SPLIT ESTIMATION OF PARAMETERS IN FUNCTIONAL GEODETIC MODELS

Zbigniew WIŚNIEWSKI

Institute of Geodesy, University of Warmia and Mazury in Olsztyn, Poland

Abstract: The method of estimation presented in the paper is based on the assumption that every measurement result can be a realization of either of two different, random variables (differing from each other in expected values). Supposing it, the functional model \( \mathbf{v} = \mathbf{y} - \mathbf{A}\mathbf{X} \) is split into two competitive ones \( \mathbf{v}_\alpha = \mathbf{y} - \mathbf{A}\mathbf{X}_\alpha \) and \( \mathbf{v}_\beta = \mathbf{y} - \mathbf{A}\mathbf{X}_\beta \), that concern the same vector of observation \( \mathbf{y} \) (\( \mathbf{A} \) is a common coefficient matrix, \( \mathbf{v}_\alpha \) and \( \mathbf{v}_\beta \) are competitive vectors of random errors, \( \mathbf{X}_\alpha \) and \( \mathbf{X}_\beta \) - competitive parameter vectors, respectively). The estimation process proposed here is based on the principle of crossing (mutual) weighting of competitive random errors \( v_{\alpha i} \) and \( v_{\beta i} \) (concerning the same observation \( y_i \)). That important rule is realized with such convex weigh functions \( w_\alpha(v_\beta) \) and \( w_\beta(v_\alpha) \) that:

\[
\sup_{v_{\alpha i} \leq \delta_v} w_\alpha(v_\beta) \Leftrightarrow \min_{v_{\alpha i} \leq \delta_v} w_\beta(v_\alpha)
\]

and

\[
\sup_{v_{\beta i} \leq \delta_v} w_\beta(v_\alpha) \Leftrightarrow \min_{v_{\beta i} \leq \delta_v} w_\alpha(v_\beta),
\]

where \( \delta_v = v_\beta - v_\alpha \). According to the above principle and referring to the \( M \)-estimation theory, the following optimisation problem: \( \min_{\mathbf{X}_\alpha} \sum_i v_{\alpha i}^2 w_\alpha(v_{\beta i}) \) and \( \min_{\mathbf{X}_\beta} \sum_i v_{\beta i}^2 w_\beta(v_{\alpha i}) \), will be formulated and solved in the paper.

The proposed method is essential extension of \( M \)-estimation class. However, its practical application is not limited to a robust estimation of the parameter \( \mathbf{X} \) (\( \hat{\mathbf{X}}_\alpha \) estimator) and extended with estimator \( \hat{\mathbf{X}}_\beta \) (concerning outliers). The presented method can be also applied to a joint adjustment of two observation sets measured in two, different epochs. Differences between competitive estimates \( \hat{\mathbf{X}}_\alpha \) and \( \hat{\mathbf{X}}_\beta \) can indicate displacements of network points.

The paper presents some basic, numerical examples that illustrate principles of the split estimation in functional geodetic models.
1. Introduction

