

COMPARISON OF THE METHOD OF LEAST SQUARES AND THE SIMPLEX METHOD FOR PROCESSING GEODETIC SURVEY RESULTS

POROVNANIE METÓDY NAJMENŠÍCH ŠTVORCOV A SIMPLEXOVEJ METÓDY PRI SPRACOVANÍ VÝSLEDKOV GEODETICKÝCH MERANÍ

Silvia GAŠINCOVÁ¹, Juraj GAŠINEC²

¹ *Assoc. prof., Ing., PhD., Institute of Geodesy, Cartography and Geographic Information Systems, Faculty of Mining, Ecology, Process Control and Geotechnologies, Technical University of Košice Park Komenského 19, 043 84 Košice, Slovak Republic, +421 55 602 2846
e-mail: silvia.gasincova@tuke.sk*

² *Assoc. prof., Ing., PhD., Institute of Geodesy, Cartography and Geographic Information Systems, Faculty of Mining, Ecology, Process Control and Geotechnologies, Technical University of Košice, Park Komenského 19, 043 84 Košice, Slovak Republic, +421 55 602 2846
e-mail: juraj.gasinec@tuke.sk*

Abstract

The present paper is devoted to the use of the simplex method in the processing of results from geodetic measurements as compared with the standard used method of least squares. Using the simplex method, a minimization problem is usually solved in a standard tabular form by rearranging lines and columns in order to find an optimal solution. The paper points out the simpler, more stable and more efficient way to solve a problem of linear programming through a matrix of relations.

Abstrakt

Predložený príspevok je venovaný použitiu simplexovej metódy pri spracovaní výsledkov geodetických meraní v porovnaní so štandardne používanou metódou najmenších štvorcov (MNS). Minimalizačný problém je pri simplexovej metóde zvyčajne riešený tabuľkovou formou, preskupovaním stĺpcov a riadkových operácií, ktorých cieľom je nájsť optimálne riešenie. V príspevku je poukázané na numericky nenáročnejšiu, stabilnejšiu a efektívnejšiu cestu riešenia problému lineárneho programovania pomocou maticových vzťahov.

Key words: method of least squares, simplex method, L1 norm, L2 norm, outlier detection

1 INTRODUCTION

In geodetic practice, the least square method (LSM) ranks among standard used processing methods. This method is based on the vector of corrections of the L2 norm which together with the L1 norm is most often applied to the processing of results of geodetic measurements. However, a prerequisite for the proper functioning of the LSM is a normal distribution of errors; otherwise the created probabilistic model is not correct. For this reason, the geodetic practice started to use variant processing methods as well. Such methods include, for example, robust estimation procedures that preserve their function in a certain neighbourhood of a normal distribution, i.e. not fail in case of a moderate failure to comply with this requirement. The more the method is resistant, the more it is robust. Several types of these methods are known, from the robust M-estimates [5],[6],[7],[8] to linear programming methods to which undoubtedly the simplex method belongs. The LSM and the simplex method will be demonstrated on two examples; on the examples of a regression line and a geodetic network. In the first case, the models of a regression line and a geodetic network are loaded by the normal distribution of measurements, in the other the line and the geodetic network are loaded by the value of an experimental outlier.

2 METHOD OF LEAST SQUARES

The essence of this method lies in minimizing the sum of squares of deviations in measurements of the behaviour of any quantity or physical phenomenon (3). The least square method is based on the condition of so-called *L2-norm*; the norm is the number assigned to each n-dimensional vector of residual deviations

$v = (v_1, v_2, \dots, v_n)$ that in some sense characterizes its size [3], [13]. In geodesy, the most commonly used types of objective functions are as follows:

$$\rho(v) = \left(\sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}} = \min. \quad i \in \langle 1, n \rangle \quad (1)$$

Where:

p - a parameter defining a special type of an objective function,

v_i - a vector of residual deviations (vector of corrections).

Assuming that $p = 2$ ($L2$ norm), the objective function is as follows:

$$\rho(v) = \left(\sum_{i=1}^n |v_i|^2 \right)^{\frac{1}{2}} = \min \quad (2)$$

which leads to the least square method that leads under certain conditions to the most reliable estimates of unknown quantities, and hence it is the most commonly used method in geodetic practice. The least square method will be explained on a one-dimensional linear model, while all the estimation methods will be demonstrated on the example of a regression line.

Let us assume that the following linear relationship exists between a variable y , variables X (Fig. 1) and a random component u_i :

$$\begin{aligned} y_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + u_i \quad \text{for } i = 1, 2, \dots, n \\ y &= \beta_0 + \beta X + u \end{aligned} \quad (3)$$

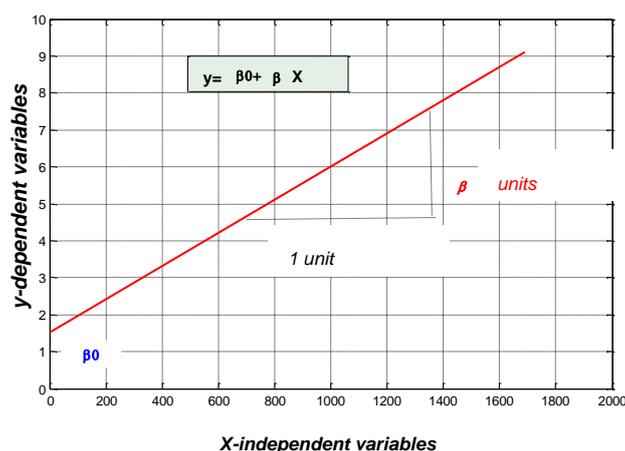


Fig. 1 One-dimensional linear model

Where:

y_i - dependent variables,

X_1, X_2, \dots, X_k - independent variables,

β_0 - the parameter indicating a value of the variable y provided that the variable X is equal to 0, the so-called intercept

β - the regression coefficient defining the slope of the regression line,

u_i - the random component.

