Jan Ševčík

Weak nonlinearity in a model which arises from the Helmert transformation

Applications of Mathematics, Vol. 48 (2003), No. 3, 161–174

Persistent URL: http://dml.cz/dmlcz/134525

WEAK NONLINEARITY IN A MODEL WHICH ARISES FROM THE HELMERT TRANSFORMATION

Jan Ševčík, Olomouc

(Received June 21, 2000)

Abstract. Nowadays, the algorithm most frequently used for determination of the estimators of parameters which define a transformation between two coordinate systems (in this case the Helmert transformation) is derived under one unreal assumption of errorless measurement in the first system. As it is practically impossible to ensure errorless measurements, we can hardly believe that the results of this algorithm are "optimal".

In 1998, Kubáček and Kubáčková proposed an algorithm which takes errors in both systems into consideration. It seems to be closer to reality and at least in this sense better. However, a partial disadvantage of this algorithm is the necessity of linearization of the model which describes the problem of the given transformation. The defence of this simplification especially with respect to the bias of linear functions of the final estimators, or better to say the specification of conditions under which such a modification is statistically insignificant is the aim of this paper.

Keywords: Helmert transformation, linear regression model, nonlinearity measures, weak nonlinearity

MSC 2000: 62J05

1. Introduction

The problem to establish the "optimum" transformation or its parameters between two geodetical networks (coordinate systems) which preserves some of their special features is still an open mathematical problem which we can find in geodesy.

One of the transformation most often used is the linear conform transformation—Helmert transformation—which is the composition of three simple ones: the shift, the rotation and changing of the scale. If we denote the coordinates of the "important"—the so called identical—points in the first system by $\boldsymbol{\eta}_{I,i} = (x_i, y_i)'$ and in the second by $\boldsymbol{\eta}_{II,i} = (X_i, Y_i)'$, $i = 1, \ldots, n$, then this transformation can be written in the following way:

$$\boldsymbol{\eta}_{II,i} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} + \begin{pmatrix} \varphi_3, & \varphi_4 \\ -\varphi_4, & \varphi_3 \end{pmatrix} \boldsymbol{\eta}_{I,i}, \quad i = 1, \ldots, n, \tag{1.1}$$

where φ_1, φ_2, φ_3 and φ_4 are the parameters of this transformation.

Measurement by any apparatus is naturally influenced by its random error. This fact makes determining these parameters difficult and changes an at first sight deterministic problem into a more complicated stochastic one. Now the problem can be formulated as determining the optimum parameters within the following linear regression model with nonlinear constraints—see [3]:

$$\begin{aligned} &(\mathbf{Y} - \boldsymbol{\beta}_{1,0}) \sim \mathcal{N}_{k_1}(\boldsymbol{\delta\beta}_1, \boldsymbol{\Sigma}), \\ &\begin{pmatrix} \boldsymbol{\delta\beta}_1 \\ \boldsymbol{\delta\beta}_2 \end{pmatrix} \in \left\{ \begin{pmatrix} \boldsymbol{\delta\beta}_1 \\ \boldsymbol{\delta\beta}_2 \end{pmatrix} : \ \mathbf{b} + \mathbf{B}_1 \boldsymbol{\delta\beta}_1 + \mathbf{B}_2 \boldsymbol{\delta\beta}_2 + \frac{1}{2}\boldsymbol{\omega}(\boldsymbol{\delta\beta}_1, \boldsymbol{\delta\beta}_2) = \mathbf{0} \right\}. \end{aligned} \tag{1.2}$$

Remark 1.1. We have used the following notation in (1.2):

$$\begin{aligned}
\boldsymbol{\beta}_1 &= (\boldsymbol{\eta}_I', \boldsymbol{\eta}_{II}')', \\
\boldsymbol{\beta}_2 &= (\varphi_1, \varphi_2, \varphi_3, \varphi_4)', \\
\boldsymbol{\beta}_{1,0} &= (\boldsymbol{\eta}_{I,0}', \boldsymbol{\eta}_{II,0}')' \text{— approximation of the actual value of } \boldsymbol{\beta}_1, \\
\boldsymbol{\beta}_{2,0} &= (\varphi_{1,0}, \varphi_{2,0}, \varphi_{3,0}, \varphi_{4,0})' \text{— approximation of the actual value of } \boldsymbol{\beta}_2, \\
\boldsymbol{\delta\beta}_1 &= (\boldsymbol{\eta}_I', \boldsymbol{\eta}_{II}')' - (\boldsymbol{\eta}_{I,0}', \boldsymbol{\eta}_{II,0}')' = (\boldsymbol{\delta\eta}_I', \boldsymbol{\delta\eta}_{II}')', \\
\boldsymbol{\delta\beta}_2 &= (\varphi_1, \varphi_2, \varphi_3, \varphi_4)' - (\varphi_{1,0}, \varphi_{2,0}, \varphi_{3,0}, \varphi_{4,0})' = (\delta\varphi_1, \delta\varphi_2, \delta\varphi_3, \delta\varphi_4)', \\
\mathbf{b} &= \boldsymbol{\eta}_{I,0}\varphi_{3,0} + \left[\mathbf{I} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right] \boldsymbol{\eta}_{I,0}\varphi_{4,0}, \\
\mathbf{B}_1 &= \left(\varphi_{3,0}\mathbf{I} + \varphi_{4,0} \left[\mathbf{I} \otimes \begin{pmatrix} 0, & 1 \\ -1, & 0 \end{pmatrix} \right], -\mathbf{I} \right), \\
\mathbf{B}_2 &= \left(\mathbf{1} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{\eta}_{I,0}, \left[\mathbf{I} \otimes \begin{pmatrix} 0, & 1 \\ -1, & 0 \end{pmatrix} \right] \boldsymbol{\eta}_{I,0} \right), \\
\mathbf{1} &= (1, 1, \ldots, 1)', \\
\frac{1}{2}\boldsymbol{\omega}(\boldsymbol{\delta\beta}_1, \boldsymbol{\delta\beta}_2) &= \boldsymbol{\delta\eta}_I \delta\varphi_3 + \left\{ \left[I \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right] \boldsymbol{\delta\eta}_I \right\} \delta\varphi_4.
\end{aligned}$$

