Mathematica Slovaca

DOI: 10.2478/s12175-013-0122-5
Math. Slovaca **63** (2013), No. 3, 611–620

TWO STAGE REGRESSION MODEL WITH CONSTRAINTS; ADMISSIBLE SECOND STAGE PARAMETER ESTIMATION

Lubomír Kubáček

(*Communicated by Gejza Wimmer*)

ABSTRACT. If a regression experiment is realized in two stages, then two possibilities can occur in the second stage. Estimates of the first stage parameters either may be corrected by use of second stage measurements or they must stay unchanged. In the latter case, this requirement must be taken into account when estimating the second stage parameters. The situation is a little more complicated when constraints on both groups of parameters are imposed.

1. Introduction

A frequently occurring problem, mainly in geodesy, can be described as follows. In the first stage of the experiment some parameters (the first stage parameters) are estimated. In the second stage experiment some other parameters (the second stage parameters) must be estimated, however some first stage parameters have to be included in the second stage measurement process (in more detail see [2]). After the second stage experiment it can happen that some constraints on the first stage and also on the second stage parameters must be taken into account. In this situation a frequent requirement is that the estimates of the first stage parameters must not be corrected after the second stage experiment, i.e., the second stage measurement must not be used for a more precise estimation of the first stage parameters (in more detail see [3] and [4]). The usual

2010 Mathematics Subject Classification: Primary 62J05.
Keywords: two stage regression model, unbiased linear estimator, constraints, generalized inverse of a matrix.
Supported by the Council of the Czech Government MSM 6 198 959 214.

reason is the fact that in the time between the first stage experiment and the second stage experiment some engineering objects (buildings, bridges, tunnels, etc.) were constructed and their position in the state coordinate system is given by the first stage parameters. They cannot be changed during the existence of the objects.

In practice the second stage parameters are estimated by the formula given in Remark 1, however, it is not known, whether these estimators are optimal in some sense or not.

In this situation the goal is to find an optimal estimator (in some sense) for the second stage parameters.

Some solutions of the problem are given except in [3] and [4] mainly in [1], [5], [6], [7], [8]. Here the algorithm is developed which enables us to determine the best estimator of a given linear function of the second stage parameters. This algorithm, called the H-optimum estimation, is given as follows. The H-optimum estimator $\widehat{\boldsymbol{\beta}}$ minimizes the quantity $\operatorname{Tr}\left[\mathbf{H}\operatorname{Var}(\widehat{\boldsymbol{\beta}})\right]$ under the condition $\boldsymbol{a} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \mathbf{B}\widehat{\boldsymbol{\beta}} = \mathbf{0}$, which holds with probability one. Here $\mathbf{H}$ is a positive semidefinite matrix chosen by a statistician. If $\mathbf{H} = \boldsymbol{h}\boldsymbol{h}'$, $\boldsymbol{h} \in R^k$, then the H-optimum estimator $\widehat{\boldsymbol{\beta}}$ minimizes the dispersion $\boldsymbol{h}'\operatorname{Var}(\widehat{\boldsymbol{\beta}})\boldsymbol{h}$ of the estimator of the function $\boldsymbol{h}'\boldsymbol{\beta}$.

A disadvantage of this algorithm is that the dispersions of the estimators of other linear functions of the second stage parameters can be inadmissible large. This occurs also in the case that the second stage measurement is much more precise than the first stage measurement.

Thus it seems that an estimator which would be uniformly optimum, with respect to all linear functions, will be useful. Unfortunately such estimator does not exist. Therefore some reasonable admissible estimator is sought.

2. Symbols and preliminaries

The following two stage model will be under consideration

$$\begin{pmatrix} \widehat{\boldsymbol{\Theta}} \\ \boldsymbol{Y} \end{pmatrix} \sim_{l+n} \left[\begin{pmatrix} \mathbf{I}, & \mathbf{0} \\ \mathbf{D}, & \mathbf{X} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Theta} \\ \boldsymbol{\beta} \end{pmatrix}, \begin{pmatrix} \mathbf{W}, & \mathbf{0} \\ \mathbf{0}, & \boldsymbol{\Sigma} \end{pmatrix} \right], \qquad \boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{B}\boldsymbol{\beta} = \mathbf{0}, \tag{1}$$

where $\boldsymbol{\Theta}$ is the l-dimensional first stage parameter, $\widehat{\boldsymbol{\Theta}}$ is its estimator from the first stage, $\boldsymbol{Y}$ is the n-dimensional random vector (observation vector of the second stage), $\mathbf{I}$ is the n-dimensional identity matrix, $\mathbf{D}$ is an $n \times l$ given matrix, $\mathbf{X}$ is an $n \times k$ given matrix with the full rank in columns, i.e. $r(\mathbf{X}) = k \le n$, $\boldsymbol{\beta}$ is the second stage unknown k-dimensional parameter, $\mathbf{W}$ is the covariance matrix of the estimator $\widehat{\boldsymbol{\Theta}}$ and $\boldsymbol{\Sigma}$ is a covariance matrix of the observation vector $\boldsymbol{Y}$.

