NUMERICAL SOLUTION OF SOME SINGULAR UNCONSTRAINED MINIMIZATION PROBLEMS

ABRAM ZHUZHUNASHVILI * AND YURI TVALODZE †

Abstract. Numerical solution of a singular unconstrained minimization problem is divided into two steps. Firstly, we consider a continuous analogue of the steepest gradient descent, which leads to a system of Volterra integral equations, and obtain convergence rate estimates. Secondly, the system is solved by an approximate-iterative method. An estimate of the total absolute error is given as a sum of the inherent error, the error of the numerical method, and the round-off error. The estimate enables one to determine parameters of the computational process and to solve the initial problem with pre-assigned accuracy.

Key words. Unconstrained minimization, singular problems, Volterra integral equation, rate of convergence, error estimate.

AMS(MOS) subject classifications. 65K05, 65R20, 65G05

1. Introduction. We consider the following unconstrained minimization problem: find $x_* \in R^n$, for which

$$f_* \equiv f(x_*) = \min_{x \in R^n} f(x), \tag{1}$$

where R^n is the n-dimensional real Euclidean space. We assume that $f : R^n \to R$ is a continuously differentiable one-extremal function, i.e. the solution x_* exists and is unique, and its gradient $\nabla f(x)$ satisfies the Lipschitz condition

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|, \quad x, y \in R^n, \tag{2}$$

where $L > 0$ is a constant.

Methods of reducing problem (1) to the Cauchy problem for systems of ordinary differential equations or for a system of Volterra equations have been considered by various authors; see, e.g., [1,5-7,10,11,15,18] and references therein.

In the present paper, we use the following method based on a continuous analogue of steepest gradient descent:

$$\frac{dx}{dt} = -\nabla f(x), \qquad x(0) = x_0$$

or

$$x(t) = x_0 - \int_0^t \nabla f(x(\tau))d\tau, \quad t \in [0, T], \tag{3}$$

where x_0 is the initial value of x, $T > 0$.

* Muskhelishvili Institute of Computational Mathematics of the Georgian Academy of Sciences, Tbilisi, Georgia. Email: abram@gw.acnet.ge.

† Muskhelishvili Kutaisi Technical University.

This research is supported by the grant N 1.21 of the Georgian Academy of Sciences.

Problem (1) is reduced to (3) in the sense that a solution of problem (1) can be obtained by solving equation (3) and finding

$$x_* = \lim_{T\to\infty} x(T), \tag{4}$$

if the above-mentioned limit exists.

Without loss of generality we assume that $x(0) = 0$, $f(0) = f_0$, and $x_* \neq 0$.

2. Estimates for a Continuous Analogue of the Steepest Gradient Descent. Commonly, some of the following additional information is known:

$$||\nabla f(x)|| = ||\nabla f(x) - \nabla f(x_*)|| \geq m||x - x_*||^\alpha, \tag{A1}$$

$$||\nabla f(x) - \nabla f(x_*)|| \leq M||x - x_*||^\alpha, \tag{A2}$$

$$(\nabla f(x),\ x - x_*) \geq \mu||\nabla f(x)||^2, \tag{A3}$$

$$f(x) - f_* \geq \kappa||x - x_*||^{1+\alpha} \tag{A4}$$

for all x such that $f(x) \leq f(x_0)$, where $\alpha > 1$, and M, m, μ, κ are positive constants.

If f is a convex function satisfying condition (2), then assumption $(A3)$ is fulfilled for any $\mu \leq L^{-1}$ and $x \in R^n$; see, e.g., [16,18].

As the stopping criterion for the choice of finite T, we take the condition

$$||\nabla f_l(\widetilde{x}(T))|| \leq \delta, \tag{5}$$

where $\widetilde{x}(t),\ t \geq 0$ is the numerical solution of system (3) obtained on a computer, ∇f_l denotes a computer representation of the vector function ∇f with l binary digits for mantissa of numbers, and $\delta > 0$ is a given number.

Lemma 1. *Let (A1) and (5) be satisfied and suppose* $||\nabla f_l(\widetilde{x}(\tau)) - \nabla f(\widetilde{x}(\tau))|| \leq C_{\nabla f}$ *for* $0 \leq \tau \leq T$, *where* $C_{\nabla f}$ *is a constant depending on* ∇f *and* l. *Then*

$$||\widetilde{x}(T) - x_*|| \leq \left(\frac{\delta + C_{\nabla f}}{m}\right)^{1/\alpha}. \tag{6}$$

The proof is obvious.