2. Model with constraints of type II

Due to the nonlinearity of the constraints it is not possible to determine the optimum estimators—in this paper we mean the best linear unbiased ones (BLUEs)—of the unknown parameters $\boldsymbol{\delta\beta}_1$ and $\boldsymbol{\delta\beta}_2$ within the model (1.2). That is why it is necessary (by neglecting the quadratic term $\boldsymbol{\omega}_{(\delta\beta_1,\delta\beta_2)}$) to make its linearization and turn it in that way into the so called linear regression model with constraints of type II:

$$(2.1)\qquad \begin{aligned}&(\mathbf{Y}-\boldsymbol{\beta}_{1,0})\sim\mathcal{N}_{k_1}(\boldsymbol{\delta\beta}_1,\boldsymbol{\Sigma}),\\ &\begin{pmatrix}\boldsymbol{\delta\beta}_1\\ \boldsymbol{\delta\beta}_2\end{pmatrix}\in\left\{\begin{pmatrix}\boldsymbol{\delta\beta}_1\\ \boldsymbol{\delta\beta}_2\end{pmatrix}:\ \mathbf{b}+\mathbf{B}_1\boldsymbol{\delta\beta}_1+\mathbf{B}_2\boldsymbol{\delta\beta}_2=\mathbf{0}\right\}.\end{aligned}$$

Theorem 2.1. *Let us consider the model* (2.1), *where Y is a k_1-dimensional observation vector, $\boldsymbol{\beta}_1\in\mathbb{R}^{k_1}$ and $\boldsymbol{\beta}_2\in\mathbb{R}^{k_2}$ are unknown parameters, $\boldsymbol{\Sigma}$ is a known positive definite matrix, $\mathbf{b}\in\mathbb{R}^q$ is a known vector*

$$\mathbf{b}\in\mathcal{M}(\mathbf{B}_1,\mathbf{B}_2)=\left\{(\mathbf{B}_1,\mathbf{B}_2)\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix}:\ \mathbf{u}\in\mathbb{R}^{k_1},\ \ \mathbf{v}\in\mathbb{R}^{k_2}\right\}$$

and $\mathbf{B}_1$ and $\mathbf{B}_2$ are known matrices satisfying

$$\mathbf{B}_1\sim q\times k_1,\ \ \mathbf{B}_2\sim q\times k_2,\ \ r(\mathbf{B}_1,\mathbf{B}_2)=q<k_1+k_2,\ \ r(\mathbf{B}_2)=k_2<q.$$

Then within this model the BLUE*s of the estimators $\boldsymbol{\delta\beta}_1$ and $\boldsymbol{\delta\beta}_2$ are*

$$(2.2)\qquad \begin{aligned}\widehat{\boldsymbol{\delta\beta}}_1(\mathbf{Y})&=(\mathbf{I}-\boldsymbol{\Sigma}\mathbf{B}_1'[\mathbf{M}_{B_2}(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^+\mathbf{B}_1)(\mathbf{Y}-\boldsymbol{\beta}_{1,0})\\ &\quad-\boldsymbol{\Sigma}\mathbf{B}_1'[\mathbf{M}_{B_2}(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^+\mathbf{b},\\ \widehat{\boldsymbol{\delta\beta}}_2(\mathbf{Y})&=\ -[(\mathbf{B}_2')^-_{m(B_1\Sigma B_1')}]'\mathbf{B}_1(\mathbf{Y}-\boldsymbol{\beta}_{1,0})-[(\mathbf{B}_2')^-_{m(B_1\Sigma B_1')}]'\mathbf{b},\end{aligned}$$

the covariance matrix of these estimators is

$$(2.3)\qquad \operatorname{var}\begin{pmatrix}\widehat{\boldsymbol{\delta\beta}}_1(\mathbf{Y})\\ \widehat{\boldsymbol{\delta\beta}}_2(\mathbf{Y})\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\mathbf{B}_1'\mathbf{Q}_{1,1}\mathbf{B}_1\boldsymbol{\Sigma}, & -\boldsymbol{\Sigma}\mathbf{B}_1'\mathbf{Q}_{1,2}\\ -\mathbf{Q}_{2,1}\mathbf{B}_1\boldsymbol{\Sigma}, & \mathbf{Q}_{2,2}\end{pmatrix},$$

where

$$(2.4)\qquad \begin{aligned}\mathbf{Q}_{1,1}&=[\mathbf{M}_{B_2}(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^+,\\ \mathbf{Q}_{1,2}&=\mathbf{Q}_{2,1}'=(\mathbf{B}_2')^-_{m(B_1\Sigma B_1')},\\ \mathbf{Q}_{2,2}&=[\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1'+\mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2]^{-1}-\mathbf{I},\\ \mathbf{M}_{B_2}&=\mathbf{I}-\mathbf{B}_2(\mathbf{B}_2'\mathbf{B}_2)^-\mathbf{B}_2'.\end{aligned}$$

P r o o f. See [1], Lemma 2.3.1. □

R e m a r k 2.1. In the previous theorem we have used the following statements for general matrices $\mathbf{C}$, $\mathbf{X}$:

$$[\mathbf{M}_X\mathbf{C}\mathbf{M}_X]^+ = \begin{cases} (\mathbf{C}+\mathbf{X}\mathbf{X}')^+ - (\mathbf{C}+\mathbf{X}\mathbf{X}')^+\mathbf{X}(\mathbf{X}'(\mathbf{C}+\mathbf{X}\mathbf{X}')^+\mathbf{X})^- & \\ \quad \times\mathbf{X}'(\mathbf{C}+\mathbf{X}\mathbf{X}')^+ & \\ \mathbf{C}^+ - \mathbf{C}^+\mathbf{X}(\mathbf{X}'\mathbf{C}^+\mathbf{X})^-\mathbf{X}'\mathbf{C}^+ & \mathcal{M}(\mathbf{X})\subset\mathcal{M}(\mathbf{C}) \\ \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{C}^{-1} & \mathbf{C}\text{—regular} \end{cases}$$

$$(\mathbf{X}')^-_{m(C)} = \begin{cases} (\mathbf{C}+\mathbf{X}\mathbf{X}')^-\mathbf{X}(\mathbf{X}'(\mathbf{C}+\mathbf{X}\mathbf{X}')^-\mathbf{X})^- & \\ \mathbf{C}^-\mathbf{X}(\mathbf{X}'(\mathbf{C})^-\mathbf{X})^- & \mathcal{M}(\mathbf{X})\subset\mathcal{M}(\mathbf{C}) \\ \mathbf{C}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^- & \mathbf{C}\text{—regular} \end{cases}$$

where $\mathbf{A}^+$ denotes the Moore-Penrose generalized inversion if it is unique, $\mathbf{A}^-$ denotes any generalized inversion in the case of nonuniqueness, $\mathcal{M}(\mathbf{A})$ denotes the space generated by the columns of matrix $\mathbf{A}$.