Both matrices $\mathbf{W}$ and $\boldsymbol{\Sigma}$ are known and positive definite. The $q \times l$ matrix $\mathbf{C}$ and $q \times k$ matrix $\mathbf{B}$ are known and $r(\mathbf{B}) = q < k$. The q-dimensional vector $\boldsymbol{a}$ is known. No assumptions on the rank of the matrices $\mathbf{D}$ and $\mathbf{C}$ are given.

The symbol $\mathcal{B}^-$ denotes the class of all generalized inverse $\mathbf{B}^-$ of the matrix $\mathbf{B}$, i.e. $\mathbf{B}\mathbf{B}^-\mathbf{B} = \mathbf{B}$. More detail about generalized inverse of the matrix see in [9]. Analogous meaning has the symbol $\mathcal{X}^-$. The symbol $(\mathbf{X}')^-_{m(\Sigma)}$ means the minimum $\boldsymbol{\Sigma}$-norm generalized inverse of the matrix $\mathbf{X}'$, i.e.

$$\mathbf{X}'(\mathbf{X}')^-_{m(\Sigma)}\mathbf{X}' = \mathbf{X}', \qquad \boldsymbol{\Sigma}(\mathbf{X}')^-_{m(\Sigma)}\mathbf{X}' = \left[\boldsymbol{\Sigma}(\mathbf{X}')^-_{m(\Sigma)}\mathbf{X}'\right]'.$$

If $\mathbf{X}'\boldsymbol{s} = \boldsymbol{h}$, then $\boldsymbol{s} = (\mathbf{X}')^-_{m(\Sigma)}\boldsymbol{h}$ is of the properties $\mathbf{X}'(\mathbf{X}')^-_{m(\Sigma)}\boldsymbol{h} = \boldsymbol{h}$ and

$$\forall\{\mathbf{X}' \in (\mathcal{X}')^-\} \;\; \|(\mathbf{X}')^-_{m(\Sigma)}\boldsymbol{h}\|_\Sigma \le \|(\mathbf{X}')^-\boldsymbol{h}\|_\Sigma.$$

Here $\forall\{\boldsymbol{u} \in R^n\} \;\; \|\boldsymbol{u}\|_\Sigma = \sqrt{\boldsymbol{u}'\boldsymbol{\Sigma}\boldsymbol{u}}$, where R^n is n-dimensional linear vector space. One version of $(\mathbf{X}')^-_{m(\Sigma)}$ is

$$(\mathbf{X}')^-_{m(\Sigma)} = \boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}.$$

Analogous meaning has the symbol $(\mathbf{B}')^-_{m(CWC')}$; one version of this matrix is

$$(\mathbf{B}')^-_{m(CWC')} = (\mathbf{B}\mathbf{B}' + \mathbf{C}\mathbf{W}\mathbf{C}')^{-1}\mathbf{B}\big[\mathbf{B}'(\mathbf{B}\mathbf{B}' + \mathbf{C}\mathbf{W}\mathbf{C}')^{-1}\mathbf{B}\big]^-.$$

The symbol $\mathbf{M}_{B'}$ means the projection matrix (in the Euclidean norm) on the orthogonal complement of the subspace $\mathcal{M}(\mathbf{B}') = \{\mathbf{B}'\boldsymbol{x} : \boldsymbol{x} \in R^q\}$, $\mathbf{M}_{B'} = \mathbf{I} - \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}\mathbf{B}$.

The matrix $(\mathbf{M}_{B'}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\mathbf{M}_{B'})^+$ is the Moore-Penrose generalized inverse of the matrix $\mathbf{M}_{B'}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\mathbf{M}_{B'}$, i.e. a matrix $\mathbf{A}^+$ must satisfy the following equalities

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}, \quad \mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+, \quad \mathbf{A}\mathbf{A}^+ = \big(\mathbf{A}\mathbf{A}^+\big)', \quad \mathbf{A}^+\mathbf{A} = \big(\mathbf{A}^+\mathbf{A}\big)'.$$

In our case

$$\begin{aligned}&(\mathbf{M}_{B'}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\mathbf{M}_{B'})^+\\ &= (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1} - (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\big[\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\big]^{-1}\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}.\end{aligned}$$

Since the estimator $\widehat{\boldsymbol{\Theta}}$ cannot be changed after the second stage experiment, the aim is to find the linear unbiased estimator $\tilde{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, such that $\boldsymbol{a} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \mathbf{B}\tilde{\boldsymbol{\beta}} = \mathbf{0}$, on the basis of $\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}$, $\widehat{\boldsymbol{\Theta}}$, $\mathbf{W}$ and $\boldsymbol{\Sigma}$. The random constraints must hold with probability one.