Lemma 2. *Let assumption (A3) be fulfilled. Then*

$$\int_0^t ||\nabla f(x(\tau))||^2 d\tau \leq \frac{C^2}{\mu}, \tag{7}$$

where $C \geq ||x_*||$ *is a constant.*

Proof. To prove the lemma, it should only be noted that

$$\frac{d}{dt}||x(t) - x_*|| = \frac{(x'(t), x(t) - x_*)}{||x(t) - x_*||} = -\frac{(\nabla f(x(t)), x(t) - x_*)}{||x(t) - x_*||}.$$

□

The next two Lemmas use the following inequalities:

$$||x(\tau) - \widetilde{x}(\tau)|| \leq \varepsilon_1, \quad 0 \leq \tau \leq T \tag{8}$$

and

$$||\nabla f_l(\widetilde{x}(\tau))|| > \delta, \quad 0 \leq \tau < T\,. \tag{9}$$

Lemma 3. *If inequalities (8) and (9) are satisfied, then the inequality*

$$||\nabla f(x(\tau))|| \geq \delta - C_{\nabla f} - L\varepsilon_1, \quad 0 \leq \tau < T \tag{10}$$

is valid.

The proof is trivial.

In the future, we assume $\delta > C_{\nabla f} + L\varepsilon_1$.

Lemma 4. *Let (8) and (9) hold. Then the stopping criterion (5) will be fulfilled at least in the interval* $(0, T]$, *where*

$$T = \frac{C^2}{\mu(\delta - C_{\nabla f} - L\varepsilon_1)^2}. \tag{11}$$

Proof. Since (10) holds under the conditions of the lemma, the proof of (11) easily follows from (10) and (7). □

We denote $f(t) \equiv f(x(t))$. Using the additional assumptions (A1)—(A4), let us now estimate $\varphi(t) = f(t) - f_*$, $||x(t) - x_*||$, and $||\nabla f(x(t))||$.

Theorem 1. *Let assumptions* $(A1)$ *and* $(A2)$ *be fulfilled. Then the inequality*

$$\varphi(t) \leq [\zeta(t)]^{\frac{1+\alpha}{1-\alpha}}, \quad t > 0, \tag{12}$$

holds, where

$$\zeta(t) = \varphi_0^{\frac{1-\alpha}{1+\alpha}} + \frac{\alpha - 1}{\alpha + 1}\rho t, \;\; \rho = m^2 \left(\frac{\alpha+1}{M}\right)^{\frac{2\alpha}{\alpha+1}},$$

and $\varphi_0 = \varphi(0)$.

Proof. Since

$$f(x) = f(x_*) + \int_0^1 (\nabla f(x_* + \tau(x - x_*)), x - x_*)d\tau,$$

we obtain, on the basis of $(A2)$,

$$\varphi(x) = f(x) - f_* \leq ||x - x_*|| \int_0^1 ||\nabla f(x_* + \tau(x - x_*))||d\tau$$

$$\leq ||x - x_*|| \int_0^1 M||\tau(x - x_*)||^{\alpha} d\tau = \frac{M}{\alpha + 1}||x - x_*||^{\alpha+1}.$$

From the above relation and $(A1)$ it follows that

$$||\nabla f(x)|| \geq m \left(\frac{\alpha+1}{M}\right)^{\frac{\alpha}{\alpha+1}} [\varphi(x)]^{\frac{\alpha}{\alpha+1}}. \tag{13}$$

Taking into account (13) in the relation

$$\frac{d\varphi(t)}{dt} = -||\nabla f(t)||^2$$

we obtain

$$\frac{d\varphi(t)}{dt} \leq -\rho[\varphi(t)]^{\frac{2\alpha}{\alpha+1}},$$

whence, using the Chaplygin–Gronwall inequality, we arrive at what is to be demonstrated. □

Remark 1. *If besides the assumptions of Theorem 1 condition (A4) is fulfilled, then we can obtain for $\varphi(t)$ the lower bound*

$$\varphi(t) \geq [\zeta_1(t)]^{\frac{1+\alpha}{1-\alpha}}, \quad t>0,$$

where

$$\zeta_1(t) = \varphi_0^{\frac{1-\alpha}{1+\alpha}} + \frac{\alpha-1}{\alpha+1}\rho_1 t, \;\; \rho_1 = M^2 \kappa^{-\frac{2\alpha}{1+\alpha}}$$

.