The first term, i.e. $[\mathbf{M}_X\mathbf{C}\mathbf{M}_X]^+$, does not depend on the choice of the generalized inversions which are used on the right-hand side. As for the second, i.e. $(\mathbf{X}')^-_{m(C)}$, this term depends on the choice of the generalized inversions used, but no terms in this paper containing it (e.g. (2.2)) do.

However, the estimators (2.2) are unbiased only within the linear model (2.1). Within the "correct" nonlinear model (1.2) they become biased. Just here the problem arises how to ensure validity of the results from the linear model in the nonlinear one.

The nonlinear (quadratic) term in the constraints of the model (1.2) depends on the choice of the approximation $\boldsymbol{\beta}_0 = (\boldsymbol{\beta}'_{1,0}, \boldsymbol{\beta}'_{2,0})'$ of the accurate values of the unknown estimated parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2)'$, or better to say it depends on their mutual (and of course also unknown) difference $\boldsymbol{\delta\beta} = (\boldsymbol{\delta\beta}'_1, \boldsymbol{\delta\beta}'_2)'$. That is why naturally the biases of the given estimators (2.2) are also functions of these differences as will be shown in the theorems below.

From now on let us suppose that the approximation $\boldsymbol{\beta}_0$ was chosen such that $\mathbf{b} = 0$ in the constraints of the model (1.2).

To derive the biases of the estimators (2.2) we have to rewrite the model (1.2) into a model without constraints. For the sake of simplicity we write $\boldsymbol{\omega}_{K_B\delta s} \equiv \boldsymbol{\omega}(\mathbf{K}_{B_1}\boldsymbol{\delta}s, \mathbf{K}_{B_2}\boldsymbol{\delta}s)$.

Theorem 2.2. *Model* (1.2) *is equivalent (up to terms of order* 2) *to the model without constraints*

$$\mathbf{Y} \sim \mathcal{N}_{k_1}\Big(\mathbf{K}_{B_1}\boldsymbol{\delta}s - \frac{1}{2}\mathbf{T}\boldsymbol{\omega}_{K_B\delta s}, \boldsymbol{\Sigma}\Big), \quad \boldsymbol{\delta}s \in \mathbb{R}^{k_1+k_2-q}, \tag{2.5}$$

where $\mathbf{K}_B = \begin{pmatrix} \mathbf{K}_{B_1} \\ \mathbf{K}_{B_2} \end{pmatrix}$ *is a* $(k_1+k_2)\times(k_1+k_2-q)$ *matrix such that* $r(\mathbf{K}_B) = k_1+k_2-q$, $(\mathbf{B}_1, \mathbf{B}_2)\mathbf{K}_B = 0$ *and the relation between* $\boldsymbol{\delta\beta}$ *and* $\boldsymbol{\delta} s$ *is given by*

$$\begin{pmatrix} \boldsymbol{\delta\beta}_1 \\ \boldsymbol{\delta\beta}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{K}_{B_1} \\ \mathbf{K}_{B_2} \end{pmatrix} \boldsymbol{\delta} s - \frac{1}{2} \begin{pmatrix} \mathbf{T} \\ \mathbf{U} \end{pmatrix} \boldsymbol{\omega}_{K_B \delta s},$$

where

$$\begin{pmatrix} \mathbf{T} \\ \mathbf{U} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}\mathbf{B}_1'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1} \\ \mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1} \end{pmatrix}.$$

P r o o f. See [1], Remark 2.3.3. □

Using this theorem we can prove the assertions of the next theorem which concerns the biases of the estimators (2.2).

Theorem 2.3. *Let us consider the model* (1.2). *Then the biases of the estimators* (2.2) *(up to terms of order* 2*) satisfy*

$$\begin{aligned} (2.6) \qquad \mathbf{b}_1 &:= E(\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1) - \boldsymbol{\delta\beta}_1 = \frac{1}{2}\mathbf{P}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}} \mathbf{T}\boldsymbol{\omega}_{K_B\delta s} \\ \mathbf{b_2} &:= E(\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_2) - \boldsymbol{\delta\beta}_2 \\ &= \frac{1}{2}(\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2)^{-1}\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\boldsymbol{\omega}_{K_B\delta s}, \end{aligned}$$

where $\mathbf{P}_A^{C^{-1}} = \mathbf{A}[\mathbf{A}'\mathbf{C}^{-1}\mathbf{A}]^{+}\mathbf{A}\mathbf{C}^{-1}$ *when* $\mathcal{M}(\mathbf{A}) \subset \mathcal{M}(\mathbf{C})$ *for general matrices* $\mathbf{A}$, $\mathbf{C}$ *(*$\mathbf{C}$ *regular).*

P r o o f. See [1], Corollary 2.3.5 and Theorem 2.3.14. □

R e m a r k 2.2. The remark "up to terms of order 2" from the preceding theorems has to be understood in the way that the argument in the term $\boldsymbol{\omega}_{K_B\delta s}$ has to be more precisely not $(\mathbf{K}_{B_1}\boldsymbol{\delta} s, \mathbf{K}_{B_2}\boldsymbol{\delta} s)$, but $(\boldsymbol{\delta\beta}_1, \boldsymbol{\delta\beta}_2)$. Nonetheless, as the terms of order higher then 2 are beyond our interest and the term $\boldsymbol{\omega}(\boldsymbol{\delta\beta}_1, \boldsymbol{\delta\beta}_2)$ is quadratic itself, it is possible to use instead of the correct argument its linear approximation.

