GEODEZJA I KARTOGRAFIA
GEODESY AND CARTOGRAPHY
Vol. 58, No 1, 2009, pp. 23-48

M_{split} estimation. Part II: Squared M_{split} estimation and numerical examples

Zbigniew Wiśniewski

Institute of Geodesy, University of Warmia and Mazury
1 Oczapowskiego St., 10-957 Olsztyn, Poland
e-mail: zbyszekw @uwm.edu.pl

Received: 27 November 2008/Accepted: 30 March 2009

Abstract: This part of the paper presents particular case of M_{split} estimation called a squared M_{split} estimation whose target function is based on convex squared functions. One can find here theoretical foundations and algorithm of the squared M_{split} estimation as well as some numerical examples.

Keywords: Geodetic adjustment, M estimation, split potential of an observation set

---

1. Introduction

The first part of the paper presented the theory of M_{split} estimation. Let $\boldsymbol{\theta}_\alpha = [\theta_{1\alpha}, \theta_{2\alpha}, \cdots, \theta_{n\alpha}]^{\mathrm{T}}$ and $\boldsymbol{\theta}_\alpha = [\theta_{1\alpha}, \theta_{2\alpha}, \cdots, \theta_{n\alpha}]^{\mathrm{T}}$ be parameters of the random variable $\mathbf{y} = [y_1, y_2, \ldots, y_n]^{\mathrm{T}}$, which probability distributions belong to either of two competitive families

$$\mathcal{P}_\alpha = \{P_{\boldsymbol{\theta}_\alpha} \succ f_\alpha(\mathbf{y}; \boldsymbol{\theta}_\alpha) : \boldsymbol{\theta}_\alpha \in \boldsymbol{\Theta}_\alpha\} \text{ or } \mathcal{P}_\beta = \{P_{\boldsymbol{\theta}_\beta} \succ f_\beta(\mathbf{y}; \boldsymbol{\theta}_\beta) : \boldsymbol{\theta}_\beta \in \boldsymbol{\Theta}_\beta\}$$

It was shown that estimates $\hat{\boldsymbol{\theta}}_\alpha$ and $\hat{\boldsymbol{\theta}}_\beta$ are M_{split} estimators of these parameters when

$$\min_{\boldsymbol{\theta}_\alpha, \boldsymbol{\theta}_\beta} [-k_{\alpha,\beta}(\mathbf{y}; \boldsymbol{\theta}_\alpha, \boldsymbol{\theta}_\beta)] = -k_{\alpha,\beta}(\mathbf{y}; \hat{\boldsymbol{\theta}}_\alpha, \hat{\boldsymbol{\theta}}_\beta) \tag{1}$$

($f\ (y; \theta)$ – density function). The term $k_{\alpha,\beta}(\mathbf{y}; \boldsymbol{\theta}_\alpha, \boldsymbol{\theta}_\beta)$ is the logarithmic split potential and can be written as follows

$$k_{\alpha,\beta}(\mathbf{y}; \boldsymbol{\theta}_\alpha, \boldsymbol{\theta}_\beta) = \sum_{i=1}^{n} I_\alpha^f(y_i; \theta_{i\alpha}) I_\beta^f(y_i; \theta_{i\beta}) = -\sum_{i=1}^{n} \ln f_\alpha(y_i; \theta_{i\alpha}) \ln f_\beta(y_i; \theta_{i\beta}) \tag{2}$$

If f-information $I_\alpha^f(y; \theta_\alpha) = -\ln f_\alpha(y; \theta_\alpha)$ and also $I_\beta^f(y; \theta_\beta) = -\ln f_\beta(y; \theta_\beta)$ are replaced with at least double differentiable and positive convex functions $\rho_\alpha(y; \theta_\alpha)$ and $\rho_\beta(y; \theta_\beta)$, respectively, then the general target function of M_{split} estimation can be written in the following form

$$\varphi(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta)=\sum_{i=1}^{n}\rho_\alpha(y_i;\theta_{i\alpha})\rho_\beta(y_i;\theta_{i\beta}) \tag{3}$$

Thus $\hat{\boldsymbol{\theta}}_\alpha$ and $\hat{\boldsymbol{\theta}}_\beta$ are M_{split} estimates of the parameters $\boldsymbol{\theta}_\alpha$, $\boldsymbol{\theta}_\beta$ when

$$\min_{\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta}\varphi(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta)=\varphi(\mathbf{y};\hat{\boldsymbol{\theta}}_\alpha,\hat{\boldsymbol{\theta}}_\beta) \tag{4}$$

If $\mathbf{y}$ is the vector of geodetic observations described by the functional model $\mathbf{v}=\mathbf{y}-\mathbf{AX}$ that is split into the models $\mathbf{v}_\alpha=\mathbf{y}-\mathbf{AX}_\alpha$ and $\mathbf{v}_\beta=\mathbf{y}-\mathbf{AX}_\beta$, then the target function is written as follows

$$\varphi(\mathbf{y};\mathbf{X}_\alpha,\mathbf{X}_\beta)=\sum_{i=1}^{n}\rho_\alpha(y_i;\mathbf{X}_\alpha)\rho_\beta(y_i;\mathbf{X}_\beta)=[\boldsymbol{\rho}_\alpha(\mathbf{y};\mathbf{X}_\alpha)]^{\mathrm{T}}\boldsymbol{\rho}_\beta(\mathbf{y};\mathbf{X}_\beta) \tag{5}$$

and M_{split} estimates of the parameters $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$ are such quantities $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$ that fulfill the condition

$$\min_{\mathbf{X}_\alpha,\mathbf{X}_\beta}\varphi(\mathbf{y};\mathbf{X}_\alpha,\mathbf{X}_\beta)=\varphi(\mathbf{y};\hat{\mathbf{X}}_\alpha,\hat{\mathbf{X}}_\beta) \tag{6}$$

The functions $\rho_\alpha(\cdot)$ and $\rho_\beta(\cdot)$ can be assumed arbitrarily as long as they have earlier mentioned theoretical properties and as long as they can guarantee assumed properties of the estimators proposed. In the paper (Wiśniewski, 2009) it was considered some case of M_{split} estimation where functions $\rho_\alpha(\cdot)$ and $\rho_\beta(\cdot)$ are squared ones. The following papers (Wiśniewski, 2008, 2009) presented some numerical analyses of such estimation method called the squared M_{split} estimation. It was proved that the squared M_{split} estimation is not only a good efficient alternative for robust against outliers methods of estimation.

This part presents extended theory of the squared M_{split} estimation, some new numerical examples illustrating its properties and pointing at some possible practical applications.

2. Theoretical foundations of the squared M_{split} estimation

2.1. Split potential

Let the functions $\rho_\alpha(y_i;\ \theta_{i\alpha})$ and $\rho_\beta(y_i;\ \theta_{i\beta})$ from the target function Eq. (3) be as follows: $\rho_\alpha(y_i;\ \theta_{i\alpha})$

$$\rho_\alpha(y_i;\theta_{i\alpha})=\sigma_i^{-2}(y_i-\theta_{i\alpha})^2=v_{i\alpha}^2 \tag{7}$$

$$\rho_\beta(y_i;\theta_{i\beta}) = \sigma_i^{-2}(y_i - \theta_{i\beta})^2 = v_{i\beta}^2 \tag{8}$$

where σ_i is the standard deviation of the variable y_i, and $v_{i\alpha}$, $v_{i\beta}$ are two variants of standardized random errors of measurements. Thus one can write

$$\begin{aligned}\varphi(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta) &= \sum_{i=1}^{n} \eta(y_i;\theta_{i\alpha},\theta_{i\beta}) = \\ &= \sum_{i=1}^{n} \rho_\alpha(y_i;\theta_{i\alpha})\rho_\beta(y_i;\theta_{i\beta}) = \sum_{i=1}^{n} v_{i\alpha}^2 v_{i\beta}^2\end{aligned} \tag{9}$$

By taking

$$\sum_{i=1}^{n} v_{i\alpha}^2 v_{i\beta}^2 = (\mathbf{v}_\alpha * \mathbf{v}_\alpha)^{\mathrm{T}}(\mathbf{v}_\alpha * \mathbf{v}_\beta)$$

the target function presented above can be rewritten as

$$\varphi(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta) = \sum_{i=1}^{n} v_{i\alpha}^2 v_{i\beta}^2 = (\mathbf{v}_\alpha * \mathbf{v}_\alpha)^{\mathrm{T}}(\mathbf{v}_\beta * \mathbf{v}_\beta) \tag{10}$$

where $\mathbf{v} = [v_1, v_2, ..., v_n]^{\mathrm{T}}$ is a vector of random errors and $*$ is the Hadamard product (e.g. Rao, 1973). Such created function Eq. (10) is the target function of the squared M_{split} estimation.

By taking into consideration the theory of the split potential presented earlier and the assumed functions Eqs. (7)-(8), the following elementary split potential can be assigned to the observation y_i (in the logarithmic form)

$$k_{\alpha,\beta}(y_i;\theta_{i\alpha},\theta_{i\beta}) = -I_\alpha^f(y_i;\theta_{i\alpha})I_\beta^f(y_i;\theta_{i\beta}) = -v_{i\alpha}^2 v_{i\beta}^2 \tag{11}$$

According to this theory, the quantities that were applied in the above formula, i.e.

$$I_\alpha^f(y_i;\theta_{i\alpha}) = v_{i\alpha}^2 \tag{12}$$

$$I_\beta^f(y_i;\theta_{i\beta}) = v_{i\beta}^2 \tag{13}$$

are f-informations that the observation y_i can provide under assumption of either of two competitive models $v_{i\alpha} = y_i - \theta_{i\alpha}$ and $v_{i\beta} = y_i - \theta_{i\beta}$. Therefore, the following global logarithmic split potential

$$\begin{aligned}k_{\alpha,\beta}(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta) &= \sum_{i=1}^{n} k_{\alpha,\beta}(y_i;\theta_{i\alpha},\theta_{i\beta}) = -\sum_{i=1}^{n} I_\alpha^f(y_i;\theta_{i\alpha})I_\beta^f(y_i;\theta_{i\beta}) = \\ &= -(\mathbf{v}_\alpha * \mathbf{v}_\alpha)^{\mathrm{T}}(\mathbf{v}_\beta * \mathbf{v}_\beta)\end{aligned} \tag{14}$$

can be assigned to the observation vector $\mathbf{y}$.

Heretofore, no probabilistic assumption was made, however, one can assign some families of probability distributions to the target function Eq. (10) and the global split potential Eq. (14).

