Effective Minimization of Nonsmooth Functions in the Limit Analysis Problem for Dielectrics in Powerful Electric Fields

Igor A. Brigadnov, Ekaterina M. Fedotova

Department of Computer Science
National Mineral Resources University "Gorniy",
St. Petersburg, Russian Federation
brigadnov@mail.ru

Abstract

The problem of minimization of ill-conditioned functions is considered. This problem arises as a result of finite-element approximation of the limit analysis problem for dielectrics in powerful electric fields. The objective function is nonsmooth therefore a smooth regularization of finite-dimensional problem is used. As a result distinct ravine of objective function is acquired. Convergence of the gradient and the heave-ball methods in relation to its internal and optimization parameters are studied inside the numerical computing environment and fourth-generation programming language Matlab.

Keywords: *nonsmooth functions, ravine functions, smooth regularization, ill-conditioned problem, numerical minimization.*

1 Introduction

Investigation of electrical strength of dielectrics has particular interest in both theory and practice [1, 2]. It is stimulated by significance and practical needs applications in field of electrical engineering and microelectronics. When dielectrics are located in powerful electric fields it appears there is the effect of a loss of global electrostatic balance state between external electric field and dielectric. This effect is interpreted as the electric breakdown of dielectric [3-7].

The electrical state of a medium in a given domain is characterized by the bulk ρ and surface g density of charges and by the vectors of electric tensity E, electric induction D and electric current density J. The vector D is introduced by the rule

$D = \varepsilon_0 E + P$, where $\varepsilon_0 \approx 0.85 * 10^{-12}$ is the dielectric constant of vacuum, and P is the vector of polarization. For the electric tensity the scalar potential u is introduced such as $E = -\nabla u$.

In weak electric fields the currents of conductivity in dielectric media are practically absent, i.e. $J \approx 0$, and the simplest linear constitutive relation $E \rightarrow D$ is used [1]. As a result, for the solution of the appropriate linear boundary value problem, various effective analytical and numerical methods have been worked out.

2 The Limit Analysis Problem in Electrostatics

In powerful electric fields the essentially nonlinear phenomena of polarization saturation $\left(|P| \leq P_* < +\infty \right)$ and ionization $\left(J \neq 0 \right)$ must be taken into account [1] as shown in Fig. 1. As a result, the integral model of bounded electric induction can be used, where $|D| \leq \lambda < +\infty$ and $\lambda > 0$ is the complementary physical parameter of dielectric medium (*the parameter of saturation*) which is easily calculated [3-7].
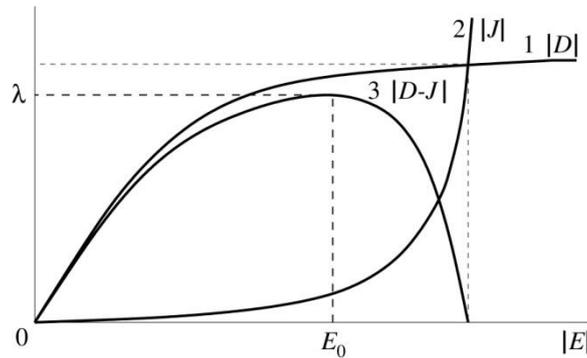
Figure 1: Dependence of electric induction (curve 1)
and current of conductivity (curve 2) on tension of electric field.
Curve 3 describes the integral model of a limit electric induction.

From the physical point of view this effect is interpreted as a loss of the global electrostatic balance state between external electric field and dielectric by the analogy of a loss of the load-carrying capability of solids in mechanics [7].

Let a dielectric medium occupy a domain $\Omega \in R^3$. It's described by two vector-functions $D = \overset{\prime}{D}(x, E) : \Omega \times R^3 \rightarrow R^3$ and $J = \overset{\prime}{J}(x, E) : \Omega \times R^3 \rightarrow R^3$.

We consider the following boundary-value problem. The quasi-static electric influences acting on the dielectric are: a bulk charge with density ρ in Ω, a surface charge with density g on a portion Γ^2 of the boundary, and a potential

u_0 on a portion Γ^1 of the boundary is also given. Here $\Gamma^1 \cup \Gamma^2 = \partial\Omega$, $\Gamma^1 \cap \Gamma^2 = 0$ and $area(\Gamma^1) > 0$. Point charges absent.