In estimation theories, it is usually assumed that measurement results are a set of realizations of one random variable $\xi$ with the expected value $E(\xi) = \mathbf{A} \mathbf{X}$ ($\mathbf{A} \in \mathbb{R}^{n \times r}$ - known coefficient matrix, $\mathbf{X} \in \mathbb{R}^{r \times 1}$ - parameter vector). Then, the following functional model $v_i = y_i - a_{(i)} \mathbf{X}$, $i = 1, K, n$, can be formulated for every observation $y_i$ ($v_i$ - random error, $a_{(i)}$ - $ith$ row of the matrix $A$). In geodetic practice, it is sometimes possible that measurement results are realizations of either of $\xi_\alpha$ or $\xi_\beta$ different, random variables differing from each other in expected values $E(\xi_\alpha) = \mathbf{A} \mathbf{X}_\alpha$ and $E(\xi_\beta) = \mathbf{A} \mathbf{X}_\beta$. For example, it can occur that an observation set contains some outliers. Then, one can assume that “good” observations are realizations of the variable $\xi_\alpha$ and outliers, the other one, $\xi_\beta$. The essential problem is how to identify a particular measurement result with either of random variables objectively. Robust $M$-estimation solves the problem by applying specially designed attenuation functions. Such functions decrease weighs of observations, which are supposed to be realizations of the variable $\xi_\beta$ (such suspicion is based on values of residuals $v_i$) (e.g. Huber 1981, Hampel et. al. 1986, Yang 1994, Yang et al 2002;). In that case, estimation results depend on arbitrarily assumed attenuation functions or more generally on assumed weigh functions (various examples of attenuation or weigh functions are presented in e. g. Hampel et. al. 1986; Koch 1996, 1999; Gui, Zhang 1998; Caspary, Hean 1990; Wiśniewski 1999). Similarly, results of geodetic network measurements obtained in two different epochs can be an example of a set that consists of realizations of two random variables differing from each other in expected values. Up to now, in the classical estimation methods, results obtained in the first epoch are arbitrarily taken as realizations of the first variable $\xi_\alpha$, and in the second one as realizations of the other $\xi_\beta$. Thus, it is necessary to assume two coexisting functional models: for one part of the observation set $v_{\alpha i} = y_i - a_{(i)} \mathbf{X}_\alpha$, $i = 1, K, n_\alpha$ and for the other $v_{\beta i} = y_i - a_{(i)} \mathbf{X}_\beta$, $i = 1, K, n_\beta$, where $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$ are two states of network point coordinates. Now, let us assume that some network points are not displaced. Then, some observations described classically with the model $v_{\beta i} = y_i - a_{(i)} \mathbf{X}_\beta$, can be interpreted as realizations of the variable $\xi_\alpha$ and can support estimation of the parameter $\mathbf{X}_\alpha$.

Let us consider the above examples or other similar situation when a set of measurement results may be a realization of either of two random variables. It is assumed here that to identify observations objectively, which is equivalent with objective estimation of parameters $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$, the random variables should rival each other. It means that, two functional models as well as two competitive random errors $v_{\alpha i}$ and $v_{\beta i}$ should exist for each observation $y_i$. The concept described here is explained with the following elemental and rather "idealistic" example.

Let $y = 1, 2, 3, 6, 8$ be measurement results of some quantity $X$ with the model $v_i = y_i - X$. One can assume $\hat{X}_{\alpha} = 2.0$ as the estimator that accepts realizations of the random variable $\xi_\alpha = 1, 2, 3$ (in the robust estimation $\hat{\xi}_\alpha$ would be a “good” variable) and ignores realizations
of the variable $\xi_\beta = 6.8$ (in the robust estimation – “strange” variable). The following residuals: $\hat{v}_{\alpha,1} = -1$, $\hat{v}_{\alpha,2} = 0$, $\hat{v}_{\alpha,3} = 1$, $\hat{v}_{\alpha,4} = 4$, $\hat{v}_{\alpha,5} = 6$ respond to that estimator. Another possible and competitive solution is an estimate $\hat{X}_\beta = 7$. This time one assumes that $\xi_\alpha$ is the “strange” variable. Thus, another version of residuals can be obtained: $\hat{v}_{\beta,1} = -6$, $\hat{v}_{\beta,2} = -5$, $\hat{v}_{\beta,3} = -4$, $\hat{v}_{\beta,4} = -1$, $\hat{v}_{\beta,5} = 1$.

If two competitive assumptions about identifying measurement results with either of random variables is considered then the classical functional model of geodetic observations must be split. This paper proposes the way how to estimate parameters of such, split models. The optimisation problem formulated here refers to the principles of M-estimation but it is also an important extension of this estimation class. The proposed method is illustrated with two numerical examples: The first one concerns the observation set presented earlier (with the solutions indicated) and the second one refers to estimation of coordinates of geodetic points in a network with a “mixed” set of measurement results (measurement results obtained in two epochs).