It is assumed in the general theory of M_{split} estimation (Part I of the paper) that an observation y_i can be a realization of either of two random variables Y_α or Y_β with the probability distributions belonging to the following families $\mathcal{P}_\alpha = \{P_{\theta_\alpha} \succ f_\alpha(y;\theta_\alpha) : \theta_\alpha \in \Theta_\alpha\}$ and $\mathcal{P}_\beta = \{P_{\theta_\beta} \succ f_\beta(y;\theta_\beta) : \theta_\beta \in \Theta_\beta\}$, respectively. Let these families be families of normal distributions with two variants of the expected value $\theta_\alpha = E(y)$ and $\theta_\beta = E(y)$, respectively, and with the same standard deviation σ, i.e. let:

$$\mathcal{P}_\alpha = \{P_{\theta_\alpha} \succ f(y;\theta_\alpha) : \theta_\alpha \in \Theta_\alpha\} = \{N[\theta_\alpha,\sigma]\} \tag{15}$$

$$\mathcal{P}_\beta = \{P_{\theta_\beta} \succ f(y;\theta_\beta) : \theta_\beta \in \Theta_\beta\} = \{N[\theta_\beta,\sigma]\} \tag{16}$$

where $f(y;\theta) = (\sigma\sqrt{2\pi})^{-1}\exp\{-\frac{1}{2}\sigma^{-2}(y-\theta)^2\}$. Thus if

$$y_i \sim P_{\theta_\alpha,\theta_\beta} \in \{\mathcal{P}_\alpha,\mathcal{P}_\beta\} = \{N[\theta_{i\alpha},\sigma_i], N[\theta_{i\beta},\sigma_i]\}$$

then the following quantities:

$$I_\alpha^f(y_i;\theta_{i\alpha}) = -\ln c_1 f(y_i;\theta_{i\alpha}) = \frac{1}{2}\sigma_i^{-2}(y_i-\theta_{i\alpha})^2 = \frac{1}{2}v_{i\alpha}^2 \tag{17}$$

$$I_\alpha^f(y_i;\theta_{i\beta}) = -\ln c_2 f(y_i;\theta_{i\beta}) = \frac{1}{2}\sigma_i^{-2}(y_i-\theta_{i\beta})^2 = \frac{1}{2}v_{i\beta}^2 \tag{18}$$

are two competitive f-informations assigned to the observation y_i ($c_1 = c_2 = \sigma\sqrt{2\pi}$). Therefore, the logarithmic elementary split potential can be written as follows

$$\begin{aligned} k_{\alpha,\beta}(y_i;\theta_{i\alpha},\theta_{i\beta}) &= -I_\alpha^f(y_i;\theta_{i\alpha})I_\beta^f(y_i;\theta_{i\beta}) = \\ &= -\frac{1}{4}\sigma_i^{-4}(y_i-\theta_{i\alpha})^2(y_i-\theta_{i\beta})^2 = -\frac{1}{4}v_{i\alpha}^2 v_{i\beta}^2 \end{aligned} \tag{19}$$

The forms of f-informations as well as the logarithmic elementary split potential presented above differ from those in Eqs. (11)-(13) only in coefficients. Values of these coefficients do not influence the solution of the optimization problem and final values of M_{split} estimates.

To verify such equivalence also for an observation vector $\mathbf{y} = [y_1, y_2, ..., y_n]^{\mathrm{T}}$, let it be a realization of either of two random variables $\mathbf{Y}_\alpha \sim P_{\boldsymbol{\theta}_\alpha}$ or $\mathbf{Y}_\beta \sim P_{\boldsymbol{\theta}_\beta}$. The probability distributions $P_{\boldsymbol{\theta}_\alpha}$ and $P_{\boldsymbol{\theta}_\beta}$ belong to the following respective families of normal distributions:

$$\begin{aligned} \mathcal{P}_\alpha &= \{P_{\boldsymbol{\theta}_\alpha} \succ f(\mathbf{y},\boldsymbol{\theta}_\alpha)\} = \{N[\boldsymbol{\theta}_\alpha,\mathbf{C}]\} \\ \mathcal{P}_\beta &= \{P_{\boldsymbol{\theta}_\beta} \succ f(\mathbf{y},\boldsymbol{\theta}_\beta)\} = \{N[\boldsymbol{\theta}_\beta,\mathbf{C}]\} \end{aligned}$$

with density functions: $f(\mathbf{y};\boldsymbol{\theta}_\alpha)=\prod_{i=1}^{n} f(y_i;\theta_{i\alpha})$, $f(\mathbf{y};\boldsymbol{\theta}_\beta)=\prod_{i=1}^{n} f(y_i;\theta_{i\beta})$. These families share the covariance matrix $\mathbf{C}=\operatorname{diag}(\sigma_1^2,\sigma_2^2,\ldots,\sigma_n^2)$ and differ from each other in the expected values:

$$
\begin{aligned}
E(\mathbf{Y}_\alpha)&=[E(Y_{1\alpha}),E(Y_{2\alpha}),\cdots,E(Y_{n\alpha})]^{\mathrm{T}}=[\theta_{1\alpha},\theta_{2\alpha},\cdots,\theta_{n\alpha}]^{\mathrm{T}}=\boldsymbol{\theta}_\alpha\\
E(\mathbf{Y}_\beta)&=[E(Y_{1\beta}),E(Y_{2\beta}),\cdots,E(Y_{n\beta})]^{\mathrm{T}}=[\theta_{1\beta},\theta_{2\beta},\cdots,\theta_{n\beta}]^{\mathrm{T}}=\boldsymbol{\theta}_\beta
\end{aligned}
$$

Thus if

$$
\mathbf{y}\sim P_{\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta}\in\{\mathcal{P}_\alpha,\mathcal{P}_\beta\}=\{N[\boldsymbol{\theta}_\alpha,\mathbf{C}_y],N[\boldsymbol{\theta}_\beta,\mathbf{C}_y\}
$$

then for $c_{1i}=c_{2i}=\sigma_i\sqrt{2\pi}$ one can write

$$
\begin{aligned}
k_{\alpha,\beta}(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta)&=-\sum_{i=1}^{n}k_{\alpha,\beta}(y_i;\theta_{i\alpha},\theta_{i\beta})=-\sum_{i=1}^{n}\ln c_{1i}f(y_i;\theta_{i\alpha})\ln c_{2i}f(y_i;\theta_{i\beta})=\\
&=-\frac{1}{4}\sum_{i=1}^{n}v_{i\alpha}^2v_{i\beta}^2=-\frac{1}{4}(\mathbf{v}_\alpha*\mathbf{v}_\alpha)^{\mathrm{T}}(\mathbf{v}_\beta*\mathbf{v}_\beta)
\end{aligned}
\tag{20}
$$

2.2. *Weight functions*

Weight functions (e.g. Kadaj, 1988; Yang, 1991) and also the influence curve IC (e.g. Hampel et al., 1986) play a significant role in designing of estimators and afterwards in their theoretical analyses. Let us remind that for the following target function

$$
\varphi(\mathbf{y};\boldsymbol{\theta})=\sum_{i=1}^{n}\rho(y_i;\theta_i)=\sum_{i=1}^{n}\rho(y_i-\theta_i)=\sum_{i=1}^{n}\rho(v_i)
\tag{21}
$$

where $v_i=y_i-\theta_i$ and the weight function is defined as $w(v)=\partial\rho(v)/\partial(v^2)$. The target function of M_{split} estimation is written in the form

$$
\begin{aligned}
\varphi(\mathbf{y};\boldsymbol{\theta}_\alpha,\boldsymbol{\theta}_\beta)&=\sum_{i=1}^{n}\eta(y_i;\theta_{i\alpha},\theta_{i\beta})=\sum_{i=1}^{n}\rho_\alpha(y_i;\theta_{i\alpha})\rho_\beta(y_i;\theta_{i\beta})=\\
&=\sum_{i=1}^{n}\eta(v_{i\alpha},v_{i\beta})=\sum_{i=1}^{n}\rho_\alpha(v_{i\alpha})\rho_\beta(v_{i\beta})
\end{aligned}
\tag{22}
$$

where $v_{i\alpha}=y_i-\theta_{i\alpha}$, $v_{i\beta}=y_i-\theta_{i\beta}$. Therefore, two weight functions can be derived, the first one in relation to v_α

$$w_\alpha(v_\alpha, v_\beta) = \frac{\partial}{\partial(v_\alpha^2)}\eta(v_\alpha, v_\beta) = \rho_\beta(v_\beta)\frac{\partial\rho_\alpha(v_\alpha)}{\partial(v_\alpha^2)} = \rho_\beta(v_\beta)w_\alpha(v_\alpha) \tag{23}$$

and the second one in relation to v_β

$$w_\beta(v_\alpha, v_\beta) = \frac{\partial}{\partial(v_\beta^2)}\eta(v_\alpha, v_\beta) = \rho_\alpha(v_\alpha)\frac{\partial\rho_\beta(v_\beta)}{\partial(v_\beta^2)} = \rho_\alpha(v_\alpha)w_\beta(v_\beta) \tag{24}$$

where

$$w_\alpha(v_\alpha) = \frac{\partial\rho_\alpha(v_\alpha)}{\partial(v_\alpha^2)}, \quad w_\beta(v_\beta) = \frac{\partial\rho_\beta(v_\beta)}{\partial(v_\beta^2)}$$

Weight functions Eqs. (23)-(24) play a special role in M_{split} estimation. Their forms point out that the competitive target functions $\rho_\alpha(v_\alpha)$ and $\rho_\beta(v_\beta)$ are mutually cross-weighted, i.e. the weight function in relation to v_α is actually equal to the competitive target function $\rho_\beta(v_\beta)$ that is strengthened or weakened by the "classic" weight function . The similar situation applies to the second weight function, in relation to v_β. This very important property of M_{split} estimators is especially distinct in case of the squared M_{split} estimation where

$$\rho_\alpha(v_\alpha) = v_\alpha^2, \quad \rho_\beta(v_\beta) = v_\beta^2$$

and futhermore

$$\eta(v_\alpha, v_\beta) = \rho_\alpha(v_\alpha)\rho_\beta(v_\beta) = v_\alpha^2 v_\beta^2$$

Since

$$w_\alpha(v_\alpha) = \frac{\partial\rho_\alpha(v_\alpha)}{\partial(v_\alpha^2)} = \frac{\partial v_\alpha^2}{\partial(v_\alpha^2)} = 1 \quad \text{and} \quad w_\beta(v_\beta) = \frac{\partial\rho_\beta(v_\beta)}{\partial(v_\beta^2)} = \frac{\partial v_\beta^2}{\partial(v_\beta^2)} = 1$$

therefore the weight functions of the squared M_{split} estimation can be written as follows:

$$w_\alpha(v_\alpha, v_\beta) = \rho_\beta(v_\beta)w_\alpha(v_\alpha) = \rho_\beta(v_\beta) = v_\beta^2 \tag{25}$$

$$w_\beta(v_\alpha, v_\beta) = \rho_\alpha(v_\alpha)w_\beta(v_\beta) = \rho_\alpha(v_\alpha) = v_\alpha^2 \tag{26}$$

The functions Eq. (25) and Eq. (26) will be denoted, respectively, as

$$w_\alpha(v_\alpha, v_\beta) = w_\alpha(v_\beta) = v_\beta^2 \tag{27}$$

$$w_\beta(v_\alpha, v_\beta) = w_\beta(v_\alpha) = v_\alpha^2 \tag{28}$$

Thus if two competitive functional models $v_\alpha = y - \theta_\alpha$ and $v_\beta = y - \theta_\beta$ result from the $split(v = y - \theta)$, then the two following relations:

$$\min_{\theta_\alpha}\{\rho_\alpha(v_\alpha) = v_\alpha^2\} \quad \Leftrightarrow \quad \sup_{\theta_\alpha \in \delta_\theta}\{w_\beta(v_\beta, v_\beta) = w_\beta(v_\alpha) = v_\beta^2\}$$

$$\min_{\theta_\beta}\{\rho_\beta(v_\beta) = v_\beta^2\} \quad \Leftrightarrow \quad \sup_{\theta_\alpha \in \delta_\theta}\{w_\alpha(v_\alpha, v_\beta) = w_\alpha(v_\beta) = v_\alpha^2\}$$

become very important for searching competitive estimates θ_α and θ_β (see, also Fig. 1)

