USING CONJUGATE GRADIENTS PROJECTION METHOD FOR SOLVING ILL-POSED PROBLEMS ON THE SPECIAL SETS

Viera MASLEJ KREŠŇÁKOVÁ
Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice,
Letná 9, 042 00 Košice, Slovak Republic, E-mail: viera.kresnakova@student.tuke.sk

ABSTRACT

The paper deals with numerical methods for the solution of ill-posed problems on special sets. A definition of well-posed problem is given. A conjugate gradients projection method and a program written in the programming language Matlab, which solve the problem on 28 special sets of correctness are briefly described. Some results for the numerical solution of Fredholm's integral equation of the first kind are presented. The method for solving systems of linear algebraic equations on special sets has been used also for the functions smoothing for the experimental measurement data.

Keywords: *systems of linear algebraic equations, special sets, ill-posed problems, conjugate gradients projection method, stabilization methods, quasisolution, Fredholm's integral equation of the first kind*

1. INTRODUCTION

Solving practical tasks formulated in the form of operator equations (e.g., integral equations or systems of linear algebraic equations)

$$A \cdot z = u, \qquad u \in U, \quad z \in Z, \tag{1}$$

where U and Z are some metric spaces, we usually assume, that we don't know accurate data of a task, but we know only approximate system matrix and right-hand side $\{A_h, u_\delta\}$, which approximate accurate data. Errors (deviation) can be interpreted as inadequacy of the idealized mathematical model, or as a result of measurement uncertainty of the input data, or errors, which are made in the way of discretization or in numerical solution of a task [3].

On the basis of input data $\{A_h, u_\delta\}$ it is necessary to determine the solution of $z_{h,\delta}$, under the assumption that $\|A_h - A\| \leqq h$ and $\|u_\delta - u\| \leqq \delta$ [7].

According to Hadamard, the problem (1) is said to be *correct*, *correctly posed* or (Hadamard) *well-posed* if the following two conditions hold:

1. For each $u \in U$ the equation (1) has a unique solution.

2. The solution of (1) is stable under perturbation of the right-hand side of this equation, i.e., the operator A^{-1} is defined on all of U and it is continuous.

Otherwise the problem (1) is said to be *incorrectly posed* or *ill-posed* [7].

It was assumed, that ill-posed problems doesn't have physical sense, and there is no solution of such problems. It turned out, that such problems arise in modeling of various real physical sense, in geophysics, hydrodynamics, spectroscopy [3].

The first work that showed the importance of the solution of unstable problems and how to deal with such problems was the work [4]. A. N. Tikhonov have formulated new definition of the correctness, now called *correctness in Tikhonov sense* [3]. The problem is *conditionally correct* if:

1. It is apriori known, that the solution z of the task (1), belongs to a certain subset M of the base space Z.

2. The solution is unique on the set M.

3. The solution is stable on the set M, i. e., if the right-hand side u_δ belongs to the operator A image of a set M — $u_\delta \in \{u \in U \mid u = Az, z \in M \subset Z\}$.

Definition: A quasisolution to the equation (1) is called an element of $\bar{z} \in M$, which minimize the size of the residuum:

$$\|A \cdot \bar{z} - u\| = \inf_{z \in M} \|A \cdot z - u\| \tag{2}$$

on the set M [2].

In the case of continuous operator A and the compact set M, the infinum in (2) is equal to the minimum, so there exists an element $\bar{z} \in M$ minimizing the residuum norm.

Authors of [7] stated that it's not necessary to specify an exact solution to the minimization task, and that any element $z_{h,\delta}$, for which the following applies:

$$\|A_h \cdot z_{h,\delta} - u_\delta\| \leqq h \cdot \|z_{h,\delta}\| + \delta, \tag{3}$$

may be used instead of the quasisolution.

This choice ensures stability, i. e., the approximate solution $z_{h,\delta}$ will be converge to the exact solution z if the inaccuracies h and δ tend to zero.

As a special sets M discrete analogs of compact subsets in corresponding functional spaces (a set of monotone non-increasing/nondecreasing convex/concave functions which could be bounded from below, upper or from both sides) are considered [7].

To solve problem (2) on some sets listed above, authors of [7] have designed FORTRAN programs based on the *conjugate gradients projection method* and on the *conditional gradient method*. Focusing on the *conjugate gradients projection method* we have created a Matlab/Octave program and spread it on the 28 special sets of functions, which are listed in Table 1.

