Technical Sciences, 2021, 24, 39–49

UNIWERSYTET WARMIŃSKO-MAZURSKI W OLSZTYNIE

Yearbook peer-reviewed scientific journal
ISSN 1505-4675
e-ISSN 2083-4527
TECHNICAL SCIENCES
Homepage: https://czasopisma.uwm.edu.pl/index.php/ts/

DOI: https://doi.org/10.31648/ts.6555

OPTIMIZATION METHOD BASED ON MINIMIZATION *M*-ORDER CENTRAL MOMENTS USED IN SURVEYING ENGINEERING PROBLEMS

***Sławomir Cellmer*[1], *Andrzej Bobojć*[2]**

[1]ORCID: 0000-0002-2614-8017
[2]ORCID: 0000-0003-0228-5962
Department of Geodesy
University of Warmia and Mazury in Olsztyn

Received 23 March 2021, accepted 26 May 2021, available online 31 May 2021.

K e y w o r d s: the Least Squares method, the Newton method, objective function, *m*-estimation, surveying engineering.

A b s t r a c t

A new optimization method presented in this work – the Least m-Order Central Moments method, is a generalization of the Least Squares method. It allows fitting a geometric object into a set of points in such a way that the maximum shift between the object and the points after fitting is smaller than in the Least Squares method. This property can be very useful in some engineering tasks, e.g. in the realignment of a railway track or gantry rails. The theoretical properties of the proposed optimization method are analyzed. The computational problems are discussed. The appropriate computational techniques are proposed to overcome these problems. The detailed computational algorithm and formulas of iterative processes have been derived. The numerical tests are presented, in order to illustrate the operation of proposed techniques. The results have been analyzed, and the conclusions were then formulated.

Correspondence: Sławomir Cellmer, Katedra Geodezji, Uniwersytet Warmińsko-Mazurski, ul. Oczapowskiego 1, 10-719 Olsztyn, e-mail: slawomir.cellmer@uwm.edu.pl, altair@uwm.edu.pl

Introduction

Optimization techniques are sometimes applied in engineering tasks. In such applications, the problem is usually formulated as fitting a geometrical figure into a set of points in 2D or 3D space. In contrast to many estimation problems where the solution has to be free of outliers (CASPARY 1990, CHANG, GUO 2005, HAMPEL et al. 1986, HUBER 1981, KAMIŃSKI, WIŚNIEWSKI 1992, KOCH 1996, YANG 1999, YANG et al. 2002, XU 1989, ZHONG 1997, ZHU 1996), the opposite problem is presented here: an optimization method preferring outliers. This feature is beneficial in engineering applications, where constraints concerning maximum shifts appear. In some cases, the maximum shifts cannot exceed a critical value. Such constraints can appear e.g. during the realignment of a railway track (SKAŁA-SZYMAŃSKA et al. 2014). These constraints result from a limited structure gauge, i.e. the width of tunnels, bridges or distances to railway platforms, buildings or other objects. The constraints, mentioned above, have to be taken into account during the construction process when some elements of a structure are installed in a limited space. An example of such civil engineering task is to fit elevator guide rails inside the elevator shaft. One of the tools for solving the problem of limited shifts can be the min-max algorithm from game theory (VON NEUMANN 1947). Also, the Least Squares (LS) method with constraints can be applied (LIEW 1976, MEAD, RENAUT 2010, WERNER 1990). An alternative method is presented here. In the proposed method, the special form of the objective function is applied. This form is related to the m-order central moments (with $m \geq 2$). Thus, the proposed method is a generalization of the LS method and the Least Fourth Powers (LFP) method (CELLMER 2014). Special attention must be paid to the optimization technique. So far, a lot of interesting optimization techniques were proposed e.g. (FLETCHER 1987, NOCEDAL, WRIGHT 1999, AVRIEL 2003). It was shown in (CELLMER 2014) that if an inappropriate technique of searching for the minimum of the objective function is applied, then the computational process is not convergent. In such a case, two alternate solutions are obtained in consecutive iterations like in the M-split estimation method (DUCHNOWSKI, WIŚNIEWSKI 2012, WIŚNIEWSKI 2009, 2010). In the *M*-split estimation method, this effect was obtained deliberately – it resulted from the theoretical foundations of this method. However, in the method considered here, a single, unique solution is required. The technique of searching for a solution should have the property of skipping local minima and pursue the global one. This property was described e.g. in (MARTINS, TSUZUKI 2009).