It follows from Remark 1.1 that if we used the approximation $\boldsymbol{\beta}_0 = \boldsymbol{\beta}$ (i.e. the unknown actual value) then the bias would naturally equal zero. Moreover it is clear (as the term $\boldsymbol{\omega}(\boldsymbol{\delta\beta}_1, \boldsymbol{\delta\beta}_2)$ is quadratic) that the biases decrease when the approximation $\boldsymbol{\beta}_0$ tends to the actual value $\boldsymbol{\beta}$. On the basis of these facts let us try to formulate some criteria—restrictions on the regions for $\boldsymbol{\beta}_0$ or $\boldsymbol{\delta\beta}$—under which it is possible to justify neglecting of these biases, i.e. to take the estimators (2.2) as practically unbiased even in the nonlinear model (1.2).

The following definition of the so called *nonlinearity measures* provides an important tool for determining such criteria.

Definition 3.1. Within the model (1.2) let us define the following nonlinearity measures:

(3.1)

$$C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boldsymbol{\omega}'_{K_B\delta s}\mathbf{T}'(\mathbf{P}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}})'\boldsymbol{\Sigma}^{-1}\mathbf{P}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}\mathbf{T}\boldsymbol{\omega}_{K_B\delta s}}}{\boldsymbol{\delta s}'\mathbf{K}'_{B_1}\boldsymbol{\Sigma}^{-1}\mathbf{K}_{B_1}\boldsymbol{\delta s}} \colon \boldsymbol{\delta s} \in \mathbb{R}^{k_1+k_2-q}\right\},$$

$$C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boldsymbol{\omega}'_{K_B\delta s}(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{P}^{(B_1\Sigma B_1' + B_2 B_2')^{-1}}_{B_2}\boldsymbol{\omega}_{K_B\delta s}}}{\boldsymbol{\delta s}'\mathbf{K}'_{B_1}\boldsymbol{\Sigma}^{-1}\mathbf{K}_{B_1}\boldsymbol{\delta s}} \colon \boldsymbol{\delta s} \in \mathbb{R}^{k_1+k_2-q}\right\}.$$

Remark 3.1. These definitions have arisen from the procedure which is used in the proof of Theorem 3.1 below. Their construction was motivated by the Bates and Watts measures of curvature which have been used in the theory of nonlinear regression models, but is not the same.

Theorem 3.1. *Using the Hölder inequality together with the preceding definition we can prove the following implications:*

$$\boldsymbol{\delta s}'\mathbf{K}'_{B_1}\boldsymbol{\Sigma}^{-1}\mathbf{K}_{B_1}\boldsymbol{\delta s} \leqslant \frac{2\varepsilon}{C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)} \Rightarrow \forall\{\mathbf{h}_1 \in \mathbb{R}^{k_1}\}|\mathbf{h}_1'\mathbf{b}_1| \leqslant \varepsilon\sqrt{\mathbf{h}_1'\boldsymbol{\Sigma}\mathbf{h}_1}, \tag{3.2}$$

$$\boldsymbol{\delta s}'\mathbf{K}'_{B_1}\boldsymbol{\Sigma}^{-1}\mathbf{K}_{B_1}\boldsymbol{\delta s} \leqslant \frac{2\varepsilon}{C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)} \tag{3.3}$$

$$\Rightarrow \forall\{\mathbf{h}_2 \in \mathbb{R}^{k_2}\}\,|\mathbf{h}_2'\mathbf{b}_2| \leqslant \varepsilon\sqrt{\mathbf{h}_2'(\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2)^{-1}\mathbf{h}_2}.$$

Proof. See [1], Theorem 2.3.7 and 2.3.16. □

R e m a r k 3.2. As $\boldsymbol{\delta\beta}_1 = \boldsymbol{\beta}_1 - \boldsymbol{\beta}_{1,0} \doteq \mathbf{K}_{B_1}\boldsymbol{\delta s}$, this theorem declares that if $\boldsymbol{\delta\beta}_1$ moves inside the ellipsoid which is determined by the quadratic form with the matrix $\mathbf{\Sigma}^{-1}$ and the right side $2\varepsilon/C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$ or $2\varepsilon/C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$ then the bias of the estimator of an arbitrary linear function respectively of $\boldsymbol{\delta\beta}_1$ or $\boldsymbol{\delta\beta}_2$ is covered by an ε-multiple of the term which as will be shown represents in some sense the standard error of this linear function. It means that this bias can be considered practically neglectable. In simple words we can say that the bias of the given linear function is "drown" in its dispersion.

Lemma 3.1. *It simply follows from Theorem* 2.1 *that*

$$\begin{aligned}(3.4)\qquad &\forall \mathbf{h}_1 \in \mathbb{R}^{k_1}:\ \mathrm{var}(\mathbf{h}_1'\boldsymbol{\delta\beta}_1) = \mathbf{h}_1'(\mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{B}_1'[\mathbf{M}_{B_2}(\mathbf{B}_1\mathbf{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^+\mathbf{B}_1\mathbf{\Sigma})\mathbf{h}_1,\\ &\forall \mathbf{h}_2 \in \mathbb{R}^{k_2}:\ \mathrm{var}(\mathbf{h}_1'\boldsymbol{\delta\beta}_2) = \mathbf{h}_2'([\mathbf{B}_2'(\mathbf{B}_1\mathbf{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2]^{-1} - \mathbf{I})\mathbf{h}_2.\end{aligned}$$

P r o o f. It is a direct consequence of Theorem 2.1. □

R e m a r k 3.3. As the quadratic forms

$$\begin{aligned}&\mathbf{h}_1'\mathbf{\Sigma}\mathbf{B}_1'[\mathbf{M}_{B_2}(\mathbf{B}_1\mathbf{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^+\mathbf{B}_1\mathbf{\Sigma}\mathbf{h}_1, &&\mathbf{h}_1 \in \mathbb{R}^{k_1},\\ &\mathbf{h}_2'\mathbf{h}_2, &&\mathbf{h}_2 \in \mathbb{R}^{k_2}\end{aligned}$$

are p.s.d., if follows from the preceding lemma that the right-hand side terms of inequalities (3.2) and (3.3) are not actual ε-multiples of the relevant standard errors but their upper estimators.