Thus, if (A1), (A2) and (A4) are fulfilled, then the last inequality together with (12) yields an unimprovable (in order t) estimate for $\varphi(t)$.

The following corollary is also valid.

Corollary 1. *Let (A1), (A2) and (A4) be fulfilled. Then the inequalities*

$$\|x(t) - x_*\| \leq \kappa^{-\frac{1}{1+\alpha}} \left[\zeta(t)\right]^{\frac{1}{1-\alpha}}, \tag{14}$$

and

$$\|\nabla f(x(t))\| \leq M\kappa^{-\frac{\alpha}{1+\alpha}} \left[\zeta(t)\right]^{\frac{\alpha}{1-\alpha}}, \;\; t>0 \tag{15}$$

hold. The proof readily follows from (12).

Remark 2. *Estimates for $\|x(t) - x_*\|$ and $\|\nabla f(x(t))\|$ can easily be obtained without assumption (A4). In particular, the following inequality, see [16], can be established:*

$$\|\nabla f(x)\|^2 \leq 2\,L(f(x) - f_*),$$

i.e.

$$\|\nabla f(x(t))\| \leq \sqrt{2L\varphi(x(t))}.$$

Using (12) in the last inequality, we obtain

$$\|\nabla f(x(t))\| \leq \sqrt{2L}\left[\,\zeta(t)\right]^{\frac{1+\alpha}{2(1-\alpha)}}$$

and

$$\|x(t) - x_*\| \leq (m^{-1}\sqrt{2L})^{\frac{1}{\alpha}}\left[\,\zeta(t)\right]^{\frac{1+\alpha}{2\alpha(1-\alpha)}}, \;\; t>0.$$

However, these estimates, as is easily seen, are worse than (15) and (14), respectively, as was to be expected.

Remark 3. *Let us now find out what the obtained estimates turn into in the limiting case as $\alpha \to 1$. Introducing the notation*

$$\chi(t,\alpha) = \left[\,\zeta(t)\right]^{\frac{1+\alpha}{1-\alpha}}$$

and assuming without loss of generality that $\varphi_0 = 1$, we can easily show that

$$\lim_{\alpha \to 1_+} \chi(t, \alpha) = exp\,(-(2m^2/M)t).$$

Thus, in this case, on the basis of (12), (14) and (15) we can make conclusions about the exponential decay of $\varphi(t)$, $\|x(t) - x_*\|$, and $\|\nabla f(x(t))\|$, which are in agreement with earlier obtained results for strongly convex functions; see, e.g., [5,16,18].

Finally, we will explain the essence of the relation (A3) in the degenerate case. For the sake of simplicity, let $f(x) = \frac{1}{2}(Ax, x) - (b, x)$, where A is a symmetric positive semidefinite matrix. Denote by $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m, \quad m < n$ all its positive eigenvalues and by $u_1, ..., u_m$ the corresponding orthonormal eigenvectors:

$$Au_i = \lambda_i u_i, \quad (u_i, u_j) = \delta_{ij}, \quad i, j \in \{1, \dots, m\},$$

where δ_{ij} is the Kronecker symbol. Further, let $u_{m+1}, ..., u_n$ be an orthonormal basis for the null space of A, so that $u_1, ..., u_n$ is a complete orthonormal basis for R^n. Then

$$x - x_* = \sum_{k=1}^{n} \alpha_k u_k, \quad \alpha_k = (x - x_*, u_k), \quad Ax_* = b,$$

and we have the relations

$$\frac{(\nabla f(x), x - x_*)}{\|\nabla f(x)\|^2} = \frac{(A(x - x_*), x - x_*)}{(A(x - x_*), A(x - x_*))} =$$

$$\frac{\sum_{i=1}^{m} \lambda_i \alpha_i^2}{\sum_{i=1}^{m} \lambda_i^2 \alpha_i^2} \geq \frac{\lambda_1}{\lambda_m^2} \geq \mu.$$

3. Statement of the Algorithm and Basic Results. To solve system (3), where T is chosen according to conditions (5), (10)–(15), we apply the approximate-iterative method of [9,14,20].