2. Assumptions

2.1. Split functional model

The following optimisation problem:

$$\min_X \varphi(X) = \sum_{i=1}^{n} \rho(\hat{v}_i), \quad \hat{v}_i = y_i - a_{i(j)} \hat{X}$$

is the basis for many estimation methods that belong to the M-estimation class; where $\rho(v_i)$, $i = 1, K, n$, are some symmetric, convex, arbitrarily assumed functions (Huber 1981; Hampel et. al. 1986; Krarup, Kubik 1983; Huang, Maricas 1995; Koch 1996; Zhu 1996; Xu 1989; Yang 1994; Yang et. al. 2002). In the least squares method, which is a neutral M-estimation, it is assumed that

$$\rho(v_i) = v_i^2 p_i$$

where $p_i$ is the weight of $ith$ observation assigned, within the function $\rho(v_i)$, to the random error $v_i$. According to the robust M-estimation, the weights $p_i$ are replaced with equivalent weights $\hat{p}_i$, which values are equal to values of concave or quasi-concave weight functions $w(v_i)$ whereat $|v_i| \neq 0$: $w(v_i) \leq w(0)$. The above properties of weight functions correspond to the assumption that the M-estimator should be function of realizations of “good” variable $\xi_\alpha$ only. Outliers, realizations of “strange” variable $\xi_\beta$, should be rejected.

The concept of equivalent weights follows the assumptions that variables $\xi_\alpha$ and $\xi_\beta$ share one expected value $E(\xi_\alpha) = E(\xi_\beta) = E(\xi) = AX$ and differ from each other much in standard deviations $\sigma$. Values of the standard deviation $\sigma_\alpha$ (variable $\xi_\alpha$) are reasonable and depend on measurement accuracy, technique etc. In contrast, values of the standard deviation $\sigma_\beta$
(“strange” variable $\xi_\beta$ ) are assumed as adequately bigger than $\sigma_\alpha$ hence not to influence the $M$-estimator computation.

In practice, weight functions, which follow the concept described above, are not so restrictive (influence of realizations identified as outliers is “eliminated” softly). It is because observation residuals are known only (real errors stay unknown) and sets of realizations $\xi_\alpha$ and $\xi_\beta$ can share some intersections. Huber’s function is a good example of such practical weight function (Huber 1986)

$$w(v) = \begin{cases} 1 & \text{for } |v| \leq c \\ \frac{c}{|v|} & \text{for } |v| > c \end{cases}$$

where $p_i = 1/\sigma_\alpha^2$ are original weights while equivalent weights $\hat{p}_i = c/|v_i|$ approach some small value referred to standard deviation $\sigma_\beta$ (usually $c = 2\sigma_\alpha$).

One can notice that values of $\sigma_\beta$ should be at an adequately high level. It is because of the assumption concerning expected values $E(\xi)$ of the random variables $\xi_\alpha$ and $\xi_\beta$.

This paper proposes more natural concept that accepts expected value $E(\xi_\beta)$ different from $E(\xi_\alpha)$ and that makes standard deviation $\sigma_\beta$ reasonable, too. Such assumptions lead to the split of the random variable realizations (concerning $\xi_\alpha$ and $\xi_\beta$, respectively) and to other following consequences, described later on.

Thus, if a set of random variable realizations is split into two competitive ones, which are identified with variables $\xi_\alpha$ and $\xi_\beta$, then two competitive, functional models $v_{ai} = y_i - a_{(i)}X_a$ and $v_{\beta i} = y_i - a_{(i)}X_\beta$ (where $E(y_i) = a_{(i)}X_a$ and $E(y_i) = a_{(i)}X_\beta$) are assigned to every observation $y_i$. The split of the functional model that concerns whole observation set can be written as follows:

$$split (v = y - AX) = \begin{cases} v_\alpha = y - AX_a \\ v_\beta = y - AX_\beta \end{cases}$$

