Mathematica Slovaca

Lubomír Kubáček
Regression model with estimated covariance matrix

Mathematica Slovaca, Vol. 33 (1983), No. 4, 395--408

Persistent URL: http://dml.cz/dmlcz/136344

Math. Slovaca 33, 1983, No. 4, 395—408

REGRESSION MODEL WITH ESTIMATED COVARIANCE MATRIX

LUBOMÍR KUBÁČEK

Introduction

The result of direct or indirect observations of parameters $\beta_1, \ldots, \beta_k$ is a realization of a random vector $\boldsymbol{Y}$. If the mean value $E_{\beta}(\boldsymbol{Y})$ of the random vector $\boldsymbol{Y}$ is $E_{\beta}(\boldsymbol{Y}) = \mathbf{X}\boldsymbol{\beta}$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$, ($\mathbf{X}$ is a given matrix of the type $n \times k$) and the covariance matrix $\boldsymbol{\Sigma}$ of the random vector $\boldsymbol{Y}$ does not depend on the vector $\boldsymbol{\beta}$, then the process of observations can be characterized by the regression model $(\boldsymbol{Y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, $\boldsymbol{\beta} \in \mathcal{R}^k$ (k-dimensional vector space).

If the covariance matrix $\boldsymbol{\Sigma}$ is not known a priori, but it is possible to obtain stochastically independent repeated realizations of the random vector $\boldsymbol{Y}$, i.e. a realization of the N-tuple stochastically independent random vectors $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_N$ with the same distribution is available, then it is possible to estimate the covariance matrix by the matrix

$$\mathbf{S} = (1/(N-1)) \sum_{i=1}^{N} (\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})', \quad \text{where} \quad \bar{\boldsymbol{Y}} = (1/N) \sum_{i=1}^{N} \boldsymbol{Y}_i.$$

In the case of normally distributed vectors $\boldsymbol{Y}_i$, $\boldsymbol{Y}_i \sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, $i = 1, \ldots, N$ $(>n)$ the vector $\bar{\boldsymbol{Y}}$ and the matrix $\mathbf{S}$ are stochastically independent, $\bar{\boldsymbol{Y}} \sim N_n(\mathbf{X}\boldsymbol{\beta}, (1/N)\boldsymbol{\Sigma})$, $(N-1)\mathbf{S} \sim W_n(N-1, \boldsymbol{\Sigma})$ (Wishart distribution with $N-1$ degrees of freedom). C. R. Rao (1967 [10]) utilized this fact for investigating stochastic properties of the least-squares estimator (LSE) of the vector $\boldsymbol{\beta}$, in which the unknown matrix $\boldsymbol{\Sigma}$ was substituted by the estimate $\mathbf{S}$ under condition of regularity of the regression model (i.e. $R(\mathbf{X})$ (rank of the matrix $\mathbf{X}$) $= k \leqq n$ and $R(\boldsymbol{\Sigma}) = n$; see the 2nd model in the 3rd section of this paper).

The aim of the paper is to show statistical properties of the above mentioned estimate in the case when the conditions of the regularity are not prescribed. The solution enables to calculate a larger class of problems from the theory of estimation (see section 3. Special cases).

1. Preliminaries

Let the random vector $\boldsymbol{Y}$ be normally distributed $\boldsymbol{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, $\boldsymbol{\beta} \in \mathcal{R}^k$ and let $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_f$ be stochastically independent random vectors with normal distribution $N_n(\boldsymbol{O}, \boldsymbol{\Sigma})$, where $f \geqq R(\boldsymbol{\Sigma})$. Let vector $\boldsymbol{Y}$ and vectors $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_f$ be stochastically independent. Let $\mathbf{S} = (1/f) \sum_{i=1}^{f} \boldsymbol{U}_i \boldsymbol{U}_i'$ and $\mathcal{M}(\boldsymbol{\Sigma})$ denote a subspace generated by columns of the matrix $\boldsymbol{\Sigma}$.

Lemma 1.1. *If $f \geqq R(\boldsymbol{\Sigma})$, then $\mathcal{M}(\boldsymbol{\Sigma}) = \mathcal{M}(\mathbf{S})$ with probability one. Consequently $R(\boldsymbol{\Sigma}) = R(\mathbf{S})$.*

Proof. See [8], Theorem 3.2.1 and Remark 3.2.1.

Lemma 1.2. *The random variable $T^2 = (\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{S}^- (\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta})$ does not depend on the choice of the g-inversion $\mathbf{S}^-$ of the matrix $\mathbf{S}$ and has the same distribution as $\{fR(\boldsymbol{\Sigma})/[f - R(\boldsymbol{\Sigma}) + 1]\} F_{R(\boldsymbol{\Sigma}), f-R(\boldsymbol{\Sigma})+1}$, where $F_{R(\boldsymbol{\Sigma}), f-R(\boldsymbol{\Sigma})+1}$ is the Fisher—Snedecor random variable with $R(\boldsymbol{\Sigma})$ and $f - R(\boldsymbol{\Sigma}) + 1$ degrees of freedom.*

Proof. See [9], Theorem 1.

Lemma 1.3. *The class of all unbiasedly estimable linear function of the parameter $\boldsymbol{\beta}$ is characterized by the vector $\mathbf{X}\boldsymbol{\beta}$. If the matrix $\boldsymbol{\Sigma}$ is a priori known, then the BLUE of this vector is $\widehat{\mathbf{X}\boldsymbol{\beta}} = \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]' \boldsymbol{Y}$, where $(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}$ is the minimum $\boldsymbol{\Sigma}$-seminorm g-inversion of the matrix $\mathbf{X}'$ (this type of g-inversion is a solution of the equations $\mathbf{X}'(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}\mathbf{X}' = \mathbf{X}'$ and $[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}\mathbf{X}']'\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}\mathbf{X}'$, see [11], p. 46). The estimate $\widehat{\mathbf{X}\boldsymbol{\beta}}$ does not depend on the choice of the g-inversion of that type with probability one. $\mathcal{D}(\widehat{\mathbf{X}\boldsymbol{\beta}}) = \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{\Sigma}$ (covariance matrix of the estimate $\widehat{\mathbf{X}\boldsymbol{\beta}}$; it does not depend on the choice of the g-inversion of that type either).*

Proof. See [14], Theorem 1.

In the following the symbol $\mathbf{Z}$ denotes a matrix of the type $n \times s$, $s \geqq n - R(\mathbf{X})$, which satisfies the condition $\mathcal{M}(\mathbf{Z}) = \operatorname{Ker}(\mathbf{X}') = \{\boldsymbol{u} : \mathbf{X}'\boldsymbol{u} = \boldsymbol{0}\}$. The vector $\boldsymbol{T}_2 = \mathbf{Z}'\boldsymbol{Y}$ characterizes the class of all unbiased linear estimators of the zero.

Lemma 1.4. *A statistic $\boldsymbol{L}'\boldsymbol{Y}$ estimates its mean value with minimal variance iff* $\operatorname{cov}(\boldsymbol{L}'\boldsymbol{Y}, \mathbf{Z}'\boldsymbol{Y}) = \boldsymbol{0}$.

Proof. The statement is a consequence of Theorem 5.3 of [6].