Instead the model (1), the following model

$$\begin{pmatrix} \widehat{\boldsymbol{\Theta}} \\ \boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}} \end{pmatrix} \sim \left[\begin{pmatrix} \mathbf{I}, & \mathbf{0} \\ 0, & \mathbf{X} \end{pmatrix}, \begin{pmatrix} \mathbf{W}, & -\mathbf{W}\mathbf{D}' \\ -\mathbf{D}\mathbf{W}, & \boldsymbol{\Sigma} + \mathbf{D}\mathbf{W}\mathbf{D}' \end{pmatrix}\right], \qquad \boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{B}\boldsymbol{\beta} = \mathbf{0}, \tag{2}$$

will be considered. The models (1) and (2) are equivalent, since the regular matrix $\begin{pmatrix} \mathbf{I}, & \mathbf{0} \\ \mathbf{D}, & \mathbf{I} \end{pmatrix}$ is used for the transformation of the model (1) into the model (2).

3. $\mathcal{U}_\beta$-admissible estimators of $\boldsymbol{\beta}$

The linear estimator of $\boldsymbol{\beta}$ in the model (2) is of the form

$$\tilde{\boldsymbol{\beta}} = \mathbf{T}_1\widehat{\boldsymbol{\Theta}} + \mathbf{T}_2(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) + \boldsymbol{t}.$$

Lemma 1. *The class of all linear unbiased estimators* $\tilde{\boldsymbol{\beta}}$ *of* $\boldsymbol{\beta}$ *which satisfy the constraints* $\boldsymbol{a} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \mathbf{B}\tilde{\boldsymbol{\beta}} = \mathbf{0}$, *is*

$$\mathcal{U}_\beta = \left\{ -\mathbf{B}^-\mathbf{C}\widehat{\boldsymbol{\Theta}} + (\mathbf{I} - \mathbf{B}^-\mathbf{B})\mathbf{X}^-(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) - \mathbf{B}^-\boldsymbol{a} : \ \mathbf{B}^- \in \mathcal{B}^-, \ \ \mathbf{X} \in \mathcal{X}^- \right\}.$$

P r o o f. The estimator $\tilde{\boldsymbol{\beta}}$ is unbiased iff

$$\forall\{\boldsymbol{\Theta}, \boldsymbol{\beta} : \ \boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{B}\boldsymbol{\beta} = \mathbf{0}\} \ \ E_{\Theta,\beta}\big[\mathbf{T}_1\widehat{\boldsymbol{\Theta}} + \mathbf{T}_2(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) + \boldsymbol{t}\big] = \boldsymbol{\beta},$$

i.e.

$$\forall\{\boldsymbol{\Theta}, \boldsymbol{\beta} : \ \boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{B}\boldsymbol{\beta} = \mathbf{0}\} \ \ \mathbf{T}_1\boldsymbol{\Theta} + (\mathbf{T}_2\mathbf{X} - \mathbf{I})\boldsymbol{\beta} + \boldsymbol{t} = \mathbf{0}$$
$$\iff \exists\{\mathbf{E} : \ k \times q \ \text{ matrix}\} \ \ \mathbf{T}_1 = \mathbf{E}\mathbf{C} \ \& \ (\mathbf{T}_2\mathbf{X} - \mathbf{I}) = \mathbf{E}\mathbf{B} \ \& \ \boldsymbol{t} = \mathbf{E}\boldsymbol{a},$$

i.e.

$$\mathbf{T}_1 = \mathbf{E}\mathbf{C}, \quad \mathbf{T}_2 = (\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-, \quad \boldsymbol{t} = \mathbf{E}\boldsymbol{a}.$$

Thus the estimator is of the form $\tilde{\boldsymbol{\beta}} = \mathbf{E}\mathbf{C}\widehat{\boldsymbol{\Theta}} + (\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) + \mathbf{E}\boldsymbol{a}$ and it must satisfy the constraints, i.e.

$$\boldsymbol{a} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \mathbf{B}\big[\mathbf{E}\mathbf{C}\widehat{\boldsymbol{\Theta}} + (\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) + \mathbf{E}\boldsymbol{a}\big] = \mathbf{0},$$
$$\iff \boldsymbol{a} + \mathbf{B}\mathbf{E}\boldsymbol{a} + \big[\mathbf{C} + \mathbf{B}\mathbf{E}\mathbf{C} - \mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-\mathbf{D}\big]\big)\widehat{\boldsymbol{\Theta}} + \mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-\boldsymbol{Y} = \mathbf{0}.$$

Since $\widehat{\boldsymbol{\Theta}} = \boldsymbol{\Theta} + \boldsymbol{\varepsilon}_{\widehat{\Theta}}$, $\boldsymbol{\varepsilon}_{\widehat{\Theta}} \sim_l (0, \mathbf{W})$, $\boldsymbol{Y} = \mathbf{D}\boldsymbol{\Theta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_Y$, $\boldsymbol{\varepsilon}_Y \sim_n (0, \boldsymbol{\Sigma})$, the constraints can be rewritten as

$$\begin{aligned}\boldsymbol{a} + \mathbf{B}\mathbf{E}\boldsymbol{a} + \big[\mathbf{C} + \mathbf{B}\mathbf{E}\mathbf{C} - \mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-\mathbf{D}\big]\boldsymbol{\Theta} + \mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-\mathbf{D}\boldsymbol{\Theta}\\ + \mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-\mathbf{X}\boldsymbol{\beta} + (\mathbf{C} + \mathbf{B}\mathbf{E}\mathbf{C})\boldsymbol{\varepsilon}_{\widehat{\Theta}} + \mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^-\boldsymbol{\varepsilon}_Y = \mathbf{0}.\end{aligned} \tag{3}$$