By analogy with [14] and [20], we put

$$\begin{aligned}
x^k(t) &= -\int_0^{k\Delta T} \nabla f(\widetilde{x}(\tau))\, d\tau - \int_{k\Delta T}^{t} \nabla f(x^k(\tau))\, d\tau, \\
x^{k,\nu}(t) &= -\int_0^{k\Delta T} \nabla f(\widetilde{x}(\tau))\, d\tau - \int_{k\Delta T}^{t} \nabla f(x^{k,\nu-1}(\tau))\, d\tau, \quad \nu \in \{1, \dots, N\}, \\
x^{k,0}(t) &= -\int_0^{k\Delta T} \nabla f(\widetilde{x}(\tau))\, d\tau, \quad \nu \in \{1, \dots, N\}; \\
\widetilde{x}_j^{k,\nu} &= \widetilde{x}_j^{k,0} - \int_{k\Delta T}^{t_j^k} \psi(\widetilde{x}^{k,\nu-1}(\tau))\, d\tau, \\
\widetilde{x}_j^{k,0} &= -\int_0^{k\Delta T} \psi(\widetilde{x}(\tau)) d\tau, \quad t_j^k = k\Delta T + j\Delta t, \\
& t \in [\, k\Delta T, (k+1)\Delta T], \quad k = \in \{1, \dots, R-1\}; \quad R\Delta T = T, \\
& j = \in \{1, \dots, K\}, \quad \Delta t = \Delta T/K,
\end{aligned} \tag{16}$$

where $\widetilde{x}^{k,\nu}(\tau)$ is the piecewise constant vector function equal to $\widetilde{x}^{k,\nu}_s$ on the segment (t^k_s, t^k_{s+1}),

$$\widetilde{x}(t) = \frac{t - t^k_j}{\Delta t}\,\widetilde{x}(t^k_j) + \frac{t^k_{j+1} - t}{\Delta t}\,\widetilde{x}(t^k_{j+1}),\quad t \in [t^k_j,\, t^k_{j+1}]\,,$$

and ψ is the vector function whose components are polygons of the corresponding components of the function ∇f; see, e.g., [3,11].

The following theorem is valid.

Theorem 2. *The error of numerical method (16) is given by*

$$\Delta^{(2)}_T = \max_{t\in[0,T]} ||x(t) - \widetilde{x}(t)|| \le e^{LT}\left[\frac{3L_0T}{2N!} + \frac{dT}{4}(\Delta t)^2 + \frac{||x''(t)||}{16}(\Delta t)^2\right]$$

$$\le e^{LT}\left[\frac{3L_0T}{2N!} + \frac{dT}{4}(\Delta t)^2 + \frac{\sqrt{n}}{16}(3L_1 + L_0L_1 + (L_1^2 + L_2)\Delta T)(\Delta t)^2\right], \tag{17}$$

where

$$d = L_2 + L_2L_0^2 + 3L_1^2 + 2L_0L_2 + L_0L_1^2 + (2L_0L_1L_2 + L_1^3 + 3L_1L_2)\Delta T + L_1^2L_2(\Delta T)^2,$$

$$L_i = \max_{x,\xi} ||g^{(i)}_t(0,x)||,\quad i = \overline{0,2},\quad g(t,x) = \nabla f(x + t\xi),\quad \xi \in R^n,\quad ||\xi|| = 1,$$

where the maximum is taken with respect to all x in the domain of existence of a solution of system (3), and thus $L \le \max_{x,\xi} ||g'_t(0,x)||$. Moreover, the number of evaluations of ∇f will not exceed $(K+2)^2NR/2$.

The proof is similar to the proof of the corresponding results from [14,20].

We emphasize that, in the same manner as in [14,20], we can estimate an inherent error $\Delta^{(1)}_T$ of the solution of system of equations (3), and a round-off error $\Delta^{(3)}_T$ of algorithm (16) implemented on an electronic computer, and hence a total absolute error of the approximate solution (1):

$$\Delta_T \le \Delta^{(1)}_T + \Delta^{(2)}_T + \Delta^{(3)}_T. \tag{18}$$

Indeed, let

$$\delta(\nabla f) = \int_0^t [\nabla f(\widehat{x}(\tau)) - \nabla \widehat{f}(\widehat{x}(\tau))]\, d\tau,\quad \delta x = x(t) - \widehat{x}(t), \tag{19}$$

where $\nabla \widehat{f}$ is a given approximation to ∇f, and $\widehat{x}$ is a solution of system (3) in which $\nabla \widehat{f}$ is taken instead of ∇f. Then under condition (2), from (19) we easily get

$$\Delta^{(1)}_T = ||\delta x||_T \le \int_0^T ||\delta x||_\tau d\tau + ||\delta(\nabla f)||_T,$$

whence on the basis of Chaplygin's inequality it follows that

$$\Delta^{(1)}_T \le L\int_0^T ||\delta(\nabla f)||_\tau e^{L(t-\tau)} d\tau + ||\delta(\nabla f)||_T \le \max_{0\le\tau\le T} ||\delta(\nabla f)||_\tau e^{LT}. \tag{20}$$

