Lubomír Kubáček

On a generalization of the orthogonal regression

Aplikace matematiky, Vol. 20 (1975), No. 2, 87–95

Persistent URL: `http://dml.cz/dmlcz/103573`

ON A GENERALIZATION OF THE ORTHOGONAL REGRESSION

LUBOMÍR KUBÁČEK

(Received March 8, 1972)

1. INTRODUCTION

The estimation of parameters a, b in the equation of a straight line $y - ax - b = 0$ from N pairs $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$, $i = 1, \ldots, N > 2$ is a frequent task in many different experiments. If the pair $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$ is considered a realization of a random vector $\begin{pmatrix} \xi_i \\ \eta_i \end{pmatrix}$ with the mean value $E\begin{pmatrix} \xi_i \\ \eta_i \end{pmatrix} = \begin{pmatrix} \mu_i \\ \nu_i \end{pmatrix}$ and with a side condition $\nu_i - a\mu_i - b = 0$, then this problem is solved under different additional assumptions and generalizations for example in [2, 4, 5, 6, 7, 9, 11, 12, 13, 14, 15]. If the covariance matrix $\mathbf{\Sigma}_i = E\{[\xi_i - \mu_i, \eta_i - \nu_i]' [\xi_i - \mu_i, \eta_i - \nu_i]\}$ is of the form $\begin{pmatrix} \sigma_x^2 & 0 \\ 0 & 0 \end{pmatrix}$, then we have the generally known regression problem. If the matrix $\mathbf{\Sigma}_i$ is regular, then the problem is called the orthogonal regression problem.

We shall now treat the following generalization: The random variables ξ_i and η_i will be substituted by normally distributed two-dimensional random vectors $\boldsymbol{\xi}_i, \boldsymbol{\eta}_i$ ([1] § 2.3) and instead of the equation of the straight line the equation $E(\boldsymbol{\eta}_i) - \mathbf{M}E(\boldsymbol{\xi}_i) - \boldsymbol{\vartheta} = \mathbf{0}$ will be considered, where the matrix $\mathbf{M}$ is assumed to be in the form $\mathbf{M} = \begin{pmatrix} \Theta_1, & \Theta_2 \\ -\Theta_2, & \Theta_1 \end{pmatrix}$ (a product of an orthogonal matrix with a number $m > 0$) and the vector $\boldsymbol{\vartheta}$ is a column vector with the components ϑ_1 and ϑ_2. At the same time, the vectors $\boldsymbol{\xi}_i$ and $\boldsymbol{\eta}_j$, $i, j = 1, \ldots, N > 2$ are assumed to be independent; however, the stochastical dependence among the vectors $\boldsymbol{\xi}_i$, $i = 1, \ldots, N$ and also among the vectors $\boldsymbol{\eta}_i$, $i = 1, \ldots, N$ is admitted. The task is to estimate the vectors $\boldsymbol{\vartheta}$ and $\boldsymbol{\Theta} = \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix}$ from the realization of the random vectors $(\boldsymbol{\xi}_i', \boldsymbol{\eta}_i')'$, $i = 1, \ldots, N$ by the maximum-likelihood method. At the same time it is assumed that for each index $i \in \{1, \ldots, N\}$ the vector $(\xi_i', \eta_i')'$ is realized n-times which means that the values $(\mathbf{x}_i', \mathbf{y}_i')^{(1)}, \ldots, (\mathbf{x}_i', \mathbf{y}_i')^{(n)}$ and $(\mathbf{x}_{i0}', \mathbf{y}_{i0}') = (1/n) \sum_{j=1}^{n} (\mathbf{x}_i', \mathbf{y}_i')^{(j)}$ (by the upper right index j we express that the j-th realization is dealt with) are available and that the covariance matrices Σ_1 and Σ_2 of the random hypervectors $(\xi_1', \ldots, \xi_N')'$ and $(\eta_1', \ldots, \eta_N')'$ are regular and known.

The motivation of the above problem is the following one. When connecting astronomical-geodetic networks, e.g. of two neighbouring countries, N of the so called identical points – creating the boundary between the two networks and belonging to both of them – have coordinates determined in the projection planes of both networks. Let the coordinates of the point P_i, $i = 1, \ldots, N$ in the first system be given by the vector $\mathbf{x}_{i0}$ and in the second system by the vector $\mathbf{y}_{i0}$. The connecting of both networks can be carried out only in such a way that the coordinates of the first network are changed by translation, rotation and possibly by change of the size of the first network, and at the same time these changed coordinates of the points $P_1, \ldots, P_N$ should be, as far as possible, identical with the coordinates of the same points in the second system. It is possible to carry this out by a linear transformation (with a matrix $\mathbf{M}$ and vector ϑ) of the first system into the second system. The matrix $\mathbf{M}$ and vector ϑ have to be determined from the vectors $\mathbf{x}_{10}, \ldots, \mathbf{x}_{N0}$ and $\mathbf{y}_{10}, \ldots, \mathbf{y}_{N0}$. Since the vectors $\mathbf{x}_{i0}$, $\mathbf{y}_{i0}$ are results of an n-times repeated measurement, they can be considered to be the arithmetical mean of the n-times independently repeated realizations of the random vectors ξ_i, η_i. With regard to the procedure of measurement which is used in this kind of work, we can assume that the nondiagonal. covariance matrices Σ_1 and Σ_2 of the hypervectors $\xi = (\xi_1', \ldots, \xi_N')'$ and $\eta = (\eta_1', \ldots, \eta_N')'$ are known and that the vectors ξ and η are stochastically independent.