If we denote

$$\mathbf{M}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}} = \mathbf{I} - \mathbf{\Sigma}\mathbf{B}_1'[\mathbf{M}_{B_2}(\mathbf{B}_1\mathbf{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^+\mathbf{B}_1,$$

the estimator $\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1$ can be written as

$$(3.5)\qquad \boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1 = \mathbf{M}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}\mathbf{Y}.$$

Taking into account Theorem 2.3 and the fact that the space $\mathbb{R}^{k_1}$ can be decomposed into the sum of two subspaces $\mathcal{M}([\mathbf{P}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}]')$ and $\mathcal{M}([\mathbf{M}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}]')$, i.e.

$$\forall\, \mathbf{h}_1 \in \mathbb{R}^{k_1}\ \exists\, \mathbf{h}_{1,1} \in \mathcal{M}([\mathbf{P}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}]'),\ \exists\, \mathbf{h}_{1,2} \in \mathcal{M}([\mathbf{M}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}]'):\ \mathbf{h}_1 = \mathbf{h}_{1,1} + \mathbf{h}_{1,2},$$

we can formulate the following lemma:

Lemma 3.2.
a) $\forall\, \mathbf{h}_1 \in \mathcal{M}([\mathbf{P}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}]'):\ \mathbf{h}_1'\boldsymbol{\delta\beta}_1 = 0$, *which also means that* $\mathrm{var}(\mathbf{h}_1'\boldsymbol{\delta\beta}_1) = 0$*;*
b) $\forall\, \mathbf{h}_1 \in \mathcal{M}([\mathbf{M}^{\Sigma^{-1}}_{\Sigma B_1' M_{B_2}}]'):\ \mathbf{h}_1'\mathbf{b}_1 = 0.$

P r o o f. Assertion easily follows from (3.5) and Theorem 2.3. □

Corollary 3.1. *It follows from the preceding lemma that every estimator* $\mathbf{h}_1'\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1$ *of the linear function* $\boldsymbol{h}_1'\boldsymbol{\delta\beta}_1$ *can be written as the sum of the estimator* $\mathbf{h}_{1,1}'\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1$ *which estimates with no bias the parameter* $\mathbf{h}_{1,1}'\boldsymbol{\delta\beta}_1$ *and the estimator* $\mathbf{h}_{1,2}'\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1$ *which is identically equal to zero vector (but as the estimator of the parameter* $\mathbf{h}_{1,2}'\boldsymbol{\delta\beta}_1$ *it is biased!).*

On the basis of Theorem 3.1 we can formulate the regions of $\boldsymbol{\delta\beta}_1$ which make the bias of linear functions of the estimator $\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1$ neglectable in the previously mentioned sense.

Definition 3.2. The ellipsoids defined by the relations

$$L_{b_1}^{\beta_1}(\boldsymbol{\beta}_0) = \left\{\boldsymbol{\delta\beta}_1 \in \mathbb{R}^{k_1} :\ \boldsymbol{\delta\beta}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta\beta}_1 \leqslant \frac{2\varepsilon}{C_{II,\delta\beta_1}^{(\mathrm{par})}(\boldsymbol{\beta}_0)}\right\},$$
$$L_{b_2}^{\beta_1}(\boldsymbol{\beta}_0) = \left\{\boldsymbol{\delta\beta}_1 \in \mathbb{R}^{k_1} :\ \boldsymbol{\delta\beta}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta\beta}_1 \leqslant \frac{2\varepsilon}{C_{II,\delta\beta_2}^{(\mathrm{par})}(\boldsymbol{\beta}_0)}\right\}$$

are called *the linearization regions* for $\mathbf{b}_1$ and $\mathbf{b}_2$, respectively, in $\boldsymbol{\beta}_0$ with respect to $\boldsymbol{\delta\beta}_1$.

To define the linearization region with respect to $\boldsymbol{\delta\beta}_2$ it is necessary to formulate the following lemma.

Lemma 3.3. *Let* $\mathbf{M}$ *be a symmetric p.d. matrix of the type* $n\times n$*, and let* $c \in \mathbb{R}$*. Then for all matrices* $\mathbf{L}$ *of the type* $k \times n$*,* $r(\mathbf{L}) = k$ *we have*

$$\begin{aligned}\{\mathbf{u} \in \mathbb{R}^k :\ \mathbf{u} = \mathbf{L}\boldsymbol{\delta\beta},\ \ \boldsymbol{\delta\beta}'\mathbf{M}\boldsymbol{\delta\beta} \leqslant c^2,\ \ \boldsymbol{\delta\beta} \in \mathbb{R}^n\}\\ = \{\mathbf{u} \in \mathbb{R}^k :\ \mathbf{u}'(\mathbf{L}\mathbf{M}^{-1}\mathbf{L}')^{-1}\mathbf{u} \leqslant c^2\}.\end{aligned} \tag{3.6}$$

P r o o f. See [2], Theorem 2.2. □

Using the fact that the matrix $\mathbf{B}_2$ is of full rank in columns we can simply prove the next lemma:

Lemma 3.4. *We have*

$$\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_2 = -[(\mathbf{B}_2')^-_{m(B_1\Sigma B_1')}]'\boldsymbol{B}_1\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1. \tag{3.7}$$

P r o o f. See [1], proof of Theorem 2.3.14. □

The matrix $[(\mathbf{B}_2')^-_{m(B_1\Sigma B_1')}]'\mathbf{B}_1$ is of the type $k_2 \times k_1$ and its rank is equal to k_2. Hence using Lemma 3.3 the linearization region with respect to $\boldsymbol{\delta\beta}_2$ can be defined.

Definition 3.3. The ellipsoids defined by the relations

$$L_{b_1}^{\beta_2}(\boldsymbol{\beta}_0) = \left\{\boldsymbol{\delta\beta}_2 \in \mathbb{R}^{k_2}:\ \boldsymbol{\delta\beta}_2'(\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2)^{-1}\boldsymbol{\delta\beta}_2 \leqslant \frac{2\varepsilon}{C_{II,\delta\beta_1}^{(\mathrm{par})}(\boldsymbol{\beta}_0)}\right\},$$

$$L_{b_2}^{\beta_2}(\boldsymbol{\beta}_0) = \left\{\boldsymbol{\delta\beta}_2 \in \mathbb{R}^{k_2}:\ \boldsymbol{\delta\beta}_2'(\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2)^{-1}\boldsymbol{\delta\beta}_2 \leqslant \frac{2\varepsilon}{C_{II,\delta\beta_2}^{(\mathrm{par})}(\boldsymbol{\beta}_0)}\right\}$$

are called *the linearization regions* for $\mathbf{b}_1$ and $\mathbf{b}_2$, respectively, in $\boldsymbol{\beta}_0$ with respect to $\boldsymbol{\delta\beta}_2$.