Provided that we have n pairs of the independent variable X and the dependent variable Y $[x_1, y_1], [x_2, y_2], \dots, [x_n, y_n]$ and are sufficiently convinced of the linear dependence, we can construct the line that best describes this fact:

$$\hat{y}_i = b_0 + b_1 x_i \quad \text{kde } i = 1, 2, \dots, n \dots \dots \text{vyrovnávajúca regresná priamka} \quad (4)$$

adjusting regression line

Where:

\hat{y}_i - the adjusted one (theoretical value) of the dependent variable,

x_i - the dependent variable for the i -th observation,

b_0 – the point estimate of the parameter β_0 ,

b – the point estimate of the parameter β .

For residual deviations (in the geodesy vector of corrections) $\mathbf{v} = \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, the relation $v_i = y_i - \hat{y}_i$ applies.

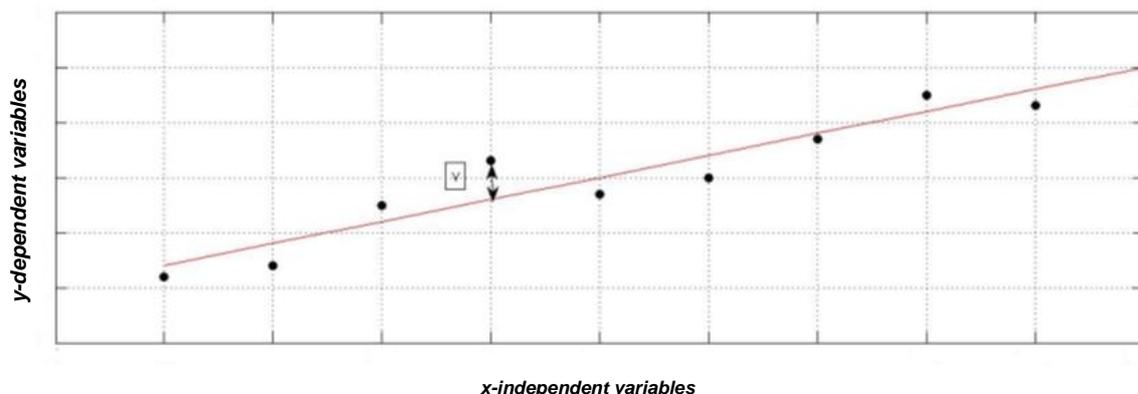


Fig. 2 Residual deviations

Since the residual deviations may take a positive as well as negative value, they null each other. The nullifying problem can be solved just by applying the LSM, the principle of which consists in the sum of squares of deviations and not of the deviations themselves. The LSM formulation is then as follows:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 = \sum_{i=1}^n (v_i)^2 = \min. \quad (5)$$

The minimum of a function of two variables can be found by placing its partial derivatives under both variables (coefficients b_0 and b_1) equal to zero:

$$\frac{\partial \left(\sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 \right)}{\partial (b_1)} = 2 \left(\sum_{i=1}^n (y_i - b_0 - b_1 x_i)(-x_i) \right) = 0, \quad (6)$$

$$\frac{\partial \left(\sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 \right)}{\partial (b_0)} = 2 \left(\sum_{i=1}^n (y_i - b_0 - b_1 x_i)(-1) \right) = 0.$$

By an appropriate algebraic transformation, we get a system of normal equations of two variables b_0 and b_1 :

$$\sum_{i=1}^n y_i = n b_0 + b_1 \sum_{i=1}^n x_i, \quad (7)$$

$$\sum_{i=1}^n y_i x_i = b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2.$$

For the variables b_0 and b_1 it is true:

$$b_0 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad b_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (8)$$

In the processing of geodetic measurements, methods of adjustment are used, by which the most probable value of a quantity and its accuracy characteristics are determined. In determining the parameters of the regression line, a mediating adjustment was used [1],[10],[9],[12], where the relationship between the measured and unknown variables is expressed by the intermediating function of the searched unknown parameters, the so-called estimates:

$$y_i = l_i + v_i = f(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k) \quad (9)$$

For the vector of residues, the following is valid:

$$\mathbf{v} = \mathbf{A}\hat{\Theta} - \mathbf{l}, \quad (10)$$

leading to the Gauss-Markov model:

$$\begin{aligned} \mathbf{v} &= \mathbf{A}\hat{\Theta} - \mathbf{l} \\ \sum_{i=1}^n \sigma_0^2 \mathbf{Q}_i \end{aligned} \quad (11)$$

It applies for LSM:

$$\rho = \sum_{i=1}^n (p_i v_i)^2 = \sum_{i=1}^n v^T \mathbf{P} v = \min., \quad \mathbf{P} \dots \dots \text{matica váhových koeficientů} \quad (12)$$

matrix of weight coefficients

The minimum of the function of two variables can be found again by placing the partial derivatives of the function (12) equal to zero $\frac{\partial(v^T \mathbf{P} v)}{\partial \hat{\Theta}}$ and so getting the system of two equations, from which the searched variables (estimates) can be determined:

$$(\mathbf{A}^T \mathbf{P} \mathbf{A} \hat{\Theta}) - (\mathbf{A}^T \mathbf{P} \mathbf{l}) = 0 \rightarrow (\mathbf{N} \hat{\Theta}) - (\mathbf{A}^T \mathbf{P} \mathbf{l}) = 0, \quad \mathbf{N} - \text{matica koeficientů normálních rovnic} \quad (13)$$

matrix of coefficients of normal equations

When using the LSM for searching the estimates, (indices), the following applies after adjusting the parameters:

$$\hat{\Theta} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{P} \mathbf{l}) = \mathbf{N}^{-1} (\mathbf{A}^T \mathbf{P} \mathbf{l}). \quad (14)$$