The case $\Sigma_1 = 0$ is solved in [8]. If the accuracy in the determination of the coordinates of the points $P_1, \ldots, P_N$ in both systems is characterized respectively by the matrices Σ_1 and Σ_2, then the assumption $\Sigma_1 = \mathbf{0}$ could significantly prefer the results of the measurement in the first system to the second. The aim of this paper is to find the estimate of the matrix $\mathbf{M}$ and the vector ϑ under the assumption $\Sigma_1 \neq \mathbf{0}$.

2. SYMBOLS AND BASIC STATEMENT

The fact that the random vector ξ is normally distributed is denoted by $\xi \sim N(\mu, \Sigma)$. The vector of mean values is denoted by μ and the covariance matrix by the symbol Σ ([1] § 2.3). In our case $\xi \sim N(\mu, \Sigma_1)$, $\eta \sim N(\nu, \Sigma_2)$, $\xi = (\xi_1', \ldots, \xi_N')'$ and at the same time it is assumed that the rank of the matrices Σ_1 and Σ_2 is $h(\Sigma_1) = h(\Sigma_2) = 2N$. Our assumption implies $E(\eta) - (\mathbf{I} \otimes \mathbf{M})\, E(\xi) - \mathbf{i} \otimes \vartheta$. The symbol $\otimes$ denotes the tensor multiplication of matrices ([3] Chpt. III § 1.6), the symbol $\mathbf{I}$ denotes the unit matrix of the type $N \times N$ and the vector $\mathbf{i}' = (1, \ldots, 1)$ has N components. The density function of the normal vector is denoted by $n(\mathbf{x}, \mu, \Sigma)$.

Basic statement. Let $[(\mathbf{x}', \mathbf{y}')^{(i)}]'$, $i = 1, 2, \ldots$ be the i-th realization of the random vector $(\boldsymbol{\xi}', \boldsymbol{\eta}')'$ which has the density function of the form

$$f(\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}) = n(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_1)\, n(\mathbf{y}, \boldsymbol{i} \otimes \boldsymbol{\vartheta} + (\mathbf{I} \otimes \mathbf{M})\, \boldsymbol{\mu}, \boldsymbol{\Sigma}_2)\,.$$

The symbol $\boldsymbol{\alpha}$ denotes the $4 + 2N$-dimensional vector of the unknown parameters ϑ_1, ϑ_2, Θ_1, Θ_2 and $\boldsymbol{\mu}$. Let the matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ be known and regular. Under these conditions the solutions $\tilde{\boldsymbol{\alpha}}_{(n)}$ of the equations

$$\sum_{i=1}^{n} \frac{\partial}{\partial \boldsymbol{\alpha}} \ln f(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}, \boldsymbol{\alpha}) = \mathbf{0}_{4+2N,1}\,,$$

due to the special form of the function $f(\mathbf{x}, \mathbf{y}, \boldsymbol{\alpha})$, converge almost everywhere with respect to n to the actual value $\boldsymbol{\alpha}_0$ and they are asymptotically normally distributed with the vector of mean values equal to the vector $\boldsymbol{\alpha}_0$ and with the covariance matrix $n^{-1}\mathbf{F}^{-1}$, where

$$\mathbf{F} = E\left(- \frac{\partial^2}{\partial \boldsymbol{\alpha}\, \partial \boldsymbol{\alpha}'} \ln f(\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\alpha})\bigg|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}\right). \tag{2.1}$$

This statement is a consequence of the nonsubstantially modified theorems 12.7.2 and 12.7.3 in [16] p. 379 and 380 (see also the comment in 12.6. (a) p. 376), the regularity in the sense of [10] Chpt. 12 of the probability density function being used.