So now we know the regions where the vectors $\boldsymbol{\delta\beta}_1$ a $\boldsymbol{\delta\beta}_2$ should occur. That is why it is necessary to compare these regions with the ones which delimitate the locus of their real occurrence. If we found out that the regions of real occurrence are covered by the linearization ones it could be considered a strong argument for "practical" unbiasness of the estimators (2.2) even within the nonlinear model (1.2) (Remark 3.2).

It follows from Theorem 2.1 that the random vectors $\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1$ and $\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_2$ satisfy

$$\begin{aligned}
\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1 &\sim \mathcal{N}_{k_1}(\boldsymbol{\delta\beta}_1 + \mathbf{b}_1, \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{B}_1'[\mathbf{M}_{B_2}(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1')\mathbf{M}_{B_2}]^{+}\mathbf{B}_1\boldsymbol{\Sigma}),\\
\boldsymbol{\delta}\hat{\boldsymbol{\beta}}_2 &\sim \mathcal{N}_{k_2}(\boldsymbol{\delta\beta}_2 + \mathbf{b}_2, (\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2)^{-1} - \mathbf{I}).
\end{aligned}$$

That is why $\boldsymbol{\delta\beta}_1$ and $\boldsymbol{\delta\beta}_2$ are covered with probability near to $1-\alpha$ by the ellipsoids

$$\begin{aligned}
E^{\beta_1} &= \{\boldsymbol{\delta\beta}_1:\ (\boldsymbol{\delta\beta}_1 - \boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\delta\beta}_1 - \boldsymbol{\delta}\hat{\boldsymbol{\beta}}_1) \leqslant \chi^2_{k_1}(0; 1-\alpha)\},\\
E^{\beta_2} &= \{\boldsymbol{\delta\beta}_2:\ (\boldsymbol{\delta\beta}_2 - \boldsymbol{\delta}\hat{\boldsymbol{\beta}}_2)'(\mathbf{B}_2'(\mathbf{B}_1\boldsymbol{\Sigma}\mathbf{B}_1' + \mathbf{B}_2\mathbf{B}_2')^{-1}\mathbf{B}_2)^{-1}(\boldsymbol{\delta\beta}_2 - \boldsymbol{\delta}\hat{\boldsymbol{\beta}}_2)\\
&\qquad \leqslant \chi^2_{k_2}(0; 1-\alpha)\}.
\end{aligned}$$

Their comparison with the linearization regions from Definitions 3.2 and 3.3 made the next definition reasonable.

Definition 3.4. The model (1.2) is said to be *weakly nonlinear* in $\boldsymbol{\beta}_0$ with respect to $\mathbf{b}_i$, $i \in \{1,2\}$, if the relation

$$\chi^2_{k_1}(0; 1-\alpha) \ll \frac{2\varepsilon}{C_{II,\delta\beta_i}^{(\mathrm{par})}(\boldsymbol{\beta}_0)} \tag{3.8}$$

holds.

Simply said: weak nonlinearity of the model (1.2) with respect to the preceding assertions means that the estimators which were derived from the linearization of this model, i.e. (2.1), can by considered unbiased in some sense even within this nonlinear model.

R e m a r k 3.4. However, it is important to mention one thing. Definition 3.4 concerns only the comparison of the areas of the given ellipsoids. It would be ideal and correct in the sense of the aim we want to reach (i.e. finding out whether our $\boldsymbol{\delta\beta}_1$ lies in the linearization region or not) to compare not only their areas but also their positions. But these ellipsoids have not the same center which makes finding some easy criterion for comparison of their positions difficult. That is why Definition 3.4 does not express exactly what we originally wanted. To temper this fact we have formulated the relation (3.8) in this definition as a "sharp" inequality instead of a "simple" one. That is why we can believe that this definition practically ensures the original purpose.

4. Example

As a model of the identical points let us use the grids of the square whose sides are 300 metres long. In System I we have located this square at the points

$$\boldsymbol{\eta}_{I,1} = (100, 100)', \ \ \boldsymbol{\eta}_{I,2} = (400, 100)', \ \ \boldsymbol{\eta}_{I,3} = (400, 400)', \ \ \boldsymbol{\eta}_{I,4} = (100, 400)'$$

and in System II, for an easy verification of the results, we have transformed its grids to the points $\boldsymbol{\eta}_{II,i}$ satisfying the relation

$$\boldsymbol{\eta}_{II,i} = \begin{pmatrix} 200 \\ 200 \end{pmatrix} + \begin{pmatrix} \cos\frac{\pi}{3} & \sin\frac{\pi}{3} \\ -\sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{pmatrix} \boldsymbol{\eta}_{I,i}, \quad i = 1, \dots, 4.$$

These relations correspond to the Helmert transformation (1.1) with parameters

$$\big(\varphi_1, \varphi_2, \varphi_3, \varphi_4\big) = \big(200, 200, \cos\tfrac{\pi}{3}, \sin\tfrac{\pi}{3}\big).$$

Let us consider the measurements of the particular grids and also of their coordinates to be independent in both systems. So we can construct the initial estimators of the coordinate vectors, i.e. $\mathbf{Y}_I$ and $\mathbf{Y}_{II}$, by generating the normally distributed errors with zero mean values and with dispersions σ_I^2 and σ_{II}^2 to the grids $\boldsymbol{\eta}_I$ and $\boldsymbol{\eta}_{II}$, respectively, i.e.

$$\mathbf{Y}_I \sim \mathcal{N}_8(\boldsymbol{\eta}_I, \sigma_I^2 \mathbf{I}), \quad \mathbf{Y}_{II} \sim \mathcal{N}_8(\boldsymbol{\eta}_{II}, \sigma_{II}^2 \boldsymbol{I}).$$

The transmission of this situation to the model (1.2) and its partial solution is shown in the paper [3]. For the matrices $\mathbf{B}_1$, $\mathbf{B}_2$, vector $\mathbf{b}$ and the quadratic term $\boldsymbol{\omega}(\boldsymbol{\delta\beta}_1, \boldsymbol{\delta\beta}_2)$ see Remark 1.1.