In the next section, the objective function of the Least m-Order Central Moments (LmOCM) method is presented. The justification for applying this form of the objective function was carried out using certain concepts of estimation theory and was illustrated using the plot of weight function for different values of the m-exponent. The third section contains the derivation of the formulas

of the computational process. Two different techniques of optimization of the LmOCM objective function have been discussed and two numerical examples were analyzed in the fourth section. The conclusions have been formulated on the basis of the results of tests.

Objective function of the LmOCM method and its properties

The optimization method proposed in this article is based on minimizing the objective function $\Psi(\mathbf{v})$:

$$\min_{\mathbf{v}}\left(\Psi(\mathbf{v})=\sum_{i=1}^{n} v_i^m\right) \tag{1}$$

where v_i are elements of the **v** vector in the simple, linear model:

$$\mathbf{y + v = Ap} \tag{2}$$

where:

y – vector of entries that are fitted into the **A** model,
v – vector of corrections (disclosures),
A – design matrix,
p – parameter vector.

The optimization method based on criterion (1), is a generalization of the LS or the LFP method (Cellmer 2014). In Cellmer (2014), the properties of the LFP method have been described using selected concepts of estimation theory. The considered estimation methods belong to the *m*-estimation class. These methods are based on the minimization:

$$\min_{\mathbf{v}}\left(\Psi_G(\mathbf{v})=\sum_{i=1}^{n} \rho_i(\mathbf{v})\right) \tag{3}$$

The ρ_i function is a component of the objective function $\Psi_G(\mathbf{v})$. The objective function $\Psi(\mathbf{v})$ in the formula (1) is a special case of the the objective function $\Psi_G(\mathbf{v})$. The form of the component ρ_i determines the properties of the results. In the method proposed here, the ρ_i function takes the following form:

$$\rho_i(\mathbf{v}) = v_i^m \tag{4}$$

One of the properties of the optimization method can be described by the weight-function (Kadaj 1988). The form of this function is:

$$\mathrm{w}(v) = \frac{\partial \rho(v)}{\partial v^2} \tag{5}$$

The plots of the weight function with the values of the m-parameter: 2, 4, 6, 8, 10 and 12 are depicted in Figure 1. The black line depicts the weight function of the LS method (m = 2). As follows from the definition of the weight function (5), in the case of the LS method, it is constant. All observations are treated equally. In the case of large values of the m-parameter, there is a range around the zero value, where the trajectory of the plot of the function is almost horizontal. Outside this range, the function value rapidly increases. This means that if there are no outliers in the data set, this optimization method provides results similar to the results obtained with the LS method. However, if there are some outliers in the data set, they have more impact on the solution than other observations. The consequence of this is a smaller maximal value of the residuals in the optimization process.

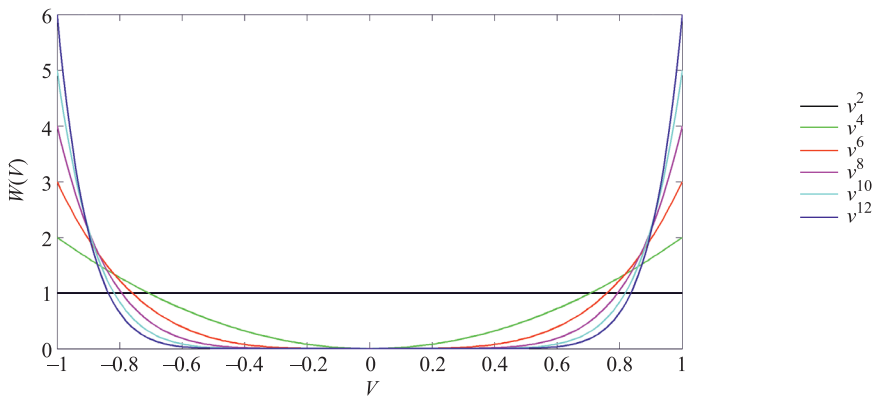

Fig. 1. Weight functions. The plots represent the weight functions of the objective functions, which are in the form of the central moments of various orders