3. ESTIMATE OF THE MATRIX M AND VECTOR ϑ

Due to our assumptions, $\boldsymbol{\alpha}' = (\boldsymbol{\vartheta}', \boldsymbol{\Theta}', \boldsymbol{\mu}')$ and $\prod_{i=1}^{n} f(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}, \boldsymbol{\alpha}) = \prod_{i=1}^{n} n(\mathbf{x}^{(i)}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_1)\,.$ $n(\mathbf{y}^{(i)}, \boldsymbol{i} \otimes \boldsymbol{\vartheta} + (\mathbf{I} \otimes \mathbf{M})\, \boldsymbol{\mu}, \boldsymbol{\Sigma}_2) = L.$

The maximum-likelihood estimate of the vector parameter $\boldsymbol{\alpha}$ is given, with regard to the basic statement, as a solution of the following likelihood equations (3.1), (3.2) and (3.3):

$$\left(-2 \frac{\partial \ln L}{\partial \boldsymbol{\vartheta}} =\right) -2 \sum_{j=1}^{N} \boldsymbol{\Sigma}_2^{j\cdot} \sum_{i=1}^{n} \mathbf{y}^{(i)} + 2n \sum_{j=1}^{N} \boldsymbol{\Sigma}_2^{j\cdot} (\mathbf{I} \otimes \mathbf{M})\, \boldsymbol{\mu} + 2n \sum_{j=1}^{N} \boldsymbol{\Sigma}_2^{j\cdot} \boldsymbol{\vartheta} = \mathbf{0}_{21}\,, \tag{3.1}$$

where the symbol $\boldsymbol{\Sigma}_2^{j\cdot}$ denotes the j-th double-row of the matrix $\boldsymbol{\Sigma}_2^{-1}$;

$$\left(- 2 \frac{\partial \ln L}{\partial \boldsymbol{\Theta}} =\right) - 2 \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\mu}'(\mathbf{I} \otimes \mathbf{A}') \end{bmatrix} \boldsymbol{\Sigma}_2^{-1} \sum_{i=1}^{n} \mathbf{y}^{(i)} + 2n \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\mu}'(\mathbf{I} \otimes \mathbf{A}') \end{bmatrix} \boldsymbol{\Sigma}_2^{-1}\,. \tag{3.2}$$

$$.\,(\boldsymbol{\mu}, (\mathbf{I} \otimes \mathbf{A})\, \boldsymbol{\mu})\, \boldsymbol{\Theta} + 2n \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\mu}'(\mathbf{I} \otimes \mathbf{A}') \end{bmatrix} \boldsymbol{\Sigma}_2^{-1} (\boldsymbol{i} \otimes \boldsymbol{\vartheta}) = \mathbf{0}_{21}\,,$$

where $\mathbf{A} = \begin{pmatrix} 0, 1 \\ -1, 0 \end{pmatrix}$ and the equality $(\mathbf{I} \otimes \mathbf{M})\,\boldsymbol{\mu} = (\boldsymbol{\mu}, (\mathbf{I} \otimes \mathbf{A})\,\boldsymbol{\mu})\,\boldsymbol{\Theta}$ is applied;

$$(3.3) \quad \left(-2\frac{\partial \ln L}{\partial \boldsymbol{\mu}} = \right) -2\Sigma_1^{-1}\sum_{i=1}^{n}\mathbf{x}^{(i)} + 2n\Sigma_1^{-1}\boldsymbol{\mu} - 2(\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{-1}\sum_{i=1}^{n}\mathbf{y}^{(i)} +$$

$$+2n(\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{-1}(\mathbf{I} \otimes \mathbf{M})\,\boldsymbol{\mu} + 2n(\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{-1}(\boldsymbol{i} \otimes \boldsymbol{\vartheta}) = \mathbf{0}_{2N,1}\,.$$

In the solution of the equations (3.1), (3.2) and (3.3) the following symbols will be used:

$$\mathbf{x}_0 = (1/n)\sum_{i=1}^{n}\mathbf{x}^{(i)}\,, \quad \mathbf{y}_0 = (1/n)\sum_{i=1}^{n}\mathbf{y}^{(i)}\,, \quad \mathbf{D}_1 = \boldsymbol{\mu} - \mathbf{x}_0\,,$$

$$\mathbf{D}_2 = \boldsymbol{\nu} - \mathbf{y}_0 = (\mathbf{I} \otimes \mathbf{M})\,\boldsymbol{\mu} + \boldsymbol{i} \otimes \boldsymbol{\vartheta} - \mathbf{y}_0\,.$$

This notation makes it easy to see that the solution of the likelihood equations minimizes the value of the expression $\mathbf{D}_1'\Sigma_1^{-1}\mathbf{D}_1 + \mathbf{D}_2'\Sigma_2^{-1}\mathbf{D}_2$ under the side condition $\mathbf{y}_0 + \mathbf{D}_2 - (\mathbf{I} \otimes \mathbf{M})\,(\mathbf{x}_0 + \mathbf{D}_1) - \boldsymbol{i} \otimes \boldsymbol{\vartheta} = \mathbf{0}$. This fact will be utilized for the modification of the equations (3.1) to (3.3). At the same time the vector $\mathbf{D}_1$ will be considered a new unknown instead of the vector $\boldsymbol{\mu}$.