The statistic $\boldsymbol{T}_1 = \mathbf{X}\mathbf{X}^-\boldsymbol{Y}$ is the unbiased estimate of the vector $\mathbf{X}\boldsymbol{\beta}$ for an arbitrary choice of the g-inversion of the matrix $\mathbf{X}$. Further

$$\begin{bmatrix} \boldsymbol{T}_1 \\ \boldsymbol{T}_2 \end{bmatrix} \sim N_{n+s} \left[\begin{pmatrix} \mathbf{X}\boldsymbol{\beta} \\ \boldsymbol{0} \end{pmatrix}, \boldsymbol{\Lambda} \right],$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_{11}, & \boldsymbol{\Lambda}_{12} \\ \boldsymbol{\Lambda}_{21}, & \boldsymbol{\Lambda}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{X}\mathbf{X}^-\boldsymbol{\Sigma}(\mathbf{X}^-)'\mathbf{X}', & \mathbf{X}\mathbf{X}^-\boldsymbol{\Sigma}\mathbf{Z} \\ \mathbf{Z}'\boldsymbol{\Sigma}(\mathbf{X}^-)'\mathbf{X}', & \mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z} \end{bmatrix}.$$

Lemma 1.5. *The expression* $\boldsymbol{\tau}_1^* = \boldsymbol{T}_1 - \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^- \boldsymbol{T}_2 = \mathbf{XX}^-\boldsymbol{Y} - \mathbf{XX}^-\boldsymbol{\Sigma}\mathbf{Z}(\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z})^-\mathbf{Z}'\boldsymbol{Y}$, *which is invariant as regards the choice of the g-inversion* $\boldsymbol{\Lambda}_{22}^-$, *is the BLUE of the vector* $\mathbf{X}\boldsymbol{\beta}$, *thus with probability one* $\boldsymbol{\tau}_1^* = \widehat{\mathbf{X}\boldsymbol{\beta}}$, *i.e.* $\mathbf{XX}^-\boldsymbol{Y} - \mathbf{XX}^-\boldsymbol{\Sigma}\mathbf{Z}\ (\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z})^-\mathbf{Z}'\boldsymbol{Y} = \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{Y}$.

Proof. As $\boldsymbol{T}_2 \in \mathcal{M}(\boldsymbol{\Lambda}_{22})$ with probability one and $\mathcal{M}(\boldsymbol{\Lambda}_{12}') \subset \mathcal{M}(\boldsymbol{\Lambda}_{22})$, $\boldsymbol{\tau}_1^*$ does not depend on the choice of the g-inversion $\boldsymbol{\Lambda}_{22}^-$. Evidently $\operatorname{cov}(\boldsymbol{\tau}_1^*, \mathbf{Z}'\boldsymbol{Y}) = \mathbf{0}$ and this with respect to Lemmas 1.4 and 1.3 is sufficient for the proof.

Lemma 1.6. *Let* $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m$ *be stochastically independent random vectors,* $\boldsymbol{Z}_i \sim N_n(\mathbf{A}\mathbf{w}_i, \boldsymbol{\Sigma})$, $i = 1, \ldots, m$, *where* $\mathbf{A}$ *is a matrix of the type* $n \times t$ *and* $\mathbf{w}_i$, $i = 1, \ldots, m$, *is a t-dimensional vector. If* $\mathbf{H} = \sum_{i=1}^{m} \mathbf{w}_i\mathbf{w}_i'$ *and* $R(\mathbf{H}) = r$, *then*

$$\sum_{i=1}^{m} \boldsymbol{Z}_i\boldsymbol{Z}_i' - \sum_{j=1}^{m} \boldsymbol{Z}_j\mathbf{w}_j'\mathbf{H}^-\left(\sum_{k=1}^{m} \boldsymbol{Z}_k\mathbf{w}_k'\right)' = \sum_{i=1}^{m-r} \boldsymbol{V}_i\boldsymbol{V}_i',$$

where $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_{m-r}$ *are stochastically independent random vectors with the same distribution* $N_n(\boldsymbol{0}, \boldsymbol{\Sigma})$ *and the matrices* $(\boldsymbol{V}_1, \ldots, \boldsymbol{V}_{m-r})$ *and*

$$\sum_{j=1}^{m} \boldsymbol{Z}_j\mathbf{w}_j'\mathbf{H}^-\left(\sum_{k=1}^{m} \boldsymbol{Z}_k\mathbf{w}_k'\right)'$$

are stochastically independent.

Proof. It is an unsubstantial modification of the proof of Theorem 4.3.2 in [1].

Further let

$$\hat{\boldsymbol{\Lambda}} = \begin{bmatrix} \mathbf{XX}^- \\ \mathbf{Z}' \end{bmatrix} \mathbf{S}[(\mathbf{X}^-)'\mathbf{X}', \mathbf{Z}] = (1/f)\sum_{i=1}^{f} \begin{bmatrix} \mathbf{XX}^- \\ \mathbf{Z}' \end{bmatrix} \boldsymbol{U}_i \left[\begin{bmatrix} \mathbf{XX}^- \\ \mathbf{Z}' \end{bmatrix} \boldsymbol{U}_i\right]'.$$

The distribution of the random matrix $f\hat{\boldsymbol{\Lambda}}$ is a Wishart one: $f\hat{\boldsymbol{\Lambda}} \sim W_{n+s}(f, \boldsymbol{\Lambda})$ (the assumption $f \geqq R(\boldsymbol{\Sigma})$ implies $f \geqq R(\boldsymbol{\Lambda})$ $(\leqq R(\boldsymbol{\Sigma}))$, which enables to define correctly the distribution of the matrix $f\hat{\boldsymbol{\Lambda}}$; for details see [8] chapt. 3).

In the next section the statistical properties of the estimator $\hat{\boldsymbol{\tau}} = \boldsymbol{T}_1 - \hat{\boldsymbol{\Lambda}}_{12}\hat{\boldsymbol{\Lambda}}_{22}^-\boldsymbol{T}_2$ are investigated.

2. Statistical properties of the estimator $\hat{\boldsymbol{\tau}}$

For the sake of simplicity the following denotation is used. All random vectors and matrices conditioned by the matrix $(\boldsymbol{T}_2, \hat{\boldsymbol{\Lambda}}_{22})$ are denoted by a right upper index (p), e.g. $\hat{\boldsymbol{\tau}}^{(p)}$.

Theorem 2.1. *The random vector* $\hat{\boldsymbol{\tau}}^{(p)}$ *and the matrix* $\hat{\boldsymbol{\Lambda}}_{11.2}^{(p)} = \hat{\boldsymbol{\Lambda}}_{11}^{(p)} - \hat{\boldsymbol{\Lambda}}_{12}^{(p)}\hat{\boldsymbol{\Lambda}}_{22}^-\hat{\boldsymbol{\Lambda}}_{21}^{(p)}$ *are stochastically independent and*

$$\hat{\boldsymbol{\tau}}^{(p)} \sim N_n(\mathbf{X}\boldsymbol{\beta}, [1+(1/f)\boldsymbol{T}_2'(\mathbf{Z}'\mathbf{S}\mathbf{Z})^- \boldsymbol{T}_2]\boldsymbol{\Lambda}_{11.2}); \tag{2.1}$$

$$f\hat{\boldsymbol{\Lambda}}_{11.2}^{(p)} \sim W_n(f-R(\boldsymbol{\Lambda}_{22}), \boldsymbol{\Lambda}_{11.2}); \tag{2.2}$$

all given expressions are independent of the used g-inversion of matrices.

Proof. Independence from the choice of the g-inversion is implied by Lemma 1.1 and by the fact that with probability one $\boldsymbol{T}_2 \in \mathcal{M}(\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z})$. Let further

$$\boldsymbol{V}_\alpha = \begin{bmatrix} \boldsymbol{V}_{\alpha 1} \\ \boldsymbol{V}_{\alpha 2} \end{bmatrix} = \begin{bmatrix} \mathbf{X}\mathbf{X}^- \boldsymbol{U}_\alpha \\ \mathbf{Z}'\boldsymbol{U}_\alpha \end{bmatrix}, \quad \alpha = 1, \ldots, f.$$

With this denotation the vector

$$\hat{\boldsymbol{\tau}}^{(p)} = \boldsymbol{T}_1^{(p)} - \sum_{\alpha=1}^{f} \boldsymbol{V}_{\alpha 1}^{(p)} \boldsymbol{V}_{\alpha 2}' \left(\sum_{\beta=1}^{f} \boldsymbol{V}_{\beta 2} \boldsymbol{V}_{\beta 2}' \right) \boldsymbol{T}_2,$$

where

$$\boldsymbol{V}_{\alpha 1}^{(p)} \sim N_n(\boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^- \boldsymbol{V}_{\alpha 2}, \boldsymbol{\Lambda}_{11.2})$$

and

$$\boldsymbol{T}_1^{(p)} \sim N_n(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^- \boldsymbol{T}_2, \boldsymbol{\Lambda}_{11.2}).$$