3 SIMPLEX METHOD

The simplex method is an iterative computational procedure that is used to find optimal solutions whereas the objective (minimized) function must be in a canonical form. The minimization problem is typically solved by means of methods of linear programming in a tabular form by rearranging columns and line unless the objective function optimization is reached. The paper presents a simpler procedure of processing based on the principle of matrix solution [2],[4]. This method will again be presented on the example of a regression line. The model line was chosen not only because of its simplicity, but especially considering the fact that in the processing of results from geodetic measurements there is a need, quite often, to determine the accuracy of the measured length which can be expressed by the following relationship:

$$STD = appm + b \dots \text{modified form of regression line} \quad (15)$$

Where:

STD - the standard deviation of a measured length,

a - the parameter reflecting the impact of errors dependent on the measured length which take into account the influence of the physical environment,

b - the parameter reflecting the impact of the errors independent of the measured length.

The functional relationship for the correction of the intermediate variable can be expressed as follows:

$$\mathbf{v}_{(n,1)} = \mathbf{A}_{n,k} \cdot \hat{\Theta}_{k,1} - \mathbf{f}_{n,1}, \quad (16)$$

where *v* is the vector of corrections, *A* the matrix of coefficients, $\hat{\Theta}$ is the vector of unknown parameters (estimates), *f* is the vector of observations, *n* is a number of measurements, and *k* is a number of necessary measurements (determined parameters). Within the search of an optimal solution, first the measured values *f* are divided to the **basic variables** (needed measurements) and **non-basic variables** ($r = n - k$, *r* - a number of redundant measurements).

$$\begin{bmatrix} \mathbf{v}_{(1)} \\ \mathbf{v}_{(2)} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_{(1)} \\ \mathbf{A}_{(2)} \end{bmatrix} \cdot \hat{\Theta} = \begin{bmatrix} \mathbf{f}_{(1)} \\ \mathbf{f}_{(2)} \end{bmatrix}. \quad (17)$$

The vector *f* can be broken down as follows:

$$\begin{bmatrix} \mathbf{f}_{(1)} \\ \mathbf{f}_{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{l}_{(1)}^o \\ \mathbf{l}_{(2)}^o \end{bmatrix} - \begin{bmatrix} \mathbf{l}_{(1)} \\ \mathbf{l}_{(2)} \end{bmatrix}. \quad (18)$$

In the case of observations with varying accuracy, the L1 norm can be derived from the objective function L2 norm:

$$\rho_{L2} = \mathbf{v}^T \mathbf{P} \mathbf{v} = \sum_{i=1}^n \mathbf{v}_i^2 p_{ii} = \sum_{i=1}^n (\mathbf{v}_i \sqrt{p_{ii}})(\mathbf{v}_i \sqrt{p_{ii}}) = \min. \quad (19)$$

To determine the individual elements of the weighting matrix \mathbf{P} , the Cholesky decomposition is used:

$$\mathbf{P} = \mathbf{S} \mathbf{S}^T. \quad (20)$$

For the functional relationships of measured values, it applies:

$$\mathbf{S} \mathbf{A} \hat{\Theta} \approx \mathbf{S} \mathbf{f}. \quad (21)$$

The relationship for the objective function for the L1 norm after the adjustment takes the form:

$$\rho_{L1} = (\text{vecd}(\mathbf{S}))^T |\mathbf{v}| \rightarrow \min. \quad (22)$$

Where $\text{vecd}()$ is a column vector and the matrix $|\mathbf{v}|$ has the following form:

$$|\mathbf{v}| = \begin{bmatrix} |v_1| \\ |v_2| \\ \vdots \\ |v_n| \end{bmatrix} \text{ with elements on diagonal.} \quad (23)$$

For calculating the unknown parameters, it applies:

$$\hat{\Theta} = (\mathbf{S}_{(1)} \mathbf{A}_{(1)})^{-1} \mathbf{S}_{(1)} \mathbf{f}_{(1)} = \mathbf{A}_{(1)} \mathbf{f}_{(1)}, \quad (24)$$

Where $\mathbf{S}_{(1)}$ is the diagonal matrix with dimensions $(k \times k)$.

The vector of corrections is divided into two parts:

$$\mathbf{v}_{(1)} = \mathbf{f}_{(1)} - \mathbf{A}_{(1)} \hat{\Theta} = \mathbf{f}_{(1)} - \mathbf{A}_{(1)} \mathbf{A}_{(1)}^{-1} \mathbf{f}_{(1)} = \mathbf{0}, \dots \text{ residues are always equal 0} \quad (25)$$

For the non-zero vector of corrections $\mathbf{v}_{(2)}$, it applies:

$$\mathbf{v}_{(2)} = \mathbf{f}_{(2)} - \mathbf{A}_{(2)} \hat{\Theta} = \mathbf{f}_{(2)} - \mathbf{A}_{(2)} \mathbf{A}_{(1)}^{-1} \mathbf{f}_{(1)}. \quad (26)$$

The assignment is solved by linear programming methods and is defined as follows:

$$\begin{aligned} \mathbf{r}^T \mathbf{x} &= \rho_{L1}, \\ \mathbf{A} \mathbf{x} &= \mathbf{g}. \end{aligned} \quad (27)$$

where $\mathbf{x} \geq 0$, \mathbf{r} is a *minimizing funkcion* also called „*cost vector*“ of size of $t \times 1$, $t=2k+2n$, \mathbf{x} is the vector of non-negative variables, \mathbf{A} is the matrix of coefficients.