Thus

$$\mathbf{B}(\mathbf{I} + \mathbf{E}\mathbf{B})\mathbf{X}^- = \mathbf{0} \implies \mathbf{B} + \mathbf{B}\mathbf{E}\mathbf{B} = \mathbf{0}, \ \ \boldsymbol{a} + \mathbf{B}\mathbf{E}\boldsymbol{a} = \mathbf{0},$$

and from (3)

$$\mathbf{C} + \mathbf{B}\mathbf{E}\mathbf{C} = \mathbf{0} \implies \mathbf{E} = -\mathbf{B}^-, \ \ \mathbf{T}_2 = (\mathbf{I} - \mathbf{B}^-\mathbf{B})\mathbf{X}^-, \ \ \mathbf{T}_1 = -\mathbf{B}^-\mathbf{C},$$

since

$$P\{\boldsymbol{\varepsilon}_Y \in \mathcal{M}(\boldsymbol{\Sigma}) = R^n\} = 1, \quad P\{\boldsymbol{\varepsilon}_{\widehat{\Theta}} \in \mathcal{M}(\mathbf{W}) = R^l\} = 1,$$

the constraints must be satisfied for every $\boldsymbol{\varepsilon}_{\widehat{\Theta}}$ and $\boldsymbol{\varepsilon}_Y$ and

$$\begin{aligned} &\boldsymbol{a} + \mathbf{BE}\boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{BEC}\boldsymbol{\Theta} + \mathbf{B}(\mathbf{I} + \mathbf{EB})\boldsymbol{\beta} \\ &\quad = \boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{B}\boldsymbol{\beta} + \mathbf{BE}(\boldsymbol{a} + \mathbf{C}\boldsymbol{\Theta} + \mathbf{B}\boldsymbol{\beta}) = \mathbf{0}. \end{aligned}$$

□

Lemma 2. *Let* $\mathbf{A}$ *be any* $q \times l$ *matrix. Then*

$$\forall\{\boldsymbol{h} \in R^k\}\ \forall\{\mathbf{B}^- \in \mathcal{B}^-\} \quad \operatorname{Var}\Big\{\boldsymbol{h}'\big[(\mathbf{B}')^-_{m(AWA')}\big]'\mathbf{A}\boldsymbol{\varepsilon}_{\widehat{\Theta}}\Big\} \le \operatorname{Var}(\boldsymbol{h}'\mathbf{B}^-\mathbf{A}\boldsymbol{\varepsilon}_{\widehat{\Theta}}).$$

P r o o f. Let $\boldsymbol{h} \in \mathcal{M}(\mathbf{B}')$, i.e. $\exists\{\boldsymbol{s}:\ \boldsymbol{s} \in R^q\}\mathbf{B}'\boldsymbol{s} = \boldsymbol{h}$. If $\boldsymbol{s} = (\mathbf{B}')^-_{m(AWA')}\boldsymbol{h}$, then regarding the properties of the matrix $(\mathbf{B}')^-_{m(AWA')}$, we have

$$\begin{aligned} \operatorname{Var}\Big\{\boldsymbol{h}'\big[(\mathbf{B}')^-_{m(AWA')}\big]'\mathbf{A}\boldsymbol{\varepsilon}_{\widehat{\Theta}}\Big\} &= \|(\mathbf{B}')^-_{m(AWA')}\boldsymbol{h}\|^2_{AWA'} \\ &\le \|(\mathbf{B}')^-\boldsymbol{h}\|^2_{AWA'} = \operatorname{Var}\big[\boldsymbol{h}'(\mathbf{B}')^-\mathbf{A}\boldsymbol{\varepsilon}_{\widehat{\Theta}}\big]. \end{aligned}$$

□

Definition. The estimator

$$\widehat{\boldsymbol{\beta}} \in \mathcal{U}_\beta = \Big\{ - \big[\mathbf{B}^-\mathbf{C} + (\mathbf{I} - \mathbf{B}^-\mathbf{B})\mathbf{X}^-\mathbf{D}\big]\widehat{\boldsymbol{\Theta}} + (\mathbf{I} - \mathbf{B}^-\mathbf{B})\mathbf{X}^-\boldsymbol{Y} - \mathbf{B}^-\boldsymbol{a} : \\ \mathbf{B} \in \mathcal{B}^-,\ \ \mathbf{X}^- \in \mathcal{X}^-\Big\}$$

is $\mathcal{U}_\beta$-admissible if

$$\exists\{\boldsymbol{h}_0 \in R^k\}\ \forall\{\tilde{\boldsymbol{\beta}} \in \mathcal{U}_\beta\} \quad \operatorname{Var}(\boldsymbol{h}_0'\widehat{\boldsymbol{\beta}}) \le \operatorname{Var}(\boldsymbol{h}_0'\tilde{\boldsymbol{\beta}}).$$