Thus

$$\mathrm{E}(\hat{\boldsymbol{\tau}}^{(p)}) = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^- \boldsymbol{T}_2 - \sum_{\alpha=1}^{f} \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^- \boldsymbol{V}_{\alpha 2} \boldsymbol{V}_{\alpha 2}' \left(\sum_{\beta=1}^{f} \boldsymbol{V}_{\beta 2} \boldsymbol{V}_{\beta 2}' \right) \boldsymbol{T}_2 = \mathbf{X}\boldsymbol{\beta}$$

and

$$\mathscr{D}(\hat{\boldsymbol{\tau}}^{(p)}) = \boldsymbol{\Lambda}_{11.2} + \sum_{\alpha=1}^{f} \boldsymbol{T}_2' \left[\left(\sum_{\beta=1}^{f} \boldsymbol{V}_{\beta 2} \boldsymbol{V}_{\beta 2}' \right)^- \right]' \boldsymbol{V}_{\alpha 2} \boldsymbol{V}_{\alpha 2}' \left(\sum_{\beta=1}^{f} \boldsymbol{V}_{\beta 2} \boldsymbol{V}_{\beta 2}' \right) \boldsymbol{T}_2 \boldsymbol{\Lambda}_{11.2} =$$

$$= [1+(1/f)\boldsymbol{T}_2'(\mathbf{Z}'\mathbf{S}\mathbf{Z})^- \boldsymbol{T}_2]\boldsymbol{\Lambda}_{11.2},$$

which proves (2.1).

Further

$$f\hat{\boldsymbol{\Lambda}}_{11.2}^{(p)} = f\hat{\boldsymbol{\Lambda}}_{11}^{(p)} - f\hat{\boldsymbol{\Lambda}}_{12}^{(p)} \hat{\boldsymbol{\Lambda}}_{22}^- \hat{\boldsymbol{\Lambda}}_{21}^{(p)} =$$

$$= \sum_{\alpha=1}^{f} \boldsymbol{V}_{\alpha 1}^{(p)} \boldsymbol{V}_{\alpha 1}^{(p)\prime} - \sum_{\alpha=1}^{f} \boldsymbol{V}_{\alpha 1}^{(p)} \boldsymbol{V}_{\alpha 2}' \left(\sum_{\beta=1}^{f} \boldsymbol{V}_{\beta 2} \boldsymbol{V}_{\beta 2}' \right)^- \sum_{\gamma=1}^{f} \boldsymbol{V}_{\gamma 2} \boldsymbol{V}_{\gamma 1}^{(p)}.$$

As $\boldsymbol{V}_{\alpha 1}^{(p)} \sim N_n(\boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^- \boldsymbol{V}_{\alpha 2}, \boldsymbol{\Lambda}_{11.2})$, we can substitute the matrix $\boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^-$ for the matrix $\mathbf{A}$ in Lemma 1.6 and the vector $\boldsymbol{V}_{\alpha 2}$ for the vector $\boldsymbol{w}_\alpha$ as well. It implies:

$$f\hat{\boldsymbol{\Lambda}}_{11.2}^{(p)} = \sum_{\alpha=1}^{f-R(\boldsymbol{\Lambda}_{22})} \boldsymbol{S}_\alpha \boldsymbol{S}_\alpha', \quad \text{where} \quad \boldsymbol{S}_1, \ldots, \boldsymbol{S}_{f-R(\boldsymbol{\Lambda}_{22})}$$

are stochastically independent random vectors with the same distribution $N_n(\boldsymbol{0}, \boldsymbol{\Lambda}_{11.2})$. It proves (2.2).

Stochastical independence of the vector $\hat{\boldsymbol{\tau}}^{(p)}$ and the matrix $\hat{\boldsymbol{\Lambda}}_{11.2}^{(p)}$ follows from Lemma 1.6, namely the expression $\sum_{\alpha=1}^{f-R(\boldsymbol{\Lambda}_{22})} \boldsymbol{S}_\alpha \boldsymbol{S}_\alpha'$ does not depend on the second term of the expression for the vector $\hat{\boldsymbol{\tau}}^{(p)}$; independence from the first term $\boldsymbol{T}_1^{(p)}$ is an obvious consequence of our assumptions.

Remark 2.1. In the course of the proof conditioning by the matrix $(\boldsymbol{T}_2, \mathbf{Z}'(\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_f))$ was used. As in the resulting conditioned distributions the matrix $(\boldsymbol{T}_2, \hat{\boldsymbol{\Lambda}}_{22})$ appears, the latter was used in the formulation of the theorem.

Remark 2.2. Lemma 1.1 and the identity $\mathbf{X}\mathbf{X}^-\boldsymbol{Y} - \mathbf{X}\mathbf{X}^-\boldsymbol{\Sigma}\mathbf{Z}(\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z})^-\mathbf{Z}'\boldsymbol{Y} = \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{Y}$ from Lemma 1.5 imply $\hat{\boldsymbol{\tau}} = \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}$. We denoted by $\widehat{\mathbf{X}\boldsymbol{\beta}} = \boldsymbol{\tau}^* = \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{Y}$ the BLUE of $\mathbf{X}\boldsymbol{\beta}$ (this estimate is used in the case of the a priori known matrix $\boldsymbol{\Sigma}$; see Lemma 1.3); analogously we denote $\widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}} = \hat{\boldsymbol{\tau}} = \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}$.

Lemma 2.1. *For the quantity* $\boldsymbol{T}_2'\hat{\boldsymbol{\Lambda}}_{22}^-\boldsymbol{T}_2 = \boldsymbol{Y}'\mathbf{Z}(\mathbf{Z}'\mathbf{S}\mathbf{Z})^-\mathbf{Z}'\boldsymbol{Y}$ *it is true that* $\boldsymbol{T}_2'\hat{\boldsymbol{\Lambda}}_{22}^-\boldsymbol{T}_2 = (\boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}})'\mathbf{S}^-(\boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}})$.

Proof. Without loss of generality the matrix $\mathbf{Z}$ can be expressed in the form $\mathbf{Z} = \mathbf{I} - (\mathbf{X}')^-_{m(\mathbf{S})}\mathbf{X}'$, thus $\boldsymbol{T}_2 = \mathbf{Z}'\boldsymbol{Y} = \boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}}$. Using the identity $[(\mathbf{X}')^-_{m(\mathbf{S})}\mathbf{X}']'\mathbf{S} = \mathbf{S}(\mathbf{X}')^-_{m(\mathbf{S})}\mathbf{X}'$, which is valid for a minimum $\mathbf{S}$-seminorm g-inversion of the matrix $\mathbf{X}'$, we get

$$\hat{\boldsymbol{\Lambda}}_{22} = \mathbf{Z}'\mathbf{S}\mathbf{Z} = \{\mathbf{I} - \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\}\mathbf{S}[\mathbf{I} - (\mathbf{X}')^-_{m(\mathbf{S})}\mathbf{X}'] = \{\mathbf{I} - \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\}\mathbf{S}.$$

The last expression does not depend on the choice of the g-inversion and therefore we use the matrix $(\mathbf{X}')^+_{\mathbf{S},\mathbf{I}}$ (minimum $\mathbf{S}$-seminorm $\mathbf{I}$-least squares g-inversion; for details see [11]) for the matrix $(\mathbf{X}')^-_{m(\mathbf{S})}$. Then the matrix $\mathbf{X}[(\mathbf{X}')^+_{\mathbf{S},\mathbf{I}}]'$ is a Euclidean projector on the subspace $\mathcal{M}(\mathbf{X})$. As a Euclidean projector is its own g-inversion, we get

$$\hat{\boldsymbol{\Lambda}}_{22}^- = \langle\{\mathbf{I} - \mathbf{X}[(\mathbf{X}')^+_{\mathbf{S},\mathbf{I}}]'\}\mathbf{S}\rangle^- = \mathbf{S}^-\{\mathbf{I} - \mathbf{X}[(\mathbf{X}')^+_{\mathbf{S},\mathbf{I}}]'\}$$

and thus

$$\boldsymbol{T}_2'\hat{\boldsymbol{\Lambda}}_{22}^-\boldsymbol{T}_2 = (\boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}})\mathbf{S}^-\{\mathbf{I} - \mathbf{X}[(\mathbf{X}')^+_{\mathbf{S},\mathbf{I}}]'\}(\boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}}).$$

As

$$\{\mathbf{I} - \mathbf{X}[(\mathbf{X}')^+_{\mathbf{S},\mathbf{I}}]'\}(\boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}}) = \boldsymbol{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}},$$

the lemma is proved.