Dealing with stabilization of ill-posed problems the regularization method developed in 1963 by A.N. Tikhonov should be mentioned. This method has become a kind of basis for solution to the issues of ill-posed problems.

"We propose constructive methods for obtaining regularizing algorithms for the problem (1) in case not only

Table 1 Special classes list

Notion	Title	ITASK
$M\downarrow$	Non-increasing functions	1
$M\downarrow_{c_l}$	Non-increasing functions bounded from below	1.1
$M\downarrow^{c_u}$	Non-increasing functions bounded from above	1.2
$M\downarrow^{c_u}_{c_l}$	Non-increasing functions bounded from above and below	1.3
$M\uparrow$	Non-decreasing functions	2
$M\uparrow_{c_l}$	Non-decreasing functions bounded from belowe	2.1
$M\uparrow^{c_u}$	Non-decreasing functions bounded from above	2.2
$M\uparrow^{c_u}_{c_l}$	Non-decreasing functions bounded from above and below	2.3
$\widehat{M}$	Concave functions	3
$\widehat{M}_{c_l}$	Concave functions bounded from belowe	3.1
$\widehat{M}\downarrow$	Concave non-increasing functions	4
$\widehat{M}\downarrow_{c_l}$	Concave non-increasing functions bounded from belowe	4.1
$\widehat{M}\downarrow^{c_u}$	Concave non-increasing functions bounded from above	4.2
$\widehat{M}\downarrow^{c_u}_{c_l}$	Concave non-increasing functions bounded from above and below	4.3
$\widehat{M}\uparrow$	Concave non-decreasing functionsw	5
$\widehat{M}\uparrow_{c_l}$	Concave non-decreasing functions bounded from below	5.1
$\widehat{M}\uparrow^{c_u}$	Concave non-decreasing functions bounded from above	5.2
$\widehat{M}\uparrow^{c_u}_{c_l}$	Concave non-decreasing functions bounded from above and below	5.3
$\breve{M}$	Convex functions	6
$\breve{M}^{c_u}$	Convex functions bounded from above	6.1
$\breve{M}\downarrow_{c_l}$	Convex non-increasing functions	7
$\breve{M}\downarrow_{c_l}$	Convex non-increasing functions bounded from below	7.1
$\breve{M}\downarrow^{c_u}$	Convex non-increasing functions bounded from above	7.2
$\breve{M}\downarrow^{c_u}_{c_l}$	Convex non-increasing functions bounded from above and below	7.3
$\breve{M}\uparrow$	Convex non-decreasing functions	8
$\breve{M}\uparrow_{c_l}$	Convex non-decreasing functions bounded from below	8.1
$\breve{M}\uparrow^{c_u}$	Convex non-decreasing functions bounded from above	8.2
$\breve{M}\uparrow^{c_u}_{c_l}$	Convex non-decreasing functions bounded from above and below	8.3

the right-hand side but also the operator involves an error. Suppose we are given an element u_δ and a linear operator A_h such that $||u_\delta - \bar{u}||_U \leqq \delta$, $||A_h - A|| \leqq h$. In other words, the initial information consists of $\{u_\delta, A_h, \delta, h\}$. We are required to construct from this data an element $z_\eta \in Z$, $\eta = \{\delta, h\}$, such that $z_\eta \to \bar{z}$ as $(\delta, h) \to (0, 0)$. The following construction for solving this problem is widely used. Consider the functional

$$M^\alpha[z] = ||A_h z - u_\delta||_U^2 + \alpha ||z||_Z^2. \tag{4}$$

Let z_η^α be an extremal of the functional $M^\alpha[z]$, i.e., an ele-

ment minimizing $M^\alpha[z]$ on Z. If the regularization parameter $\alpha = \alpha(\eta)$ matches in a certain sense the set $\eta = \{\delta, h\}$, then in a certain sense $z_\eta^{\alpha(\eta)}$ will be a solution of (1)" [7].

Next in our paper we compare the numerical solution of Fredholm's integral equation of the first kind with our method of conjugate gradients projection with the direct method of solution of an unexact system and with two regularization methods based on two different regularization functionals – the first using the L_2 norm, the second using the Sobolev's W_2^1 norm $||z||_Z$.

2. METHODS

In our function `PTILRB.m`, the corresponding special set may be chosen using a variable `ITASK`. All `ITASK` values are listed in the Table1.