We assume that from the classical Thomson's principle [1, 2] it follows that the free energy of the electric field in dielectric has the global minimum on the real potential, i.e. the potential u is a solution of the following variational problem:

$$u_* = \arg\inf\{I(u) : u \in V\} \tag{1}$$

where:

$$I(u) = \int_\Omega \Phi(x, \Delta u(x)) dx - A(u), \quad A(u) = \int_\Omega \rho u dx + \int_{\Gamma^2} g u d\gamma,$$

where $V = \{u : \Omega \to R; u(x) = u_0, x \in \Gamma^1\}$ is the set of admissible electric potentials, Φ is the specific energy of the electric field such that $D_i(x, E) = \partial\Phi(x, E)/\partial E_i$ for every $E \in R^3$ and almost every $x \in \Omega$, $A(u)$ is the work of the electric field on external charges.

In the general case, according to the Thomson and Joule-Lenz laws specific potential energy of electric field with regards to noninvertible losses from the current of conductivity is searched by the following formula:

$$\Phi(x, E) = E \cdot \int_0^1 \left[\hat{D}(x, pE) - \hat{J}(x, pE) \right] dP, \tag{2}$$

where point is the scalar product in R^3.

In the work [3] it is shown that there is always exists such parameter of electrostatic saturation $\lambda > 0$, that for every $E \in R^3$ and almost every $x \in \Omega$ the following estimation is faithful:

$$\Phi(x, E) \leq \lambda(x)|E|. \tag{3}$$

In this case a solution of the problem (1) can be absent. Physical sense of this effect is that a dielectric can perceive a limited amount of the electric fields that is interpreted as a loss of electrostatic equilibrium between externals source of electric field and dielectric. This effect we call as the electrical breakdown of dielectric.

In the work [3] it is shown that analysis of electrostatic equilibrium in a dielectric is reduced in order to searching of a parameter $t_* > 0$, such that for $t \in (0, t_*)$ functional

$$I_t(u) = \int_\Omega \Phi(x, \Delta u(x)) dx - tA(u) \tag{4}$$

is limited on V, and it is proved that for t_* an estimation above $t_* < t_+$ is correct, where

$$t_+ = \inf \left\{ \int_\Omega |\nabla u(x)| \lambda(x) \, dx : u \in V, A(u) = 1 \right\} \tag{5}$$

and

$$\lambda(x) = \max \left\{ \left| \hat{D}(x,E) - \hat{J}(x,E) \right| : E \in R^3 \right\}.$$

Through the variational-difference method based on finite-element piecewise linear approximation of the required solution, the initial continuous problem (5) is reduced to the next finite-dimensional problem of minimization of uneven function of kind:

$$f(x) = \sum_{i=2}^{n} |\psi_i x_i - \psi_{i-1} x_{i-1}| \to \min. \tag{6}$$

The function (6) is convex, but it is nonsmooth because it contains a sum of absolute value functions, that does not allow the applying the standard methods of minimization, for example, the gradient or Newton's methods. Even more, the simplest coordinate descent minimization requires a continuously differentiable (smooth) objective function for the convergence [8-10].

For application of standard methods of minimization it is necessary to smooth out this function, i.e. to carry out its regularization (Fig. 2).
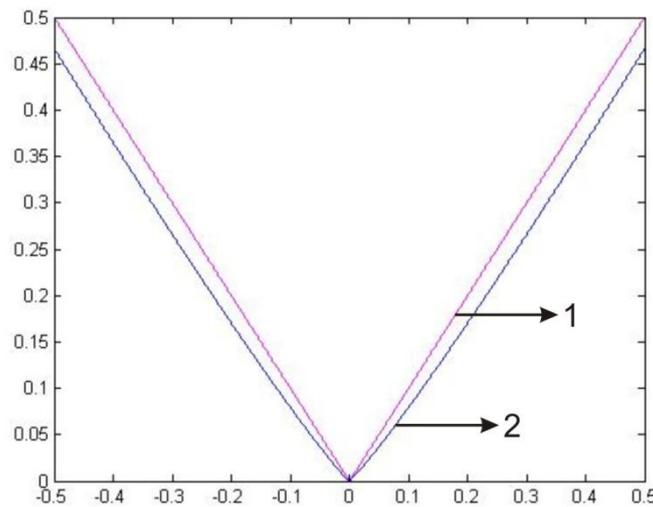
Figure 2: Regularization of the absolute function $y = |x|$ (curve 1)

by the smooth function $y = (x)^{1+\varepsilon}$ with ε=0.01 (curve 2).