Lemma 2.2. *The rank of the matrices* $\boldsymbol{\Lambda}$, $\boldsymbol{\Lambda}_{11.2}$ *and* $\boldsymbol{\Lambda}_{22}$ *respectively is:*

$$R(\boldsymbol{\Lambda}) = R(\boldsymbol{\Sigma}), \quad R(\boldsymbol{\Lambda}_{11.2}) = R[\boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{X}')^-\mathbf{X}], \quad R(\boldsymbol{\Lambda}_{22}) = R(\boldsymbol{\Lambda}) - R(\boldsymbol{\Lambda}_{11.2}).$$

Proof. The matrix $\boldsymbol{\Sigma}$ is positive semidefinite and therefore there exists a matrix $\mathbf{J}$ of the type $n \times R(\boldsymbol{\Sigma})$ such that $\boldsymbol{\Sigma} = \mathbf{J}\mathbf{J}'$. As for every matrix $\mathbf{A}$ $R(\mathbf{A}) = R(\mathbf{A}\mathbf{A}')$, we have

$$R(\boldsymbol{\Lambda}) = R\left[\begin{pmatrix}\mathbf{X}\mathbf{X}^- \\ \mathbf{Z}'\end{pmatrix}\mathbf{J}\mathbf{J}'[(\mathbf{X}^-)'\mathbf{X}', \mathbf{Z}]\right] = R\left[\begin{pmatrix}\mathbf{X}\mathbf{X}^- \\ \mathbf{Z}'\end{pmatrix}\mathbf{J}\right].$$

Using Lemma 7.1.2 from [11] we obtain

$$R\begin{bmatrix}\mathbf{XX}\\ \mathbf{Z}'\end{bmatrix}=R[\mathbf{XX}^{-}\operatorname{Ker}(\mathbf{Z}')]+R(\mathbf{Z}')=$$

$$=R(\mathbf{XX}^{-}\mathbf{X})+R(\mathbf{Z}')=R(\mathbf{X})+n-R(\mathbf{X})=n,$$

thus the matrix $\begin{pmatrix}\mathbf{XX}^{-}\\ \mathbf{Z}\end{pmatrix}$ has full rank in its columns. This fact implies

$$R(\boldsymbol{\Lambda})=R\left[\begin{pmatrix}\mathbf{XX}^{-}\\ \mathbf{Z}'\end{pmatrix}\mathbf{J}\right]=R(\mathbf{J})=R(\boldsymbol{\Sigma}).$$

The identity $R(\boldsymbol{\Lambda}_{11.2})=R[\boldsymbol{\Sigma}(\boldsymbol{\Sigma}+\mathbf{XX}')^{-}\mathbf{X}]$ follows from the identity

$$\boldsymbol{\Lambda}_{11.2}=\mathbf{X}[(\mathbf{X}')^{-}_{m(\boldsymbol{\Sigma})}]'\boldsymbol{\Sigma}(\mathbf{X}')^{-}_{m(\boldsymbol{\Sigma})}\mathbf{X}'=\boldsymbol{\Sigma}(\mathbf{X}')^{-}_{m(\boldsymbol{\Sigma})}\mathbf{X}'$$

(this is the consequence of Lemmas 1.5 and 1.3 and of the properties of the minimum $\boldsymbol{\Sigma}$-seminorm g-inversion) and from Theorem 2.1 in [13] which states the identity $R[\boldsymbol{\Sigma}(\mathbf{X}')^{-}_{m(\boldsymbol{\Sigma})}\mathbf{X}']=R[\boldsymbol{\Sigma}(\boldsymbol{\Sigma}+\mathbf{XX}')^{-}\mathbf{X}]$.

The identity

$$\begin{bmatrix}\mathbf{I}, & -\boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-}\\ \mathbf{0}, & \mathbf{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Lambda}_{11}, & \boldsymbol{\Lambda}_{12}\\ \boldsymbol{\Lambda}_{21}, & \boldsymbol{\Lambda}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{I}, & \mathbf{0}\\ -\boldsymbol{\Lambda}_{22}^{-}\boldsymbol{\Lambda}_{21}, & \mathbf{I}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Lambda}_{11.2}, & \mathbf{0}\\ \mathbf{0}, & \boldsymbol{\Lambda}_{22}\end{bmatrix}$$

implies the last affirmation $R(\boldsymbol{\Lambda}_{22})=R(\boldsymbol{\Lambda})-R(\boldsymbol{\Lambda}_{11.2})$ of the lemma.

Now the following symbols will be used: $C=R(\boldsymbol{\Sigma})$, $v_1=R[\boldsymbol{\Sigma}(\boldsymbol{\Sigma}+\mathbf{XX}')^{-}\mathbf{X}]$, $v_2=f-R(\boldsymbol{\Sigma})+1$ and $\hat{\boldsymbol{v}}=\boldsymbol{Y}-\widehat{\mathbf{X}\boldsymbol{\beta}}$; the vector $\hat{\boldsymbol{v}}$ is an approximation of the error vector $\boldsymbol{Y}-\mathbf{X}\boldsymbol{\beta}$.

With respect to Lemma 1.1 in all the relations for the rank of the above mentioned matrices the matrix $\mathbf{S}$ can be used for the matrix $\boldsymbol{\Sigma}$.

Theorem 2.2. *The random variable*

$$T^2=\left\langle(\widehat{\mathbf{X}\boldsymbol{\beta}}-\mathbf{X}\boldsymbol{\beta})'\{\mathbf{X}[(\mathbf{X}')^{-}_{m(\mathbf{S})}]'\mathbf{S}\}^{-}(\widehat{\mathbf{X}\boldsymbol{\beta}}-\mathbf{X}\boldsymbol{\beta})\Big/\left(1+\frac{1}{f}\,\hat{\boldsymbol{v}}'\mathbf{S}^{-}\hat{\boldsymbol{v}}\right)\right\rangle[f-(C-v_1)]/f$$

has the same distribution as the random variable $[f-(C-v_1)](v_1/v_2)F_{v_1,v_2}$, *where* F_{v_1,v_2} *is the Fisher—Snedecor random variable with* v_1 *and* v_2 *degrees of freedom.*

Proof. Regarding Lemma 1.2 and Theorem 2.1 the random variable

$$T^2=(\hat{\boldsymbol{\tau}}^{(p)}-\mathbf{X}\boldsymbol{\beta})'\left(\frac{f}{f-R(\boldsymbol{\Lambda}_{22})}\right)^{-1}\hat{\boldsymbol{\Lambda}}_{11.2}^{-}(\hat{\boldsymbol{\tau}}^{(p)}-\mathbf{X}\boldsymbol{\beta})\left(1+\frac{1}{f}\,\boldsymbol{T}_2'\hat{\boldsymbol{\Lambda}}_{22}^{-}\boldsymbol{T}_2\right)^{-1}$$

has the same distribution as the random variable

$$\{[f-R(\boldsymbol{\Lambda}_{22})]R(\boldsymbol{\Lambda}_{11.2})/[f-R(\boldsymbol{\Lambda}_{22})-R(\boldsymbol{\Lambda}_{11.2})+1]\}F_{R(\boldsymbol{\Lambda}_{11.2}),\,f-R(\boldsymbol{\Lambda}_{22})-R(\boldsymbol{\Lambda}_{11.2})+1}.$$

The distribution of the last random variable does not depend on the matrix $(\boldsymbol{T}_2, \hat{\boldsymbol{A}}_{22})$. An application of Lemmas 2.1 and 2.2 respectively is sufficient for concluding the proof.