Here, $||\cdot||$ is an arbitrary norm of a vector function, possessing the property

$$\Big\|\int_0^t x(\tau)d\tau\Big\|_t \le \int_0^t \|x\|_\tau d\tau \qquad x \in X, \tag{21}$$

where X is the normed space of vector functions, defined on $[0, T]$, with the norm $||\cdot||_t$, $0 < t \leq T$. We can show, see [14], that in case $X = L_p(0,T)$, $1 \leq p < \infty$ and $C([0,T])$ property (21) is fulfilled.

Thus, the following theorem is valid.

Theorem 3. *Under conditions (2) and (21), the inherent absolute error estimate (20) is valid.*

It should also be remarked that on the basis of results obtained in [14,20], we can achieve better estimates for $\Delta_T^{(1)}$ under some other assumptions. Note here that the problem under consideration is specific in the sense that if the initial function f is given approximately, then to obtain $\nabla\widehat{f}$ one should solve an ill-posed problem of numerical differentiation by applying well-known methods of regularization; see, e.g. [13]. Moreover, in cases where the gradient can be calculated in the same number of operations as the function itself, computational effort can be minimized by using methods described in [12].

To estimate the round-off error $\Delta_T^{(3)}$, we apply the well-known estimates [19] of the type

$$|\Big(\sum_{k=1}^{m} a_k\Big)_l - \sum_{k=1}^{m} a_k| \leq 1.06 \cdot 2^{-l} \max_{1\leq k\leq m} |a_k| \cdot \frac{(m+2)^2}{2},$$

$$|(a_k b_k)_l - a_k b_k| \leq 2^{-l}|a_k b_k|, \; 2^{-l}m \leq 0.1$$

to scheme (16). As a result, we obtain the maximal round-off error per step by one iteration

$$\Delta_{\Delta T}^{(3)} \leq 1.06 \cdot 2^{-l} L_0[(K+1)R+2]^2/2 + C_{\nabla f}(K+1)R, \quad 2^{-l}(K+1)R \leq 0.1.$$

It is assumed that we take the Chebyshev norm of the vector $x^{k,\nu}(t) - (x^{k,\nu}(t))_l$. As the iterative process in (16) converges with respect to ν with the rate $L\Delta T \leq \frac{1}{2}$, the final error on one step will not exceed $2\Delta_{\Delta T}^{(3)}$ [13]. This error will take place on every next step as an inherent error, and therefore by virtue of (20) the total accumulated error satisfies

$$\Delta_T^{(3)} \leq [1.06 \cdot 2^{-l} L_0[(K+1)R+2]^2 + 2C_{\nabla f}(K+1)R]e^{LT}\,. \tag{22}$$

Thus, the following theorem holds.

Theorem 4. *Under conditions (2) and (6), the absolute round-off error (in the Chebyshev norm) of an approximate solution of system (3) by method (16) is given by estimate (22).*

By combining estimates (17), (20), and (22), we obtain the estimate for Δ_T. Further, by analogy with [14,20], we can deduce the estimate of the total absolute error of solution (3) by using the above-considered method. The most essential difference of the given case is that the value T itself must be defined with regard to relations (5), (10)–(15).

Let $\widetilde{x}$ be an approximate solution of system (3) with the error Δ_T, and let (5) be fulfilled. Then using (6) we obtain

$$\delta \leq m\varepsilon^{\alpha} - C_{\Delta f}\,, \tag{23}$$

where $\varepsilon > 0$ is the pre-assigned accuracy of the solution of our original problem (1). From (23) and (5) we deduce certain requirements on T and on the error Δ_T from (18), i.e. on the parameters of the method K, N, R, as well as on the value of the inherent error $||\delta(\nabla f)||$ of input data and the parameter l of the electronic computer.

It is also not difficult to show that the numerical method under consideration is optimal both for a needed number of operations and for stability, i.e. it satisfies the corresponding hypotheses about the properties of optimal methods indicated in [2].

4. A Connection Between Problems of Minimization and Problems of Solving a System of Volterra Integral Equations of General Type. Here we consider systems of equations of the type

$$x(t) = x_0 - \int_0^t \nabla_x f(\tau, x(\tau)) d\tau, \tag{24}$$

or

$$x(t) = x_0 - \int_0^t \nabla_x f(t, \tau, x(\tau)) d\tau\,. \tag{25}$$

Equation (24) is, obviously, equivalent to

$$\frac{dx}{dt} = -\nabla_x f(t, x(t)),$$

which can be interpreted as a continuous analogue of the steepest gradient descent, when the objective function $f(t, x)$ varies in time, see [16].