Theorem 1. *The estimator*

$$\begin{aligned} \widehat{\boldsymbol{\beta}} = &-\big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{C}\widehat{\boldsymbol{\Theta}} + \Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\mathbf{X}^-\mathbf{D}\widehat{\boldsymbol{\Theta}} \\ &+ \Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\mathbf{X}^-\boldsymbol{Y} - \big[(\mathbf{B}')^-_{m(CWC')}\big]\boldsymbol{a} \end{aligned}$$

is $\mathcal{U}_\beta$*-admissible for any* $\mathbf{X}^- \in \mathcal{X}^-$.

P r o o f. Let $\boldsymbol{h}_0 \in \mathcal{M}(\mathbf{B}')$. Then $\exists\{\boldsymbol{u} \in R^k\}$ $\boldsymbol{h}_0 = \mathbf{B}'\boldsymbol{u}$ and

$$\boldsymbol{h}_0'\widehat{\boldsymbol{\beta}} = -\boldsymbol{u}'\mathbf{B}\big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{C}\widehat{\boldsymbol{\Theta}} - \boldsymbol{u}'\boldsymbol{a} = -\boldsymbol{u}'\mathbf{C}\widehat{\boldsymbol{\Theta}} - \boldsymbol{u}'\boldsymbol{a},$$

since $\mathbf{B}\Big[(\mathbf{B}')^-_{m(CWC')}\Big]' = \mathbf{I}$ and $\mathbf{B}\Big\{\mathbf{I} - \Big[(\mathbf{B}')^-_{m(CWC')}\Big]'\mathbf{B}\Big\} = \mathbf{0}$.

If $\boldsymbol{u}$ is chosen as

$$\boldsymbol{u} = (\mathbf{B}')^-_{m(CWC')}\boldsymbol{h}_0,$$

then

$$\forall\{\mathbf{B}^- \in \mathcal{B}^-\} \ \ \|(\mathbf{B}')^-_{m(CWC')}\boldsymbol{h}_0\|^2_{CWC'} = \boldsymbol{h}_0'\big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{CWC}'(\mathbf{B}')^-_{m(CWC')}\boldsymbol{h}_0$$
$$= \mathrm{Var}(\boldsymbol{h}_0'\widehat{\boldsymbol{\beta}}) \le \|(\mathbf{B}')^-\boldsymbol{h}_0\|^2_{CWC'}$$
$$= \boldsymbol{h}_0'\big[(\mathbf{B}')^-\big]'\mathbf{CWC}'(\mathbf{B}')^-\boldsymbol{h}_0 = \mathrm{Var}(\boldsymbol{h}_0'\tilde{\boldsymbol{\beta}}),$$

where

$$\tilde{\boldsymbol{\beta}} = -\mathbf{B}^-\mathbf{C}\widehat{\boldsymbol{\Theta}} - (\mathbf{I} - \mathbf{B}^-\mathbf{B})\mathbf{X}^-\mathbf{D}\widehat{\boldsymbol{\Theta}} + (\mathbf{I} - \mathbf{B}^-\mathbf{B})\mathbf{X}^-\boldsymbol{Y} - \mathbf{B}^-\boldsymbol{a} \in \mathcal{U}_\beta.$$

□

If $\boldsymbol{h}_1 \in \mathcal{M}(\mathbf{I} - \mathbf{B}'(\mathbf{B}')^-_{m(CWC')})$, then

$$\boldsymbol{h}_1'\widehat{\boldsymbol{\beta}} = \Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\mathbf{X}^-(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}).$$

The choice

$$\mathbf{X}^- = \big[(\mathbf{X}')^-_{m(\Sigma+DWD')}\big]'$$

makes the dispersion $\mathrm{Var}(\boldsymbol{h}_1'\widehat{\boldsymbol{\beta}})$ of $\boldsymbol{h}_1\widehat{\boldsymbol{\beta}}$ minimum, i.e.

$$\forall\{\mathbf{X} \in \mathcal{X}^-\} \ \ \mathrm{Var}\Big(\Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\big[(\mathbf{X}')^-_{m(\Sigma+DWD')}\big](\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}})\Big)$$
$$\le \mathrm{Var}\Big(\Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\mathbf{X}^-(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}).$$

Thus

$$\widehat{\boldsymbol{\beta}} = -\big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{C}\widehat{\boldsymbol{\Theta}} + \Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\big[(\mathbf{X}')^-_{m(\Sigma+DWD')}\big]'$$
$$\times (\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\boldsymbol{a} \qquad (4)$$

is $\mathcal{U}_\beta$-admissible estimator and in addition it is valid that for any function $\boldsymbol{h}'\boldsymbol{\beta}$, where $\boldsymbol{h} \in \mathcal{M}\big[\mathbf{I} - \mathbf{B}'(\mathbf{B}')^-_{M(CWC'}\big]$ the estimator $\boldsymbol{h}'\widehat{\boldsymbol{\beta}}$ has the smallest dispersion. The covariance matrix of $\widehat{\boldsymbol{\beta}}$ from (4) is