2.2. Weight functions and target function components.

Estimation of competitive parameters $X_a$ and $X_\beta$ by using the same vector of observations $y$, requires specially formulated target function of the optimisation problem. This paper proposes to replace function $\rho(v)$ with functions $\rho_\alpha(v_\alpha)$ and $\rho_\beta(v_\beta)$, according to the model in Eq. (4) and in compliance with the principle of crossing, mutual weighting of competitive residuals $v_\alpha$ and $v_\beta$. That cross-weighting is guaranteed when the weight functions $w_\alpha(v_\beta)$ and $w_\beta(v_\alpha)$ are obtained as
The functions $\rho_\alpha$ and $\rho_\beta$, as well as the weight functions, should also own the following properties (considering the optimisation target):

$$
\begin{align*}
\min_{v_\alpha} \rho_\alpha(v_\alpha) & \Leftrightarrow \sup_{v_\alpha \leq \delta_v} w_\alpha(v_\beta), & \min_{v_\beta} \rho_\beta(v_\beta) & \Leftrightarrow \sup_{v_\beta \leq \delta_v} \rho_\alpha(v_\alpha) \\
\min_{v_\beta} \rho_\beta(v_\beta) & \Leftrightarrow \sup_{v_\alpha \leq \delta_v} w_\beta(v_\alpha), & \min_{v_\alpha} \rho_\alpha(v_\alpha) & \Leftrightarrow \sup_{v_\alpha \leq \delta_v} \rho_\beta(v_\beta)
\end{align*}
$$

(6)

where $\delta_v = v_\beta - v_\alpha$. The following functions of standardized random errors $v_\alpha, v_\beta$:

$$
\begin{align*}
\rho_\alpha(v_\alpha) &= v_\alpha^2, & w_\alpha(v_\beta) &= \frac{\partial \rho_\alpha(v_\alpha)}{\partial (v_\alpha^2)} = v_\alpha^2 \\
\rho_\beta(v_\beta) &= v_\beta^2, & w_\beta(v_\alpha) &= \frac{\partial \rho_\beta(v_\beta)}{\partial (v_\beta^2)} = v_\beta^2
\end{align*}
$$

(7)

own mentioned properties in natural way (Fig. 1).

---

**Fig. 1- Illustration to cross-weighting**
3. Optimisation Problem and Solution

The following optimisation problem is formulated on the basis of the model Eq. (4) and functions from Eq. (7):

\[
\begin{aligned}
\mathbf{v}_\alpha &= \mathbf{y} - \mathbf{AX}_\alpha \\
\mathbf{v}_\beta &= \mathbf{y} - \mathbf{AX}_\beta \\
\min_{\mathbf{X}_\alpha} \varphi(\mathbf{X}_\alpha) &= \varphi(\hat{\mathbf{X}}_\alpha) \\
\min_{\mathbf{X}_\beta} \varphi(\mathbf{X}_\beta) &= \varphi(\hat{\mathbf{X}}_\beta)
\end{aligned}
\]  

(8)

where

\[
\begin{aligned}
\varphi(\mathbf{X}_\alpha) &= \sum_{i=1}^{n} \rho_{\alpha}(v_{ai}) = \sum_{i=1}^{n} v_{ai}^2 w_{\alpha}(v_{\beta i}) = \mathbf{v}_\alpha^T \mathbf{w}_\alpha(\mathbf{v}_\beta) \mathbf{v}_\alpha \\
\varphi(\mathbf{X}_\beta) &= \sum_{i=1}^{n} \rho_{\beta}(v_{\beta i}) = \sum_{i=1}^{n} v_{\beta i}^2 w_{\beta}(v_{ai}) = \mathbf{v}_\beta^T \mathbf{w}_\beta(\mathbf{v}_\alpha) \mathbf{v}_\beta
\end{aligned}
\]  