For the solution of the conditioned extreme the Lagrange method of undetermined multipliers will be used. These multipliers $k_1, \ldots, k_{2N}$ will be arranged in the column vector $\boldsymbol{k}$. The symbol $\boldsymbol{k}_i$ denotes the vector $\begin{pmatrix} k_{2i-1} \\ k_{2i} \end{pmatrix}$, $i = 1, \ldots, N$. Instead of the equations (3.1) to (3.3) we have:

$$(3.1)' \qquad \sum_{i=1}^{N}\boldsymbol{k}_i = \mathbf{0}_{21}\,,$$

$$(3.2)' \qquad [\mathbf{N} + (\mathbf{D}_1, (\mathbf{I} \otimes \mathbf{A})\,\mathbf{D}_1)]'\,\boldsymbol{k} = \mathbf{0}_{21}\,, \quad \text{where} \quad \mathbf{N} = (\mathbf{x}_0, (\mathbf{I} \otimes \mathbf{A})\,\mathbf{x}_0)\,,$$

$$(3.3)' \qquad \Sigma_1^{-1}\mathbf{D}_1 + (\mathbf{I} \otimes \mathbf{M})\,\boldsymbol{k} = \mathbf{0}_{2N,1}$$

and with regard to the introducing of the new unknown vector $\boldsymbol{k}$ we have the equation

$$(3.4)' \qquad \mathbf{y}_0 + \Sigma_2\boldsymbol{k} - (\boldsymbol{\mu}, (\mathbf{I} \otimes \mathbf{A})\,\boldsymbol{\mu})\,\boldsymbol{\Theta} - \boldsymbol{i} \otimes \boldsymbol{\vartheta} = \mathbf{0}_{2N,1}\,.$$

In the sequel it will be assumed that the approximate values $\vartheta_{10}, \vartheta_{20}, \Theta_{10}, \Theta_{20}$ of the parameters $\vartheta_1, \vartheta_2, \Theta_1, \Theta_2$ are known in such a degree of accuracy that in Taylor's series of the expressions in (3.1)' to (3.4)' the members $(\vartheta_1 - \vartheta_{10}), \ldots$ $\ldots, (\Theta_2 - \Theta_{20})$ with the powers greater then first can be neglected. We shall use the notation

$$\delta\Theta = \begin{bmatrix} \Theta_1 - \Theta_{10} \\ \Theta_2 - \Theta_{20} \end{bmatrix}, \quad \delta\vartheta = \begin{bmatrix} \vartheta_1 - \vartheta_{10} \\ \vartheta_2 - \vartheta_{20} \end{bmatrix}, \quad \mathbf{M}_0 = \begin{bmatrix} \Theta_{10}, & \Theta_{20} \\ -\Theta_{20}, & \Theta_{10} \end{bmatrix},$$

$\mathbf{R} = \Sigma_2 + (\mathbf{I} \otimes \mathbf{M}_0)\,\Sigma_1(\mathbf{I} \otimes \mathbf{M}_0')$ and $\delta\mathbf{M} = \mathbf{M} - \mathbf{M}_0$.

In this notation the equation (3.4)′ can be written in the form

$$\mathbf{y}_0 + \Sigma_2 \boldsymbol{k} - (\mathbf{I} \otimes \mathbf{M})(\mathbf{x}_0 + \boldsymbol{D}_1) - \boldsymbol{i} \otimes \boldsymbol{\vartheta} = \mathbf{0}_{2N,1} .$$

If the vector $\boldsymbol{D}_1$ from (3.3) is substituted here then the vector $\boldsymbol{k}$ satisfies

$$\boldsymbol{k} = -[\Sigma_2 + (\mathbf{I} \otimes \mathbf{M})\Sigma_1(\mathbf{I} \otimes \mathbf{M}_1')]^{-1}(\mathbf{y}_0 - (\mathbf{I} \otimes \mathbf{M})\overline{\mathbf{x}} - \boldsymbol{i} \otimes \boldsymbol{\vartheta}) .$$