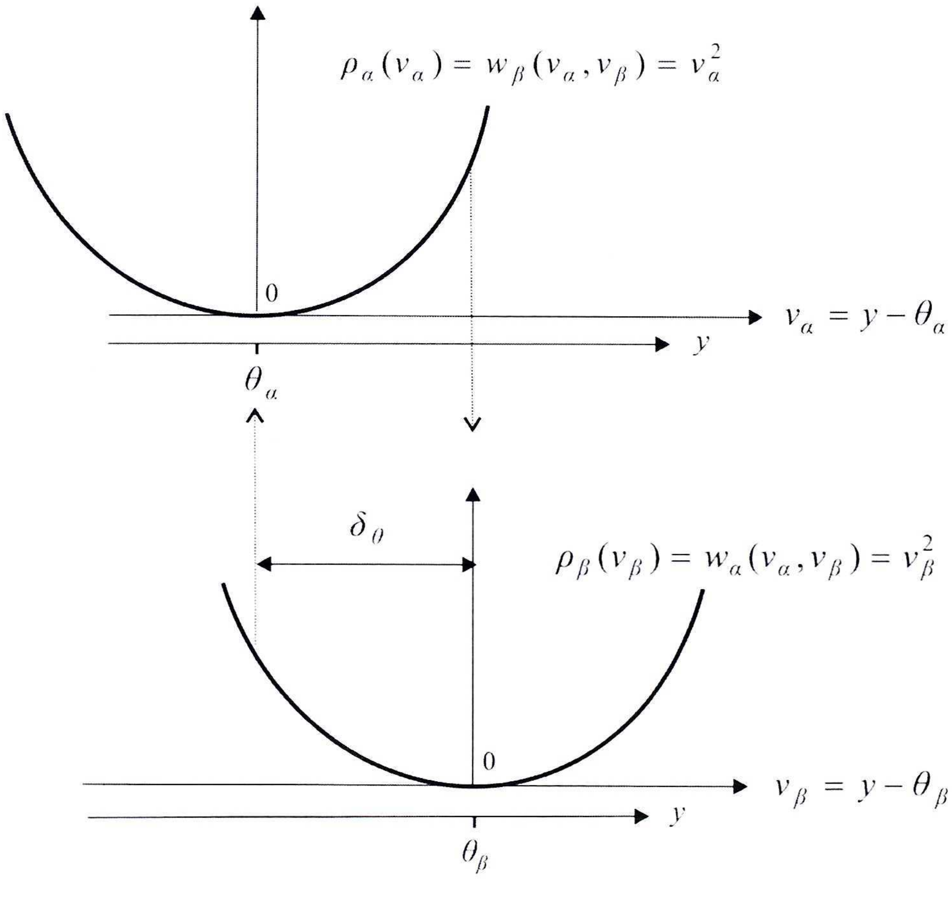


Fig. 1. Weight functions and target functions of the squared M_{split} estimation

2.3. Optimization problem and its solution

Let the squared M_{split} estimation is applied to estimate parameters $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$ from split functional model of geodetic observations: $\mathbf{v}_\alpha = \mathbf{y} - \mathbf{A}\mathbf{X}_\alpha$, $\mathbf{v}_\beta = \mathbf{y} - \mathbf{A}\mathbf{X}_\beta$. Then the target function can be written in the following form

$$\varphi(\mathbf{y};\mathbf{X}_\alpha,\mathbf{X}_\beta)=\sum_{i=1}^{n} v_{i\alpha}^2 v_{i\beta}^2=(\mathbf{v}_\alpha * \mathbf{v}_\alpha)^{\mathrm{T}}(\mathbf{v}_\beta * \mathbf{v}_\beta) \tag{29}$$

In the part I of the present paper it was shown that if the functions $\rho_\alpha(y_i;\ \mathbf{X}_\alpha)$ and $\rho_\beta(y_i;\ \mathbf{X}_\beta)$ are convex and twice differentiable, then to solve the optimization problem

$$\min_{\mathbf{X}_\alpha,\mathbf{X}_\beta}\varphi(\mathbf{y};\mathbf{X}_\alpha,\mathbf{X}_\beta)=\varphi(\mathbf{y};\hat{\mathbf{X}}_\alpha,\hat{\mathbf{X}}_\beta)$$

one can apply the Newton method. Therefore, the following iterative process can be used (Teunissen, 1990; Wiśniewski, 2008, 2009)

$$\left.\begin{array}{l}\mathbf{X}_\alpha^j=\mathbf{X}_\alpha^{j-1}+d\mathbf{X}_\alpha^{j-1}=\mathbf{X}_\alpha^{j-1}-\{\mathbf{H}_\alpha(\mathbf{X}_\alpha^{j-1},\mathbf{X}_\beta^{j-1})\}^{-1}\mathbf{g}_\alpha(\mathbf{X}_\alpha^{j-1},\mathbf{X}_\beta^{j-1})\\ \mathbf{X}_\beta^j=\mathbf{X}_\beta^{j-1}+d\mathbf{X}_\beta^{j-1}=\mathbf{X}_\beta^{j-1}-\{\mathbf{H}_\beta(\mathbf{X}_\alpha^{j},\mathbf{X}_\beta^{j-1})\}^{-1}\mathbf{g}_\beta(\mathbf{X}_\alpha^{j},\mathbf{X}_\beta^{j-1})\end{array}\right\} \tag{30}$$

Let us remind that the necessary gradients $\mathbf{g}_\alpha(\mathbf{X}_\alpha,\ \mathbf{X})$, and $\mathbf{g}_\beta(\mathbf{X}_\alpha,\ \mathbf{X}_\beta)$ has the following general forms (in case of M_{split} estimation, see Part I, Eqs. (38) and (39)):

$$\mathbf{g}_\alpha(\mathbf{X}_\alpha,\mathbf{X}_\beta)=\frac{\partial}{\partial\mathbf{X}_\alpha}\varphi(\mathbf{y};\mathbf{X}_\alpha,\mathbf{X}_\beta)=-\mathbf{A}^{\mathrm{T}}\mathrm{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v}_\beta)\}\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha)$$

$$\mathbf{g}_\beta(\mathbf{X}_\alpha,\mathbf{X}_\beta)=\frac{\partial}{\partial\mathbf{X}_\beta}\varphi(\mathbf{y};\mathbf{X}_\alpha,\mathbf{X}_\beta)=-\mathbf{A}^{\mathrm{T}}\mathrm{diag}\{\boldsymbol{\rho}_\alpha(\mathbf{v}_\alpha)\}\mathbf{g}_{M\beta}(\mathbf{v}_\beta)$$

where

$$\mathrm{diag}\{\boldsymbol{\rho}_\alpha(\mathbf{v})\}=\mathrm{diag}\{\rho_\alpha(v_{1\alpha}),\rho_\alpha(v_{2\alpha}),\cdots,\rho_\alpha(v_{n\alpha})\}$$

$$\mathrm{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v})\}=\mathrm{diag}\{\rho_\beta(v_{1\beta}),\rho_\beta(v_{2\beta}),\cdots,\rho_\beta(v_{n\beta})\}$$

and

$$\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha)=[\frac{\partial\rho_\alpha(v_{1\alpha})}{\partial v_{1\alpha}},\frac{\partial\rho_\alpha(v_{2\alpha})}{\partial v_{2\alpha}},\cdots,\frac{\partial\rho_\alpha(v_{n\alpha})}{\partial v_{n\alpha}}]^{\mathrm{T}}$$

$$\mathbf{g}_{M\beta}(\mathbf{v}_\beta)=[\frac{\partial\rho_\beta(v_{1\beta})}{\partial v_{1\beta}},\frac{\partial\rho_\beta(v_{2\beta})}{\partial v_{2\beta}},\cdots,\frac{\partial\rho_\beta(v_{n\beta})}{\partial v_{n\beta}}]^{\mathrm{T}}$$

Since, in the squared M_{split} estimation

$$\rho_\alpha(y_i;\mathbf{X}_\alpha)=\rho_\alpha(v_{i\alpha})=v_{i\alpha}^2$$

$$\rho_\beta(y_i;\mathbf{X}_\beta)=\rho_\beta(v_{i\beta})=v_{i\beta}^2$$

then

$$\mathrm{diag}\{\boldsymbol{\rho}_\alpha(\mathbf{v})\}=\mathrm{diag}\{v_{1\alpha}^2,v_{2\alpha}^2,\cdots,v_{n\alpha}^2\}$$

$$\mathrm{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v})\}=\mathrm{diag}\{v_{1\beta}^2,v_{2\beta}^2,\cdots,v_{n\beta}^2\}$$

and

$$\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha) = 2[v_{1\alpha}, v_{2\alpha}, \cdots, v_{n\alpha}]^{\mathrm{T}}$$
$$\mathbf{g}_{M\beta}(\mathbf{v}_\beta) = 2[v_{1\beta}, v_{2\beta}, \cdots, v_{n\beta}]^{\mathrm{T}}$$

It can also be written that (on the basis of Eqs. (27)-(28))

$$\operatorname{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v}_\beta)\} = \operatorname{diag}\{v_{1\beta}^2, v_{2\beta}^2, \cdots, v_{n\beta}^2\} = \mathbf{w}_\alpha(\mathbf{v}_\beta) \tag{31}$$

and

$$\operatorname{diag}\{\boldsymbol{\rho}_\alpha(\mathbf{v}_\alpha)\} = \operatorname{diag}\{v_{1\alpha}^2, v_{2\alpha}^2, \cdots, v_{n\alpha}^2\} = \mathbf{w}_\beta(\mathbf{v}_\alpha) \tag{32}$$

For the squared functions $\rho_\alpha(v_\alpha)$ and $\rho_\beta(v_\beta)$, one can write

$$\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha) = 2\mathbf{v}_\alpha, \quad \mathbf{g}_{M\beta}(\mathbf{v}_\beta) = 2\mathbf{v}_\beta$$

In such case, i.e. the squared M_{split} estimation, gradients $\mathbf{g}_\alpha(\mathbf{X}_\alpha, \mathbf{X}_\beta)$ and $\mathbf{g}_\beta(\mathbf{X}_\alpha, \mathbf{X}_\beta)$ can be presented in the forms (see, also Wiśniewski, 2009):

$$\mathbf{g}_\alpha(\mathbf{X}_\alpha, \mathbf{X}_\beta) = -\mathbf{A}^{\mathrm{T}}\operatorname{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v}_\beta)\}\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha) = -2\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta)\mathbf{v}_\alpha \tag{33}$$

$$\mathbf{g}_\beta(\mathbf{X}_\alpha, \mathbf{X}_\beta) = -\mathbf{A}^{\mathrm{T}}\operatorname{diag}\{\boldsymbol{\rho}_\alpha(\mathbf{v}_\alpha)\}\mathbf{g}_{M\beta}(\mathbf{v}_\beta) = -2\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha)\mathbf{v}_\beta \tag{34}$$

Hessians of the general target function have the following forms (see Part I, Eqs. (41) and (42)):

$$\mathbf{H}_\alpha(\mathbf{X}_\alpha, \mathbf{X}_\beta) = \frac{\partial^2}{\partial\mathbf{X}_\alpha\partial\mathbf{X}_\alpha^{\mathrm{T}}}\varphi(\mathbf{y}; \mathbf{X}_\alpha, \mathbf{X}_\beta) = \mathbf{A}^{\mathrm{T}}\operatorname{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v}_\beta)\}\mathbf{H}_{M\alpha}(\mathbf{v}_\alpha)\mathbf{A}$$
$$\mathbf{H}_\beta(\mathbf{X}_\alpha, \mathbf{X}_\beta) = \frac{\partial^2}{\partial\mathbf{X}_\beta\partial\mathbf{X}_\beta^{\mathrm{T}}}\varphi(\mathbf{y}; \mathbf{X}_\alpha, \mathbf{X}_\beta) = \mathbf{A}^{\mathrm{T}}\operatorname{diag}\{\boldsymbol{\rho}_\alpha(\mathbf{v}_\alpha)\}\mathbf{H}_{M\beta}(\mathbf{v}_\beta)\mathbf{A}$$

where

$$\mathbf{H}_{M\alpha}(\mathbf{v}_\alpha) = \frac{\partial\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha)}{\partial\mathbf{v}_\alpha^{\mathrm{T}}}, \quad \mathbf{H}_{M\beta}(\mathbf{v}_\beta) = \frac{\partial\mathbf{g}_{M\beta}(\mathbf{v}_\beta)}{\partial\mathbf{v}_\beta^{\mathrm{T}}}$$