(9)

and

\[
\begin{aligned}
\mathbf{w}_\alpha(\mathbf{v}_\beta) &= \text{Diag}(v_{\beta 1}^2, v_{\beta 2}^2, \Lambda, v_{\beta n}^2) \\
\mathbf{w}_\beta(\mathbf{v}_\alpha) &= \text{Diag}(v_{ai}^2, v_{ai}^2, \Lambda, v_{ai}^2)
\end{aligned}
\]

The competitive estimators \(\hat{\mathbf{X}}_\alpha\) and \(\hat{\mathbf{X}}_\beta\) are solutions of the problem Eq. (8) when gradients \(\mathbf{g}_\alpha\) and \(\mathbf{g}_\beta\) are zero vectors, i.e., when the following equation system is fulfilled (for \(\mathbf{v}_\alpha = \hat{\mathbf{v}}_\alpha = \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_\alpha\) and \(\mathbf{v}_\beta = \hat{\mathbf{v}}_\beta = \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_\beta\)):

\[
\begin{aligned}
\mathbf{g}_\alpha^T(X_\alpha, \mathbf{X}_\alpha) &= \frac{\partial \varphi(\mathbf{X}_\alpha)}{\partial \mathbf{X}_\alpha} = 2\mathbf{v}_\alpha^T \mathbf{w}_\alpha(\mathbf{v}_\beta) \mathbf{A} = \mathbf{0} \\
\mathbf{g}_\beta^T(X_\alpha, \mathbf{X}_\beta) &= \frac{\partial \varphi(\mathbf{X}_\beta)}{\partial \mathbf{X}_\beta} = 2\mathbf{v}_\beta^T \mathbf{w}_\beta(\mathbf{v}_\alpha) \mathbf{A} = \mathbf{0}
\end{aligned}
\]  

(10)

Because for the functions \(\varphi(\mathbf{X}_\alpha)\) and \(\varphi(\mathbf{X}_\beta)\) the following Hessians exist:
then Newton’s method can be applied to compute $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$ (e.g. Teunissen 1990). Let us consider necessary conditions Eq. (10) (gradients are referred to each other by shared variable $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$), the iterative formula can be written as ($j=1,K,k$)

$$
\begin{align*}
\mathbf{X}_\alpha^j &= \mathbf{X}_\alpha^{j-1} + d\mathbf{X}_\alpha^{j-1}, & d\mathbf{X}_\alpha^j &= H_{\alpha}^{-1}(\mathbf{X}_\beta^{j-1})g_{\alpha}(\mathbf{X}_\alpha^{j-1}, \mathbf{X}_\beta^{j-1}) \\
\mathbf{X}_\beta^j &= \mathbf{X}_\beta^{j-1} + d\mathbf{X}_\beta^{j-1}, & d\mathbf{X}_\beta^j &= H_{\beta}^{-1}(\mathbf{X}_\alpha^j)g_{\beta}(\mathbf{X}_\alpha^j, \mathbf{X}_\beta^j)
\end{align*}
$$

(12)

The iterative process Eq. (12) is ended for such $j=k$, where $g_{\alpha}(\mathbf{X}_\alpha^{k-1}, \mathbf{X}_\beta^{k-1})=\mathbf{0}$ and $g_{\beta}(\mathbf{X}_\alpha^{k}, \mathbf{X}_\beta^{k-1})=\mathbf{0}$. Thus $\hat{\mathbf{X}}_\alpha = \mathbf{X}_\alpha^k$, $\hat{\mathbf{X}}_\beta = \mathbf{X}_\beta^k$ and $\hat{\mathbf{v}}_\alpha = \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_\alpha$, $\hat{\mathbf{v}}_\beta = \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_\beta$. The estimates of the least squares method $\hat{\mathbf{X}}_{LS} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$ and $\hat{\mathbf{v}}_{LS} = \mathbf{y} - \mathbf{A}\hat{\mathbf{X}}_{LS}$ can be a starting point for such iterative process.