As a result, a finite-dimensional problem takes the following form:

$$f(x) = \sum_{i=2}^{n} \left(\psi_i x_i - \psi_{i-1} x_{i-1} \right)^{1+\varepsilon} \to \min,$$

where ε is a small positive parameter of regularization ($0 < \varepsilon \ll 1$). In this case the object function is ravine.

A function is named as ravine if its hessian $A = \partial^2 f$ is ill-conditioned.

We remind that the condition number of a matrix A is

$$cond(A) = \frac{\max |\gamma_i|}{\min |\gamma_i|},$$

possessing properties $cond(A) \geq 1$ and $cond(A^k) = cond^k(A)$, where γ_i are eigenvalues of the matrix A.

Matrix A is named ill-conditioned if $cond(A) \gg 1$. In practically important problems the condition number can be evaluated too large as (10^6-10^9).

The original objective function is attributed to the class of functions of ravine type, for minimization of which a number of special methods has been worked out. We have chosen the heavy-ball method, since this method is master of the certain set of properties allowing effectively resolve problems of minimization of ill-conditioned functions [8, 11].

The heavy-ball two-step method is a lookup method of the global minimum and its modification of the gradient method.

The solution of the problem is performed as follows. Starting from some dot x^n consequent steps to some other dots are realized unless acceptable solution of the problem will be found.

The gradient method for solution of the global minimization problem is

$$f(x) \to \min, \tag{7}$$

where $f: R^m \to R$, can be interpreted in terms of usual differential equations in such a manner. The corresponding trajectory of minimisation is presented in Fig. 3.

An example, consider the follows differential equation

$$p\dot{x} + f'(x) = 0, \tag{8}$$

(here the dot x is the derivative with respect to the independent variable t, and $f'(x)$ means the gradient of mapping: $R^m \to R$; it is assumed that $p > 0$). The simplest finite-difference analogue of equation (9) has the view of obvious chart of Euler

$$p \frac{x^{n-1} - x^n}{h} + f'(x^n) = 0.$$

The step of the gradient method for problem (7) is written as

$$x^{n+1} = x^n - \frac{h}{p} f'\left(x^n\right). \tag{9}$$

Consider now instead of equation (9) the following equation

$$m\ddot{x} + p\dot{x} + f'\left(x\right) = 0, \tag{10}$$

describing motion of a ball of mass m in the potential field f' in presence of a friction force. The loss of energy due to friction will force a ball to get down pithily a minimum of potential of f, and forces of inertia will not allow oscillating. It is presented in Fig 4. It allows hoping that change of the equation (8) by introduction an inertial member will improve the convergence of the gradient method (9). The finite-difference analogue of equation, describing motion of a ball is

$$m\frac{x^{n+1} - 2x^n + x^{n-1}}{h^2} + p\frac{x^n - x^{n-1}}{h} + f'\left(x^n\right) = 0 \ .$$

After simple transformations and obvious labelling we get the following iteration scheme:

$$x^{n+1} = x^n - \alpha f'\left(x^n\right) + \beta\left(x^n - x^{n-1}\right). \tag{11}$$
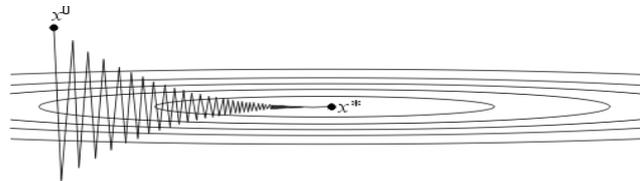
Figure 3: Trajectory of minimization by the gradient method.
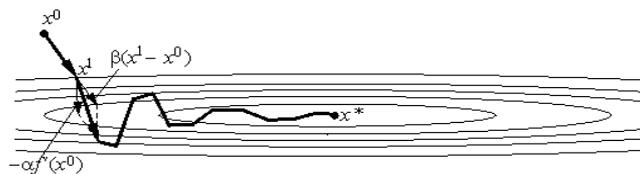
Figure 4: Trajectory of minimization by the heavy-ball method.

3 Numerical Experiments

Numerical experiments were carried out inside the computing environment and fourth-generation programming language Matlab [12, 13]. The input includes the numerical vector of coefficients ψ, the parameter of regularization ε, parameters α and β for calculations according to the heavy-ball method and the precision of calculations δ.