Optimization techniques

The objective function in minimization problem (1) can be presented in the following matrix notation:

$$\Psi(\mathbf{v}) = \mathbf{v}_r^{\mathrm{T}}\mathbf{v}_r \tag{6}$$

where:

$$r = \frac{m}{2} \tag{7}$$

$$\mathbf{v}_{\mathrm{r}} = [v_1^{\mathrm{r}}, v_2^{\mathrm{r}}, \ldots, v_n^{\mathrm{r}}]^T \tag{8}$$

The solution, which is at the point of a minimum of the objective function (6) can be found by zeroing the gradient of this function. As was shown in (CELLMER 2014), this technique cannot be employed in its simple, classic form. The gradient of the objective function (6) is derived, based on (2) and (7):

$$\frac{\partial \Psi}{\partial \mathbf{p}} = \frac{\partial \Psi}{\partial \mathbf{v}_r}\frac{\partial \mathbf{v}_r}{\partial \mathbf{v}}\frac{\partial \mathbf{v}}{\partial \mathbf{p}} = 2\mathbf{v}_r^{\mathrm{T}} \cdot r \cdot \mathrm{diag}(\mathbf{v}_{r\text{-}1}) \cdot \mathbf{A} = 2r \cdot \mathbf{v}_{2r\text{-}1}^{\mathrm{T}}\,\mathbf{A} = 2r \cdot \mathbf{v}^{\mathrm{T}}\mathrm{diag}(\mathbf{v}_{m\text{-}2})\,\mathbf{A} \tag{9}$$

where $\mathrm{diag}(\mathbf{v}_{m\text{-}2})$ denotes a diagonal matrix, containing elements of the $\mathbf{v}_{m\text{-}2}$ – vector on a diagonal. Thence the system of normal equations is obtained:

$$\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}\mathbf{p} - \mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{y} = \mathbf{0} \tag{10}$$

where:

$$\mathbf{W} = \mathrm{diag}(\mathbf{v}_{m\text{-}2}) \tag{11}$$

The solution can be obtained in an iterative process:

$$\mathbf{p}_{(i)} = (\mathbf{A}^{\mathrm{T}}\mathbf{W}_{(i-1)}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}_{(i-1)}\mathbf{y}$$

$$\mathbf{v}_{(i)} = \mathbf{A}\mathbf{p}_{(i)} - \mathbf{y}, \text{ for } i = 1, 2, \ldots \text{ and } \mathbf{W}_{(0)} = \mathbf{I} \tag{12}$$

The **W** matrix is formed according to (11) in each iteration. However, as was shown in (CELLMER 2014), in the case of the objective function (6), process (12) does not work properly. This technique does not ensure proper convergence to the correct, unique solution. As a result of using formulas (12), two alternate solutions are obtained. They appear alternately in consecutive iterations. A similar effect is obtained when using the *M*-split estimation proposed by (WIŚNIEWSKI 2009, 2010). However, the unique solution of the minimization process of the objective function (6) exists. CELLMER (2014) proposed applying one of two optimization techniques for obtaining a unique solution in the considered problem. The first one is based on the modification of the process (12). The second is based on applying the Newton technique of optimization. In the first technique, the iterative process (12) is modified to the form:

$$\begin{aligned}
\mathrm{d}\mathbf{y}_{(i)} &= \mathbf{y} - \mathbf{A}\mathbf{p}_{(i-1)} \\
\mathrm{d}\mathbf{p}_{(i)} &= k(\mathbf{A}^{\mathrm{T}}\mathbf{W}_{(i-1)}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}_{(i-1)}\mathrm{d}\mathbf{y}_{(i)} \\
\mathbf{p}_{(i)} &= \mathbf{p}_{(i-1)} + \mathrm{d}\mathbf{p}_{(i)} \\
\mathbf{v}_{(i)} &= \mathbf{A}\mathrm{d}\mathbf{p}_{(i)} - \mathrm{d}\mathbf{y}_{(i)} \\
\mathbf{W}_{(i)} &= \mathrm{diag}\left(\mathbf{v}_{m\text{-}2(i)}\right) \text{ for } i = 1, 2, \ldots \text{ and } \mathbf{p}_{(0)} = \mathbf{0},\ \mathbf{W}_{(0)} = \mathbf{I}
\end{aligned} \tag{13}$$

