nu \in \boldsymbol{\nu}_*$, and with respect to Lemma 1.4 for $\mathbf{A}_0 \in \mathscr{C}^\perp$ we have $\text{cov}\,[\text{Tr}\,(\mathbf{V}_i\mathbf{S})/\text{Tr}\,(\mathbf{V}_i),\ \text{Tr}\,(\mathbf{A}_0\mathbf{S})] = (2\nu_i^2/k)\,\text{Tr}\,(\mathbf{V}_i\mathbf{A}_0)/ /\text{Tr}\,(\mathbf{V}_i)$; $\mathbf{A}_0 \in \mathscr{C}^\perp \Rightarrow \text{Tr}\,(\mathbf{V}_i\mathbf{A}_0) = 0$. Thus with respect to Lemmas 1.5 and 1.6, respectively, it can be seen that the statistic $\text{Tr}\,(\mathbf{S}\mathbf{V}_i)/\text{Tr}\,(\mathbf{V}_i)$ estimates its mean value ν_i with a minimal dispersion at each point ν of the set $\boldsymbol{\nu}_*$.

The assumption $k\mathbf{S} \sim W_n\left(k, \sum_{i=1}^{m} \nu_i\mathbf{V}_i\right)$ implies $\hat{\nu}_i = \text{Tr}\,(\mathbf{V}_i\mathbf{S})/\text{Tr}\,(\mathbf{V}_i) = [1/(kr_i)] \sum_{\alpha=1}^{m} \mathbf{Z}'_\alpha\mathbf{V}_i\mathbf{Z}_\alpha$. Regarding Lemma 1.8 the random variable $\mathbf{Z}'_\alpha\mathbf{V}_i\mathbf{Z}_\alpha$ has the same distribution as the random variable $\nu_i\chi^2(r_i)$. For $\alpha \neq \beta$ the random variables $\mathbf{Z}'_\alpha\mathbf{V}_i\mathbf{Z}_\alpha$ and $\mathbf{Z}'_\beta\mathbf{V}_i\mathbf{Z}_\beta$ are stochastically independent. This fact and the additivity of the chi-square distribution imply that $\hat{\nu}_i$ is a random variable with the identical distribution as that of the random variable $\nu_i\chi^2(kr_i)/(kr_i)$.

For $i \neq j$ $\mathbf{\Sigma}\mathbf{V}_i\mathbf{\Sigma}\mathbf{V}_j\mathbf{\Sigma} = \mathbf{0}$. Thus, regarding Lemma 1.7, $\hat{\nu}_i$ and $\hat{\nu}_j$ are stochastically independent.

In our case the Fisher information matrix $\mathbf{F}(\nu)$ is $\mathbf{F}(\nu) = (k/2)\,\text{diag}\,[(1/ /\nu_1^2)\,\text{Tr}(\mathbf{V}_1), \ldots, (1/\nu_m^2)\text{Tr}\,(\mathbf{V}_m)]$ and thus $\{\mathbf{F}^{-1}(\nu)\}_{i,i} = (2/k)\nu_i^2/\text{Tr}\,(\mathbf{V}_i)$. With respect to Lemma 1.4 $\mathscr{D}[\text{Tr}\,(\mathbf{V}_i\mathbf{S})/\text{Tr}\,(\mathbf{V}_i)] = (2/k)\nu_i^2/\text{Tr}\,(\mathbf{V}_i) = \{\mathbf{F}^{-1}(\nu)\}_{i,i}$, which proves the last part of the statement.

Remark 3.5. If $\text{Tr}\,(\mathbf{V}_i) = R(\mathbf{V}_i) = r_i \geq 2$, then $\mathbf{V}_i = \sum_{j=1}^{r_i} \mathbf{f}_j\mathbf{f}'_j$, where $\mathbf{f}_j \in \mathscr{R}^n$, $\mathbf{f}'_j\mathbf{f}_j = 1$, $j = 1, \ldots, r_i$ and for $j \neq l$ $\mathbf{f}'_j\mathbf{f}_l = 0$, $j, l = 1, \ldots, r_i$. In this case it can be easily verified that the estimators $\hat{\nu}_i^{(j)} = \text{Tr}\,(\mathbf{f}_j\mathbf{f}'_j\mathbf{S})$ and $\hat{\nu}_i^{(l)} = \text{Tr}\,(\mathbf{f}_l\mathbf{f}'_l\mathbf{S})$ are stochastically independent and they have the same dispersion $\mathscr{D}(\hat{\nu}_i^{(j)}) = \mathscr{D}(\hat{\nu}_i^{(l)}) = (2/k)\nu_i^2$. Combining these estimators we obtain $(1/r_i)(\hat{\nu}_i^{(1)} + \ldots + \hat{\nu}_i^{(r_i)}) = \text{Tr}\,(\mathbf{V}_i\mathbf{S})/\text{Tr}\,(\mathbf{V}_i)$.