The main aim of the paper [3] has been to compare two algorithms which lead to derivation of estimators of the parameters of the Helmert transformation between two coordinate systems. Namely, standardly used algorithm which unrealistically assumes errorless measurement in the first coordinate system and the algorithm which we consider in this paper, i.e. the algorithm which is based on the linearization of the model (1.2), see Section 2. Within the paper [3] several arguments were found which support the use of the second algorithm. But these arguments represent "only" the empirical point of view.

Therefore, let's try to support the second algorithm not only by simulation but also theoretically. Let us try to show that this is the question of the model with weak nonlinearity and so the linearization we used to obtain the model (2.1) has not any statistically important influence on the bias of the estimators which we are looking for.

The following tables contain the nonlinearity measures (3.1) from Definition 3.1. We have used the actual values of the vectors $\boldsymbol{\eta}_I$, $\boldsymbol{\eta}_{II}$ and $\boldsymbol{\beta}_2$, as $\boldsymbol{\eta}_{I,0}$, $\boldsymbol{\eta}_{II,0}$, and $\boldsymbol{\beta}_{2,0}$.

	$C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$				
$\sigma_I^2\backslash\sigma_{II}^2$	0.01	0.1	0.2	0.5	1
0.01	$1.667\cdot10^{-4}$	$2.247\cdot10^{-4}$	$2.300\cdot10^{-4}$	$2.334\cdot10^{-4}$	$2.344\cdot10^{-4}$
0.1	$2.247\cdot10^{-4}$	$5.270\cdot10^{-4}$	$6.086\cdot10^{-4}$	$6.804\cdot10^{-4}$	$7.107\cdot10^{-4}$
0.2	$2.300\cdot10^{-4}$	$6.086\cdot10^{-4}$	$7.454\cdot10^{-4}$	$8.909\cdot10^{-4}$	$9.623\cdot10^{-4}$
0.5	$2.334\cdot10^{-4}$	$6.804\cdot10^{-4}$	$8.909\cdot10^{-4}$	$11.785\cdot10^{-4}$	$13.609\cdot10^{-4}$
1	$2.345\cdot10^{-4}$	$7.107\cdot10^{-4}$	$9.623\cdot10^{-4}$	$13.608\cdot10^{-4}$	$16.667\cdot10^{-4}$

Table 1. Nonlinearity measure $C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$ with respect to σ_I^2 and σ_{II}^2.

	$2/C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$				
$\sigma_I^2\backslash\sigma_{II}^2$	0.01	0.1	0.2	0.5	1
0.01	$1.200\cdot10^{4}$	$0.890\cdot10^{4}$	$0.870\cdot10^{4}$	$0.857\cdot10^{4}$	$0.853\cdot10^{4}$
0.1	$0.890\cdot10^{4}$	$0.380\cdot10^{4}$	$0.329\cdot10^{4}$	$0.294\cdot10^{4}$	$0.281\cdot10^{4}$
0.2	$0.870\cdot10^{4}$	$0.329\cdot10^{4}$	$0.269\cdot10^{4}$	$0.225\cdot10^{4}$	$0.208\cdot10^{4}$
0.5	$0.857\cdot10^{4}$	$0.294\cdot10^{4}$	$0.225\cdot10^{4}$	$0.170\cdot10^{4}$	$0.147\cdot10^{4}$
1	$0.853\cdot10^{4}$	$0.281\cdot10^{4}$	$0.208\cdot10^{4}$	$0.147\cdot10^{4}$	$0.120\cdot10^{4}$

Table 2. Corresponding values $2/C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$, $\varepsilon=1$.

	$C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$				
$\sigma_I^2 \backslash \sigma_{II}^2$	0.01	0.1	0.2	0.5	1
0.01	$0.515 \cdot 10^{-4}$	$0.906 \cdot 10^{-4}$	$1.125 \cdot 10^{-4}$	$1.434 \cdot 10^{-4}$	$1.653 \cdot 10^{-4}$
0.1	$4.151 \cdot 10^{-4}$	$4.767 \cdot 10^{-4}$	$5.233 \cdot 10^{-4}$	$6.040 \cdot 10^{-4}$	$6.668 \cdot 10^{-4}$
0.2	$7.871 \cdot 10^{-4}$	$8.419 \cdot 10^{-4}$	$8.857 \cdot 10^{-4}$	$9.686 \cdot 10^{-4}$	$10.353 \cdot 10^{-4}$
0.5	$17.562 \cdot 10^{-4}$	$17.835 \cdot 10^{-4}$	$18.132 \cdot 10^{-4}$	$18.639 \cdot 10^{-4}$	$19.066 \cdot 10^{-4}$
1	$30.505 \cdot 10^{-4}$	$30.548 \cdot 10^{-4}$	$30.568 \cdot 10^{-4}$	$30.575 \cdot 10^{-4}$	$30.583 \cdot 10^{-4}$

Table 3. Nonlinearity measure $C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$ with respect to σ_I^2 and σ_{II}^2.

	$2/C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$				
$\sigma_I^2 \backslash \hat{\sigma}_{II}^2$	0.01	0.1	0.2	0.5	1
0.01	$3.883 \cdot 10^4$	$2.208 \cdot 10^4$	$1.778 \cdot 10^4$	$1.395 \cdot 10^4$	$1.210 \cdot 10^4$
0.1	$0.482 \cdot 10^4$	$0.420 \cdot 10^4$	$0.382 \cdot 10^4$	$0.331 \cdot 10^4$	$0.300 \cdot 10^4$
0.2	$0.254 \cdot 10^4$	$0.238 \cdot 10^4$	$0.226 \cdot 10^4$	$0.206 \cdot 10^4$	$0.193 \cdot 10^4$
0.5	$0.114 \cdot 10^4$	$0.112 \cdot 10^4$	$0.110 \cdot 10^4$	$0.107 \cdot 10^4$	$0.105 \cdot 10^4$
1	$0.066 \cdot 10^4$	$0.065 \cdot 10^4$	$0.065 \cdot 10^4$	$0.065 \cdot 10^4$	$0.065 \cdot 10^4$

Table 4. Corresponding values $2/C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$, $\varepsilon = 1$.