Corollary 1. *If $F_{\nu_1,\nu_2}(1-\alpha)$ is a $(1-\alpha)$ quantile of the Fisher—Snedecor random variable, then the $(1-\alpha)$-confidence ellipsoid of the vector $\mathbf{X}\boldsymbol{\beta}$ is given by the set*

$$\left\{\boldsymbol{u}: (\boldsymbol{u}-\widehat{\mathbf{X}\boldsymbol{\beta}})'\{\mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{S}\}^-(\boldsymbol{u}-\widehat{\mathbf{X}\boldsymbol{\beta}}) \leqq \right.$$
$$\left. \leqq (f\nu_1/\nu_2)\left(1+\frac{1}{f}\,\hat{\boldsymbol{v}}'\mathbf{S}^-\hat{\boldsymbol{v}}\right) F_{\nu_1,\nu_2}(1-\alpha)\right\}.$$

Corollary 2. *If the function $f(\boldsymbol{\beta})=\boldsymbol{p}'\boldsymbol{\beta}$ is unbiasedly estimable, i.e. if $\boldsymbol{p}\in\mathcal{M}(\mathbf{X}')$ $(\Leftrightarrow\exists\{\boldsymbol{u}\in\mathcal{R}^k\}\boldsymbol{p}=\mathbf{X}'\boldsymbol{u})$, then the interval $[\boldsymbol{p}'[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}-\varkappa,\ \boldsymbol{p}'[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}+\varkappa]$, where*

$$\varkappa=t_\varphi\left(1-\frac{\alpha}{2}\right)\sqrt{f/\varphi}\,\sqrt{\left(1+\frac{1}{f}\,\hat{\boldsymbol{v}}'\mathbf{S}^-\hat{\boldsymbol{v}}\right)\boldsymbol{p}'[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{S}(\mathbf{X}')^-_{m(\mathbf{S})}\boldsymbol{p}},$$

$\varphi=f-(C-\nu_1)$ and $t_\varphi\left(1-\frac{\alpha}{2}\right)$ is the $\left(1-\frac{\alpha}{2}\right)$ quantile of the Student random variable with φ degrees of freedom, covers the value $\boldsymbol{p}'\boldsymbol{\beta}$ with probability $1-\alpha$.

Proof. Taking into acount the relations

$$\hat{\boldsymbol{A}}_{11.2}=\mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{S}=\mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{S}(\mathbf{X}')^-_{m(\mathbf{S})}\mathbf{X}'$$

and $\boldsymbol{p}=\mathbf{X}'\boldsymbol{u}$, $\boldsymbol{u}\in\mathcal{R}^k$ we have

$$\boldsymbol{u}'\hat{\boldsymbol{A}}_{11.2}\boldsymbol{u}=\boldsymbol{p}'[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{S}(\mathbf{X}')^-_{m(\mathbf{S})}\boldsymbol{p}.$$

Theorem 2.1 implies stochastical independence of the random variables $\boldsymbol{u}'\hat{\boldsymbol{\tau}}^{(p)}$ and $f\boldsymbol{u}'\hat{\boldsymbol{A}}^{(p)}_{11.2}\boldsymbol{u}$, where

$$\boldsymbol{u}'\hat{\boldsymbol{\tau}}^{(p)}\sim N_1\left(\boldsymbol{p}'\boldsymbol{\beta},\left(1+\frac{1}{f}\,\hat{\boldsymbol{v}}'\mathbf{S}^-\hat{\boldsymbol{v}}\right)\boldsymbol{u}'\boldsymbol{A}_{11.2}\boldsymbol{u}\right)$$

and

$$f\boldsymbol{u}'\hat{\boldsymbol{A}}^{(p)}_{11.2}\boldsymbol{u}\sim W_1[(f-R(\boldsymbol{A}_{22})),\ \boldsymbol{u}'\boldsymbol{A}_{11.2}\boldsymbol{u}]\equiv\chi^2_{f-R(\boldsymbol{A}_{22})}\boldsymbol{u}'\boldsymbol{A}_{11.2}\boldsymbol{u}.$$

Symbol $\chi^2_{f-R(\boldsymbol{A}_{22})}$ denotes the random variable with the chi-square distribution with $f-R(\boldsymbol{A}_{22})$ degrees of freedom. Taking into acount the definition of the Student variable and its independence from the conditioning matrix $(\boldsymbol{T}_2, \hat{\boldsymbol{A}}_{22})$ we conclude the proof.

Lemma 2.3. *The random variable $\boldsymbol{T}_2'(\mathbf{Z}'\mathbf{S}\mathbf{Z})^-\boldsymbol{T}_2=(\boldsymbol{Y}-\widehat{\mathbf{X}\boldsymbol{\beta}})'\mathbf{S}^-(\boldsymbol{Y}-\widehat{\mathbf{X}\boldsymbol{\beta}})$ has the same distribution as $[f(C-\nu_1)/(\nu_1+\nu_2)]F_{C-\nu_1,\nu_1+\nu_2}$.*

Proof. With respect to our assumptions the random vector $\boldsymbol{T}_2$ and the random matrix $\mathbf{Z}'\mathbf{S}\mathbf{Z}$ are stochastically independent and $\boldsymbol{T}_2 \sim N_s(\boldsymbol{0}, \mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z})$, $f\mathbf{Z}'\mathbf{S}\mathbf{Z} \sim W_s(f, \mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z})$. Lemmas 1.2 and 2.2 imply val dity of the affirmation.

Theorem 2.3. $\widehat{\mathbf{X}\boldsymbol{\beta}} = \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}$ *(see Remark 2.2) is an unbiased estimate of the vector* $\mathbf{X}\boldsymbol{\beta}$ *and*

$$\mathscr{D}(\widehat{\mathbf{X}\boldsymbol{\beta}}) = \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{\Sigma}(C + \nu_2 - 2)/(\nu_1 + \nu_2 - 2).$$

Proof. The unbiasedness of the estimate $\widehat{\mathbf{X}\boldsymbol{\beta}}$ is an obvious consequence of (2.1). Using this relation and Lemmas 2.3 and 1.3 we get

$$\mathscr{D}(\widehat{\mathbf{X}\boldsymbol{\beta}}) = \mathrm{E}\,\{\mathscr{D}(\widehat{\mathbf{X}\boldsymbol{\beta}}|(\boldsymbol{T}_2, \hat{\boldsymbol{\Lambda}}_{22}))\} =$$

$$= \mathrm{E}\left[1 + \frac{1}{f}\,\boldsymbol{T}_2'(\mathbf{Z}'\mathbf{S}\mathbf{Z})^-\boldsymbol{T}_2\right]\boldsymbol{\Lambda}_{11.2} =$$

$$= \mathbf{X}[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{\Sigma}\left\{1 + \frac{C - \nu_1}{\nu_1 + \nu_2}\,\mathrm{E}(F_{C-\nu_1,\,\nu_1+\nu_2})\right\};$$

$\mathrm{E}(F_{C-\nu_1,\,\nu_1+\nu_2}) = (\nu_1 + \nu_2)/(\nu_1 + \nu_2 - 2)$ see [2] relation (16.28).

Corollary 3. *The variance of the estimate* $\boldsymbol{p}'[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}$ *of the unbiasedly estimable function* $f(\boldsymbol{\beta}) = \boldsymbol{p}'\boldsymbol{\beta}$ *is*

$$\boldsymbol{p}'[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\boldsymbol{\Sigma}(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}\boldsymbol{p}(C + \nu_2 - 2)/(\nu_1 + \nu_2 - 2).$$

Theorems 2.1, 2.2 and 2.3 characterize basic statistical properties of the LSE in which the empirical covariance matrix $\mathbf{S}$ with the Wishart distribution is used instead of the matrix $\boldsymbol{\Sigma}$.