If considering the squared M_{split} estimation and the gradient forms $\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha) = 2\mathbf{v}_\alpha$ and $\mathbf{g}_{M\beta}(\mathbf{v}_\beta) = 2\mathbf{v}_\beta$, then $\mathbf{H}_{M\alpha}(\mathbf{v}_\alpha)$ and $\mathbf{H}_{M\beta}(\mathbf{v}_\beta)$ are as follows:

$$\mathbf{H}_{M\alpha}(\mathbf{v}_\alpha) = \frac{\partial\mathbf{g}_{M\alpha}(\mathbf{v}_\alpha)}{\partial\mathbf{v}_\alpha^{\mathrm{T}}} = \frac{\partial(2\mathbf{v}_\alpha)}{\partial\mathbf{v}_\alpha^{\mathrm{T}}} = 2\mathbf{I}$$
$$\mathbf{H}_{M\beta}(\mathbf{v}_\beta) = \frac{\partial\mathbf{g}_{M\beta}(\mathbf{v}_\beta)}{\partial\mathbf{v}_\beta^{\mathrm{T}}} = \frac{\partial(2\mathbf{v}_\beta)}{\partial\mathbf{v}_\beta^{\mathrm{T}}} = 2\mathbf{I}$$

where **I** is the identity matrix. Furthermore, considering relations Eq. (29) and Eq (30), Hessians have got the forms (see, also Wiśniewski, 2009):

$$\begin{aligned}\mathbf{H}_\alpha(\mathbf{X}_\alpha,\mathbf{X}_\beta) &= \mathbf{A}^{\mathrm{T}}\mathrm{diag}\{\boldsymbol{\rho}_\beta(\mathbf{v}_\beta)\}\mathbf{H}_{M\alpha}(\mathbf{v}_\alpha)\mathbf{A} = \\ &= 2\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta)\mathbf{A} = \mathbf{H}_\alpha(\mathbf{X}_\beta)\end{aligned} \tag{35}$$

$$\begin{aligned}\mathbf{H}_\beta(\mathbf{X}_\alpha,\mathbf{X}_\beta) &= \mathbf{A}^{\mathrm{T}}\mathrm{diag}\{\rho_\alpha(\mathbf{v}_\alpha)\}\mathbf{H}_{M\beta}(\mathbf{v}_\beta)\mathbf{A} = \\ &= 2\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha)\mathbf{A} = \mathbf{H}_\beta(\mathbf{X}_\alpha)\end{aligned} \tag{36}$$

By assuming the forms of the gradients and Hessians presented above, i.e. the squared M_{split} estimation case, the iterative formula Eq. (30) can be now written as follows ($j = 1, \ldots, k$)

$$\left.\begin{aligned}\mathbf{X}_\alpha^j &= \mathbf{X}_\alpha^{j-1} - \{\mathbf{H}_\alpha(\mathbf{X}_\alpha^{j-1},\mathbf{X}_\beta^{j-1})\}^{-1}\mathbf{g}_\alpha(\mathbf{X}_\alpha^{j-1},\mathbf{X}_\beta^{j-1}) = \\ &= \mathbf{X}_\alpha^{j-1} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^{j-1})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^{j-1})\,\mathbf{v}_\alpha^{j-1} \\[1em] \mathbf{X}_\beta^j &= \mathbf{X}_\beta^{j-1} - \{\mathbf{H}_\beta(\mathbf{X}_\alpha^{j},\mathbf{X}_\beta^{j-1})\}^{-1}\mathbf{g}_\beta(\mathbf{X}_\alpha^{j},\mathbf{X}_\beta^{j-1}) = \\ &= \mathbf{X}_\beta^{j-1} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^{j-1})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\beta^{j-1})\,\mathbf{v}_\beta^{j-1}\end{aligned}\right\} \tag{37}$$

where

$$\left.\begin{aligned}\mathbf{v}_\alpha^{j-1} &= \mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^{j-1} \\ \mathbf{v}_\beta^{j-1} &= \mathbf{y} - \mathbf{A}\mathbf{X}_\beta^{j-1}\end{aligned}\right\} \tag{38}$$

The squared M_{split} estimates of the parameters $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$ are such quantities $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$, respectively, that fulfill equations:

$$\hat{\mathbf{X}}_\alpha = \mathbf{X}_\alpha^k = \mathbf{X}_\alpha^{k-1}, \qquad \hat{\mathbf{X}}_\beta = \mathbf{X}_\beta^k = \mathbf{X}_\beta^{k-1}$$

and

$$\mathbf{g}_\alpha(\hat{\mathbf{X}}_\alpha,\hat{\mathbf{X}}_\beta) = -2\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\hat{\mathbf{v}}_\beta)\hat{\mathbf{v}}_\alpha = \mathbf{0}$$

$$\mathbf{g}_\beta(\hat{\mathbf{X}}_\alpha,\hat{\mathbf{X}}_\beta) = -2\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\hat{\mathbf{v}}_\alpha)\hat{\mathbf{v}}_\beta = \mathbf{0}$$

Initialization of the iterative process Eq. (30) may be a significant problem of M_{split} estimation. However, as for the squared M_{split} estimation with the iterative formula Eq. (37), the estimation process can be started with application of the *LS*-method estimates, i.e.

$$\left.\begin{aligned}\hat{\mathbf{X}}_{LS} &= (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y} \\ \hat{\mathbf{v}}_{LS} &= \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_{LS}\end{aligned}\right\} \tag{39}$$

Then

$$\mathbf{X}_\alpha^0 = \hat{\mathbf{X}}_{LS}, \quad \mathbf{v}_\alpha^0 = \hat{\mathbf{v}}_{LS}$$

and

$$\mathbf{X}_\beta^0 = \hat{\mathbf{X}}_{LS} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\hat{\mathbf{v}}_{LS})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\hat{\mathbf{v}}_{LS})\hat{\mathbf{v}}_{LS}, \quad \mathbf{v}_\beta^0 = \mathbf{y} - \mathbf{A}\mathbf{X}_\beta^0$$

where

$$\mathbf{w}_\beta(\hat{\mathbf{v}}_{LS}) = \mathrm{diag}(\hat{v}_{1LS}^2, \hat{v}_{2LS}^2, \cdots, \hat{v}_{nLS}^2)$$

Thus the next iterative step should result in

$$\mathbf{X}_\alpha^1 = \hat{\mathbf{X}}_{LS} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^0)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^0)\hat{\mathbf{v}}_{LS}, \quad \mathbf{v}_\alpha^1 = \mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^1$$
$$\mathbf{X}_\beta^1 = \mathbf{X}_\beta^0 + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^1)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^1)\mathbf{v}_\beta^0, \quad \mathbf{v}_\beta^1 = \mathbf{y} - \mathbf{A}\mathbf{X}_\beta^1$$

where

$$\mathbf{w}_\alpha(\mathbf{v}_\beta^0) = \mathrm{diag}\{(v_\alpha^0)^2, (v_\alpha^0)^2, \cdots, (v_\alpha^0)^2\}$$
$$\mathbf{w}_\beta(\mathbf{v}_\alpha^1) = \mathrm{diag}\{(v_\alpha^1)^2, (v_\alpha^1)^2, \cdots, (v_\alpha^1)^2\}$$

3. Examples

Example 1

In the part I of this paper, the idea of M_{split} estimation was illustrated with application of the following observation set

$$\{1.1, \ 1.3, \ 1.4\ 1.5, \ 1.7, \ 3.4, \ 3.5, \ 3.6\}$$

and with the functional model $v_i = y_i - \theta$, $i = 1, ..., 8$. It was also computed that $\hat{\theta}_{LS} = 2.19$ (the estimate of the LS-method) and $\hat{\theta}_R = 1.49$ (the robust estimate of the Danish method). Now, let the estimates of two competitive variants θ_α and θ_β of the parameter θ be computed by applying the squared M_{split} estimation:

Step 0

$$\mathbf{A} = [1_1, \cdots, 1_8]^{\mathrm{T}}, \quad \mathbf{y} = [1.1, \ 1.3, \ 1.4\ 1.5, \ 1.7, \ 3.4, \ 3.5, \ 3.6]^{\mathrm{T}}$$
$$\hat{\theta}_{LS} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y} = 2.19$$
$$\hat{\mathbf{v}}_{LS} = \mathbf{y} - \mathbf{A}\hat{\theta}_{LS} = [-1.09, \ -0.89, \ -.079, \ -0.69, \ -0.49, \ 1.21, \ 1.31, \ 1.41]^{\mathrm{T}}$$
$$\theta_\alpha^0 = \hat{\theta}_{LS}$$
$$\mathbf{v}_\alpha^0 = \mathbf{v}_{LS}$$

$$\mathbf{w}_{\beta}(\hat{\mathbf{v}}_{LS}) = \mathrm{diag}(\hat{v}^2_{1LS}, \hat{v}^2_{2LS}, \cdots, \hat{v}^2_{8LS}) =$$
$$= \mathrm{diag}(1.18, 0.79, 0.62, 0.47, 0.24, 1.47, 1.72, 2.00)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\hat{\mathbf{v}}_{LS})\mathbf{A} = 8.49, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\hat{\mathbf{v}}_{LS})\hat{\mathbf{v}}_{LS} = 3.95$$

$$\theta^0_{\beta} = \hat{\theta}_{LS} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\hat{\mathbf{v}}_{LS})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\hat{\mathbf{v}}_{LS})\hat{\mathbf{v}}_{LS} = 2.19 + 0.47 = \underline{2.65}$$
$$\mathbf{v}^0_{\beta} = \mathbf{y} - \mathbf{A}\theta^0_{\beta} = [-1.55, -1.35, -1.25, -1.15, -0.95, 0.75, 0.85, 0.95]^{\mathrm{T}}$$

Step 1

$$\mathbf{w}_{\alpha}(\mathbf{v}^0_{\beta}) = \mathrm{diag}\{(v^0_{1\beta})^2, (v^0_{2\beta})^2, \cdots, (v^0_{8\beta})^2\} =$$
$$= \mathrm{diag}(2.41, 1.83, 1.57, 1.33, 0.91, 0.56, 0.72, 0.90)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^0_{\beta})\mathbf{A} = 10.22, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^0_{\beta})\hat{\mathbf{v}}_{LS} = -3.95$$

$$\theta^1_{\alpha} = \hat{\theta}_{LS} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^0_{\beta})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^0_{\beta})\hat{\mathbf{v}}_{LS} = 2.19 - 0.39 = \underline{1.80}$$
$$\mathbf{v}^1_{\alpha} = \mathbf{y} - \mathbf{A}\theta^1_{\alpha} = [-0.70, -0.50, -0.40, -0.30, -0.10, 1.60, 1.70, 1.80]^{\mathrm{T}}$$

$$\mathbf{w}_{\beta}(\mathbf{v}^1_{\alpha}) = \mathrm{diag}\{(v^1_{1\alpha})^2, (v^1_{2\alpha})^2, \cdots, (v^1_{8\alpha})^2\} =$$
$$= \mathrm{diag}(0.49, 0.25, 0.16, 0.09, 0.01, 2.56, 2.89, 3.24)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\mathbf{v}^1_{\alpha})\mathbf{A} = 9.68, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\mathbf{v}^1_{\alpha})\mathbf{v}^0_{\beta} = 6.00$$

$$\theta^1_{\beta} = \theta^0_{\beta} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\mathbf{v}^1_{\alpha})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\beta}(\mathbf{v}^1_{\alpha})\mathbf{v}^0_{\beta} = 2.65 + 0.62 = \underline{3.27}$$
$$\mathbf{v}^1_{\beta} = \mathbf{y} - \mathbf{A}\theta^1_{\beta} = [-2.17, -1.97, -1.87, -1.77, -1.57, 0.13, 0.23, 0.33]^{\mathrm{T}}$$