where the $\mathbf{v}_{m\text{-}2(i)}$ vector in the last row is formed according to (8). In this technique, in contrast to process (12), the results obtained from consecutive iterations are accepted as the starting point (approximate value of the parameter) in the next iteration. This is performed by updating the d**y** and **p** vectors. Additionally, the reduction parameter k is introduced. This parameter reduces the length of every single step in the iterative process of searching for the solution. This was imposed to eliminate the effect of jumping over the global solution in consecutive iterations. The k parameter has to take the value from the range: $(0; 1)$. Determination of the optimal value of this parameter is presented in the next section.

In the second technique, the Newton method of optimization is applied in its classic form:

$$\mathbf{p}_{(i)} = \mathbf{p}_{(i-1)} - \mathbf{H}^{-1}\left(\mathbf{p}_{(i-1)}\right)\mathbf{G}\left(\mathbf{p}_{(i-1)}\right) \tag{14}$$

where **G** and **H** are, appropriately, the gradient and Hessian of the objective function Ψ. The gradient of the objective function (8) is expressed by formula (9). Let us denote the gradient **G** as a transpose of (9):

$$\mathbf{G} = 2r\,\mathbf{A}^{\mathrm{T}}\mathbf{v}_{m\text{-}1} \tag{15}$$

or

$$\mathbf{G} = m\,\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{v} \tag{16}$$

The Hessian can then be formed as:

$$\mathbf{H} = \frac{\partial \mathbf{G}}{\partial \mathbf{p}} = \frac{\partial \mathbf{G}}{\partial \mathbf{v}_{m\text{-}1}}\frac{\partial \mathbf{v}_{m\text{-}1}}{\partial \mathbf{v}}\frac{\partial \mathbf{v}}{\partial \mathbf{p}} = m\,\mathbf{A}^{\mathrm{T}}(m-1)\mathrm{diag}(\mathbf{v}_{m\text{-}2})\,\mathbf{A} =$$

$$= m\,(m-1)\,\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A} \tag{17}$$

Thus, in this case the optimization process can be presented as follows:

$$\mathbf{p}_{(i)} = \mathbf{p}_{(i-1)} - \frac{1}{m-1}(\mathbf{A}^{\mathrm{T}}\mathbf{W}_{(i-1)}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}_{(i-1)}\mathbf{v}_{(i-1)}$$

$$\mathbf{v}_{(i)} = \mathbf{A}\mathbf{p}_{(i)} - \mathbf{y} \quad \text{for } i = 1, 2, \ldots \text{ and } \mathbf{W}_{(0)} = \mathbf{I} \tag{18}$$

The correction vector in the first iteration ($\mathbf{v}_{(0)}$) is calculated using formula (12).

Numerical examples

Table 1 contains the coordinates of the points. The axes are to be fitted into each set of these points using criterion (1). In the first example the formula of the fitted line is:

$$y = ax \tag{19}$$

whereas in the second example, it is:

$$y = ax + b \tag{20}$$

Table 1

Set of points for fitting the axes

Example 1			Example 2		
No	x [m]	y [m]	No	x [m]	y [m]
1	1.000	0.102	1	1.000	0.507
2	2.000	0.196	2	2.000	0.603
3	3.000	0.302	3	3.000	0.699
4	4.000	0.405	4	4.000	0.801
5	5.000	0.501	5	5.000	0.897
6	6.000	0.603	6	6.000	1.006
7	7.000	0.727	7	7.000	1.128
8	8.000	0.799	8	8.000	1.202
9	9.000	0.914	9	9.000	1.310
10	10.000	0.998	10	10.000	1.401