3.1 Program development

The program in Matlab was organized in the form of M-files. The modules were implemented by corresponding M-functions.

The module for the calculation of gradient was implemented as a separate M-function called in the process of calculations in the module, solving the problem according to the gradient or heavy-ball method.

The module for the calculation of the function was implemented in a body of the M-function for calculations according on the gradient method.

3.2 Testing of the program

Before using the program, it was necessary to carry out on testing with the aim of verification of all functions that the program must execute and the verification of a response of the program in different situations.

It was necessary to conduct next tests to define influence of different factors, such as:
- precision δ;
- parameter of regularization ε;
- parameters α and β for the heavy-ball method.

For the verification of the program a testing function which can be solved in a closed analytical form was used. With the aim of the comparison of the efficiency of different algorithms special test functions having essential singularities in extreme point lookup are developed. One of well known is the Rosenbrock function, that was used for testing of the complex program.

3.3 Results of numerical experiments according to the gradient method

For Rosenbrock function it was shown that with increasing of the precision the number of iterations grows up almost linearly. Extreme value with precision 10^{-5} arrived at 10969 iterations.

For model function, as a result of the numerical experiments, both taking into account boundary conditions and without their account, it was shown that at any precision δ and parameter of regularization ε the number of iterations is too large. This can be interpreted as a practical loss of convergence of the gradient method for minimization of model function.

3.4 Results of numerical experiments according to the "heavy-ball" method

For testing the program was used the vector of coefficients with dimensionality *n* generated by the randomize procedure.

For estimation of influence of the precision on the number of iterations, the parameter ε and precision of calculations δ were changed with constant dimensionality of the vector of coefficients and parameters α and β.

For estimation of the sensitivity of the heavy-ball method to the parameters α and β with the constant dimensionality of the vector of coefficients, precision and parameter of regularization ε, parameters α and β we changed itself in given range and the number of iterations was measured.

For the estimation of influence of the parameter of regularization ε on the quantity of iteration with the constant dimensionality of the vector of coefficients, the parameters α and β and the precision δ were changed parameter of regularization itself in a given range and the number of iterations were measured.

3.5 Results of numerical experiments without considering of boundary conditions

3.5.1 Dependence of the number of iterations on parameter of regularization ε

At the fixed value of precision $\delta=0.1$ and $\beta=0.1$ there was a following picture:

At $\alpha=0.12$ the number of iterations was 1 at any value of ε.

At values from $\alpha=0.15$ to $\alpha=0.13$ and $\varepsilon=0.1$ the number of iterations was 1, then at $\varepsilon<0.1$ a process loops.

At $\alpha>0.16$ the process loops at all values ε.

3.5.2 Dependence of the number of iterations on precision δ

At $\varepsilon=0.1$, $\varepsilon=0.01$, $\varepsilon=0.001$, $\varepsilon=0.0001$ dependence is identical.

At precision $\delta=0.1$ the number of iterations was 1, at $\delta<0.1$ the process loops.

3.5.3 Research of sensibility of heavy-ball method to parameters (α, β)

We see distinctly, that at $\varepsilon=0.1$ the number of iterations rises sharply and a process loops at $\alpha>0.15$, while at $\varepsilon<0.1$ the number of iterations rises sharply already at $\alpha>0.12$.

3.6 Results of numerical experiments with considering of boundary conditions

The analysis of problem showed that objective function without boundary conditions has an endless great number of points of minimum, located in a ravine area that results in the formal cycling of algorithm. As a result, for providing of unicity of decision the boundary condition of *x(1)=2* was used. We will notice that an initial variation problem in a natural way contains boundary conditions on the required function.

3.6.1 Research of sensibility of heavy-ball method to parameters (α, β)

General view of chart of dependence of number of iterations from parameters α and β at precision δ=0.1 and ε=0.01 presented in Fig. 5.
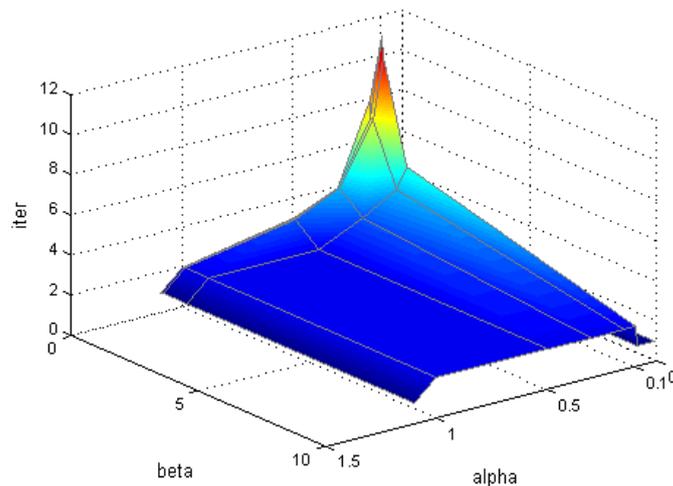
Figure 5: Sensibility of the heavy-ball method to parameters (α, β)