Step 2

$$\mathbf{w}_{\alpha}(\mathbf{v}^1_{\beta}) = \mathrm{diag}\{(v^1_{1\beta})^2, (v^1_{2\beta})^2, \cdots, (v^1_{8\beta})^2\} =$$
$$= \mathrm{diag}(4.72, 3.89, 3.51, 3.14, 2.47, 0.02, 0.05, 0.11)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^1_{\beta})\mathbf{A} = 17.91, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^1_{\beta})\mathbf{v}^1_{\alpha} = -7.56$$

$$\theta^2_{\alpha} = \theta^1_{\alpha} + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^1_{\beta})\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_{\alpha}(\mathbf{v}^1_{\beta})\mathbf{v}^1_{\alpha} = 1.80 - 0.42 = \underline{1.38}$$
$$\mathbf{v}^2_{\alpha} = \mathbf{y} - \mathbf{A}\theta^2_{\alpha} = [-0.28, -0.08, 0.02, 0.12, 0.32, 2.02, 2.12, 2.22]^{\mathrm{T}}$$

$$\mathbf{w}_\beta(\mathbf{v}_\alpha^2) = \text{diag}\{(v_{1\alpha}^2)^2, (v_{2\alpha}^2)^2, \cdots, (v_{8\alpha}^2)^2\} =$$
$$= \text{diag}(0.08, 0.01, 0.00, 0.01, 0.10, 4.08, 4.50, 4.93)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^2)\mathbf{A} = 13.72, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^2)\mathbf{v}_\beta^1 = 2.79$$

$$\theta_\beta^2 = \theta_\beta^1 + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^2)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^2)\mathbf{v}_\beta^1 = 3.27 + 021 = \underline{3.48}$$
$$\mathbf{v}_\beta^2 = \mathbf{y} - \mathbf{A}\theta_\beta^2 = [-2.38, -2.18, -2.08, -1.98, -1.78, -0.08, 0.02, 0.12]^{\mathrm{T}}$$

Step 3

$$\mathbf{w}_\alpha(\mathbf{v}_\beta^2) = \text{diag}\{(v_{1\beta}^2)^2, (v_{2\beta}^2)^2, \cdots, (v_{8\beta}^2)^2\} =$$
$$= \text{diag}(5.64, 4.73, 4.31, 3.90, 3.15, 0.01, 0.00, 0.02)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^2)\mathbf{A} = 21.77, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^2)\mathbf{v}_\alpha^2 = -0.33$$

$$\theta_\alpha^3 = \theta_\alpha^2 + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^2)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^2)\mathbf{v}_\alpha^2 = 1.38 - 0.02 = \underline{1.36}$$
$$\mathbf{v}_\alpha^3 = \mathbf{y} - \mathbf{A}\theta_\alpha^3 = [-0.26, -0.06, 0.04, 0.14, 0.34, 2.04, 2.14, 2.24]^{\mathrm{T}}$$

$$\mathbf{w}_\beta(\mathbf{v}_\alpha^3) = \text{diag}\{(v_{1\alpha}^3)^2, (v_{2\alpha}^3)^2, \cdots, (v_{8\alpha}^3)^2\} =$$
$$= \text{diag}(0.07, 0.00, 0.00, 0.02, 0.11, 4.15, 4.56, 5.00)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^3)\mathbf{A} = 13.91, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^3)\mathbf{v}_\beta^2 = 0.00$$

$$\theta_\beta^3 = \theta_\beta^2 + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^3)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^3)\mathbf{v}_\beta^2 = 3.48 + 0.00 = \underline{3.48}$$
$$\mathbf{v}_\beta^3 = \mathbf{y} - \mathbf{A}\theta_\beta^3 = [-2.38, -2.18, -2.08, -1.98, -1.78, -0.08, 0.02, 0.12]^{\mathrm{T}}$$

Step 4

$$\mathbf{w}_\alpha(\mathbf{v}_\beta^3) = \text{diag}\{(v_{1\beta}^3)^2, (v_{2\beta}^3)^2, \cdots, (v_{8\beta}^3)^2\} =$$
$$= \text{diag}(5.65, 4.74, 4.31, 3.90, 3.15, 0.01, 0.00, 0.02)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^3)\mathbf{A} = 21.77, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^4)\mathbf{v}_\beta^3 = 0.00$$

$$\theta_\alpha^4 = \theta_\alpha^3 + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^3)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^3)\mathbf{v}_\alpha^3 = 1.36 + 0.00 = \underline{1.36}$$
$$\mathbf{v}_\alpha^4 = \mathbf{y} - \mathbf{A}\theta_\alpha^3 = [-0.26, -0.06, 0.04, 0.14, 0.34, 2.04, 2.14, 2.24]^{\mathrm{T}}$$

$$\mathbf{w}_\beta(\mathbf{v}_\alpha^4) = \mathbf{w}_\beta(\mathbf{v}_\alpha^3) = \text{diag}(0.07, 0.00, 0.00, 0.02, 0.11, 4.15, 4.56, 5.00)$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^4)\mathbf{A} = 13.91, \quad \mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^4)\mathbf{v}_\beta^3 = 0.00$$

$$\theta_\beta^4 = \theta_\beta^3 + \{\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^4)\mathbf{A}\}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^4)\mathbf{v}_\beta^3 = 3.48 + 0.00 = \underline{3.48}$$
$$\mathbf{v}_\beta^4 = \mathbf{y} - \mathbf{A}\theta_\beta^4 = [-2.28, -2.18, -2.08, -1.98, -1.78, -0.08, 0.02, 0.12]^{\mathrm{T}}$$

Since $\theta_\alpha^4 = \theta_\alpha^3 = 1.36$ and $\theta_\beta^4 = \theta_\beta^3 = 3.48$ and additionally

$$g_\alpha(\theta_\alpha^3, \theta_\beta^3) = -\mathbf{A}^{\mathrm{T}}\mathbf{w}_\alpha(\mathbf{v}_\beta^3)\mathbf{v}_\alpha^3 = 0$$

$$g_\beta(\theta_\alpha^4, \theta_\beta^3) = -\mathbf{A}^{\mathrm{T}}\mathbf{w}_\beta(\mathbf{v}_\alpha^4)\mathbf{v}_\beta^3 = 0$$

then $\hat{\theta}_\alpha = 1.36$ and $\hat{\theta}_\beta = 3.48$ are the squared M_{split} estimates of the parameters θ_α and θ_β, respectively. The following residual vectors are assigned to the respective parameter estimates:

$$\hat{\mathbf{v}}_\alpha = [-0.26, -0.06, 0.04, 0.14, 0.34, 2.04, 2.14, 2.24]^{\mathrm{T}}$$

$$\hat{\mathbf{v}}_\beta = [-2.38, -2.18, -2.08, -1.98, -1.78, -0.08, 0.02, 0.12]^{\mathrm{T}}$$

The positions of the parameter estimates among all observations is presented in Figure 2

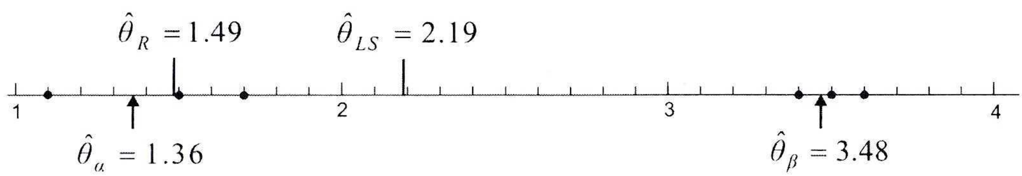


Fig. 2. Positions of the estimates $\hat{\theta}_\alpha$ and $\hat{\theta}_\beta$ among the observations (in comparison with the robust estimate $\hat{\theta}_R$ and the LS-estimate)

Example 2

The part I of this paper presented also the other observation set {1.1, 1.3, 1.4, 1.5, 1.7, 3.7} with only one quantity "3.7" that does not "suit" the others. Thus the following LS-estimates can be computed:

$$\hat{\theta}_{LS} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y} = 1.78$$
$$\hat{\mathbf{v}}_{LS} = \mathbf{y} - \mathbf{A}\hat{\theta}_{LS} = [-0.68, -0.48, -0.38, -0.28, -0.08, 1.92]^{\mathrm{T}}$$

where

$$\mathbf{A} = [1_1, \cdots, 1_6]^{\mathrm{T}}, \quad \mathbf{y} = [1.1, 1.3, 1.4, 1.5, 1.7, 3.7]^{\mathrm{T}}$$

Table 1 presents the iterative process of the squared M_{split} estimation resulting in the estimates $\hat{\theta}_\alpha$ and $\hat{\theta}_\beta$.

Table 1. The course of the iterative process (Example 2)

	Steps			
	0	*1*	*2*	*3*
$H_\alpha(\theta_\alpha^{j-1})$		16.65	24.79	24.82
$g_\alpha(\theta_\alpha^{j-1}, \theta_\beta^{j-1})$		6.53	0.65	0.00
$d\theta_\alpha^{j-1}$		–0.39	–0.03	0.00
θ_α^j	$\hat{\theta}_{LS} = 1.78$	1.39	1.36	$\underline{1.36}$
$H(\theta_\beta^{j-1})$	4.61	5.53	5.66	5.66
$g_\beta(\theta_\alpha^{j-1}, \theta_\beta^{j-1})$	–6.53	-2.31	–0.01	0.00
$d\theta_\beta^{j-1}$	1.42	0.42	0.00	0.00
θ_β^j	3.2	3.62	3.62	$\underline{3.62}$

Since

$$\theta_\alpha^3 = \theta_\alpha^2 = 1.36, \quad \theta_\beta^3 = \theta_\beta^2 = 3.62$$
$$g_\alpha(\theta_\alpha^2, \theta_\beta^2) = 0, \quad g_\beta(\theta_\alpha^3, \theta_\beta^2) = 0$$

then $\hat{\theta}_\alpha = 1.36$ and $\hat{\theta}_\beta = 3.62$ (Fig. 3).

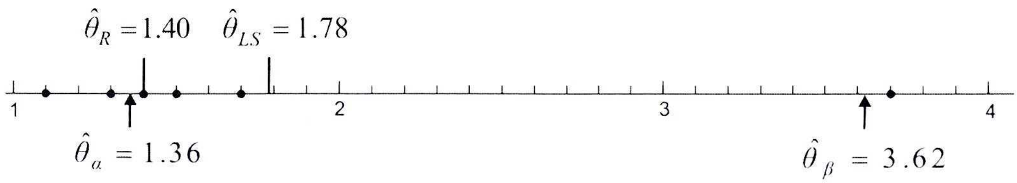


Fig. 3. Positions of the estimates $\hat{\theta}_\alpha$ and $\hat{\theta}_\beta$ among the observations (in comparison with the robust estimate $\hat{\theta}_R$ and the *LS*-estimate)

Here, the residual vectors $\hat{\mathbf{v}}_\alpha$ and $\hat{\mathbf{v}}_\beta$ are as follows, respectively:

$$\hat{\mathbf{v}}_\alpha = [-0.26, -0.06, 0.04, 0.14, 0.34, 2.34]^{\mathrm{T}}$$
$$\hat{\mathbf{v}}_\beta = [-2.52, -2.32, -2.22, -2.12, -1.92, 0.08]^{\mathrm{T}}$$

Example 3

M_{split} estimation can also be applied to split a set of points that is known to be generated by two functions belonging to the same family (it is not necessary to know detailed properties of such family on this stage of the method development).