All considered special sets are determined using a system of inequalities, e.g., the set $\breve{M}\uparrow^{c_u}$ of convex nondecreas-

ing functions with an upper bound c_u is defined by:

$$\breve{M}\uparrow^{c_u} = \left\{ z \in R^n : \begin{array}{l} z_1 \leq z_2, \\ z_{i-1} - 2z_i + z_{i+1} \geq 0; \ i = 2, \ldots, n-1, \\ z_n \leq c_u. \end{array} \right\} \tag{5}$$

All 28 cases of special sets can be described with a sim-

ilar set of linear inequalities written in the form $F \cdot z \leq b$, with the size $m_F \times n$ of a matrix F. So the problem of quasisolution of a system (1) determinantion is formulated by the following:

$$\text{find } z^* = \arg\min_{z \in M} \|A \cdot z - u\|_2^2, \quad M = \{z \in R^n \,|\, F \cdot z \leq b\}, \quad (6)$$

where $\|\cdot\|_2$ is the Euclidean norm of a vector.

Below we will use $I(z) = \{i \in \{1, \ldots, m_F\} \,|\, F_i \cdot z = b_i\}$ – a set of active inequalities, where F_i denotes the i-th row of a matrix F. The size of an actual set $I(z)$ we will denote m. We will use a submatrix F_I of the matrix F with rows corresponding to the set $I(z)$, too. E_n is the identity matrix of size n. Matrix Q is a matrix of a quadratic form

$$\varphi(z) = \|A \cdot z - u\|_2^2 = (z, Qz) + (d, z) + e.$$

The constrained minimization problem (6) can be solved by many optimization methods. Below, a brief description of an algorithm of the conjugate gradients projection method is described:

Step 1: We start the minimization at any point $z^{(0)}$ satisfying required conditions $F \cdot z \leq b$. The number m of active inequalities is set to 0. For a register of iterations methods we choose $k = 0$.

Step 1.1: If $k = n$ we continue to step 6. Since the conjugate gradients projection method reaches after n steps the minimum in $\mathbb{R}^n$.

Step 1.2: Here we determine a direction $p^{(k)}$ of a functional decrease. If $k = 0$ then we calculate it as:

$$p^{(k)} = -\operatorname{grad}\varphi(z^{(k)}), \quad (7)$$

Otherwise:

$$p^{(k)} = -\operatorname{grad}\varphi(z^{(k)}) + \frac{\|\operatorname{grad}\varphi(z^{(k)})\|^2}{\|\operatorname{grad}\varphi(z^{(k-1)})\|^2} \cdot p^{(k-1)}. \quad (8)$$

Step 1.3: A value of the optimal step a_k is determined:

$$a_k = \frac{1}{2} \cdot \frac{(\operatorname{grad}\varphi(z^{(k)}), p^{(k)})}{(Qp^{(k)}, p^{(k)})} \quad (9)$$

Step 1.4: The maximal value of a step – $a_{\max}$ in a given direction $p^{(k)}$ is determined, to move within the boundaries of the M in (6).

Step 1.5: If is $a_k \leq a_{\max}$, then

$$z^{(k+1)} = z^{(k)} + a_k \cdot p^{(k)}, \qquad k = k + 1, \quad (10)$$

and repeat algorithm from step 1.1. If $a_k > a_{\max}$ then

$$z^{(k+1)} = z^{(k)} + a_{\max} \cdot p^{(k)}, \quad (11)$$

and continue in step 2.

Step 2: New active boundary appeared, set $m = m + 1$.

Step 3: The projector on a subspace $\mathbb{R}^{n-m}$ is determined:

$$P_I = E_n - F_I^T (F_I F_I^T)^{-1} F_I. \quad (12)$$

Step 4: The new starting point will be $z^{(k)}$ and instead of $\operatorname{grad}\varphi(z^{(k)})$ its projection is used, if $k < n - m$ repeat step 1.

If $k = n - m$ and we still didn't found the minimum, we go back to step 2. Otherwise, if we found the minimum and $m = 0$, we continue at the step 6. If $m \neq 0$ we continue at the step 5.

Step 5: We have reached a minimum, and we have to remove some active bounding to get new decrease possibilities.

Step 5.1: We define a parameters vector w^0 of a length m:

$$w^0 = (F_I F_I^T)^{-1} F_I \operatorname{grad}\varphi. \quad (13)$$

Step 5.2: If $w_i^0 \geq 0$, for $i = 1, 2, \ldots, m$, we've got the problem solution. We pass by the end of the algorithm, step 6.