The **A** matrix and the **y** vector in each example are formed based on the following observation equations:

$$y_i + v_i = ax_i \text{ (example 1)} \tag{21}$$

$$y_i + v_i = ax_i + b \text{ (example 2)} \tag{22}$$

Hence, in the first example, the **A** is the column vector of x-coordinates, whereas in the second example this is a two-column matrix: the first column contains x-coordinates and the second column is a vector of ones. At the start point of optimization, the **y**-vector contains the y-coordinates of the points listed in Table 1. The first test was performed in order to determine the optimal value of the k-factor in process (13). Therefore, this process was carried out for different values of k, and afterwards, the speed of convergence was analyzed for each of them. The results are listed in Table 2. It was shown in (CELLMER 2014) that if $k = 1$, the solution splits into two results, repeating alternately in consecutive iterations. Therefore, the k-factor values have been assumed here as:

$$k = \frac{1}{k_1}, \quad \text{for } k_1 = 2, 3, \ldots, 14 \tag{23}$$

The first column contains the consecutive values of k_1. The next columns contain the number of iterations needed to stabilize the solution with a precision of 0.0001. An iterative process was terminated when the values of parameters in consecutive iterations have differed by less than 0.0001. Each column contains

results for a different m-value. The first part of Table 2 concerns the example of '$y = ax$', and the second part the example of '$y = ax + b$'. Analyzing minima values, we can come to the conclusion that for most cases the fastest convergence is observed for $k_1 = m - 1$. The k-factor corresponding to such k_1 value is identical to the factor in the formula of the Newton method (18). Tables 3 and 4 contain results of optimization using processes (13) and (18) for the case of '$y = ax$' (Tab. 3) and for '$y = ax + b$' (Tab. 4). For each 'm', the parameter values and maximum values of misclosures are given in consecutive iterations.

Table 2

Determination of the optimal value of the k_1-factor

k_1	$y = ax$					$y = ax+b$				
	m					m				
	4	6	8	10	12	4	6	8	10	12
2	10	-	-	-	-	8	-	-	-	-
3	3	15	-	-	-	3	13	-	-	-
4	4	5	15	-	-	5	5	11	-	-
5	5	3	5	12	-	6	3	4	-	-
6	6	3	4	7	-	8	5	5	11	16
7	7	4	4	5	12	9	7	5	7	7
8	9	5	4	4	8	11	8	5	6	7
9	10	6	4	4	4	12	9	6	6	7
10	11	7	5	5	5	14	11	7	5	6
11	12	7	6	5	5	16	12	8	6	5
12	13	8	6	6	4	17	13	9	7	6
13	14	9	7	7	5	19	14	10	7	7
14	15	9	7	7	5	20	16	11	8	8
$v_{1\max}$* [mm]	21					22				
$v_{2\max}$** [mm]	18	18	17	17	17	17	16	16	15	15

* maximal shift in the LS method ($m = 2$)
** maximal shift in the LmOCM method ($m > 2$)

Table 3

Results of optimization ($y = ax$). Table contains the value of the a parameter and maximum shift m after fitting axis into a set of points

Iter.	m									
	4		6		8		10		12	
	a	$v_{\max}$	a	$v_{\max}$	a	$v_{\max}$	a	$v_{\max}$	a	$v_{\max}$
1	0.1009	0.021	0.1009	0.021	0.1009	0.021	0.1009	0.021	0.1009	0.021
2	0.1013	0.018	0.1013	0.018	0.1013	0.018	0.1012	0.019	0.1011	0.019
3	**0.1012**	**0.018**	**0.1014**	**0.018**	**0.1014**	**0.017**	**0.1014**	**0.017**	0.1014	0.018
4	**0.1012**	**0.018**	**0.1014**	**0.018**	**0.1014**	**0.017**	**0.1014**	**0.017**	**0.1014**	**0.017**
5	**0.1012**	**0.018**	**0.1014**	**0.018**	**0.1014**	**0.017**	**0.1014**	**0.017**	**0.1014**	**0.017**