In order to meet the condition $\mathbf{x} \geq 0$ for linear programming, the calculated estimates and the vector of corrections is divided into two non-negative components:

$$\begin{aligned} \hat{\Theta} &= \delta - \gamma \quad \delta, \gamma \geq 0, \\ \mathbf{v} &= \mathbf{y} - \mathbf{r} \quad \mathbf{y}, \mathbf{r} \geq 0 \end{aligned} \quad (28)$$

$$\mathbf{v}_i = y_i \quad \text{ak } v_i > 0, r_i = 0, -\mathbf{v}_i = r_i \quad \text{ak } v_i < 0, y_i = 0,$$

$$\mathbf{v}_i = 0 \quad \text{ak } y_i = 0, r_i = 0, |v_i| = y_i + r_i$$

The target function (27) can be written in a matrix form as well:

$$\begin{bmatrix} \mathbf{S}_{(1)} \mathbf{A}_{(1)} & -\mathbf{S}_{(1)} \mathbf{A}_{(1)} & \mathbf{I} & \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{S}_{(2)} \mathbf{A}_{(2)} & -\mathbf{S}_{(2)} \mathbf{A}_{(2)} & \mathbf{0} & \mathbf{I} & \mathbf{0} & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \delta \\ \gamma \\ \mathbf{S}_{(1)} \mathbf{y}_{(1)} \\ \mathbf{S}_{(2)} \mathbf{y}_{(2)} \\ \mathbf{S}_{(1)} \mathbf{z}_{(1)} \\ \mathbf{S}_{(2)} \mathbf{z}_{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{(1)} \mathbf{f}_{(1)} \\ \mathbf{S}_{(2)} \mathbf{f}_{(2)} \end{bmatrix} \quad (29)$$

$$\begin{bmatrix} 0^T & 0^T & \mathbf{1}^T & \mathbf{1}^T & \mathbf{1}^T & \mathbf{1}^T \end{bmatrix} \begin{bmatrix} \delta \\ \gamma \\ \mathbf{S}_{(1)} \mathbf{y}_{(1)} \\ \mathbf{S}_{(2)} \mathbf{y}_{(2)} \\ \mathbf{S}_{(1)} \mathbf{z}_{(1)} \\ \mathbf{S}_{(2)} \mathbf{z}_{(2)} \end{bmatrix} = \phi_{L1}$$

where $\mathbf{1}$ is the all-ones vector ($\mathbf{1}^T = [1, 1 \dots 1]$). The default simplex table has the following structure:

$$T = \begin{bmatrix} \rho_{L1} & \mathbf{z}^T \\ \mathbf{g} & \mathbf{A} \end{bmatrix} =$$

$$= \begin{array}{ccccccc} 0 & 0^T & 0^T & \mathbf{1}^T & \mathbf{1}^T & \mathbf{1}^T & \mathbf{1}^T \\ \mathbf{S}_{(1)} \cdot \mathbf{f}_{(1)} & \mathbf{S}_{(1)} \cdot \mathbf{A}_{(1)} & -\mathbf{S}_{(1)} \cdot \mathbf{A}_{(1)} & \mathbf{I} & 0 & -\mathbf{I} & 0 \\ \mathbf{S}_{(2)} \cdot \mathbf{f}_{(2)} & \mathbf{S}_{(2)} \cdot \mathbf{A}_{(2)} & -\mathbf{S}_{(2)} \cdot \mathbf{A}_{(2)} & 0 & \mathbf{I} & 0 & -\mathbf{I} \end{array} \quad (30)$$

The calculation can be tabulated using the rearrangement operations of columns and lines; in the paper, the procedure of transforming the simplex table using a “pivot” matrix \mathbf{K} was used:

$$\mathbf{K} = \begin{bmatrix} \mathbf{I} & 0^T & \mathbf{1}^T \\ \mathbf{0} & \mathbf{S}_{(1)} \mathbf{A}_{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{(2)} \mathbf{A}_{(2)} & \mathbf{I}^* \end{bmatrix}, T^* = \mathbf{K}^{-1} \cdot T, \quad (31)$$

\mathbf{I}^* - the variable pivot matrix; its diagonal elements have a value +1, or -1 to meet the condition $\text{met}v_{(2)} \geq 0$.

The iterative process is carried out by multiplying the initial simplex table (30) with the pivot matrix in inverse form. The structure of the transformed simplex table after the adjustment is as follows:

$$T^* = \begin{array}{ccccccc} \mathbf{u}_1 & 0^T & 0^T & \mathbf{1}^T + \mathbf{1}^T \mathbf{U}_2 & \mathbf{1}^T - \mathbf{1}^T \mathbf{I}^* & \mathbf{1}^T - \mathbf{1}^T \mathbf{U}_2 & \mathbf{1}^T + \mathbf{1}^T \mathbf{I}^* \\ \mathbf{A}_{(1)}^{-1} \cdot \mathbf{f}_{(1)} & \mathbf{I} & -\mathbf{I} & \mathbf{A}_{(1)}^{-1} & \mathbf{0} & -\mathbf{A}_{(1)}^{-1} & \mathbf{0} \\ \mathbf{u}_3 & \mathbf{0} & \mathbf{0} & \mathbf{U}_2 & \mathbf{I}^* & \mathbf{I}^* & -\mathbf{I}^* \end{array} \quad (32)$$