If the right-hand side is developed and if the approximate relation

$$[\Sigma_2 + (\mathbf{I} \otimes \mathbf{M})\Sigma_1(\mathbf{I} \otimes \mathbf{M}')]^{-1} \doteq \mathbf{R}^{-1} - \mathbf{R}^{-1}\,\delta\mathbf{S}\mathbf{R}^{-1}$$

is taken into account, where

$$\delta\mathbf{S} = (\mathbf{I} \otimes \delta\mathbf{M})\Sigma_1(\mathbf{I} \otimes \mathbf{M}_0') + (\mathbf{I} \otimes \mathbf{M}_0)\Sigma_1(\mathbf{I} \otimes \delta\mathbf{M}') ,$$

then we have

$$\boldsymbol{k} = -(\mathbf{R}^{-1} - \mathbf{R}^{-1}\delta\mathbf{S}\mathbf{R}^{-1})(\boldsymbol{\mu} - \mathbf{N}\delta\boldsymbol{\Theta} - \boldsymbol{i} \otimes \delta\boldsymbol{\vartheta}) .$$

Here, $\boldsymbol{u} = \mathbf{y}_0 - \mathbf{N}\boldsymbol{\Theta}_0 - \boldsymbol{i} \otimes \boldsymbol{\vartheta}_0$ and $\mathbf{N}\boldsymbol{\Theta}_0 = (\mathbf{I} \otimes \mathbf{M}_0)\,\mathbf{x}_0$. The vector $\boldsymbol{u}$ characterizes the accuracy of the measurement of the coordinates of the points $P_1, \ldots, P_N$ in both coordinate systems and the a priori knowledge of the parameters $\boldsymbol{\Theta}$ and $\boldsymbol{\vartheta}$. If the number n is sufficiently great and if the a priori estimate of the parameters $\boldsymbol{\Theta}$ and $\boldsymbol{\vartheta}$ is sufficiently good (in practice it is always possible), then the vector $\boldsymbol{u}$ does not differ significantly from the zero vector. If the notation

$$\mathbf{U}_1 = \Sigma_1(\mathbf{I} \otimes \mathbf{M}_0')\,\mathbf{R}^{-1} + (\mathbf{I} \otimes \mathbf{M}_0)\Sigma_1\mathbf{R}^{-1} ,$$

$$\mathbf{U}_2 = (\mathbf{I} \otimes \mathbf{A})\Sigma_2(\mathbf{I} \otimes \mathbf{M}_0')\,\mathbf{R}^{-1} + (\mathbf{I} \otimes \mathbf{M}_0)\Sigma_1(\mathbf{I} \otimes \mathbf{A}')\,\mathbf{R}^{-1} ,$$

$$\boldsymbol{l}_1 = \mathbf{U}_1\boldsymbol{u} , \quad \boldsymbol{l}_2 = \mathbf{U}_2\boldsymbol{u}$$

is used and if the members $(\vartheta_1 - \vartheta_{10}), \ldots, (\Theta_2 - \Theta_{20})$ with higher than the first powers are neglected, then the vector $\boldsymbol{k}$ satisfies

$$\boldsymbol{k} = -\mathbf{R}^{-1}\boldsymbol{u} - \mathbf{R}^{-1}(\mathbf{N} + (\boldsymbol{l}_1, \boldsymbol{l}_2))\,\delta\boldsymbol{\Theta} + \mathbf{R}^{-1}(\boldsymbol{i} \otimes \delta\boldsymbol{\vartheta}) .$$

In this equation let us take into account the relation (3.1)′. If the notation

$$\mathbf{s}' = \Big(\sum_{j=1}^{N}\sum_{k=1}^{N}\mathbf{R}^{jk}\Big)^{-1}\sum_{k=1}^{N}\mathbf{R}^{j\bullet}$$

is used, where the symbol $\mathbf{R}^{jk}$ denotes the jk-th submatrix of the type 2×2 of the matrix $\mathbf{R}^{-1}$ and the symbol $\mathbf{R}^{j\bullet}$ denotes the j-th double-row of the matrix $\mathbf{R}^{-1}$, then we have

$$\delta\boldsymbol{\vartheta} = \mathbf{s}'\boldsymbol{u} - \mathbf{s}'(\mathbf{N} + (\boldsymbol{l}_1, \boldsymbol{l}_2))\,\delta\boldsymbol{\Theta} .$$

Since the vector $\boldsymbol{u} = \mathbf{y}_0 - \mathbf{N}\boldsymbol{\Theta}_0 - \boldsymbol{i} \otimes \boldsymbol{\vartheta}$, as already mentioned, should non-significantly differ from the zero vector, it is possible in practice to choose such a value of the vector $\boldsymbol{\vartheta}_0$ that $\mathbf{s}'\boldsymbol{u} = \mathbf{0}$ and hence we have

$$\boldsymbol{\vartheta}_0 = \mathbf{s}'(\mathbf{y}_0 - \mathbf{N}\boldsymbol{\Theta}_0). \tag{3.5}$$

In this case it holds

$$\delta\boldsymbol{\vartheta} = -\mathbf{s}'(\mathbf{N} + (\boldsymbol{l}_1, \boldsymbol{l}_2))\,\delta\boldsymbol{\Theta}. \tag{3.6}$$