Table 4

Results of optimization ($y = ax + b$). Table contains the value of the a and b parameters and maximum shift m after fitting axis into a set of points

iter.	m															
	4			6			8			10			12			
No.	a	b	v_{max}	a	b	v_{max}	a	b	v_{max}	a	b	v_{max}	a	b	v_{max}
1	0.1006	0.4021	0.022	0.1006	0.4021	0.022	0.1006	0.4021	0.022	0.1006	0.4021	0.022	0.1006	0.4021	0.022
2	0.1011	0.4036	0.017	0.1010	0.4036	0.018	0.1010	0.4025	0.019	0.1010	0.4014	0.019	0.1011	0.4006	0.020
3	0.1009	0.4042	0.017	0.1006	0.4074	0.016	0.1004	0.4090	0.016	0.1003	0.4085	0.017	0.1004	0.4074	0.018
4	**0.1009**	**0.4043**	**0.017**	0.1008	0.4065	0.016	0.1008	0.4068	0.016	0.1010	0.4052	0.016	0.1011	0.4038	0.016
5	**0.1009**	**0.4043**	**0.017**	**0.1008**	**0.4064**	**0.016**	**0.1007**	**0.4073**	**0.016**	0.1007	0.4079	0.015	0.1005	0.4089	0.015
6	**0.1009**	**0.4043**	**0.017**	**0.1008**	**0.4064**	**0.016**	**0.1007**	**0.4073**	**0.016**	**0.1007**	**0.4077**	**0.015**	**0.1007**	**0.4078**	**0.015**

In process (13), the k-factor has been assumed as:

$$k = \frac{1}{m - 1} \tag{24}$$

In the case analyzed here, both methods (13) and (18) provided the same results in each iteration. The results with a stabilized solution (with a precision of 0.0001 for parameters, and 0.001 for maximum misclosure) are in bold. The result of the first iteration is merely the Least Squares (LS) solution. The maximal shift needed for the axis alignment according to the LS method ($v_{\max}$) is above 2 cm in both examples. In the first example, it amounts to 21 mm and in the second example 22 mm. This value was reduced using the LmOCM method. In the first example this value amounts to 18 mm for m = 4 or 6, and 17 mm for m = 8, 10 or 12. In the second example, this value has been maximally reduced to 15 mm (for m = 12). In some cases, this reduction can be of critical significance (e.g. in such engineering tasks where the values of shifts are limited to a range less than 20 mm). The $v_{\max}$-values for different 'm' differ by less than 1 mm in the first example, and by less than 2 mm in the second example. It is clearly seen, that for the lower required maximal shifts, the higher m-values should be taken in the adjustment process – the lowest shift values of 15 mm for the greater m-values (10 and 12) are visible in Table 4.

Summary and Conclusions

In the paper, a generalization of the Least Squares optimization method is proposed. A new method is based on the criterion of minimization of a sum of misclosures raised to the power of 'm'. This criterion allows one to reduce the maximum misclosure in the optimization problem in comparison to the LS method. A crucial problem in the use of this method is ensuring proper convergence of the computational process. Two techniques of finding the minimum of the objective function were tested: the simple gradient zeroing method and the Newton method. Both methods employ an iterative process. The formula of the computational process of the first technique contains the k-factor. This factor warrants the proper convergence of the computational process. It is proposed here, that the value of this factor is calculated as $k = (m - 1)^{-1}$ if the sum of misclosures to the power of 'm' is the optimization criterion. In the case study considered here, the two optimization techniques provided the same results. The results of this case study confirmed the property of reducing the maximum shift in the task of the alignment of the axis into a set of points in comparison to the LS method.

References

AVRIEL M. 2003. *Nonlinear Programming: Analysis and Methods*. Dover Publishing.

CASPARY W., HAEN W. 1990. *Simultaneous estimation of location and scale parameters in the context of robust M-estimation*. Manuscripta Geodaetica, 15: 273–282.

CELLMER S. 2014. *Least fourth powers: optimisation method favouring outliers*. Survey Review, 47(345): 417. DOI: https://doi.org/10.1179/1752270614Y.0000000142.