Where the elements of the table can be determined according to the following relationship:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{1}^T \mathbf{I} \mathbf{S}_{(2)} \mathbf{A}_{(2)} \mathbf{A}_{(1)}^{-1} \mathbf{f}_{(1)} - \mathbf{1}^T \mathbf{I}^* \mathbf{S}_{(2)} \mathbf{f}_{(2)}, \\ \mathbf{U}_2 &= \mathbf{I}^* \mathbf{S}_2 \mathbf{A}_{(2)} (\mathbf{S}_{(1)} \mathbf{A}_{(1)})^{-1}, \\ \mathbf{u}_3 &= -\mathbf{I}^* \mathbf{S}_{(2)} \mathbf{A}_{(2)} \mathbf{A}_{(1)}^{-1} \mathbf{f}_{(1)} + \mathbf{I}^* \mathbf{S}_{(2)} \mathbf{f}_{(2)}. \end{aligned} \quad (33)$$

The objective function is transformed to minimize the function \mathbf{r} so-called reduced „cost“ vector (34) which can be simplified in comparison with the original form, because its elements $\mathbf{1}^T - \mathbf{1}^T \mathbf{I}^*$ and $\mathbf{1}^T - \mathbf{1}^T \mathbf{U}_2$ have a value of 0 or 2.

$$\mathbf{r}^T = [\mathbf{0}^T, \mathbf{0}^T, \mathbf{1}^T + \mathbf{1}^T \mathbf{U}_2, \mathbf{1}^T - \mathbf{1}^T \mathbf{I}^*, \mathbf{1}^T - \mathbf{1}^T \mathbf{U}_2, \mathbf{1}^T + \mathbf{1}^T \mathbf{I}^*] \quad (34)$$

The optimal solution is found when for all the elements of the vector $\mathbf{r} \geq 0$ applies which can be transformed to the following condition:

$$-\mathbf{1}^T \leq \mathbf{S}_{(2)} \mathbf{A}_{(2)} \mathbf{S}_{(1)} \mathbf{A}_{(1)} \leq \mathbf{1}^T \quad (35)$$

4 EMPIRICAL DEMONSTRATION

In the following chapter, the method of least squares and the simplex method are presented on the example of a regression line and a geodetic network. The regression line and the geodetic network in the first example are loaded by a normal distribution of measurements; due to the investigation of the properties of the used estimation methods, the line and the geodetic network are loaded by an experimental outlier before the adjustment.

4.1 Application of estimation methods using the example of a regression line loaded by a normal distribution of measurements

The regression line in this case is expressed by the equation of a rangefinder as follows:

$$STD = \sigma_D = 2. ppm d + 3[mm] \dots \text{rovnic} \text{a regresnej priamky}, \quad (36)$$

Where:

STD - a standard deviation of the measured length,
d - the measured length,
ppm - parts per million (10^{-6}).

The deterministic model of the regression line loaded by a normal distribution of measurements is presented in a graphic form in **Chyba! Nenalezen zdroj odkazů..** The experimental values shown in blue are simulated in the MATLAB environment by the function *normrnd* (normal random numbers).

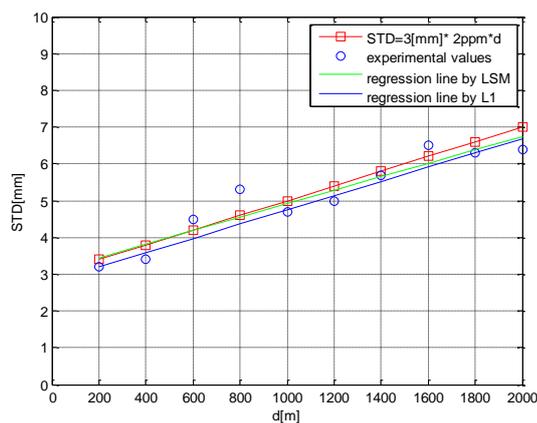


Fig. 4 Graphical interpretation of LSM and L1 norms on the example of the regression line loaded by normal distribution

From the results of the adjustment (Tab. 2) of such proposed regression line, it is evident that in case of the load of the line by normal distribution of measurements neither the LSM, nor the simplex method in the set of measured data revealed any outlier (“which go beyond”) value and the results of the adjustment are very similar. The graph of the regression line behaviour after the adjustment by individual estimate procedures is presented in **Chyba! Nenalezen zdroj odkazů.**

4.2 Application of estimation methods using the example of the regression line loaded by an experimental outlier

With regard to the fact that the present contribution is devoted to the issue of estimation methods which allow to reveal so-called outlier measurements in the set of measured data, and for which the arithmetic mean was not selected as the centrality parameter, the regression line was deliberately loaded with the only outlier (Fig. 5) in order to track the performance of such methods. In the figure, the experimental values are shown in blue; the line behaviour in deterministic form is shown in red.

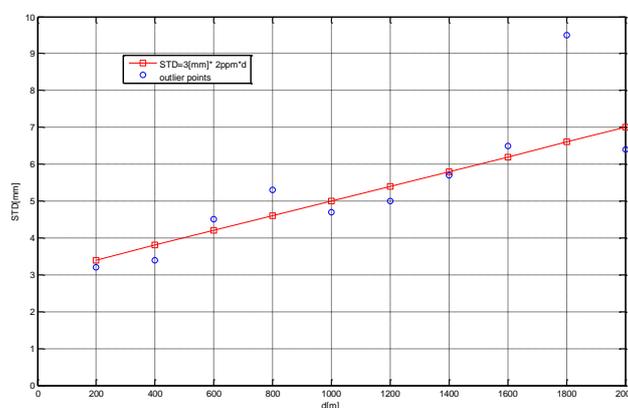


Fig. 5 Deterministic form of the regression line loaded by an experimental outlier

Such modified regression line was adjusted first by the method of least squares, and consequently by the simplex method. The results of adjustments of the regression line loaded by an experimental outlier after the adjustment by the LSM and the simplex method are interpreted in Tab. 4 and in Fig. 6.

arrive at a mutually similar results and the simplex method found an outlier measurement at a measured length of 1800 m.