To illustrate the idea presented above, let us assume the following set of points (x, y)

$$\{(4,4),(5,5),(6,6),(8,8),(10,10),(5,2),(8,4),(11,6),(14,8),(17,10)\}$$

Let the set be a random realization of the following function

$$y = ax + b \tag{40}$$

To compute the LS-estimates of the parameters a and b, the regression equation is written

$$v_i = y_i - (ax_i + b) \overset{i=1,\dots,10}{\longrightarrow} \mathbf{v} = \mathbf{y} - \mathbf{AX} \tag{41}$$

where $\mathbf{y} = [4,5,6,8,10,2,4,6,8,10]^{\mathrm{T}}$, and

$$\mathbf{A} = \begin{bmatrix} 4 & 5 & 6 & 8 & 10 & 5 & 8 & 11 & 14 & 17 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{X} = \begin{bmatrix} a \\ b \end{bmatrix}$$

Obviously, the LS-estimates can be computed as $\hat{\mathbf{X}}_{LS} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y}$ and $\hat{\mathbf{v}}_{LS} = \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_{LS}$. Thus one can obtain (Fig. 4a):

$$\hat{\mathbf{X}}_{LS} = [\hat{a}_{LS}, \hat{b}_{LS}]^{\mathrm{T}} = [0.48, 2.13]^{\mathrm{T}}$$
$$\hat{\mathbf{v}}_{LS} = [-0.02, 0.50, 1.03, 2.08, 3.13, -2.50, -1.92, -1.34, -0.76, -0.19]^{\mathrm{T}}$$

Actually, the point set is a realization of two functions that belong to the family described by Eq. (40) (to simplify the result interpretation, the realization is assumed to be free of random disturbances).Therefore, one can assume two competitive parameter vectors:

$$\mathbf{X}_\alpha = [a_\alpha, b_\alpha]^{\mathrm{T}}, \quad \mathbf{X}_\beta = [a_\beta, b_\beta]^{\mathrm{T}}$$

According to the M_{split} estimation principles, the set is split into two subsets that are assigned to the two following competitive regression equations:

$$v_{i\alpha} = y_i - (a_\alpha x_i + b_\alpha) \overset{i=1,\dots,10}{\longrightarrow} \mathbf{v}_\alpha = \mathbf{y} - \mathbf{AX}_\alpha$$
$$v_{i\beta} = y_i - (a_\beta x_i + b_\beta) \overset{i=1,\dots,10}{\longrightarrow} \mathbf{v}_\beta = \mathbf{y} - \mathbf{AX}_\beta$$

By applying the iterative process (37), the following M_{split} estimates are obtained:

$$\hat{\mathbf{X}}_\alpha = [\hat{a}_\alpha, \hat{b}_\alpha] = [0.67, -1.33]^{\mathrm{T}}, \quad \hat{\mathbf{X}}_\beta = [\hat{a}_\beta, \hat{b}_\beta] = [1.00, 0.00]^{\mathrm{T}}$$

and, respectively

$$\hat{\mathbf{v}}_\alpha = [2.67, 3.00, 3.33, 4.00, 4.67, 0.00, 0.00, 0.00, 0.00, 0.00]^{\mathrm{T}}$$
$$\hat{\mathbf{v}}_\beta = [0.00, 0.00, 0.00, 0.00, 0.00, -3.00, -4.00, -5.00, -6.00, -7.00]^{\mathrm{T}}$$

The proper iterative process is shown in Table 2 and is illustrated in Figure 4.

Table 2. The course of the iterative process (Example 3)

Steps	$\mathbf{g}_\alpha(\mathbf{X}_\alpha^{j-1}, \mathbf{X}_\beta^{j-1})$	$d\mathbf{X}_\alpha^j$	$\mathbf{X}_\alpha^j$	$\mathbf{g}_\beta(\mathbf{X}_\alpha^{j-1}, \mathbf{X}_\beta^{j-1})$	$d\mathbf{X}_\beta^j$	$\mathbf{X}_\beta^j$
0			$\hat{\mathbf{X}}_{LS} = \begin{bmatrix} 0.474 \\ 2.129 \end{bmatrix}$	–218.323 –15.334	0.654 –4.851	1.128 –2.722
1	553.362 39.503	–0.020	0.454 2.013	–138.158 –14.851	0.124 –0.507	1.252 –3.230
2	168.645 20.754	0.053 –0.870	0.507 1.144	–149.315 –20.160	–0.128 1.674	1.124 –1.556
3	209.435 29.445	0.114 –1.777	0.621 –0.634	–206.214 –30.602	–0.122 1.514	1.002 –0.042
4	80.342 12.523	0.046 –0.699	0.667 –1.333	–12.136 –1.708	–0.002 0.042	1.000 0.000
5	0.086 0.014	0.000 –0.001	0.667 –1.333	0.000 0.000	0.000 0.000	0.000 0.000
6	0.000 0.000	0.000 0.000	$\hat{a}_\alpha = \mathbf{0.667}$ $\hat{b}_\alpha = \mathbf{-1.333}$	0.000 0.000	0.000 0.000	$\hat{a}_\beta = \mathbf{1.000}$ $\hat{b}_\beta = \mathbf{0.000}$

Example 4

The method proposed in the present paper can also be applied to geodetic networks that are measured two times, e.g. to find out point displacements. However, the example presented here should be regarded just as an illustration of M_{split} estimation properties. The practical application of M_{split} estimation to such problems needs more careful theoretical as well as empirical analyses. It would also be necessary to compare the method to other well known methods of displacement estimation.

The example presented here refers to Example 4, section 4 in paper (Wiśniewski, 2009). This time, new variants of observation sets and all iterative processes that result in solutions of the M_{split} estimation optimization problem is presented.

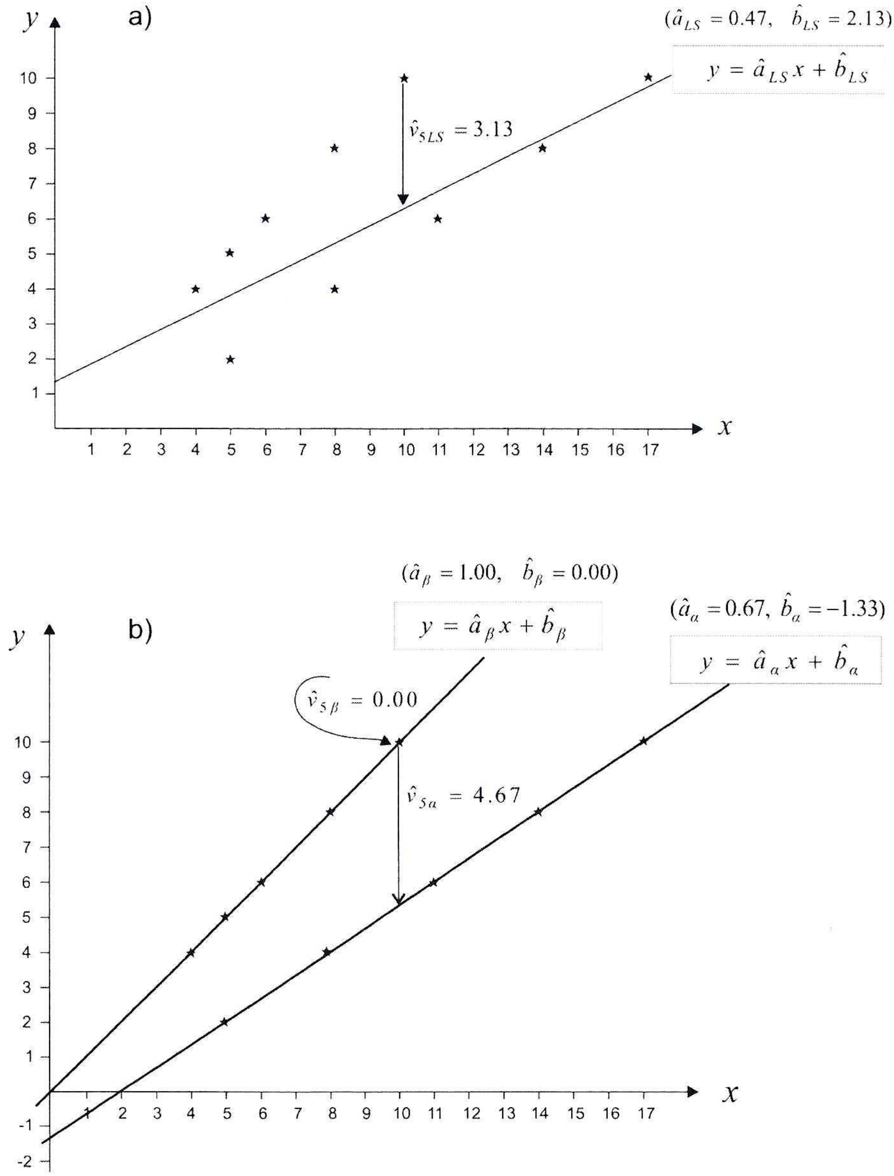
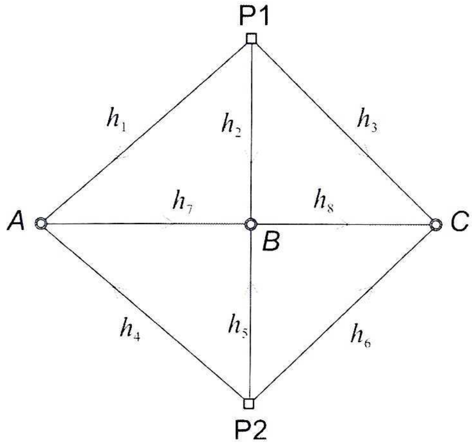


Fig. 4. Graphical interpretation of a set split

The levelling network that we are interested in is presented in Figure 5. Two network points P1 and P2 are the fixed ones with heights $H_{\mathrm{P1}} = 0.00$ and $H_{\mathrm{P2}} = 0.00$, respectively. There are also three unknown points *A*, *B* and *C*. Figure 5 presents also vectors $\mathbf{y}$ and $\mathbf{X}$, and the matrix $\mathbf{A}$ that are assigned to the functional model $\mathbf{v} = \mathbf{y} - \mathbf{AX}$.



$$\mathbf{y} = \begin{bmatrix} h_1 + H_{\mathrm{P1}} \\ h_2 + H_{\mathrm{P1}} \\ h_3 + H_{\mathrm{P1}} \\ h_4 + H_{\mathrm{P2}} \\ h_5 + H_{\mathrm{P2}} \\ h_6 + H_{\mathrm{P2}} \\ h_7 \\ h_8 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Fig. 5. Tested levelling network

Let the height differences h_i, $i = 1, \ldots, 8$, be measured two times at two different epochs α and β and let the following observation vector contains all the observation results (for $H_{\mathrm{P1}} = 0.00$ and $H_{\mathrm{P2}} = 0.00$)

$$\mathbf{y} = [h_{11}, h_{12} \vdots h_{21}, h_{22} \vdots h_{31}, h_{32} \vdots h_{41}, h_{42} \vdots h_{51}, h_{52} \vdots h_{61}, h_{62} \vdots h_{71}, h_{72} \vdots h_{81}, h_{82}]^{\mathrm{T}}$$

The two competitive functional models are assigned to such observation vector:

$$\mathbf{v}_\alpha = \mathbf{y} - \mathbf{AX}_\alpha, \quad \mathbf{v}_\beta = \mathbf{y} - \mathbf{AX}_\beta$$

where

$$\mathbf{v}_\alpha = \begin{bmatrix} v_{1,1,\alpha} \\ v_{1,2,\alpha} \\ \vdots \\ v_{8,1,\alpha} \\ v_{8,2,\alpha} \end{bmatrix}, \quad \mathbf{X}_\alpha = \begin{bmatrix} H_{A\alpha} \\ H_{B\alpha} \\ H_{C\alpha} \end{bmatrix}, \quad \mathbf{v}_\beta = \begin{bmatrix} v_{1,1,\beta} \\ v_{1,2,\beta} \\ \vdots \\ v_{8,1,\beta} \\ v_{8,2,\beta} \end{bmatrix}, \quad \mathbf{X}_\beta = \begin{bmatrix} H_{A\beta} \\ H_{B\beta} \\ H_{C\beta} \end{bmatrix}$$