CHANG X.W., GUO Y. 2005. *Huber's M-estimation in relative GPS positioning: computational aspects*. Journal of Geodesy, 79: 351–362. DOI: https://doi.org/10.1007/s00190-005-0473-y.

DUCHNOWSKI R., WIŚNIEWSKI Z. 2012. *Estimation of the shift between parameters of functional models of geodetic observations by applying M split estimation*. Journal of Surveying Engineering, 138: 1–8. DOI: 10.1061/(ASCE)SU.1943-5428.0000062.

FLETCHER R. 1987. *Practical methods of optimization*. 2nd ed. John Wiley & Sons, New York.

HAMPEL F.R., RONCHETTI E., ROUSSEEUW P.J., STAHEL W.A. 1986. *Robust statistics: the approach based on influence function*. Wiley, New York.

HUBER P.J. 1981. *Robust statistics*. Wiley, New York.

KADAJ R. 1988. *Eine verallgemeinerte Klasse von Schätzverfahren mit praktischen Anwendungen*. Z Vermessungs-wesen, 113(4): 157–166.

KAMIŃSKI W., WIŚNIEWSKI Z. 1992. *Analysis of some, robust, adjustment methods*. Geodezja i Kartografia, 41(3-4): 173-182.

KOCH K.R. 1996. *Robuste Parameterschätzung*. Allgemeine Vermessungs-Nachrichten, 103(11): 1–18.

LIEW C.K. 1976. *Inequality constrained least squares estimation*. Journal of the American Statistical Association, 71: 746–751.

MARTINS T.C., TSUZUKI M.S.G. 2009. *Placement over containers with fixed dimensions solved with adaptive neighborhood simulated annealing*. Bulletin of the Polish Academy of Sciences. Technical Sciences, 57(3): 273-280.

MEAD J.L., RENAUT R.A. 2010. *Least squares problems with inequality constraints as quadratic constraints*. Linear Algebra and its Applications, 432(8): 1936–1949.

NEUMANN J. von, MORGENSTERN O. 1947. *Theory of games and economic behavior*. Princeton Univ. Press. Princeton, New Jersey.

NOCEDAL J., WRIGHT S.J. 1999. *Numerical Optimization*. Springer-Verlag, Berlin.

SKAŁA-SZYMAŃSKA M., CELLMER S., RAPIŃSKI J. 2014. *Use of Nelder-Mead simplex method to arc fitting for railway track realignment*. The 9th International Conference Environmental Engineering, selected papers. DOI: 10.3846/enviro.2014.244.

WERNER H.J. 1990. *On inequality constrained generalized least-squares estimation*. Linear Algebra and its Applications, 27: 379–392. DOI: http://dx.doi.org/10.1179/1752270614Y.0000000142.

WIŚNIEWSKI Z. 2009. *Estimation of parameters in a split functional model of geodetic observations (M split estimation)*. Journal of Geodesy, 83: 105–120. DOI: https://doi.org/10.1007/s00190-008-0241-x.

WIŚNIEWSKI Z. 2010. *M split(q) estimation: estimation of parameters in a multi split functional model of geodetic observations*. Journal of Geodesy, 84: 355–372. DOI: https://doi.org/10.1007/s00190-010-0373-7.

XU P. 1989. *On robust estimation with correlated observations*. Bulletin Géodésique, 63: 237–252.

YANG Y.1999. *Robust estimation of geodetic datum transformation*. Journal of Geodesy, 73: 268–274. DOI: https://doi.org/10.1007/s001900050243.

YANG Y., SONG I., XU T. 2002. *Robust estimation for correlated observations based on bifactor equivalent weights*. Journal of Geodesy, 76: 353–358. DOI: https://doi.org/10.1007/s00190-002-0256-7.

ZHONG D. 1997. *Robust estimation and optimal selection of polynomial parameters for the interpolation of GPS geoid heights*. Journal of Geodesy, 71: 552–561. DOI: https://doi.org/10.1007/s001900050123.

ZHU J. 1996. *Robustness and the robust estimate*. Journal of Geodesy, 70: 586–590. DOI: https://doi.org/10.1007/BF00867867.