4.3 Adjustment of a geodetic network loaded by a normal distribution of measurements

Characteristics of the presented estimation methods will also be investigated on the example of a geodetic network. This is a simulated geodetic network where the point No. 6 is the point being determined whose position is defined by triangulation (angular) measurements (Fig. 7) . Such a network model was proposed on the grounds that the geodetic practice is often encountered with the task when it is necessary to determine the location of an inaccessible point which can be done just through triangulation measurements. The geodetic network will be processed as a binding network by both network estimative procedures, first by a normal distribution of measurements, and consequently after it is loaded by outlier experimental values.

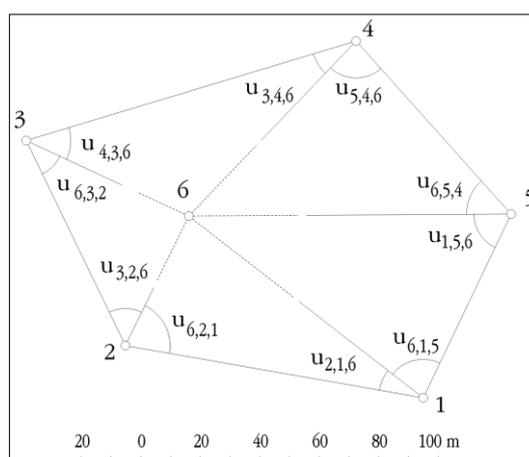


Fig. 7 Structure of the geodetic network loaded by a normal distribution of measurements

Tab. 1 Results of adjustment of the geodetic network loaded by a normal distribution through the LSM

L-S-P	l	l	l^{\wedge}	v	v	vC^{\wedge}	p	T.Baarda	T.Pope	$s(v)$	$s(l^{\wedge})$	r^*
angle	[g]	[g]	[g]	[cc]	[cc]	[cc]		-B-	-P-	[cc]		
2-1-6	31.04550	31.04500	31.04559	4.95	5.94	5.94	1.0000	1.58	1.43	3.76	4.15	0.88
6-1-5	86.04331	86.04290	86.04321	4.08	3.09	3.09	1.0000	0.82	0.74	3.76	4.15	0.88
3-2-6	57.35031	57.35090	57.35053	-5.90	-3.72	-3.72	1.0000	1.17	1.06	3.18	3.51	0.63
6-2-1	82.30387	82.30390	82.30366	-0.26	-2.45	-2.45	1.0000	0.77	0.70	3.18	3.51	0.63
4-3-6	46.63242	46.63190	46.63229	5.21	3.95	3.95	1.0000	1.18	1.07	3.33	3.68	0.69
6-3-2	43.61620	43.61630	43.61633	-0.98	0.29	0.29	1.0000	0.09	0.08	3.33	3.68	0.69
5-4-6	94.62331	94.62290	94.62322	4.11	3.15	3.15	1.0000	0.84	0.76	3.74	4.13	0.88
6-4-3	32.88348	32.88290	32.88358	5.81	6.76	6.76	1.0000	1.81	1.64	3.74	4.13	0.88
1-5-6	71.37301	71.37290	71.37303	1.07	1.28	1.28	1.0000	0.33	0.30	3.83	4.23	0.92
6-5-4	54.12859	54.12900	54.12857	-4.08	-4.29	-4.29	1.0000	1.12	1.01	3.83	4.23	0.92

Legend:

Significance level alpha	= 0.050	B - Baarda Data snooping ($N(0,1)$)	= 2.800
Basic standard deviation of angle mesur.	= 4.000 [cc]	Number of critical measure. (Baarda data-snooping)	= 0
Estimate of the accuracy of angle m. (MINQUE)	= 4.417 [cc]	P - Pope Tau-test (Tau(r, $1-\alpha/2$))	= 2.361
Posterior standard deviation	= 4.417 [cc]	Number of critical measure. Pope Tau-test)	= 0
Critical limit s_0 posterior	= 5.569 [cc]	Adjustment efficiency	= 0.800
s_0 poster ² / s_0 aprior ²	= 1.219	Redundancy	= 8.000

The results of adjustment of thus proposed geodetic network by the method of least squares are presented in

Tab. 1. In the case of the network loaded by a normal distribution of measurements, the standard deviation of the measured angle takes a value of 4.000 cc. The accuracy of the angular measurement 4.417 cc was estimated by the method MINQUE [10],[12]. The fact whether the set of measured data was

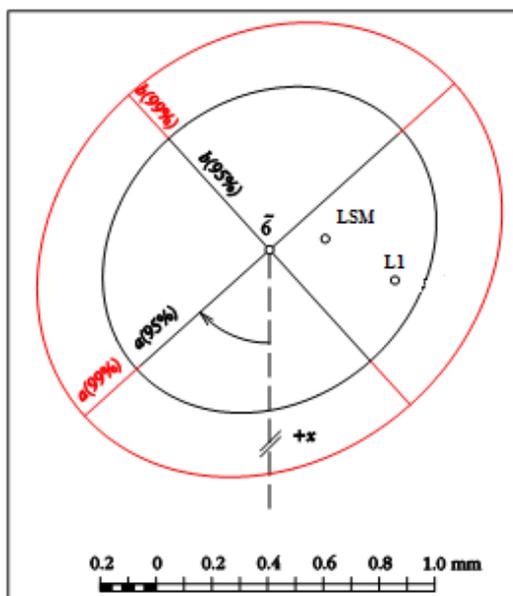


Fig. 8 Graphical interpretation of the estimation methods for adjusting the geodetic network loaded by a normal distribution of measurements

4.4 The adjustment of the geodetic network loaded by a normal distribution of measurements

To investigate the properties of alternative estimation methods enabling to detect outlier measurements that can infiltrate such a set of measured data, e.g. by a supraliminal influence of the physical environment, two measured angles were loaded in the proposed geodetic network prior to the processing (Fig. 9), once by five times the mean error of the measured angle, once by six times the estimated error. Such proposed geodetic network was again adjusted as the binding network by the (MNS) (

Tab. 5) and by the simplex method (

Tab. 7).