The matrix $\mathbf{A}$ is shared by both models and has the following form (Wiśniewski, 2009)

$$\begin{array}{cccccccccccccccc} h_{1.1} & h_{1.2} & h_{2.1} & h_{2.2} & h_{3.1} & h_{3.2} & h_{4.1} & h_{4.2} & h_{5.1} & h_{5.2} & h_{6.1} & h_{6.2} & h_{7.1} & h_{7.2} & h_{8.1} & h_{8.2} \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{array}$$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{array}{l} \leftarrow H_{A\alpha},\ H_{A\beta} \\ \leftarrow H_{B\alpha},\ H_{B\beta} \\ \leftarrow H_{C\alpha},\ H_{C\beta} \end{array}$$

The observation vector $\mathbf{y}$ is now simulated under assumption of theoretical heights of the points A, B and C at the epochs α and β. Let $H_{A\alpha} = 1.0$, $H_{B\alpha} = 1.0$, $H_{c\alpha} = 1.0$ at the epoch α, i.e. let $\mathbf{X}_\alpha = [1.0,\ 1.0,\ 1.0]^{\mathrm{T}}$. The points considered here are assumed to move downwards at the epoch β and let two following theoretical variants of such displacements be considered

$$Variant\ 1 : \mathbf{X}_\beta = [H_{A\beta}, H_{B\beta}, H_{C\beta}]^{\mathrm{T}} = [1.0, 1.0, \underline{0.0}]^{\mathrm{T}}$$
$$\delta^t_{\mathbf{X}} = \mathbf{X}_\beta - \mathbf{X}_\alpha = [0.0, 0.0, -\underline{1.0}]^{\mathrm{T}}$$

$$Variant\ 2 : \mathbf{X}_\beta = [H_{A\beta}, H_{B\beta}, H_{C\beta}]^{\mathrm{T}} = [1.0, 0.0, 0.0]^{\mathrm{T}}$$
$$\delta^t_{\mathbf{X}} = \mathbf{X}_\beta - \mathbf{X}_\alpha = [0.0, -\underline{1.0}, -\underline{1.0}]^{\mathrm{T}}$$

where $\delta_{\mathbf{X}}$ is a vector of the theoretical displacements. The results of the squared M_{split} estimation for both proposed variants are presented below.

Variant 1

Let the observation vector $\mathbf{y}$, in relation to the theoretical displacement vector $\delta^t_{\mathbf{X}} = \mathbf{X}_\beta - \mathbf{X}_\alpha = [0.0,\ 0.0,\ -1.0]^{\mathrm{T}}$, be as follows (the observations at the epoch β that differ significantly from those at the epoch are underlined)

$$\mathbf{y} = [1.01, 0.98 \vdots 1.00, 1.02 \vdots 0.98, \underline{0.01} \vdots 0.97, 0.99 \vdots$$
$$1.00, 1.01 \vdots 0.99, -\underline{0.01} \vdots 0.02, -0.01 \vdots -0.01, -\underline{1.01}]^{\mathrm{T}}$$

The natural course of the iterative process (without any possible acceleration procedure) is shown in Table 3.

Table 3. The course of the iterative process (Example 4, *Variant* 1)

Steps	$\mathbf{g}_\alpha(\mathbf{X}_\alpha^{j-1}, \mathbf{X}_\beta^{j-1})$	$d\mathbf{X}_\alpha^j$	$\mathbf{X}_\alpha^j$	$\mathbf{g}_\beta(\mathbf{X}_\alpha^{j-1}, \mathbf{X}_\beta^{j-1})$	$d\mathbf{X}_\beta^j$	$\mathbf{X}_\beta^j$
0			$\hat{\mathbf{X}}_{LS} = \begin{bmatrix} 0.991 \\ 1.004 \\ 0.493 \end{bmatrix}$	0.000 0.001 0.000	0.003 –0.003 –0.001	0.994 1.000 0.492
1	0.000 –0.001 –0.000	–0.003 0.003 0.001	0.988 1.007 0.494	0.000 0.002 0.000	0.004 –0.006 –0.002	0.998 0.994 0.489
2	0.000 –0.004 –0.001	–0.003 0.013 0.005	0.985 1.020 0.499	0.000 0.008 0.002	–0.016 –0.026 –0.001	0.982 0.968 0.479
3	0.000 –0.016 –0.003	0.043 0.050 0.019	1.028 1.070 0.518	0.000 0.030 0.006	–0.032 –0.093 –0.036	0.950 0.876 0.444
4	0.000 –0.052 –0.013	0.103 0.138 0.056	1.131 1.208 0.574	–0.002 0.060 0.028	0.007 –0.135 –0.065	0.957 0.741 0.379
5	–0.006 –0.017 –0.057	0.124 0.066 0.059	1.255 1.274 0.633	–0.008 –0.024 0.102	0.028 –0.016 –0.070	0.985 0.725 0.309
6	–0.001 0.060 –0.175	0.005 –0.005 0.106	1.261 1.269 0.739	0.000 –0.105 0.283	0.000 0.035 –0.155	0.986 0.760 0.154
7	0.000 0.171 –0.411	–0.098 –0.098 0.185	1.163 1.171 0.924	0.000 –0.228 0.458	–0.001 0.182 –0.145	0.985 0.941 0.008
8	0.000 0.200 –0.313	–0.173 –0.169 0.061	0.990 1.002 0.985	0.000 –0.076 0.092	0.014 0.069 –0.008	0.999 1.010 0.000
9	0.000 0.007 –0.007	–0.007 –0.007 0.000	0.983 0.995 0.985	0.000 0.000 0.000	0.008 0.000 0.000	1.007 1.010 0.000
10	0.000 0.000 0.000	–0.006 0.000 0.000	0.977 0.995 0.985	0.000 0.000 0.000	0.004 0.000 0.000	1.011 1.010 0.000
11	0.000 0.000 0.000	–0.001 0.000 0.000	0.976 0.995 0.985	0.000 0.000 0.000	0.000 0.000 0.000	1.011 1.010 0.000
12	0.000 0.000 0.000	0.000 0.000 0.000	**0.976** **0.995** **0.985**	0.000 0.000 0.000	0.000 0.000 0.000	**1.011** **1.010** **0.000**

The process presented above results in the vectors $\mathbf{X}_\alpha^{12} = \mathbf{X}_\alpha^{11}$ and $\mathbf{X}_\beta^{12} = \mathbf{X}_\beta^{11}$ that makes: $\mathbf{g}_\alpha(\mathbf{X}_\alpha^{12}, \mathbf{X}_\beta^{12}) = \mathbf{0}$, $\mathbf{g}_\beta(\mathbf{X}_\alpha^{12}, \mathbf{X}_\beta^{12}) = \mathbf{0}$. Therefore, the estimates $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$ of the parameters in the split functional model are as follows (with assigned residual vectors $\hat{\mathbf{v}}_\alpha$ and $\hat{\mathbf{v}}_\beta$, respectively):

$$\hat{\mathbf{X}}_\alpha = [0.98, 1.00, 0.98]^{\mathrm{T}}$$

$$\hat{\mathbf{v}}_\alpha = [0.03, 0.00 \vdots 0.00, 0.02 \vdots 0.00, -0.98 \vdots -0.01, 0.01 \vdots 0.00, 0.01 \vdots 0.00, -1.00 \vdots 0.00, -0.03 \vdots 0.00, -1.00]^{\mathrm{T}}$$

$$\hat{\mathbf{X}}_\beta = [1.01, 1.01, \underline{0.0}]^{\mathrm{T}}$$

$$\hat{\mathbf{v}}_\beta = [0.00, -0.03 \vdots -0.01, 0.01 \vdots 0.99, 0.01 \vdots -0.04, -0.02 \vdots -0.01, 0.00 \vdots 0.99, -0.01 \vdots 0.02, -0.01 \vdots 1.00, 0.00]^{\mathrm{T}}$$

Let us pay attention to the fact that

$$\delta_{\hat{\mathbf{v}}} = \hat{\mathbf{v}}_\beta - \hat{\mathbf{v}}_\alpha = [-0.03, -0.03 \vdots -0.01, -0.01 \vdots 0.99, 0.99 \vdots -0.03, -0.03 \vdots -0.01, -0.01 \vdots 0.99, 0.99 \vdots 0.02, 0.02 \vdots 1.00, 1.00]^{\mathrm{T}}$$

and (Fig. 6)

$$\delta_{\hat{\mathbf{X}}} = \hat{\mathbf{X}}_\beta - \hat{\mathbf{X}}_\alpha = [0.03, 0.01, -\underline{0.98}]^{\mathrm{T}}$$

The theoretical values of the displacements ($\delta^t_{\mathbf{X}}$) as well as their estimates obtained by applying M_{split} estimation ($\delta_{\hat{\mathbf{X}}}$) are additionally illustrated in Figure 6. (the bold lines indicate the height differences which values differ from themselves at the epochs α and β significantly)

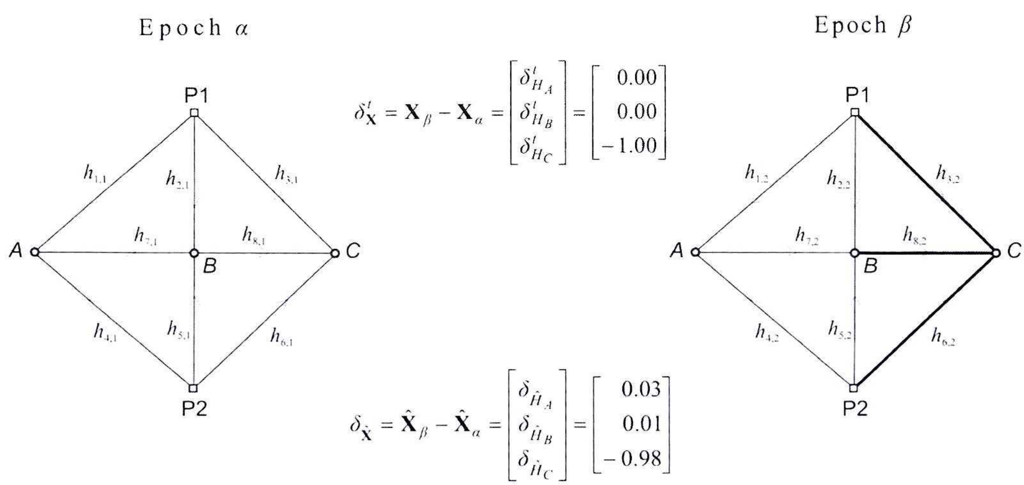


Fig. 6. Assumed and obtained values of point displacements (Example 4, *Variant* 1)

Variant 2

This time the vector of the point displacements is as follows

$$\delta_{\mathbf{X}}^{t} = \mathbf{X}_{\beta} - \mathbf{X}_{\alpha} = [0.0, -\underline{1.0}, -\underline{1.0}]^{\mathrm{T}}$$

i.e. the height of the point C (like it is in the first variant) but also the height of the point B has been changed. Thus, one can assume the following observation vector

$$\mathbf{y} = [1.01, 0.98 \vdots 1.00, \underline{0.02} \vdots 0.98, \underline{0.01} \vdots 0.97, 0.99 \vdots$$
$$1.00, \underline{0.01} \vdots 0.99, -\underline{0.01} \vdots 0.02, -1.01 \vdots -0.01, 0.01]^{\mathrm{T}}$$

The course of the iterative process of the squared M_{split} estimation resulting in the competitive estimates $\hat{\mathbf{X}}_{\alpha}$ and $\hat{\mathbf{X}}_{\beta}$ is shown in Table 4.