Step 5.3: If w_i^0 is negative for some i, we will remove the index i out from a set $I(z)$ of active inequalities and we are going to step 3.

Step 6: Finish of the algorithm.

3. PROGRAM

The main function of our program has the following structure:

```
function conjugate_gradients_projection
    [d,Q] = PTILRA(A,u);
    [F,b,z0] = PTILRB(n,ITASK,cu,cl);
    [IEND,ICI,r,ma,mask,zv]=
    =PTILR1(A,z0,u,Q,d,F,b,dgr,ICM,nn);
    [IEND,ICI,r,ma,zv,mask]=
    =PTILR2(A,z0,u,Q,d,F,b,dgr,ICM,mask,ma);
        [FJ,ma]= PTILR3(F,mask);
        [PL,P] = PTILR4(FJ,ma);
        [IED,zv,ma,mask] =
        =PTILR5(F,b,zv,mask,PL,Q,d,dgr,ma);
        [IED,mask,ma] = PTILR6(P,mask,Q,d,zv);
    [IEND,ICI,r,ma,zv,mask]=
    =PTILR2(A,zv,u,Q,d,F,b,dgr,ICM,mask,ma);
        [FJ,ma]= PTILR3(F,mask);
        [PL,P] = PTILR4(FJ,ma);
        [IED,zv,ma,mask] =
        = PTILR5(F,b,zv,mask,PL,Q,d,dgr,ma);
        [IED,mask,ma] = PTILR6(P,mask,Q,d,zv);
```

As we can see, function PTILR2 is calling two times. This function is the core of all program. For the first time it is calling through the function PTILR1, where the mask and the number of borders are set to zero. Then we enter to the function PTILR2 with initial vector $z^{(0)}$, satisfying the condition $F \cdot z \leq b$. The function PTILR2 is called for the second time with the initial vector z_v – the result of the first run. After running this program, we get the final solution vector z_v.

In a function PTILR2 another functions are called. Namely: PTILR3 – the subroutine for active boundaries according to the mask creation, PTILR4 – subroutine of a projector forming, PTILR5 – subroutine of functional minimization within a domain determined by active boundaries

and their order in `MASK`, `PTILR6` – subroutine for the boundaries removal.

The value of variable `ITASK` shall be chosen according to the Table 1.

4. NUMERICAL RESULTS

Below some numerical results of our calculations using the conjugate gradients projection method are given.

4.1. Solving Fredholm's Integral Equations on a Special Sets

We will consider *convolution Fredholm's integral equation of the first kind*:

$$\int_0^1 K(x,s) \cdot z(s)\, ds = u(x), \qquad -2 \leqq x \leqq 2, \qquad (14)$$

with the kernel

$$K(x,s) = \frac{1}{1+100(x-s)^2}. \qquad (15)$$

Two exact solutions of the problem (14) $z_1(s) = 1 - s^2$ and $z_2(s) = 4s(1-s)$, $s \in [0,1]$ were considered. The corresponding exact right-hand sides $u_1(x)$ and $u_2(x)$ of (14) were determined using the program wxMaxima [1], e.g.:

```
z1:1-s^2;
u1:integrate(1/(1+100*(x-s)^2)*z1,s,0,1);
tex(%);
```
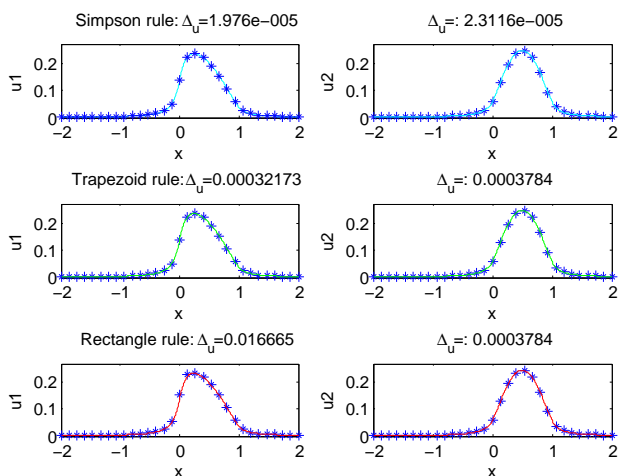

Fig. 1 Comparison of analytical and numerical determination of right-hand side $u_1(x)$ (left) and $u_2(x)$ (right)

To check the correctness of the right-hand side vector $u(x)$ determination we also used the numerical integration methods based on the Simpson rule, trapezoid rule and rectangle rule. In Figure 1 both analytical $u(x)$ and numerical $u = (u_1, \dots, u_{n_x})$ right-hand sides are compared. The difference calculated according

$$\Delta u = \max_{i=1,\dots,n_x} |u(x_i) - u_i|, \qquad (16)$$

where n_x is a number of discrete values x_i, has been used.