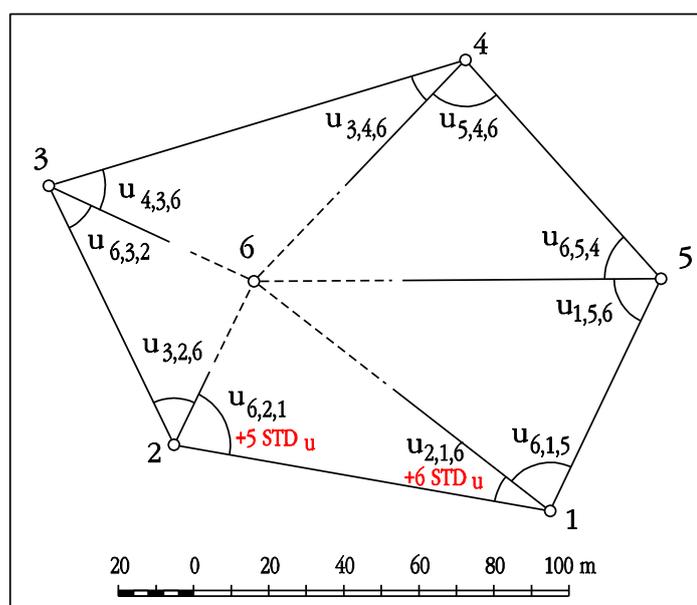


Fig. 9 Structure of the geodetic network loaded by two outlier measurements

Tab. 5 Results of adjustment of the geodetic network loaded by outlier measurements through the LSM

L-S-P	l_{\sim}	l	l^{\wedge}	v_{\sim}	v	vC^{\wedge}	p	T.Baarda	T.Pope	$s(v)$	$s(l^{\wedge})$	r^*	f
angle	[g]	[g]	[g]	[cc]	[cc]	[cc]		-B-	-P-	[cc]			
2-1-6	31.04550	31.04790	31.04581	-24.00	-20.86	-20.86	1.0000	5.55 B	2.02	3.76	10.33	0.88	66.0
6-1-5	86.04331	86.04290	86.04299	4.08	0.94	0.94	1.0000	0.25	0.09	3.76	10.33	0.88	66.0
3-2-6	57.35031	57.35090	57.34998	-5.90	-9.24	-9.24	1.0000	2.91 B	1.06	3.18	8.72	0.63	39.2
6-2-1	82.30387	82.30587	82.30421	-20.00	-16.66	-16.66	1.0000	5.25 B	1.91	3.18	8.72	0.63	39.2
4-3-6	46.63242	46.63190	46.63183	5.21	-0.71	-0.71	1.0000	0.21	0.08	3.33	9.14	0.69	44.6
6-3-2	43.61620	43.61630	43.61679	-0.98	4.94	4.94	1.0000	1.48	0.54	3.33	9.14	0.69	44.6
5-4-6	94.62331	94.62290	94.62363	4.11	7.32	7.32	1.0000	1.95	0.71	3.74	10.28	0.88	64.7
6-4-3	32.88348	32.88290	32.88316	5.81	2.60	2.60	1.0000	0.70	0.25	3.74	10.28	0.88	64.7
1-5-6	71.37301	71.37290	71.37337	1.07	4.67	4.67	1.0000	1.22	0.44	3.83	10.52	0.92	71.1
6-5-4	54.12859	54.12900	54.12823	-4.08	-7.68	-7.68	1.0000	2.01	0.73	3.83	10.52	0.92	71.1

Significance level α	= 0.050	B - Baarda Data snooping ($N(0,1)$)	= 2.800
Basic standard deviation of angular meas.	= 4.000 [cc]	Number of critical meas. (Baarda)	= 3
Estimate of precision of angular meas. (MINQUE)	= 10.984 [cc]	P - Pope Tau-test (Tau(r, $1-\alpha/2$))	= 2.361
Posterior standard deviation	= 10.984 [cc]	Number of critical meas. (Pope Tau-test)	= 0
Critical limit $s0_posterior$	= 5.569 [cc]	Redundancy	= 0.800
$s0_poster^2/s0_aprior^2$	= 7.541		
Crit. ratio $s0_poster^2/s0_aprior^2$	= 1.938		

Results of the adjustment of the geodetic network show that the LSM found two outliers just on the measured angles loaded by multiples of the median error of the measured angle. Infiltrating outlier measurements to the set of the processed data, however, was also studied through the Baarda Snooping test and the Pope test. From

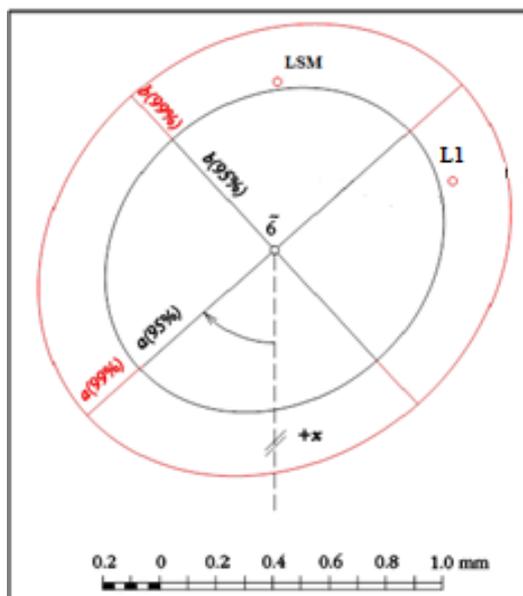


Fig. 10 Graphical interpretation of the estimation methods for adjusting the geodetic network loaded by outlier measurements