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \Big(\big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{C} + \Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\big[(\mathbf{X}')^-_{m(\Sigma+DWD')}\big]'\Big)$$
$$\times \mathbf{W}\Big\{\mathbf{C}'(\mathbf{B}')^-_{m(CWC')} + \mathbf{D}'(\mathbf{X}')^-_{m(\Sigma+DWD')}\big[\mathbf{I} - \mathbf{B}'(\mathbf{B}')^-_{m(CWC')}\big]\Big\}$$
$$+ \Big\{\mathbf{I} - \big[(\mathbf{B}')^-_{m(CWC')}\big]'\mathbf{B}\Big\}\big[(\mathbf{X}')^-_{m(\Sigma+DWD')}\big]'\boldsymbol{\Sigma}(\mathbf{X}')^-_{m(\Sigma+DWD')}$$
$$\times \big[\mathbf{I} - \mathbf{B}'(\mathbf{B}')^-_{m(CWC')}\big]. \qquad (5)$$

Remark 1. In practice the estimator

$$\tilde{\tilde{\boldsymbol{\beta}}} = \tilde{\boldsymbol{\beta}} - (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\big[\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\big]^{-1}(\mathbf{B}\tilde{\boldsymbol{\beta}} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \boldsymbol{a}),$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}})$ has been used frequently. The estimator $\tilde{\tilde{\boldsymbol{\beta}}}$ is unbiased and it satisfies the identity $\boldsymbol{a} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \mathbf{B}\tilde{\tilde{\boldsymbol{\beta}}} = \mathbf{0}$ as well. However we do not know whether this estimator is in general $\mathcal{U}_\beta$-admissible, or not.

4. Examples

Example 1. Let a levelling traverse $\{A, P_1, P_2, B\}$ between points A, B be considered. Here P_1 and P_2 are points which heights must be determined. The heights $H_A = \Theta_1$ and $H_B = \Theta_2$ of the points A and B, respectively, are measured in the first stage, i.e.

$$\begin{pmatrix} \widehat{\Theta}_1 \\ \widehat{\Theta}_2 \end{pmatrix} \sim_2 \left[\begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix}, (0.05m)^2 \mathbf{I}_3 \right]$$

and the differences $H_{P_1} - H_A = \beta_1$, $H_{P_2} - H_{P_1} = \beta_2$ and $H_B - H_{P_2} = \beta_3$ are measured in the second stage, i.e.

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \sim_3 \left(\mathbf{I}_3 \boldsymbol{\beta}, (0.01m)^2 \mathbf{I}_3 \right).$$

Let $\mathbf{1} = (1, \ldots, 1)'$ and $\boldsymbol{e}_i$ is the vector with 1 on the ith position and other entries are zero.

It must be valid that $\mathbf{1}'\boldsymbol{\beta} + (1, -1)\boldsymbol{\Theta} = 0$. The estimator $\begin{pmatrix} \widehat{\Theta}_1 \\ \widehat{\Theta}_2 \end{pmatrix}$ must not be changed after the second stage measurement.

We have

$$\mathbf{X} = \mathbf{I}_3, \quad \mathbf{B} = \mathbf{1}_3', \quad \mathbf{C} = (1, -1), \quad \mathbf{W} = (0.05m)^2 \mathbf{I}_2, \quad \boldsymbol{\Sigma} = (0.01m)^2 \mathbf{I}_3.$$

The $\mathcal{U}_\beta$-admissible estimator from (4) is

$$\begin{aligned} \widehat{\boldsymbol{\beta}} = & - \left[(\mathbf{B}')^-_{m(CWC')} \right]' \mathbf{C}\widehat{\boldsymbol{\Theta}} + \left\{ \mathbf{I} - \left[(\mathbf{B}')^-_{m(CWC')} \right]' \mathbf{B} \right\} \left[(\mathbf{X}')^-_{m(\Sigma + DWD')} \right]' \\ & \times (\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) - \left[(\mathbf{B}')^-_{m(CWC')} \right]' \boldsymbol{a} \\ = & - \frac{1}{3} \mathbf{1}(1, -1)\widehat{\boldsymbol{\Theta}} + \left(\mathbf{I}_3 - \frac{1}{3}\mathbf{1}\mathbf{1}' \right) \boldsymbol{Y} = -\frac{1}{3}\mathbf{1}(\widehat{\Theta}_1 - \widehat{\Theta}_2) + \mathbf{M}_3 \boldsymbol{Y} \end{aligned}$$

and the estimator used in practice is

$$\begin{aligned} \tilde{\tilde{\boldsymbol{\beta}}} &= \tilde{\boldsymbol{\beta}} - (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}' \left[\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}' \right]^{-1} (\mathbf{B}\tilde{\boldsymbol{\beta}} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \boldsymbol{a}) \\ &= -\frac{1}{3}\mathbf{1}(\widehat{\Theta}_1 - \widehat{\Theta}_2) + \mathbf{M}_3 \boldsymbol{Y}, \end{aligned}$$

i.e. it is also $\mathcal{U}_\beta$-admissible estimator. Here $\mathbf{M}_3 = \mathbf{I}_3 - \frac{1}{3}\mathbf{1}\mathbf{1}'$.