4. Comparison of estimators based on the repeated realizations of the vector $\boldsymbol{Y}$ with estimators based on the realization of the matrix $\mathbf{S}$

In the case of a repetion of the regression experiment $\boldsymbol{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sum_{i=1}^{m} v_i \mathbf{V}_i)$ three following situations can occur in estimating the function $g(\boldsymbol{v}) = \boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v} \in \boldsymbol{v}_*$: 1. $\boldsymbol{\lambda} \in \mathcal{M}(\mathbf{K}_0)$, 2. $\boldsymbol{\lambda} \notin \mathcal{M}(\mathbf{K}_0)$ & $\boldsymbol{\lambda} \in \mathcal{M}(\mathbf{K})$ and 3. $\boldsymbol{\lambda} \notin \mathcal{M}(\mathbf{K})$ (matrices $\mathbf{K}_0$ and $\mathbf{K}$, respectively, are mentioned in Theorem 2.1 and in its corollary).

The last situation is not interesting because the function $g(.)$ is not estimable. The second situation results in the necessity to repeat the experiment in order to be able to estimate the function $g(.)$. The first situation is interesting because of the possibility to compare the estimator based on the realization of the vector $\boldsymbol{Y}$ with the estimator based on the realization of the matrix $\mathbf{S}$.

This comparison is made only for the neighbourhood of the point $\boldsymbol{\gamma} = \boldsymbol{v}$; similarly as in Part 3 the covariance matrix $\boldsymbol{\Sigma}$ of stochastically independent, normal and equally distributed random vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}$ is assumed regular.

The Fisher information matrix of the vector $\boldsymbol{Y}_i$, the $(k+1)$-tuple of the vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}$ and the vector $\bar{\boldsymbol{Y}}$ for the parameter $(\boldsymbol{\beta}', \boldsymbol{v}')$ are denoted sequentially $\mathbf{F}_1(\boldsymbol{\beta}, \boldsymbol{v})$, $\mathbf{F}_2(\boldsymbol{\beta}, \boldsymbol{v})$ and $\mathbf{F}_3(\boldsymbol{\beta}, \boldsymbol{v})$. The Fisher information matrix of the matrix $\mathbf{S}$ is denoted $\mathbf{F}_4(\boldsymbol{v})$. Analogously to the remark 3.4 we obtain

$$\mathbf{F}_1(\boldsymbol{\beta}, \boldsymbol{v}) = \begin{bmatrix} \mathbf{X}'\left(\sum_{i=1=1}^{m} v_i \mathbf{V}_i\right)^{-1}\mathbf{X}, & \boldsymbol{0} \\ \boldsymbol{0}, & \frac{1}{2}\mathbf{M}^{(v)} \end{bmatrix}; \quad \mathbf{F}_2(\beta, \boldsymbol{v}) = (k+1)\mathbf{F}_1(\boldsymbol{\beta}, \boldsymbol{v});$$

$$\mathbf{F}_3(\boldsymbol{\beta}, \boldsymbol{v}) = \begin{bmatrix} (k+1)\mathbf{X}'\left(\sum_{i=1}^{m} v_i \mathbf{V}_i\right)^{-1}\mathbf{X}, & \boldsymbol{0} \\ \boldsymbol{0}, & \frac{1}{2}\mathbf{M}^{(v)} \end{bmatrix}; \quad \mathbf{F}_4(\boldsymbol{v}) = (k/2)\mathbf{M}^{(v)},$$

where $\mathbf{M}^{(v)}$ is the matrix mentioned in Theorem 3.2.

The values $2\boldsymbol{\lambda}'\mathbf{M}^{(v)^{-1}}\boldsymbol{\lambda}$, $[2/(k+1)]\boldsymbol{\lambda}'\mathbf{M}^{(v)^{-1}}\boldsymbol{\lambda}$, $2\boldsymbol{\lambda}'\mathbf{M}^{(v)^{-1}}\boldsymbol{\lambda}$ and $(2/k)\boldsymbol{\lambda}'\mathbf{M}^{(v)^{-1}}\boldsymbol{\lambda}$ give the Rao—Cramér lower bound for estimators based on the realization of the vector $\boldsymbol{Y}$, on the relalization of the $(k+1)$-tuple vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}$, on the realization of the vector $\bar{\boldsymbol{Y}}$ and on the realization of the matrix $\mathbf{S}$, respectively. In the last case we already know that in the sufficient small neighbourhood of the point $\boldsymbol{v}$ the dispersion of the MINUγE deviates unsubstantially from the corresponding lower bound.