5 CONCLUSIONS

The estimation methods that were used and applied to the example of a regression line and a geodetic network showed the mutual tightness of the achievements. In the present contribution, simple but all the more illustrating examples demonstrating the positive attributes of alternative estimation methods were deliberately chosen. Among the many such methods published in foreign literature [8], the simplex method was presented in this paper which also allows to solve the problem of infiltration of outlier measurements to the set of processed data. The minimization problem of the L1-norm is normally solved using a linear programming tabular form; however, the paper presents a simpler and more efficient way to solve a linear programming problem using matrix relations.

REFERENCES

- [1] BAJTALA, M. – SOKOL, Š.: Odhad variančných komponentov z meraní v geodetickej sieti. In: *Acta Montanistica Slovaca*. ISSN 1335-1788, 2005, roč. 10, č. 2. s. 68-77.
- [2] BASSET, P.: *Some properties of the least absolute error estimator*. Ph.D. Dissertation, University of Michigan, Ann Arbor, 1973.
- [3] BÖHM, J. – RADOUCH, V. – HAMPACHER, M.: *Teorie chyb a vyrovnávací počet*. 2 vydanie. Praha: Geodetický a kartografický podnik, s. p., 1990. ISBN 80-7011-056-2.
- [4] CASPARY, W.F.: *Concepts of network and deformation analysis*. First edition. Kensington: School of surveying The University of New South Wales, 1987. 187p. ISBN 0-85839-044-2.
- [5] GAŠINCOVÁ, S., GAŠINEC, J., WEISS, G., LABANT, S.: Application of robust estimation methods for the analysis of outlier measurements. In: *GeoScienceEngineering*. Vol. 57, no. 3 (2011), p. 14-29. ISSN 1802-5420.
- [6] GAŠINCOVÁ, S.: *Spracovanie 2D sietí pomocou robustných metód : doktorandská dizertačná práca / školiteľ Janka Sabová*. - Košice, 2007. - 92 s.
- [7] IŽVOLTOVÁ, J. *Teória chýb a vyrovnávací počet I. Príklady ku cvičeniam*. Žilinská univerzita, 2004, p. 121., ISBN 9788080702397.
- [8] JÄGER, R., MÜLLER, T., SALER, H. SCHVÄBLE, R.: *Klassische und robuste Ausgleichungsverfahren*, Herbert Wichmann Verlag, Heidelberg, 2005.
- [9] LABANT, S., WEISS, G., KUKUČKA, P.: Robust adjustment of a geodetic network measured by satellite technology in the Dargovských Hrdinov suburb. In: *Acta Montanistica Slovaca*. Roč. 16, č. 3 (2011), s. 229-237. - ISSN 1335-1788 Spôsob prístupu: <http://actamont.tuke.sk/pdf/2011/n3/7labant.pdf...>

- [10] SOKOL, Š., BAJTALA, M., LIPTÁK, M. Creation of a Surveying Base for an Ice Rink Reconstruction Project. In: *Acta Montanistica Slovaca*. ISSN 1335-1788, 2011, vol. 16, no. 4, p. 312-318.
- [11] STAŇKOVÁ, H., ČERNOTA, P., NOVOSAD, M. *Problematika rovinných transformací na území ovlivněném hornickou činností*. In: *GeoScience Engineering*. ISSN 1802-5420, 2012, vol. 58, no. 3, p. 1-12.
- [12] WEISS, G., LABANT, S, WEISS, E., MIXTAJ, L. : *Detection of erroneous values in the measurement of local geodetic networks*. In: *Acta Montanistica Slovaca*. Roč. 15, č. 1 (2010), s. 62-70. - ISSN 1335-1788
Spôsob prístupu: <http://actamont.tuke.sk/pdf/2010/n1/13weiss.pdf...>
- [13] WEISS, G., ŠŮTTI, J.: *Geodetické lokálne siete I*. 1st ed. Košice, Vydavateľstvo Štroffek, 1997. 130p. ISBN 80-967636-2-8 .

RESUMÉ

Záverom si dovoľujeme konštatovať, že použitie alternatívnych odhadovacích metód má nesporne svoje opodstatnenie, pretože geodetické merania sa nezriedka realizujú v náročných a neštandardných podmienkach, ktoré môžu byť príčinou skutočnosti, že do súboru meraných dát preniknú odľahlé merania. Z opakovaných meraní na stanovisku a z obmedzených doplnkových fyzikálnych meraní teploty, tlaku a vlhkosti má geodet len obmedzenú možnosť posúdiť, či a do akej miery sa do súborov meraných geometrických veličín infiltroval vplyv predovšetkým tepelného poľa, prípadne iných rušivých fenoménov, o ktorom môže prijať závery zväčša až na základe štatistického spracovania po návrate z terénu. Pokiaľ sa jedná o jedno meranie, je identifikovateľné napr. štatistickými testami, v prípade, že súbor je kontaminovaný viacerými odľahlými meraniami, štatistické testy nemusia byť úspešné. V tomto prípade alternatívne odhadovacie metódy, kde nesporne patria aj metódy lineárneho programovania dokážu identifikovať túto množinu a potlačiť ich vplyv na výsledky vyrovnania.