Example 2. Let in the preceding experiment the design of measurement be changed in the following way

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \sim_3 \left[\begin{pmatrix} 1, & 0, & 1, & 0, & 0 \\ 1, & 0, & 1, & 1, & 0 \\ 1, & 0, & 1, & 1, & 1 \end{pmatrix} \begin{pmatrix} \Theta_1 \\ \Theta_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}, (0.01m)^2 \mathbf{I}_3 \right],$$

$$\mathbf{1}'\boldsymbol{\beta} + (\Theta_1 - \Theta_2) = 0.$$

Thus

$$\mathbf{X} = \begin{pmatrix} 1, & 0, & 0 \\ 1, & 1, & 0 \\ 1, & 1, & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 1, & 0 \\ 1, & 0 \\ 1, & 0 \end{pmatrix}, \quad \mathbf{C} = (1, -1), \quad \mathbf{B} = \mathbf{1}'.$$

The $\mathcal{U}_\beta$-admissible estimator from (4) is

$$\begin{aligned} \widehat{\boldsymbol{\beta}} = & -\left[(\mathbf{B}')^-_{m(CWC')}\right]'\mathbf{C}\widehat{\boldsymbol{\Theta}} + \left\{\mathbf{I} - \left[(\mathbf{B}')^-_{m(CWC')}\right]'\mathbf{B}\right\}\left[(\mathbf{X}')^-_{m(\Sigma+DWD')}\right]' \\ & \times (\boldsymbol{Y} - \mathbf{D}\widehat{\boldsymbol{\Theta}}) - \left[(\mathbf{B}')^-_{m(CWC')}\right]'\boldsymbol{a} \\ = & -\frac{1}{3}\mathbf{1}(\widehat{\Theta}_1 - \widehat{\Theta}_2) + \left(\mathbf{I}_3 - \frac{1}{3}\mathbf{1}\mathbf{1}'\right) \begin{pmatrix} 1, & 0, & 0 \\ -1, & 1, & 0 \\ 0, & -1, & 1 \end{pmatrix} \begin{pmatrix} Y_1 - \widehat{\Theta}_1 \\ Y_2 - \widehat{\Theta}_1 \\ Y_3 - \widehat{\Theta}_1 \end{pmatrix} \\ = & \begin{pmatrix} -1, & \frac{1}{3} \\ 0, & \frac{1}{3} \\ 0, & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \widehat{\Theta}_1 \\ \widehat{\Theta}_2 \end{pmatrix} + \begin{pmatrix} 1, & 0, & -\frac{1}{3} \\ -1, & 1, & -\frac{1}{3} \\ 0, & -1, & \frac{2}{3} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}. \end{aligned}$$

and the estimator used in practice is

$$\begin{aligned} \tilde{\tilde{\boldsymbol{\beta}}} &= \tilde{\boldsymbol{\beta}} - (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\left[\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\right]^{-1}(\mathbf{B}\tilde{\boldsymbol{\beta}} + \mathbf{C}\widehat{\boldsymbol{\Theta}} + \boldsymbol{a}) \\ &= \begin{pmatrix} -1, & 0 \\ 0, & 0 \\ 0, & 1 \end{pmatrix} \begin{pmatrix} \widehat{\Theta}_1 \\ \widehat{\Theta}_2 \end{pmatrix} + \begin{pmatrix} 1, & 0, & 0 \\ -1, & 1, & 0 \\ 0, & -1, & 0 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_1 \\ Y_3 \end{pmatrix} \end{aligned}$$

(in this case $\widehat{\boldsymbol{\beta}} \neq \tilde{\tilde{\boldsymbol{\beta}}}$).

$$\operatorname{Var}(\widehat{\boldsymbol{\beta}}) = (0.05m)^2 \frac{1}{9} \begin{pmatrix} 10, & 1, & 1 \\ 1, & 1, & 1 \\ 1, & 1, & 1 \end{pmatrix} + (0.01m)^2 \frac{1}{9} \begin{pmatrix} 10, & -8, & -12 \\ -8, & 19, & -11 \\ -2, & -11, & 13 \end{pmatrix}$$

$$\operatorname{Var}(\tilde{\tilde{\boldsymbol{\beta}}}) = (0.05m)^2 \frac{1}{9} \begin{pmatrix} 9, & 0, & 0 \\ 0, & 0, & 0 \\ 0, & 0, & 9 \end{pmatrix} + (0.01m)^2 \frac{1}{9} \begin{pmatrix} 9, & -9, & 0 \\ -9, & 18, & -9 \\ 0, & -9, & 9 \end{pmatrix}$$