It is quite clear that for $f \to \infty$ $(\Rightarrow \nu_2 \to \infty)\widehat{\mathbf{X}\boldsymbol{\beta}} \to \widehat{\mathbf{X}\boldsymbol{\beta}}$ with probability one and $\mathscr{D}(\widehat{\mathbf{X}\boldsymbol{\beta}}) \to \mathscr{D}(\widehat{\mathbf{X}\boldsymbol{\beta}})$.

3. Special cases

From the practical point of view the important cases are the following regular models [4], [3]:

1st model·

$$\boldsymbol{Y} \equiv \boldsymbol{\xi} \sim N_n(\boldsymbol{i}\beta, \boldsymbol{\Sigma}), \quad \boldsymbol{i} = (1, \ldots, 1)', \quad \beta \in \mathscr{R}^1,\ R(\boldsymbol{\Sigma}) = n,\ k = 1$$

(direct measurement of the scalar parameter β; n is the number of measurements);

2nd model:

$$\boldsymbol{Y} \equiv \boldsymbol{\xi} \sim N_n(\mathbf{A}\boldsymbol{\beta}, \boldsymbol{\Sigma}),\ R(\mathbf{A}) = k \leq n,\ \boldsymbol{\beta} \in \mathscr{R}^k,\ R(\boldsymbol{\Sigma}) = n$$

(indirect measurement of the k-dimensional parameter $\boldsymbol{\beta}$);

3rd model:

$$\mathbf{Y} \equiv \begin{bmatrix} \xi \\ -\mathbf{b} \end{bmatrix} \sim N_{n+q}\left[\begin{pmatrix} \mathbf{I} \\ \mathbf{B} \end{pmatrix} \boldsymbol{\beta}, \begin{pmatrix} \boldsymbol{\Sigma}_{11}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0} \end{pmatrix}\right], \quad \mathbf{b} \in \mathcal{R}^q$$

($\mathbf{b}$ is a given vector), $\boldsymbol{\beta} \in \mathcal{R}^n$, $R(\mathbf{B}) = q \leqq n$, $R(\boldsymbol{\Sigma}_{11}) = n = k$ (direct measurement of the n-dimensional vectorial parameter $\boldsymbol{\beta}$ with q side conditions);

4th model:

$$\mathbf{Y} \equiv \begin{bmatrix} \xi \\ -\mathbf{b} \end{bmatrix} \sim N_{n+q}\left[\begin{pmatrix} \mathbf{I}, & \mathbf{0} \\ \mathbf{B}_1, & \mathbf{B}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \right.$$

$$\left.\begin{pmatrix} \boldsymbol{\Sigma}_{11}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0} \end{pmatrix}\right], \quad \mathbf{b} \in \mathcal{R}^q$$

($\mathbf{b}$ is a given vector), $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$, $\boldsymbol{\beta}_1 \in \mathcal{R}^n$, $\boldsymbol{\beta}_2 \in \mathcal{R}^l$, $R(\mathbf{B}_1, \mathbf{B}_2) = q \leqq n + l$, $R(\mathbf{B}_2) = l \leqq q$, $k = n + l$ (direct measurement of the n-dimensional subvector of the k-dimensional vectorial parameter $\boldsymbol{\beta}$ with q side conditions);

5th model:

$$\mathbf{Y} \equiv \begin{bmatrix} \xi \\ -\mathbf{b} \end{bmatrix} \sim N_{n+q}\left[\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \boldsymbol{\beta}, \begin{pmatrix} \boldsymbol{\Sigma}_{11}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0} \end{pmatrix}\right], \quad \mathbf{b} \in \mathcal{R}^q$$

($\mathbf{b}$ is a given vector), $\boldsymbol{\beta} \in \mathcal{R}^k$, $R(\mathbf{A}) = k \leqq n$, $R(\mathbf{B}) = q \leqq k$, $R(\boldsymbol{\Sigma}_{11}) = n$ (indirect measurement of the k-dimensional vectorial parameter with a system of q conditions).

Next a review of expressions for $[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y}$ and

$$\mathcal{D}\{[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y}\} = \mathcal{D}\{[(\mathbf{X}')^-_{m(\boldsymbol{\Sigma})}]'\mathbf{Y}\}(C + \nu_2 - 2)/(\nu_1 + \nu_2 - 2)$$

for the single above mentioned regular model is given (regularity of all these models enables to estimate unbiasedly the whole vector $\boldsymbol{\beta}$ and this is reason for which the formulae for $\hat{\boldsymbol{\beta}} = [(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y}$ and its dispersion instead of the formulae for $\widehat{\mathbf{X}\boldsymbol{\beta}}$ and its dispersion are given).

1st model

$$[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y} = (\mathbf{i}'\mathbf{S}^{-1}\mathbf{i})^{-1}\mathbf{i}\mathbf{S}^{-1}\xi$$

$$\mathcal{D}\{[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y}\} = (\mathbf{i}'\boldsymbol{\Sigma}^{-1}\mathbf{i})^{-1}(f-1)/(f-n)$$

$$\nu_1 = 1, \quad k = 1$$

2nd model

$$[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y} = (\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{S}^{-1}\xi$$

$$\mathcal{D}\{[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y}\} = (\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-1}(f-1)/[f-(n-k)-1]$$

$$\nu_1 = k, \quad k = k$$

3rd model

$$[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y} = [\mathbf{I} - \mathbf{SB}'(\mathbf{BSB}')^{-1}\mathbf{B}]\xi - \mathbf{SB}'(\mathbf{BSB}')^{-1}\boldsymbol{b}$$

$$\mathscr{D}\{[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}\} = [\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{11}\mathbf{B}'(\mathbf{B}\boldsymbol{\Sigma}_{11}\mathbf{B}')^{-1}\mathbf{B}\boldsymbol{\Sigma}_{11}](f-1)/(f-q-1)$$

$$v_1 = n - q, \quad k = n$$

4th model

$$[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y} = \begin{bmatrix} \mathbf{I} - \mathbf{SB}_1'\hat{\mathbf{Q}}_{11}\mathbf{B}_1 \\ -\hat{\mathbf{Q}}_{21}\mathbf{B}_1 \end{bmatrix} \xi + \begin{bmatrix} -\mathbf{SB}_1\hat{\mathbf{Q}}_{11} \\ -\hat{\mathbf{Q}}_{21} \end{bmatrix} \boldsymbol{b}$$

$$\mathscr{D}\{[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{11}\mathbf{B}_1'\mathbf{Q}_{11}\mathbf{B}_1\boldsymbol{\Sigma}_{11}, & -\boldsymbol{\Sigma}_{11}\mathbf{B}_1'\mathbf{Q}_{12} \\ -\mathbf{Q}_{21}\mathbf{B}_1\boldsymbol{\Sigma}_{11}, & -\mathbf{Q}_{22} \end{bmatrix} (f-1)/[f-(q-l)-1]$$

$$v_1 = n - (q - l), \quad k = n + l$$

$$\begin{bmatrix} \hat{\mathbf{Q}}_{11}, \hat{\mathbf{Q}}_{12} \\ \hat{\mathbf{Q}}_{21}, \hat{\mathbf{Q}}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1\mathbf{SB}_1', \mathbf{B}_2 \\ \mathbf{B}_2', \quad \mathbf{0} \end{bmatrix}^{-1};$$

$$\begin{bmatrix} \mathbf{Q}_{11}, \mathbf{Q}_{12} \\ \mathbf{Q}_{21}, \mathbf{Q}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_1\boldsymbol{\Sigma}_{11}\mathbf{B}_1', \mathbf{B}_2 \\ \mathbf{B}_2', \quad \mathbf{0} \end{bmatrix}^{-1}$$

5th model

$$\begin{aligned}[(\mathbf{X}')_{m(\mathbf{S})}]'\boldsymbol{Y} = \{(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1} - \\ -(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{B}'[\mathbf{B}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{B}']^{-1}\mathbf{B}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\}\mathbf{A}'\mathbf{S}^{-1}\xi - \\ -(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{B}'[\mathbf{B}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{B}']^{-1}\boldsymbol{b}\end{aligned}$$

$$\begin{aligned}\mathscr{D}\{[(\mathbf{X}')^-_{m(\mathbf{S})}]'\boldsymbol{Y}\} = \{(\mathbf{A}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A})^{-1} - \\ -(\mathbf{A}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A})^{-1}\mathbf{B}'[\mathbf{B}(\mathbf{A}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A})^{-1}\mathbf{B}']^{-1}\mathbf{B}(\mathbf{A}'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A})^{-1}\}(f-1)/[f-(n-k+q)-1]\end{aligned}$$

$$v_1 = k - q, \quad k = k$$

The last three models are called models with conditions; they can be rewritten into a form with explicit conditions; e.g. $\xi \sim N_n(\boldsymbol{\beta}_1, \boldsymbol{\Sigma}_{11})$, $\boldsymbol{b} + \mathbf{B}_1\boldsymbol{\beta}_1 + \mathbf{B}_2\boldsymbol{\beta}_2 = \boldsymbol{0}$ etc.