If the notation

$$(\mathbf{I} - (\boldsymbol{i} \otimes \mathbf{s}'))\,(\cdot) = (\cdot)_{red},$$

is used, then the vector $\boldsymbol{k}$ satisfies

$$\boldsymbol{k} = -\mathbf{R}^{-1}\boldsymbol{u} + \mathbf{R}^{-1}(\mathbf{N}_{red} + (\boldsymbol{l}_1, \boldsymbol{l}_2)_{red})\,\delta\boldsymbol{\Theta}. \tag{3.7}$$

After a slight modification we can get now from the relation (3.3)′ the following equality:

$$\mathbf{D}_1 = \Sigma_1(\mathbf{I} \otimes \mathbf{M}_0')\,\mathbf{R}^{-1}\boldsymbol{u} + [\Sigma_1(\mathbf{R}^{-1}\boldsymbol{u}, (\mathbf{I} \otimes \mathbf{A})\,\mathbf{R}^{-1}\boldsymbol{u}) - \Sigma_1(\mathbf{I} \otimes \mathbf{M}_0')\,\mathbf{R}^{-1}(\mathbf{N}_{red} + (\boldsymbol{l}_1, \boldsymbol{l}_2)_{red})]\,\delta\boldsymbol{\Theta}. \tag{3.8}$$

If the relations (3.7) and (3.8) are used in (3.2)′, then after a minor arrangement we obtain

$$\mathbf{Z}\delta\boldsymbol{\Theta} = \mathbf{N}_{red}'\mathbf{R}^{-1}\boldsymbol{u} + \begin{bmatrix} \boldsymbol{u}'\mathbf{R}^{-1} \\ \boldsymbol{u}'\mathbf{R}^{-1}(\mathbf{I} \otimes \mathbf{A}) \end{bmatrix} \Sigma_1(\mathbf{I} \otimes \mathbf{M}_0')\,\mathbf{R}^{-1}\boldsymbol{u}, \tag{3.9}$$

where

$$\mathbf{Z} = (\mathbf{N}_{red} + (\boldsymbol{l}_1, \boldsymbol{l}_2))'\,\mathbf{R}^{-1}(\mathbf{N}_{red} + (\boldsymbol{l}_1, \boldsymbol{l}_2)_{red}) - \begin{bmatrix} \mathbf{U}'\mathbf{R}^{-1} \\ \mathbf{U}'\mathbf{R}^{-1}(\mathbf{I} \otimes \mathbf{A}) \end{bmatrix} \Sigma_1(\mathbf{R}^{-1}\boldsymbol{u}, (\mathbf{I} \otimes \mathbf{A}')\,\mathbf{R}^{-1}\boldsymbol{u}).$$

The solution of our problem follows from the equations (3.9), (3.5), (3.6), (3.8), from the basic statement and from the relations $\tilde{\boldsymbol{\mu}} = \mathbf{x}_0 + \mathbf{D}_1$, $\tilde{\boldsymbol{v}} = \mathbf{y}_0 + \mathbf{D}_2 = \mathbf{y}_0 + \Sigma_2\boldsymbol{k}$.

Remark: It is possible to estimate the matrix $\mathbf{M}$ and the vector $\boldsymbol{\vartheta}$ also in the case when $\Sigma_1 = c_1^2\mathbf{S}_1$, $\Sigma_2 = c_2^2\mathbf{S}_2$, $c_{12}^2 = c_1^2/c_2^2$, where the matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ and the number c_{12}^2 are known but the values c_1^2 and c_2^2 are unknown. This case also occurs in the applications.

4. ASYMPTOTIC PROPERTIES OF THE ESTIMATES OF THE MATRIX $\mathbf{M}$ AND THE VECTOR $\boldsymbol{\vartheta}$

With respect to the basic statement (Chpt. 2) the estimates of the matrix $\mathbf{M}$ and the vector $\boldsymbol{\vartheta}$ are asymptotically unbiassed and the covariance matrix $\boldsymbol{\Sigma}$ of the asymptotic normal distribution is given by the relation (2.1). If the equations (3.1), (3.2) and (3.3) are considered, then we have

$$-\frac{\partial^2 \ln f}{\partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\vartheta}'} = \sum_{j=1}^{N} \sum_{k=1}^{N} \boldsymbol{\Sigma}_2^{jk} = E\left(-\frac{\partial^2 \ln f}{\partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\vartheta}'}\right),$$

$$-\frac{\partial^2 \ln f}{\partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\Theta}'} = \sum_{j=1}^{N} \boldsymbol{\Sigma}_2^{j\cdot}(\boldsymbol{\mu}, (\mathbf{I} \otimes \mathbf{A})\, \boldsymbol{\mu}) = E\left(-\frac{\partial^2 \ln f}{\partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\Theta}'}\right),$$