Using statistical tables we can easily find out that if $k_1 = 16$ (which is our case) then $\chi^2_{16}(0; 0.95) = 26.30$ and so the relation (3.8) holds for all cases mentioned in the previous tables. That is why, according to Definition 3.4, the model (1.2) involves a weak nonlinearity in $(\boldsymbol{\eta}_{I,0}, \boldsymbol{\eta}_{II,0}, \boldsymbol{\beta}_0)$ and so using linearization is, in the sense we have talked about, correct.

For a check let us verify validity of the relations (3.2) and (3.3). Let us denote:

$\mathbf{b}_1$ the biases of the estimators of coordinate vectors of the identical points in both systems, i.e. $E(\hat{\boldsymbol{\eta}}'_I, \hat{\boldsymbol{\eta}}'_{II})' - (\boldsymbol{\eta}'_{I,0}, \boldsymbol{\eta}'_{II,0})'$,

$\mathbf{b}_2$ the bias of the estimators of the transformation parameters, i.e. $E(\hat{\boldsymbol{\beta}}) - \boldsymbol{\beta}_0$,

h_1 vector from $\mathbb{R}^{16}$, $(k_1 = 16)$,

h_2 vector from $\mathbb{R}^4$, $(k_2 = 4)$,

p_{1,h_1} right-hand side of (3.2), i.e. $p_{1,h_1} = \varepsilon\sqrt{h'_1 \Sigma h_1}$,

p_{2,h_2} right-hand side of (3.3), i.e. $p_{2,h_2} = \varepsilon\sqrt{h'_2(\mathbf{B}'_2(\mathbf{B}_1\Sigma\mathbf{B}'_1 + \mathbf{B}_2\mathbf{B}'_2)^{-1}\mathbf{B}_2)^{-1}h_2}$,

ε in our case $\varepsilon = 1$,

e_i $= (0, \dots 0, 1, 0, \dots 0)' \in \mathbb{R}^{16}$, where 1 is at the i-th position,

f_i $= (0, \dots 0, 1, 0, \dots 0)' \in \mathbb{R}^4$, where 1 is at the i-th position.

Of course we do not know the actual values of the biases $\mathbf{b}_1$ a $\mathbf{b}_2$. But as the vectors $\hat{\mathbf{b}}_1 := (\hat{\boldsymbol{\eta}}_I', \hat{\boldsymbol{\eta}}_{II}')' - (\boldsymbol{\eta}_{I,0}', \boldsymbol{\eta}_{II,0}')'$ and $\hat{\mathbf{b}}_2 := \hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0$ are their unbiased estimators, we can approximate biases values by the averages from 1000 realizations of these estimators. The results for some combinations of the unit dispersions σ_I and σ_{II} are given in Table 5.

h	$\sigma_I^2 = 0.1,\ \sigma_{II}^2 = 0.1$		$\sigma_I^2 = 0.5,\ \sigma_{II}^2 = 0.01$		$\sigma_I^2 = 0.1,\ \sigma_{II}^2 = 0.5$	
	$h_1'\mathbf{b}_1$	p_{1,h_1}	$h_1'\mathbf{b}_1$	p_{1,h_1}	$h_1'\mathbf{b}_1$	p_{1,h_1}
e_1	−0.0081	0.3162	0.0152	0.7071	−0.0047	0.3162
e_2	−0.0028	0.3162	0.0157	0.7071	0.0171	0.3162
e_3	0.0015	0.3162	−0.0067	0.7071	0.0086	0.3162
e_4	−0.0073	0.3162	0.0061	0.7071	−0.0073	0.3162
e_5	0.0164	0.3162	0.0062	0.7071	0.0038	0.3162
e_6	−0.0074	0.3162	−0.0144	0.7071	0.0062	0.3162
e_7	0.0036	0.3162	0.0213	0.7071	−0.0224	0.3162
e_8	−0.0034	0.3162	−0.0083	0.7071	0.0059	0.3162
e_9	−0.0051	0.3162	0.0030	0.1000	0.0067	0.7071
e_{10}	0.0128	0.3162	0.0004	0.1000	0.0048	0.7071
e_{11}	−0.0067	0.3162	−0.0011	0.1000	0.0131	0.7071
e_{12}	0.0072	0.3162	0.0030	0.1000	−0.0202	0.7071
e_{13}	−0.0044	0.3162	−0.0008	0.1000	0.0237	0.7071
e_{14}	−0.0081	0.3162	−0.0034	0.1000	0.0115	0.7071
e_{15}	−0.0049	0.3162	−0.0031	0.1000	−0.0104	0.7071
e_{16}	−0.0002	0.3162	−0.0019	0.1000	0.0354	0.7071
	$h_2'\mathbf{b}_2$	p_{2,h_2}	$h_2'\mathbf{b}_2$	p_{2,h_2}	$h_2'\mathbf{b}_2$	p_{2,h_2}
f_1	0.0039	1.0904	−0.0270	1.2172	−0.0130	1.2517
f_2	0.0063	1.0904	0.0045	1.2172	−0.0143	1.2517
f_3	0.0000	1.0000	0.0001	1.0000	0.0001	1.0000
f_4	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000

Table 5.

Vectors e_i and f_i were chosen purposefully in their forms as the values $e_i'\mathbf{b}_1$ and $f_i'\mathbf{b}_2$ represent now the biases of the i-th components of the estimators $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$, respectively. The results from these tables expressively show that in all the mentioned cases the relations (3.2) and (3.3) hold, i.e. that the biases of all components of the given estimators are covered by their standard errors or more precisely by their upper estimators.

References

[1] *L. Kubáček, L. Kubáčková*: Regression models with a weak nonlinearity. Preprint of the Department of Mathematical Analysis and Applied Mathematics. Faculty of Science, Palacký University, 1998.

[2] *L. Kubáček, L. Kubáčková*: Testing statistical hypotheses in deformation measurement; One generalization of the Scheffé theorem. Acta Univ. Palacki Olomuc., Fac. rer. nat., Mathematica *37* (1998), 81–88.

[3] *J. Ševčík*: Generalization of the orthogonal regression on the case of the Helmert transformation. WDS '99 Proceedings of Contributed Papers, part I. MATFYZPRESS, Praha, 1999.

Author's address: J. Ševčík, Komerční banka, a. s., Ovocný trh 573, P. O. Box 839, 114 07 Prague 1, Czech Republic, e-mail: `jan_sevcik@kb.cz`.