The MINQUE based on the realization of the vector $\boldsymbol{Y}$ and respecting a priori the (approximate) value $\boldsymbol{\alpha}$ of the vector $\boldsymbol{v}$ from the sufficient small neighbourhood of the point $\boldsymbol{v}$ is $\widehat{\boldsymbol{\lambda}'\boldsymbol{v}} = \boldsymbol{Y}'\mathbf{A}_*\boldsymbol{Y}$ (see (7.1) in [4]), where the matrix $\mathbf{A}_* \in \mathscr{A}$ minizes the quantity $\mathrm{Tr}\left(\mathbf{A}\sum_{i=1}^{m}\alpha_i\mathbf{V}_i\mathbf{A}\sum_{j=1}^{m}\alpha_j\mathbf{V}_j\right)$ and satisfies the conditions $\mathrm{Tr}(\mathbf{A}_*\mathbf{V}_i)=\lambda_i$, $i=1, \ldots, m$ (unbiasedness) and $\mathbf{X}'\mathbf{A}_*=\mathbf{0}$ (invariance of the MINIQUE on the translation of the parameter β).

The dispersion of the MINIQUE at the point $\boldsymbol{\alpha}=\boldsymbol{v}$ does not attain the Rao—Cramér lower bound in general; thus $\mathscr{D}(\boldsymbol{Y}'\mathbf{A}_*\boldsymbol{Y}) \geqslant 2\boldsymbol{\lambda}'\mathbf{M}^{(v)^{-1}}\boldsymbol{\lambda} > (2/k)\boldsymbol{\lambda}'\mathbf{M}^{(v)^{-1}}\boldsymbol{\lambda} = \mathscr{D}(\tau_g(\mathbf{S}))$.

As the estimators $\boldsymbol{Y}_1'\mathbf{A}_*\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{k+1}'\mathbf{A}_*\boldsymbol{Y}_{k+1}$ are stochastically independent and have the same dispersion, we can combine them and obtain an estimator with the dispersion $\mathscr{D}(\boldsymbol{Y}'\mathbf{A}_*\boldsymbol{Y})/(k+1) \geqq (k/(k+1))\mathscr{D}(\tau_g(\mathbf{S}))$.

The estimator $\bar{\boldsymbol{Y}}'\mathbf{A}_*\bar{\boldsymbol{Y}}(k+1)$ of the function $g(\boldsymbol{v})=\boldsymbol{\lambda}'\boldsymbol{v}$, $\boldsymbol{v}\in\boldsymbol{v}_*$ has the same dispersion as the estimator $\boldsymbol{Y}_j'\mathbf{A}_*\boldsymbol{Y}_j$, but the first of them is stochastically independent on the estimator $\tau_g(\mathbf{S})$.

Thus if in the actual situation it is possible to obtain a realization of the matrix $\mathbf{S}$ from the results of a repeated regression experiment, we use it for estimating the function $g(\boldsymbol{v})=\boldsymbol{\lambda}'\boldsymbol{v}$, i.e. the estimator $\tau_g(\mathbf{S})=\mathrm{Tr}\,(\mathbf{AS})$ is to be used. This can be combined with the estimator $\bar{\boldsymbol{Y}}'(\mathbf{A})\bar{\boldsymbol{Y}}(k+1)$. The combination of estimators in this case has to be weighted, of course.

REFERENCES

[1] ANDERSON, T. W.: Introduction to Multivariate Statistical Analysis. J. Wiley, N. York 1958.

[2] KUBÁČEK, L: Regression model with estimated covariance matrix. Math. Slovaca 33, 1983, 395—408.

[3] LEHMANN, E. L.—SCHEFFÉ, H.: Completeness, similar regions and unbiased estimation — Part I. Sankhya 10, 1950, 306—340.

[4] RAO, C. R.: Estimation of variance and covariance components-MINQUE theory. Journ. Multivariat. Analysis 1, 1971, 257—275.

[5] RAO, C. R.—MITRA, K. S.: Generalized Inverse of Matrices and its Application. J. Wiley, N. York 1971.

[6] RAO, C. R.—KLEFFE, J.: Estimation of Variance Components. In: Krisnaiah, P. R., ed. Handbook of Statistics, Vol. I. 1—40, North Holland, N. York 1980.

[7] SEELY, J.: Linear spaces and unbiased estimation. Ann. Math. Statistics, 41, 1971, 1725—1734.

[8] SEELY, J.: Linear spaces and unbiased estimation — application to the mixed linear model. Ann. Math. Statistics, 41, 1970, 1735—1748.

Received February 3, 1982

Matematický ústav SAV
Obrancov mieru 49
814 73 Bratislava

ПОВТОРЕНИЕ РЕГРЕССИОННОГО ЭКСПЕРИМЕНТА И ОЦЕНКА КОМПОНЕНТ КОВАРИАЦИОННОЙ МАТРИЦЫ

Lubomír Kubáček

Резюме

Предложена несмещенная оценка минимальной нормы (MINUE) компонент $v_1, \ldots, v_m$ ковариационной матрицы случайного векторы

$$\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \mathbf{\Sigma} = v_1\mathbf{V}_1 + \ldots + v_m\mathbf{V}_m),$$

основанная на реализации матрицы

$$\mathbf{S} = (1/k) \sum_{i=1}^{k+1} (\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Y}_i - \bar{\mathbf{Y}})'.$$

Сравнивается MINUE с оценкой, основанной на реализации случайного вектора