At different precision δ the peak of the number of iterations is observed at different values α, that evidently from the brought 4th charts in the Fig. 6.
With the change of precision the characteristic peak of number of iterations evolves in the area of small values, in this process its size raises sharply.

3.6.2 Dependence of the number of iterations on parameter of precision δ

The number of iterations was measured at the fixed values of parameters α=0.01 and β=0.01.
From a chart on the Fig. 7 it is evident, that at precision δ=0.1 the number of iterations is small, at the increase of precision to δ=0.01 it increases sharply and then it changes insignificantly. Also it is obvious, that with the change of parameter ε the number of iterations changes in the narrow range of values. Thus, it is possible to conclude that the parameter of regularization ε poorly influences the convergence of the heavy-ball method.
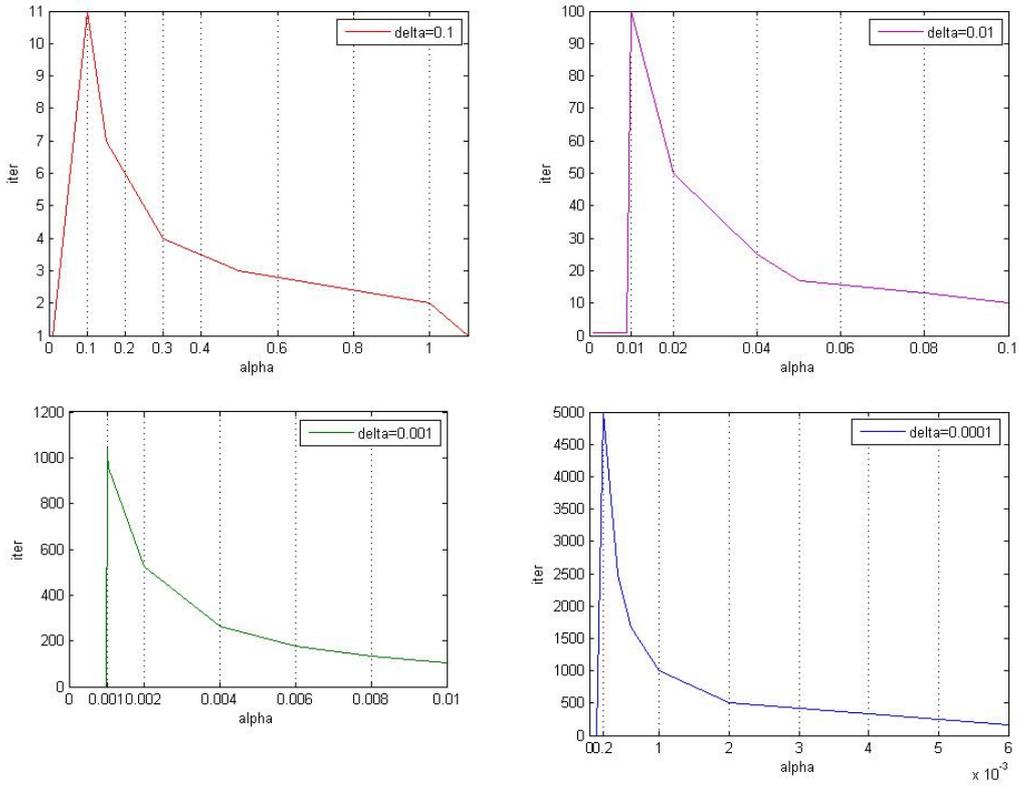
Figure 6: Dependence of the number of iterations on the parameter α at fixed
parameters $\varepsilon=0.01$ and $\beta=0.01$ with $\delta=0.1$, $\delta=0.01$, $\delta=0.001$ and $\delta=0.0001$.
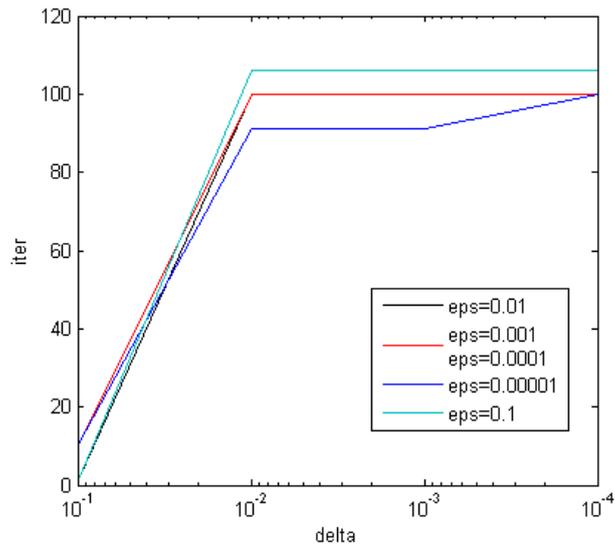
Figure 7: Dependence of the number of iterations on the parameter of precision δ.