4. Examples

A. Let, as it was in the first part of the paper, $\mathbf{y}^T = [1, 2, 3, 6, 8]$ be a vector of measurements of some quantity $X$ and let $\mathbf{v} = \mathbf{y} - \mathbf{A}\mathbf{X}$ be the classical, functional model, where $\mathbf{A}^T = [1, 1, 1, 1, 1]$. Then the $LS$ estimates are as follows: $\hat{\mathbf{X}}_{LS} = 4$ and $\hat{\mathbf{v}}_{LS}^T = [-3, -2, -1, 2, 4]$. Let now the optimisation problem Eq.(8) with split, functional model $\mathbf{v}_\alpha = \mathbf{y} - \mathbf{X}_\alpha$, $\mathbf{v}_\beta = \mathbf{y} - \mathbf{X}_\beta$ be considered. Then the following competitive estimator can be obtained (mostly in compliance with “intuitional” values from the first part of the paper):

$$
\hat{\mathbf{X}}_\alpha = 1.87, \quad \hat{\mathbf{v}}_\alpha^T = [-0.87, 0.13, 1.13, 4.13, 6.13] \\
\hat{\mathbf{X}}_\beta = 7.19, \quad \hat{\mathbf{v}}_\beta^T = [-6.19, -5.19, -4.19, -1.19, 0.81]
$$

The main results of the iterative estimation process Eq. (12), which lead to the final results presented above are listed in table 1 ($\mathbf{P} = \text{Diag}(\mathbf{P})$ is the weight matrix in the $LS$ method and $\mathbf{w} = \text{Diag}(\mathbf{w})$)
### Table 1 - Iterative process results

B. Consider a levelling network with one unknown point, four fixed points and four height differences measured in two epochs. Measurement results were simulated under assumption that \( \delta_H = H_\beta - H_\alpha = 0.100 \) is a theoretical difference of the unknown point heights in two measurement epochs \( \alpha \) and \( \beta \). Figure 2 presents the network sketch together with the simulated results of measurements.
Fig. 2 - Levelling network sketch

According to the principles of the proposed method all the observations can be written as one vector of measurement results (the observation order does not matter)

\[ y^T = [h_{a1}, h_{a2}, h_{a3}, h_{a4}, h_{a5}, h_{a6}, h_{a7}, h_{a8}] = [1.002, 1.096, 1.003, 1.101, 0.998, 1.099, 0.997, 1.104] \]

and two competitive, functional models can be formulated

\[ v_a = y - AH_a, \quad v_\beta = y - AH_\beta \]

where \( A^T = [1, 1, 1, \Lambda, 1, 1] \). Thus, the optimisation problem Eq. (8) can be solved in six steps

\[
\hat{H}_a = 1.000, \quad \hat{v}_a^T = [0.002, 0.096, 0.003, 0.101, -0.002, 0.099, -0.003, 0.104] \\
\hat{H}_\beta = 1.100, \quad \hat{v}_\beta^T = [-0.098, -0.004, -0.097, 0.009, -0.102, -0.001, -0.103, 0.004]
\]

It should be emphasised that \( \hat{H}_\beta - \hat{H}_a = \delta_H^I \) and also \( \forall i : \hat{v}_{ai} - \hat{v}_{Ii} = \delta_H^I \).

5. Conclusions

The method proposed here is an extension of M-estimation class where measurement results can be realizations of either of two different, competitive, random variables. The algorithm presented in Eq. (12) can be applied to the robust estimation of the parameters (example A) as well as to estimation of geodetic points displacements (example B).

The cross-weighting principle, concerning the competitive residuals of the same observation, is theoretical foundation of the proposed estimation method. This paper assumed that the weight functions are convex, squared ones. However, other solutions are also possible if only other assumed weight function fulfil required theoretical conditions Eq. (6).
The solution of the optimisation problem Eq. (8) proposed here refers to Newton’s method. Such algorithm is especially effective with the assumed weight functions (the examples show that satisfactory results can be obtained after few iterative steps).

**Acknowledgments.** The research is supported by the Polish Ministry of Science and Higher Education under the grant no. N N526 255 734

**References**


Corresponding author contacts

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