Below in the Figure 2, 4 methods of the solution of the problem (1) are compared. As the right-hand side vector the exact vector with a multiplicative noise of size $\delta = 10^{-3}$ has been used. Exact solution is denote as `ze`. The solution using the inverse matrix A^{-1} (or pseudoinverse matrix A^+) is called *direct solution*, and is denoted by `zd`. Two solutions based on the regularization methods mentioned above using the regularization parameter $\alpha = 10^{-6}$ are denoted as `zr1` and `zr2`, and, finally, the result of the stabilization conjugate gradients projection method is denoted by `zs`.

In Figure 3 the logarithms of the residuals at each point are shown. It is evident that in this particular case the second regularization and our method gave us the best and stable results.
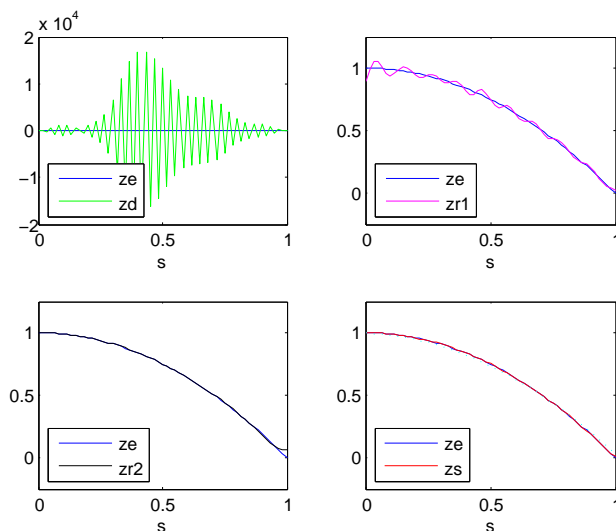

Fig. 2 Comparison of methods for numerical solution of Fredholm's equations of the first kind – $\delta = 10^{-3}$, $\alpha = 10^{-6}$, `ITASK=4.3`
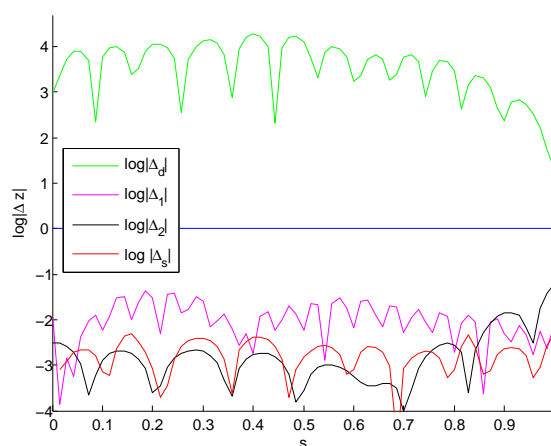

Fig. 3 The logarithms of absolute differences between the exact and numerical solutions

4.2. Functions Smoothing

Let us consider a function $u(t)$ which belongs to some class of correctness. E.g., if we consider a sum of k expo-

nential functions

$$u(t) = \sum_{i=1}^{k} A_i \cdot e^{-\alpha_i \cdot t}, \qquad (17)$$

with $A_i > 0$ and $\alpha_i > 0$, $i = 1, 2, \ldots, k$, we know, that the function $u(t)$ will be a positive decreasing (nonincreasing) convex function (case `ITASK=7.1`).

For the smoothing of the noised data we simply put $A = E_n$, where E_n is the identity matrix of size n. So, we are looking for a data vector – the nearest vector to the noised data which belongs to the set of positive nonincreasing convex functions. E.g., we use the next script:

```
ITASK=7.3; delta=0.3;cl=0.1; cu=0.9;
A=eye(length(u)); ue=...;
u=(1+(rand(size(ue))-0.5)*2*delta).*ue;
[zv,z0]=conjugate_gradients_projection
        (A,u,ITASK,cl,cu,ICM,dgr);
```
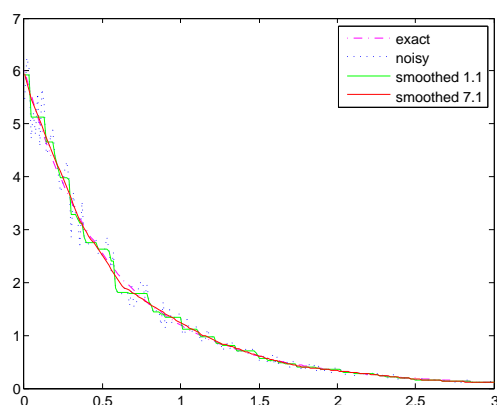