4. Remarks on the structure of the covariance matrix

The aim of the measurement is not always to get estimates of a function $f(\boldsymbol{\beta})$, many times we have to estimate parameters of the covariance matrix $\boldsymbol{\Sigma}$. The most frequent structure of the matrix $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma} = \sum_{j=1}^{p} \lambda_j \mathbf{V}_j$, where λ_j, $j = 1, \ldots, p$ are unknown parameters and matrices $\mathbf{V}_1, \ldots, \mathbf{V}_p$, are known from the design of the experiment.

Two cases have to be distinguished in dependence on the input information:

a) We know the outcome of the random vector $\boldsymbol{Y}$ only;

b) We know the outcomes of the random vector $\mathbf{Y}$ and of the matrix $\mathbf{S}$.

Interesting cases occur when the matrix $\mathbf{\Sigma}$ has the form $\mathbf{\Sigma} = \lambda\mathbf{I} + \mathbf{X\Gamma X}' + \mathbf{Z\Theta Z}'$. The importance of this structure is shown in the following lemma.

Lemma 4.1 (*modification of Lemma 5a from* [10]). *The identity* $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y} = \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{\Sigma})}]'\mathbf{Y}$ (*and thus* $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{\Sigma X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}' = \mathscr{D}[\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}] = \mathscr{D}\{\mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{\Sigma})}]'\mathbf{Y}\} = \mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{\Sigma})}]\mathbf{\Sigma}$) *holds iff there exist matrices* $\mathbf{\Gamma}$, $\mathbf{\Theta}$ *and the number* λ *satisfying the condition* $\mathbf{\Sigma} = \lambda\mathbf{I} + \mathbf{X\Gamma X}' + \mathbf{Z\Theta Z}'$.

Proof. By Lemma 1.4 $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}$ is the BLUE of its mean value $\mathbf{X}\boldsymbol{\beta}$ iff $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{\Sigma Z} = \mathbf{0}$. The matrix $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ is the Euclidean projector on the subspace $\mathscr{M}(\mathbf{X})$. By Theorem 2.3.2 from [11] the set of all solution $\mathbf{\Sigma}$ of the equation $\mathbf{P\Sigma Z} = \mathbf{0}$ is $\mathbf{\Sigma} = \mathbf{B} - \mathbf{P}_0^-\mathbf{PBZZ}_0^-$, where $\mathbf{P}_0^-$ and $\mathbf{Z}_0^-$ respectively are arbitrary but fixed g-inversions of the matrices $\mathbf{P}$ and $\mathbf{Z}$ respectively and $\mathbf{B}$ is an arbitrary matrix with proper dimension. Let $\mathbf{X}_1$ and $\mathbf{Z}_1$ be matrices with a column full rank satisfying the condition $\mathscr{M}(\mathbf{X}) = \mathscr{M}(\mathbf{X}_1)$ and $\mathscr{M}(\mathbf{Z}) = \mathscr{M}(\mathbf{Z}_1)$. The matrix $(\mathbf{X}_1, \mathbf{Z}_1)$ is regular and $\mathbf{X}_1'\mathbf{Z}_1 = \mathbf{0}$. Every matrix $\mathbf{B}$ can be expressed in the form

$$\mathbf{B} = (\mathbf{X}_1\mathbf{Z}_1)\begin{bmatrix}\mathbf{\Gamma}_1, & \mathbf{M}_1\\ \mathbf{L}_1, & \mathbf{\Theta}_1\end{bmatrix}\begin{bmatrix}\mathbf{X}_1'\\ \mathbf{Z}_1'\end{bmatrix}.$$

The matrices $\mathbf{P}_0^-$ and $\mathbf{Z}_0^-$ respectively are chosen in such a way that $\mathbf{P}_0^-\mathbf{P} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1' = \mathbf{P}$ and $\mathbf{ZZ}_0^- = \mathbf{Z}_1(\mathbf{Z}_1'\mathbf{Z}_1)^-\mathbf{Z}_1' = \mathbf{P}_1$ (the Euclidean projector on the subspace $\mathscr{M}(\mathbf{Z})$). Then with respect to the relation $\mathbf{X}_1'\mathbf{Z}_1 = \mathbf{0}$, there holds $\mathbf{\Sigma} = \mathbf{X}_1\mathbf{\Gamma}_1\mathbf{X}_1' + \mathbf{Z}_1\mathbf{L}_1\mathbf{X}_1' + \mathbf{Z}_1\mathbf{\Theta}_1\mathbf{Z}_1'$. Because of $\mathbf{P\Sigma Z} = \mathbf{0} \Leftrightarrow \mathbf{P\Sigma Z}_1 = \mathbf{0} \Leftrightarrow \mathbf{Z}_1'\mathbf{\Sigma P} = \mathbf{0}$, where $\mathbf{Z}_1'\mathbf{\Sigma P} = \mathbf{Z}_1'\mathbf{Z}_1\mathbf{L}_1\mathbf{X}_1'$, we have $\mathbf{L}_1 = \mathbf{0}$ ($\mathbf{Z}_1'\mathbf{Z}_1$ is a regular matrix and $\mathbf{X}_1$ has a row full rank). Choosing $\mathbf{\Gamma}_1 = \mathbf{\Gamma}_2 + \lambda(\mathbf{X}_1'\mathbf{X}_1)^{-1}$, $\mathbf{\Theta}_1 = \mathbf{\Theta}_2 + \lambda(\mathbf{Z}_1'\mathbf{Z}_1)^{-1}$ and taking into account the identity $\mathbf{I} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1' + \mathbf{Z}_1(\mathbf{Z}_1'\mathbf{Z}_1)^-\mathbf{Z}_1'$ we get $\mathbf{\Sigma} = \mathbf{X}_1\mathbf{\Gamma}_2\mathbf{X}_1' + \mathbf{Z}_1\mathbf{\Theta}_2\mathbf{Z}_1' + \lambda\mathbf{I}$. For the matrix $\mathbf{X}$ there exists a matrix $\mathbf{M}$ that $\mathbf{X} = \mathbf{X}_1\mathbf{M}$ and therefore $\mathbf{X}_1\mathbf{\Gamma}_2\mathbf{X}_1' = \mathbf{X}_1\mathbf{M\Gamma M}'\mathbf{X}_1' = \mathbf{X\Gamma X}'$; similarly we can reestablish the term $\mathbf{Z}_1\mathbf{\Theta}_2\mathbf{Z}_1'$. By application of Lemma 1.3 the proof is concluded.

Remark 4.1. By Lemma 4.1 the best estimate of the unknown vector $\mathbf{X}\boldsymbol{\beta}$ in the case $\mathbf{\Sigma} = \lambda\mathbf{I} + \mathbf{X\Gamma X}' + \mathbf{Z\Theta Z}'$ is $\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{Y}$. Application of the matrix $\mathbf{S}$ in the corresponding estimate $\mathbf{X}[(\mathbf{X}')^-_{m(\mathbf{S})}]'\mathbf{Y}$ results in the enlargement of dispersions with respect to the BLUE. Of course, in the case when we do not know anything about the structure of the matrix $\mathbf{\Sigma}$ we are thrown upon utilization of the matrix $\mathbf{S}$.