$$\mathrm{Var}(\mathbf{1}'\widehat{\boldsymbol{\beta}}) = \mathrm{Var}(\mathbf{1}'\tilde{\tilde{\boldsymbol{\beta}}}) = 0.0050m^2 = (0.071m)^2,$$

$$\mathrm{Var}(\boldsymbol{e}_1'\widehat{\boldsymbol{\beta}}) = 0.0029m^2 = (0.054m)^2, \quad \mathrm{Var}(\boldsymbol{e}_1'\tilde{\tilde{\boldsymbol{\beta}}}) = 0.0026m^2 = (0.051m)^2,$$

$$\mathrm{Var}(\boldsymbol{e}_2'\widehat{\boldsymbol{\beta}}) = 0.0005m^2 = (0.022m)^2, \quad \mathrm{Var}(\boldsymbol{e}_2'\tilde{\tilde{\boldsymbol{\beta}}}) = 0.0002m^2 = (0.014m)^2,$$

$$\mathrm{Var}(\boldsymbol{e}_3'\widehat{\boldsymbol{\beta}}) = 0.0004m^2 = (0.020m)^2, \quad \mathrm{Var}(\boldsymbol{e}_3'\tilde{\tilde{\boldsymbol{\beta}}}) = 0.0026m^2 = (0.051m)^2.$$

If $\boldsymbol{h}' = \left(\frac{-1}{3}, \frac{2}{3}, \frac{-1}{3}\right)$, then

$$\mathrm{Var}(\boldsymbol{h}_1'\widehat{\boldsymbol{\beta}}) = 0.000\,395m^2 = (0.020m)^2, \quad \mathrm{Var}(\boldsymbol{h}_1'\tilde{\tilde{\boldsymbol{\beta}}}) = 0.000\,461m^2 = (0.021m)^2.$$

Some linear function of $\boldsymbol{\beta}$ may be better estimated by the $\mathcal{U}_\beta$-admissible estimator (e.g. $\boldsymbol{e}_3'\boldsymbol{\beta}, \boldsymbol{h}'\boldsymbol{\beta}$), however $\boldsymbol{e}_2'\boldsymbol{\beta}$ is better estimated by the procedure used in practice. The $\mathcal{U}_\beta$-admissible estimator is not the uniformly (with respect to the class $\mathcal{U}_\beta$) best estimator and thus such behaviour of the $\mathcal{U}_\beta$ admissible estimator must be expected.

5. Conclusion

Two stage regression models with constraints occur frequently mainly in geodetical practice. A less precise measurement in the first stage can enlarge the variance of the estimators of the second stage parameters significantly. Therefore several attempts were made to introduce algorithms which suppress the influence of the not sufficient precision of the first stage measurement on the variance of the second stage parameter estimators. They were already mentioned in the text.

If a single linear function of the second stage parameters is estimated only, then the H-optimum estimator is a solution.

The aim of the research was to find the class of all $\mathcal{U}_\beta$-admissible estimators, however this aim was not attained. Until now it is even not known whether procedures frequently used in practice are in general $\mathcal{U}_\beta$-admissible. Therefore using the $\mathcal{U}_\beta$-admissible estimator is more safe.

REFERENCES

[1] KORBAŠOVÁ, M.—MAREK, J.: *Connecting measurements in surveying and its problem.* In: Proceedings of INGEO 2004 and FIG Regional Central and Eastern European Conference on Engineering Surveying, November, Bratislava, Slovakia, 2004, pp. 11–123.

[2] KUBÁČEK, L.: *Two stage regression model*, Math. Slovaca **38** (1988), 383–393.

[3] KUBÁČEK, L.: *Two stage linear model with constraints*, Math. Slovaca **43** (1993), 643–658.

[4] KUBÁČEK, L.—KUBÁČKOVÁ, L.—VOLAUFOVÁ, J.: *Statistical Models with Linear Structures*, Veda, Bratislava, 1995.

[5] KUBÁČEK, L.—MAREK, J.: *Partial optimum estimator in two stage regression model with constraints and a problem of equivalence*, Math. Slovaca **55** (2005), 477–494.

[6] MAREK, J.: *Estimation in connecting measurements*, Acta Univ. Palack. Olomuc. Fac. Rerum Natur. Math. **42** (2003), 69–86.

[7] MAREK, J.: *Estimation in connecting measurements with constraints of type II*, Acta Univ. Palack. Olomuc. Fac. Rerum Natur. Math. **43** (2004), 119–131.

[8] MAREK, J.—FIŠEROVÁ, E.: *Statistical analysis of geodetical measurementrs*. In: ROBUST 2004. Sborník 13. letní školy, Třešť, 7.–11. 6. 2004, JČMF, Praha, 2004, pp. 253–260.

[9] RAO, C. R.—MITRA, S. K.: *Generalized Inverse of Matrices and Its Applications*, John Wiley & Sons, New York, 1971.

Received 10. 11. 2010
Accepted 30. 6. 2011

Department of Mathematical Analysis
and Applications of Mathematics
Faculty of Science
17. listopadu 12
CZ–771 46 Olomouc
CZECH REPUBLIC
E-mail: lubomir.kubacek@upol.cz