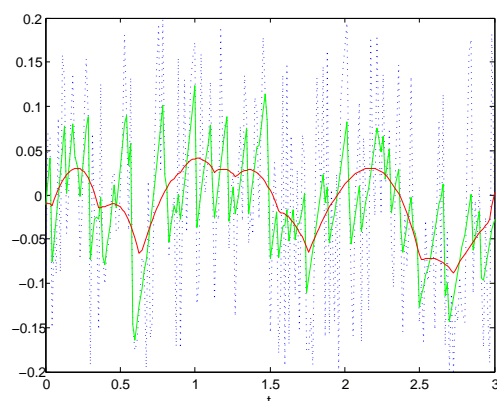**Fig. 4** Comparison of smoothing methods with `ITASK=1.1` and `ITASK=7.1`

Fig. 5 Relative residuals

For the calculation we used an example of a sum of three exponential functions with positive coefficients. In Figures 4 and 5 the corresponding result of functions smoothing and its relative residuals to the exact data vector are shown. Using Task 7.1 (solid red line in Fig. 5) led us to the maximal relative error cca. 8% for a 20% noised data.

5. CONSLUSION

The results presented in Fig. 3 show that using the stabilization methods including a solution of system of equations on a special sets of functions may be an effective tool to solve ill-posed problems. Presented method can be used also for noised data smoothing.

ACKNOWLEDGEMENT

REFERENCES

[1] BUŠA, J.: *MAXIMA. Open source systém počítačovej algebry,* [In Slovak] 1. edition, Košice, TUKE, 2006. 92 p. ISBN 80-8073-640-5.

[2] IVANOV, V. K.: *O linejnych nekorrektnych zadačach.* [In Russian] Doklady Akademii nauk SSSR, 1962, **145**, 2, pp. 270-272.

[3] IVANOV, V. K. – VASIN, V. V. – TANANA, V. P.: *Teorija linejnych nekorrektnych zadač i jejo priloženija.* [In Russian] Moskva, Nauka, 1978, 206 p.

[4] TICHONOV, A. N.: *Ob ustojčivosti obratnych zadač,* [In Russian] Doklady Akademii nauk SSSR, 1943, **39**, 5, pp. 195-198.

[5] TICHONOV, A. N. – GONČARSKIJ, A. V. – STEPANOV, V. V. – JAGOLA, A. G.: *Reguljarizujuščie algoritmy i apriornaja informacija,* [In Russian] Nauka, 1983, 200 p.

[6] TICHONOV, A. N. – GONČARSKIJ, A. V. – STEPANOV, V. V. – JAGOLA, A. G.: *Čislennye metody rešenija nekorrektnych zadač,* [In Russian] Nauka, 1990, 232 p., ISBN 5-02-014135-6.

[7] TIKHONOV, A. N. – GONCHARSKY, A. V. – STEPANOV, V. V. – YAGOLA, A. G.: *Numerical Methods for the Solution of Ill-Posed Problems,* Mathematics and Its Applications, editor HAZEWINKEL, M., Vol. 328, Kluwer Academic Publishers, 1995, 253 p., ISBN 978-90-481-4583-6.

[8] TICHONOV, A. N. – GONČARSKIJ, A. V. – STEPANOV, V. V. – JAGOLA, A. G.: *Numerické metódy riešenia nekorektných úloh,* [In Slovak] Slovak translation: BUŠA J., the 1. edition, Košice, TUKE, 2000, 268 p., ISBN 80-88922-10-0.

BIOGRAPHY

Viera Maslej Krešňáková (Ing., Bc.) was born on 21. 11. 1990. In 2014 was awarded Bc. in a study program Applied Informatics. Currently in 2017 she defended a diploma thesis entitled "The numerical methods for the solution of ill-posed problems on special sets" in a study program Business Informatics at the Department of Cybernetics and Artificial Intelligence of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, and received the engineering degree.