Lemma 4.2. *Consider the regression model*

$$\left(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{\Sigma} = \sum_{i=1}^{p}\lambda_i\mathbf{V}_i\right).$$

The function $g(\lambda_1, \ldots, \lambda_p) = \mathbf{g}'\boldsymbol{\lambda}$ *is unbiasedly estimable by the statistic* $\mathbf{Y}'\mathbf{AY}$ ($\mathbf{A}$ *is a symmetric matrix) iff* $\mathbf{g} \in \mathscr{M}(\mathbf{H})$, *where* $\mathbf{H}$ *is a matrix of the type* $p \times p$, *the elements of which are*

$$\{\mathbf{H}\}_{i,j} = \mathrm{Tr}(\mathbf{V}_i\mathbf{V}_j - \mathbf{P}\mathbf{V}_i\mathbf{P}\mathbf{V}_j), \quad \mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}', \quad i, j = 1, \ldots, p.$$

Proof. See [7], Theorem 3.

Remark 4.2. Lemma 4.2 shows the influence of the structure of the covariance matrix $\boldsymbol{\Sigma}$ on the estimability of its parameters. Consider the following example. Let $\boldsymbol{f}_i \in \mathcal{R}^n$, $i = 1, \ldots, l$ be orthonormal vectors and $\mathbf{X} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_k)$, $k \leqq l$. Let the covariance matrix be of the form

$$\boldsymbol{\Sigma} = \gamma_1 \boldsymbol{f}_1 \boldsymbol{f}_1' + \ldots + \gamma_k \boldsymbol{f}_k \boldsymbol{f}_k' + O_1 \boldsymbol{f}_{k+1} \boldsymbol{f}_{k+1}' + \ldots + \Theta_{l-k} \boldsymbol{f}_l \boldsymbol{f}_l' = \mathbf{X}\boldsymbol{\Gamma}\mathbf{X}' + \mathbf{Z}\boldsymbol{\Theta}\mathbf{Z}',$$

where $\boldsymbol{\Gamma} = \mathrm{diag}(\gamma_1, \ldots, \gamma_l)$, $\boldsymbol{\Theta} = \mathrm{diag}(\Theta_1, \ldots, \Theta_{l-k}, 0, \ldots, 0)$ and $\mathbf{Z} = (\boldsymbol{f}_{k+1}, \ldots, \boldsymbol{f}_l, \boldsymbol{f}_{l+1}, \ldots, \boldsymbol{f}_n)$; the vectors $\boldsymbol{f}_{l+1}, \ldots, \boldsymbol{f}_n$ complete the vectors $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_l$ to be a base of the space $\mathcal{R}^n$. This situation is typical for the investigation of the stochastic structure of measured geophysical potential fields (see [5]). The structure of the matrix $\mathbf{H}$ from Lemma 4.2 is

$$\mathbf{H} = \begin{bmatrix} \mathbf{0}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{I} \end{bmatrix},$$

where $\mathbf{I}$ has the dimension $(l-k) \times (l-k)$. That is the reason why it is possible to estimate only the parameters $\Theta_1, \ldots, O_{l-k}$ by the vector $\boldsymbol{Y}$. If in the just considered case there holds $\mathcal{M}(\mathbf{X}) = \mathcal{M}(\boldsymbol{\Sigma})$, then it is impossible to estimate any parameter. In the model $(\boldsymbol{Y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma} = \sigma^2\mathbf{V})$ the condition $R(\mathbf{V}, \mathbf{X}) - R(\mathbf{X}) > 0$ is sufficient for the estimability of the parameter σ^2 (for more details see [11]).

The situation changes essentially if we have at disposal a realization of the matrix $\mathbf{S}$, $f\mathbf{S} \sim W_n(f, \boldsymbol{\Sigma})$. In the case of $\boldsymbol{\Sigma} = \sum_{i=1}^{l} \lambda_i \boldsymbol{f}_i \boldsymbol{f}_i'$ it is obvious that $\hat{\lambda}_i = \boldsymbol{f}_i'\mathbf{S}\boldsymbol{f}_i = \lambda_i \chi_f^2/f$ is an unbiased estimate of the parameter λ_i and the distribution of the chi-square enables to determine the confidence interval for λ_i as well (for some details of the spectral decomposition of the matrix $\mathbf{S}$ see [8], p. 86).

REFERENCES

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

[2] KENDALL, M. G., STUART, A.: The Advanced Theory of Statistics. Griffin, London 1963.

[3] KUBÁČEK, L.: On the General Problem of Adjustement of Measured Values. Matematický časopis 19, 1969, No 4, 270—275.

[4] KUBÁČEK, L.: Universal Model for Adjusting Observed Values. Studia geoph. et geod. 20, 1978, 103—113.

[5] KUBÁČKOVÁ, L.: Algorithmization of Computations in a Discretely Observed Gravitational Field. Contr. Geophys. Inst. SAS, 13, 1982, 85—93.

[6] LEHMANN, E. L., SCHEFFÉ, H.: Completeness, Similar Regions and Unbiased Estimation — Part I. Sankhyā 10, 1950, 305—340.

[7] SEELY, J.: Linear Spaces and Unbiased Estimation — Application to Mixed Linear Model. Ann. Math. Statist. 42, 1970, 710—721.

[8] SRIVASTAVA, M. S., KHATRI, C. G.: An Introduction to Multivariate Statistics. North Holland, N. York 1979.

[9] STREIT, F.: On Generalized Inverse Matrices and Multivariate Statistical Distributions. Transaction of the Seventh Prague conference on information theory, statistical decision functions, random processes and of the 1974 European meeting of statisticians. Academia, Prague 1978, 513—519.

[10] RAO, C. R.: Least Squares Theory Using an Estimated Dispersion Matrix and Its Application to Measurements of Signals. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I.: Theory of Statistics, 1967, 355—372. University of California Press, Berkeley and Los Angeles 1967.

[11] RAO, C. R., MITRA, S. K.: Generalized Inverse of Matrices and Its Applications. J. Wiley, N. York 1971.

[12] 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.

[13] VOLAUFOVÁ, J.: On the Confidence Region of a Vector Parameter. Math. Slovaca 30, 1980, 113—120.

[14] WIMMER, G.: On Equivalence Problem in Linear Regression Models. Part I. BLUE of the Mean Value. Aplikace matematiky 25, 1980, 417—422.

Received July 29, 1981

Matematický ústav SAV
Obrancov mieru 49
814 73 Bratislava

РЕГРЕССИЯ С ОЦЕНИВАЕМОЙ КОВАРИАЦИОННОЙ МАТРИЦЕЙ

Lubomír Kubáček

Резюме

На основе реализации нормально расспределенного случайного вектора $\boldsymbol{Y} \sim N_n(\mathbf{X}\beta, \boldsymbol{\Sigma})$ и случайной матрицы $\mathbf{S}$, которая расспределена по закону Уишарта $f\mathbf{S} \sim W_n(f, \boldsymbol{\Sigma})$, получается оценка линейной функции $f(\boldsymbol{\beta}) = \boldsymbol{p}'\boldsymbol{\beta}$ параметра $\boldsymbol{\beta}(\boldsymbol{p}, \boldsymbol{\beta} \in \mathscr{R}^k$, k-размерное векторное пространство) и исследуются ее статистические свойства при следующих предположениях: Вектор $\boldsymbol{Y}$ и матрица $\mathbf{S}$ статистически независимы, f (число степеней свободы) $\geqq R(\boldsymbol{\Sigma})$ (ранг матрицы $\boldsymbol{\Sigma}$), матрица $\mathbf{X}$ известна; никакие предположения не сделаны о рангах матриц $\boldsymbol{\Sigma}$ и $\mathbf{X}$.