3.6.3 Dependence of the number of iterations on the parameter regularization ε

The number of iterations was measured at the fixed values of parameters α=0.04 and β=0.01.
The number of iterations at different values of the parameter ε changes insignificantly, as well as at different values of precision δ.

4 Main Conclusions Based on Numerical Results

As a result of the numerical experiments it was discovered, that without boundary conditions the objective function has an infinite number of the points of minimum, located in a ravine area, that result in formal cycling of algorithm. As consequence, for providing of uniqueness of decision it is necessary to use boundary conditions.
Also it was shown, that for gradient method, both taking into account boundary conditions and without their account, at any precision δ and parameter of regularization ε the number of iterations is too large. This can be interpreted as a practical loss of convergence of the gradient method for minimization of model function.
During numerical experiments for heavy-ball method, with boundary conditions for objective function it was found, that:
- At different precision δ the peak of number of iterations is observed at different values α. With the change of precision δ the characteristic peak of number of iterations evolves in the range of small parameters, there peak increases sharply and then it changes insignificantly.
- At precision δ=0.1 the number of iterations is small. With the increase of precision to δ=0.01 the number of iterations increases sharply and then changes insignificantly.
- Number of iterations at the change of parameter ε changes in the narrow range of values. As a result, it is possible to conclude that the parameter of regularization ε poorly influences the convergence of heavy-ball method.
- With a decrease of parameter of regularization the number of iterations of heavy-ball method increases nonlinearly.
- The number of iterations at different values of parameter ε changes insignificantly, as well as at the different values of precision δ.

References

[1]   I.E. Tamm, The Bases of the Electricity Theory, Nauka, Moscow, 1989.
[2]   G. G. Raju, Dielectrics in Electric Fields, CRC Press, 2003.
[3]   I.A. Brigadnov, Numerical analysis of dielectrics in powerful electrical fields, Comp. Assisted Mech. Eng. Sci. .8(4) (2001) 227-234.

[4] I.A. Brigadnov, Variational-difference method for estimation of electrical breakdown of dielectrics, Math. Models and Comp. Simulations 14 (4) (2002) 57-66.

[5] I.A. Brigadnov, Duality method for estimation of breakdown conditions for dielectrics, Math. Models and Comp. Simulations 15(5) (2003) 106-114.

[6] I.A. Brigadnov, Duality method for limit analysis of dielectrics in powerful electric fields, J. Comp. Appl. Math. 168 (2004) 87-94.

[7] I.A. Brigadnov, Limit analysis method in elastostatics and electrostatics, Math. Meth. Appl. Sci. 28 (2005) 253-273.

[8] J. Nocedal and S. J. Wright, Numerical Optimization, Springer, Berlin, 1999.

[9] R. E. Miller, Optimization: Foundations and Applications, John Wiley, Chichester, 2000.

[10] Ch. A. Floudas and P. M. Pardalos, Encyclopedia of Optimization, Vol. II, Springer, Berlin, 2001.

[11] E.M. Fedotova and I.A. Brigadnov, Effective minimization of ill-conditioned functions, Comp. Sci. Eng. 2(3) (2012) 13-17.

[12] S. Attaway, Matlab: A Practical Introduction to Programming and Problem Solving, Elsevier, 2011.

[13] G. Lindfield and J. Penny, Numerical Methods: Using Matlab, Academic Press, 2012.