$$-\frac{\partial^2 \ln f}{\partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\mu}'} = \sum_{j=1}^{N} \boldsymbol{\Sigma}_2^{j\cdot}(\mathbf{I} \otimes \mathbf{M}) = E\left(-\frac{\partial^2 \ln f}{\partial \boldsymbol{\vartheta}\, \partial \boldsymbol{\mu}'}\right),$$

$$-\frac{\partial^2 \ln f}{\partial \boldsymbol{\Theta}\, \partial \boldsymbol{\Theta}'} = \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\mu}'(\mathbf{I} \otimes \mathbf{A}') \end{bmatrix} \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\mu}, (\mathbf{I} \otimes \mathbf{A})\, \boldsymbol{\mu}) = E\left(-\frac{\partial^2 \ln f}{\partial \boldsymbol{\Theta}\, \partial \boldsymbol{\Theta}'}\right),$$

$$\begin{aligned}
-\frac{\partial^2 \ln f}{\partial \boldsymbol{\Theta}\, \partial \boldsymbol{\mu}'} &= -\begin{bmatrix} \mathbf{y}'\boldsymbol{\Sigma}_2^{-1} \\ \mathbf{y}'\boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{A}) \end{bmatrix} + \\
&+ \begin{bmatrix} \boldsymbol{\mu}'[(\mathbf{I} \otimes \mathbf{M}')\boldsymbol{\Sigma}_2^{-1} + \boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{M})] \\ \boldsymbol{\mu}'[(\mathbf{I} \otimes \mathbf{M}')\boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{A}) + (\mathbf{I} \otimes \mathbf{A}')\boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{M})] \end{bmatrix} + \\
&+ \begin{bmatrix} (\boldsymbol{i} \otimes \boldsymbol{\vartheta}')\boldsymbol{\Sigma}_2^{-1} \\ (\boldsymbol{i} \otimes \boldsymbol{\vartheta}')\boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{A}) \end{bmatrix} \Rightarrow E\left(-\frac{\partial^2 \ln f}{\partial \boldsymbol{\Theta}\, \partial \boldsymbol{\mu}'}\right) = \\
&= \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\mu}'(\mathbf{I} \otimes \mathbf{A}') \end{bmatrix} \boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{M})
\end{aligned}$$

with regard to the relation $E(\boldsymbol{\eta}) = (\mathbf{I} \otimes \mathbf{M})\, \boldsymbol{\mu} + \boldsymbol{i} \otimes \boldsymbol{\vartheta}$

$$-\frac{\partial^2 \ln f}{\partial \boldsymbol{\mu}\, \partial \boldsymbol{\mu}'} = \boldsymbol{\Sigma}_1^{-1} + (\mathbf{I} \otimes \mathbf{M}')\boldsymbol{\Sigma}_2^{-1}(\mathbf{I} \otimes \mathbf{M}) = E\left(-\frac{\partial^2 \ln f}{\partial \boldsymbol{\mu}\, \partial \boldsymbol{\mu}'}\right).$$

Hence the covariance matrix $\boldsymbol{\Sigma}$ of the asymptotic distribution of the random vector $(\tilde{\boldsymbol{\vartheta}}', \tilde{\boldsymbol{\Theta}}', \tilde{\boldsymbol{\mu}}')'$ fulfils

$$\Sigma^{-1} = n\mathbf{F} = n\,.$$

$$\cdot \begin{bmatrix} \sum_{j=1}^{N} \sum_{k=1}^{N} \Sigma_2^{jk}, & \sum_{j=1}^{N} \Sigma_2^{j\cdot}(\mu, (\mathbf{I} \otimes \mathbf{A})\,\mu), & \sum_{j=1}^{N} \Sigma_2^{j\cdot}(\mathbf{I} \otimes \mathbf{M}) \\ \begin{bmatrix} \mu' \\ \mu'\mathbf{I} \otimes \mathbf{A}' \end{bmatrix} \sum_{j=1}^{N} \Sigma_2^{\cdot j}, & \begin{bmatrix} \mu' \\ \mu'\mathbf{I} \otimes \mathbf{A}' \end{bmatrix} \Sigma_2^{-1}(\mu, (\mathbf{I} \otimes \mathbf{A})\,\mu), & \begin{bmatrix} \mu' \\ \mu'(\mathbf{I} \otimes \mathbf{A}') \end{bmatrix} \Sigma_2^{-1}(\mathbf{I} \otimes \mathbf{M}) \\ \sum_{j=1}^{N} (\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{\cdot j}, & (\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{-1}(\mu, (\mathbf{I} \otimes \mathbf{A})\,\mu), & \Sigma_1^{-1} + (\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{-1}(\mathbf{I} \otimes \mathbf{M}) \end{bmatrix}.$$