Table 4. The course of the iterative process (Example 4, *Variant* 2)

Steps	$\mathbf{g}_{\alpha}(\mathbf{X}_{\alpha}^{j-1}, \mathbf{X}_{\beta}^{j-1})$	$d\mathbf{X}_{\alpha}^{j}$	$\mathbf{X}_{\alpha}^{j}$	$\mathbf{g}_{\beta}(\mathbf{X}_{\alpha}^{j-1}, \mathbf{X}_{\beta}^{j-1})$	$d\mathbf{X}_{\beta}^{j}$	$\mathbf{X}_{\beta}^{j}$
0			$\hat{\mathbf{X}}_{LS} = \begin{bmatrix} 0.991 \\ 1.004 \\ 0.493 \end{bmatrix}$	–0.012 0.000 0.001	0.034 0.012 –0.001	1.025 0.515 0.492
1	0.012 0.000 –0.001	–0.033 –0.012 0.001	0.958 0.492 0.494	–0.023 0.000 0.002	0.065 0.023 –0.002	1.090 0.538 0.489
2	0.044 0.000 –0.005	–0.112 –0.040 0.005	0.846 0.452 0.499	–0.068 0.000 0.009	0.157 0.056 –0.009	1.247 0.595 0.480
3	0.070 0.005 –0.016	–0.112 –0.043 0.015	0.734 0.409 0.514	–0.021 –0.013 0.027	0.040 0.022 –0.026	1.287 0.617 0.454
4	–0.005 0.020 –0.048	–0.002 –0.012 0.047	0.732 0.397 0.561	0.010 –0.026 0.085	–0.005 0.011 –0.080	1.283 0.628 0.373
5	–0.012 0.022 –0.140	0.013 0.000 0.122	0.745 0.397 0.683	0.005 0.023 0.194	–0.030 –0.034 –0.155	1.253 0.594 0.218
6	0.020 –0.109 –0.211	0.041 0.094 0.153	0.786 0.491 0.836	–0.070 0.257 0.180	–0.047 –0.174 –0.132	1.205 0.420 0.087
7	0.130 –0.429 –0.153	0.056 0.257 0.113	0.842 0.749 0.949	–0.160 0.559 0.125	–0.116 –0.315 –0.081	1.089 0.105 0.006
8	0.108 –0.508 0.066	0.013 0.242 0.036	0.972 0.990 0.985	–0.027 0.198 0.011	–0.064 –0.090 –0.006	1.025 0.015 0.000
9	0.001 –0.020 0.000	0.008 0.009 0.000	0.980 1.000 0.985	0.000 0.000 0.000	0.000 0.000 0.000	1.025 0.015 0.000
10	0.000 0.000 0.000	0.000 0.000 0.000	**0.980** **1.000** **0.985**	0.000 0.000 0.000	0.000 0.000 0.000	**1.025** **0.015** **0.000**

The final results of the process presented above are the following estimates:

$$\hat{\mathbf{X}}_\alpha = [0.98, 1.00, 0.98]^{\mathrm{T}}$$

$$\hat{\mathbf{v}}_\alpha = [0.03, 0.00 \vdots 0.00, -0.98 \vdots 0.00, -0.97 \vdots -0.01, 0.01 \vdots 0.00, -0.99 \vdots 0.00, -1.00 \vdots 0.00, -1.03 \vdots 0.00, 0.00]^{\mathrm{T}}$$

and

$$\hat{\mathbf{X}}_\beta = [1.02, 0.02, 0.00]^{\mathrm{T}}$$

$$\hat{\mathbf{v}}_\beta = [-0.01, -0.04 \vdots 0.98, 0.00 \vdots 0.98, 0.01 \vdots -0.06, -0.04 \vdots 0.99, 0.00 \vdots 0.99, -0.01 \vdots 1.03, 0.00 \vdots 0.00, 0.00]^{\mathrm{T}}$$

Furthermore, the vectors $\delta_{\hat{\mathbf{v}}}$ and $\delta_{\hat{\mathbf{X}}}$ can be computed as:

$$\delta_{\hat{\mathbf{v}}} = [-0.04, -0.04 \vdots 0.98, 0.98 \vdots 0.98, 0.98 \vdots -0.05, -0.05 \vdots 0.99, 0.99 \vdots 0.99, 0.99 \vdots 1.03, 1.03 \vdots 0.00, 0.00]^{\mathrm{T}}$$

$$\delta_{\hat{\mathbf{X}}} = \hat{\mathbf{X}}_\alpha - \hat{\mathbf{X}}_\beta = [0.04, -\underline{0.98}, -\underline{0.98}]^{\mathrm{T}}$$

The displacement values assumed in the present variant $\delta^t_{\mathbf{X}}$ and those obtained by applying the squared M_{split} estimation are shown in Figure 7 (the bold lines indicate the height differences which values differ from themselves at the epochs α and β significantly).

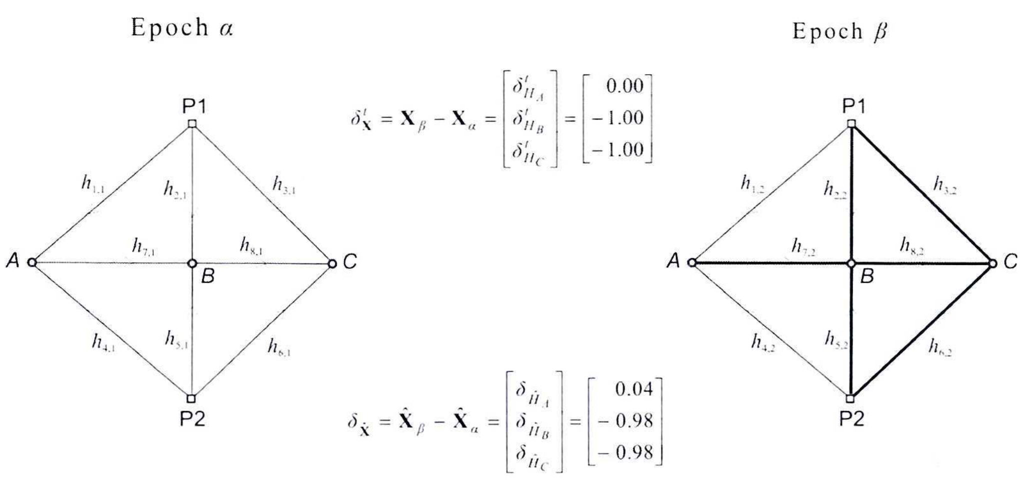


Fig. 7. Assumed and obtained values of point displacements (Example 4, *Variant* 2)

4. Conclusions

M_{split} estimation demands the functions $\rho_\alpha(\cdot)$ and $\rho_\beta(\cdot)$ to be symmetric, convex and at least twice differentiable. Thus they can also be squared functions. Such special case of M_{split} estimation is called the squared M_{split} estimation. On the other hand, such estimation can also result from some probabilistic assumptions. In such case, f-information and, consequently, the split potential is based on some normal distribution (see Eqs. (17)-(19)). However, such assumption is not necessary to create the target function of the squared M_{split} estimation (see section 2 of the present paper).

The optimization problem of M_{split} estimation can be solved by applying the Newton method. In such case, estimates of the LS-method can be a starting point of such iterative process. Let us pay attention to the values of the gradients that are presented in Table 3. and Table 4 (Example 4). Firstly, those values grow larger in most of the iterative steps that may suggest that the optimization problem cannot be solved. However, in some last steps the values grow smaller till they fulfill the necessary condition for the function minimum.

Examples 1 and 2 point out that the squared M_{split} estimation can be an alternative for robust M-estimation. However, the method proposed in the present paper enables to estimate some additional parameters related to outliers. Such opportunity would be useful and important if sources and properties of this observation type are analyzed.

Estimation of two competitive parameters assigned to one observation set has different significance when a set of observations contains realizations of two functions (not necessarily random ones) that belong to the same function family (Example 3). The natural property of M_{split} estimation is that observations are assigned to either of these two functions. It results from the split potential, see part I (let us remind that M_{split} estimates are such quantities that maximize such potential). It means that the method "chooses" itself observations that belong (or suit better) to either of two functions. Such property is especially significant when observation sets coincide partly or when such sets are realizations of random functions.

Another possible application of the method presented in the paper is the elaboration of a geodetic network that is measured at two different epochs (Example 4, *Variants* 1 and 2). The difference between the parameters $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$ can be regarded as displacements of the network points. On this stage of development, it is hard to say that M_{split} estimation can be a good alternative for other well known methods for displacement estimation. However, it is worth remembering that an assignment of observations to either of the epochs is not necessary in case of M_{split} estimation. Such assignment results from maximization of the split potential. It is an important property when an observation set is disturbed deterministically (or with outliers) or some observations are assigned to either of the epochs wrongly. An example of such method application was presented in (Wiśniewski, 2009, section 4.4, Example 4), where the observation $h_{8\beta} = -0.01$ is disturbed with the gross error $= -1.0$.

Analysis of the presented examples points out that values of estimates α and β are generally consistent with their intuitive positions in observation sets. Example 1, the estimate values are close to the centres of the mutually outlying subsets. Example 2, one outlier, the first estimate α lies in the centre of the main subset and the second one β is close to the outlier value. Example 3, the observation set is split into two subsets that are uniquely assigned to the functions that generated the observations. Example 4, the displacement obtained by applying the squared M_{split} estimators are very consistent with the theoretical values.

The examples presented in the paper are just illustrations of M_{split} estimation properties. Thus, they should not be regarded as technological suggestions concerning solutions of presented practical problems. The practical applications of M_{split} estimation need more theoretical as well as empirical analyses.

Acknowledgments

The research is supported by the Polish Ministry of Science and Higher Education under the grant "New methods of estimation of parameters in a functional model of geodetic observations (M_{split} estimation)", N N526 255 734.

References

Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A., (1986): *Robust statistics. The approach based on influence functions*, John Wiley & Sons, New York.

Kadaj R., (1988): *Eine verallgemeinerte Klasse von Schatzverfahren mit praktischen Anwendungen*, Zeitschrift für Vermessungswesen, Vol. 113, No 4, pp. 157-166.

Rao C.R., (1973): *Linear statistical inference and its applications*, John Wiley & Sons, Inc.

Teunissen P.J.G., (1990): *Nonlinear least squares*, Manuscripta Geodaetica, Vol. 15, pp. 137-150.

Wiśniewski Z., (2008): *Split estimation of parameters in functional geodetic models*, Technical Sciences, Vol. 11, pp. 202-212.

Wiśniewski Z., (2009): *Estimation of parameters in a split functional model of geodetic observations (M_{split} estimation)*, Journal of Geodesy, Vol. 83, pp. 105-120.

Yang Y., (1991): *Robust Bayesian estimation*, Bulletin Géodésique, Vol. 65, pp. 145-150.

M_{split} estymacja. Część II. Kwadratowa M_{split} estymacja i przykłady numeryczne

Zbigniew Wiśniewski

Instytut Geodezji
Uniwersytet Warmińsko-Mazurski
ul. Oczapowskiego 1, 10-957 Olsztyn
e-mail: zbyszek@uwm.edu.pl

Streszczenie

W tej części pracy zaprezentowano szczególny przypadek M_{split} estymacji, nazwany squared M_{split} estymacją. Funkcja celu jest tutaj ustalana na podstawie wypukłych funkcji kwadratowych. Przedstawiono teoretyczne podstawy squared M_{split} estymacji, jej algorytm oraz kilka przykładów numerycznych.