If the matrix $\mathbf{F}$ is divided into submatrices

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{11}, & \mathbf{F}_{12} \\ \mathbf{F}_{21}, & \mathbf{F}_{22} \end{bmatrix},$$

where $\mathbf{F}_{22} = \Sigma_1^{-1} + (\mathbf{I} \otimes \mathbf{M}')\,\Sigma_2^{-1}(\mathbf{I} \otimes \mathbf{M})$, then the covariance matrix of the asymptotic distribution of the random vector $(\tilde{\vartheta}', \tilde{\Theta}')'$ can be expressed with the help of the known relation

$$\Sigma_{\begin{bmatrix} \tilde{\vartheta} \\ \tilde{\Theta} \end{bmatrix}} = \frac{1}{n}\,(\mathbf{F}_{11} - \mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{21})^{-1}\,.$$

The use of the above result in the geometrical interpretation of the investigated transformation is similar to that in [8], Chapter 5.

Acknowledgement. I should like to express my gratitude to my colleague RNDr. A. Pázman, CSc. from the Institute of the Theory of Measurement for his valuable advice.

References

[1] *Anderson, T. W.:* An introduction to multivariate statistical analysis. J. Wiley, New York (1958).

[2] *Bartlett, M. S.:* Fitting a straight line when both variables are subject to error. Biometrics, 5 (1949), 207—212.

[3] *Bourbaki, N.:* Éléments de Mathématique. Livre II, Algebre, Ch. 3, Hermann, Paris (1958).

[4] *Creasy, M. A.:* Confidence limits for the gradient in the linear functional relationship. J. Roy. Statist. Soc. B 18, (1956), 65—69.

[5] *Dorff, M. - Gurland, J.:* Estimation of the parameters of a linear functional relation. J. Roy. Statist. Soc. B 23, (1961), 160—170.

[6] *Geary, R. C.:* Determination of linear relations between systematic parts of variables with errors observations the variances of which are unknown. Econometrica 17, (1949), 30—59.

[7] *Halperin, M.:* Fitting of straight lines and prediction when both variables are subject to error. J. Amer. Statist. Assoc. 56, (1961), 657—669.

[8] *Kubáček, L.:* Some statistical aspects of the estimation of parameters of a linear conform transformation. Aplikace matematiky 15, (1970), 190—206.

[9] *Madansky, A.:* The fitting of straight lines when both variables are subject to error. Amer. Statist. Ass. Journ. 54, (1959), 173—206.
[10] *Rao, C. R.:* Linear statistical inference and its applications. J. Wiley, New York (1965).
[11] *Villegas, C.:* Maximum likelihood estimation of a linear functional relationship. Ann. Math. Statist. 32, (1961), 1048—1062.
[12] *Villegas, C.:* On the least squares estimation of a linear relation. Bol. Fac. Ingen. Agrimens. 8, (1963), 47—63; Fac. Ingen. Montevideo. Publ. Inst. Mat. Estadíst. 3, (1963), 189—203.
[13] *Villegas, C.:* Confidence region for a linear relation. Ann. Math. Statist. 35, (1964), 780—788.
[14] *Villegas, C.:* On the asymptotic efficiency of least squares estimators. Ann. Math. Statist. 37, (1966), 1676—1683.
[15] *Wald, A.:* The fitting of straight lines if both variables are subject to error. Ann. Math. Statist. 11, (1940), 284—300.
[16] *Wilks, S. S.:* Mathematical Statistics. J. Wiley, New York (1963).

Súhrn

O JEDNOM ZOVŠEOBECNENÍ ORTOGONÁLNEJ REGRESIE

LUBOMÍR KUBÁČEK

V práci sú vyriešené rovnice vierohodnosti pre určenie odhadu parametrov lineárnej konformnej transformácie a je určený výraz pre kovariančnú maticu ich asymptotického rozdelenia.

Pritom sú uvažované nasledujúce predpoklady:

1) Je daných n realizácií odhadu súradníc tzv. identických bodov P_i, $i = 1, \ldots$ $\ldots, N > 2$ v prvej aj druhej súradnicovej sústave. (Za identický bod považujeme každý bod P_i prvej súradnicovej sústavy, pre ktorý máme k dispozícii realizáciu odhadu jeho súradníc v prvej sústave a súčasne realizáciu odhadu súradníc jeho transformácie do sústavy druhej.)

2) Odhadové štatistiky pre $2N$-rozmerné vektory súradníc v 1. a 2. súradnicovej sústave sú normálne rozdelené, nevychýlené a ich kovariančné matice sú známe, pričom sa predpokladá ich regulárnosť a nediagonálnosť.

Author's address: RNDr. Ing. *Lubomír Kubáček,* CSc. Ústav teórie merania SAV, Dúbravská cesta, 885 27 Bratislava.