Equation (25) is equivalent to

$$\frac{dx}{dt} = -\nabla_x f(t, t, x(t)) - \int_0^t \frac{\partial}{\partial t}(\nabla_x f(t, \tau, x(\tau))) d\tau\,,$$

which can be interpreted as a problem of minimization of the function $f(t, t, x) \equiv f_1(t, x)$ by the method of differential descent, when the antigradient direction $-\nabla_x f$ is corrected aside its variation in time t on the whole pre-history $(0, t)$.

To solve (24) and (25) numerically, one can apply the approximate-iterative method (16) with nearly the same results as above.

5. Concluding Remarks. A large number of papers is devoted to the numerical methods for solving singular minimization problems; see, e.g., [1,4,8,17], and references therein. However, these works deal with the method error only. They do not touch upon investigating the inherent and round-off errors, neither optimality of methods.

It should also be noted that the result coinciding with the upper bound (14) has been obtained essentially in [1]. However, an estimate similar to (13) and its natural generalization have been assumed, not deduced from other conditions as in the present paper.

Acknowledgements. The authors would like to thank Steve Billups and Andrew Knyazev for comments and help with preparing the publication.

REFERENCES

1. Ya. I. Alber. Local theorems of convergence of gradient methods in the Hilbert space, *USSR Computational Math. and Math. Physics*, (4): 829–838, 1973.

2. N. S. Bakhvalov. On the properties of the optimal methods for solving of the problems of mathematical physics, *USSR Computational Math. and Math. Physics*, (3): 555–568, 1970.
3. N. S. Bakhvalov. Numerical methods: Analysis, Algebra, Ordinary Differential Equations. *MIR Publishers,* Moscow, 1977.
4. K. N. Belash, A. A. Tret'jakov. Methods of solving the singular problems, *USSR Computational Math. and Math. Physics*, (7): 1097–1102, 1988.
5. C. A. Botsaris. A class of methods for unconstrained minimization based on stable numerical integration techniques, *J. of Math. Anal. Appl.*, 63(1978), N 3, 729–749.
6. A. A. Brown, M. C. Bartholomew-Biggs. Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations, *JOTA*, 62(1989), N 2, 211–224.
7. M. T. Chu. On the continuous realization of iterative processes, *SIAM Review*, 30(1988), N 3, 375–387.
8. Yu. M. Danilin. The Newton's method with step regulation for singular problems of minimization, *Bull. Acad. Sci. USSR*, (3): 530–532, 1981.
9. V. K. Dzyadik. An approximate-iterative method for approach of the solution of the Cauchy problem for ordinary differential equations by polynomials, *USSR Computational Math. and Math. Physics*, (3): 357–372, 1986.
10. Yu. G. Evtushenko, V. G. Zhadan. The barrier-projective methods of solving the problems of nonlinear programming, *USSR Computational Math. and Math. Physics*, (5): 669–684, 1994.
11. M. K. Gavurin. Nonlinear functional equations and continuous analogues of the iterative methods, *Izvestija VUZov, Ser. Mathematics*, (5(6)): 18–31, 1958.
12. M. Iri. Simultaneous computation of functions, partial derivatives and estimates of rounding errors — complexity and practicality, *Japan J. of Applied Mathematics*, 1(1984), N 2, 223–252.
13. V. V. Ivanov. Numerical methods, *NAUKOVA DUMKA Publishers,* Kiev, 1986.
14. V. V. Ivanov, Y. G. Tvalodze, A. Sh. Zhuzhunashvili. On solving of Volterra's integral equations with pre-assigned accuracy, *Technical Report 1886, Muskhelishvili Institute of Computational Mathematics of Georgian Acad. Sci.* Tbilisi, 1989.
15. B. T. Poljak. The gradient methods for minimization of functionals, *USSR Computational Math. and Math. Physics*, (4): 643–653, 1963.
16. B. T. Poljak. Introduction to optimization, *NAUKA Publishers*, Moscow, 1983.
17. B. T. Poljak. Iterative algorithms for singular minimization problems, *In: Nonlinear Programming, 4. Academic Press, N.Y.*, 1981, 147–166.
18. F. P. Vasil'ev. Numerical methods of solving the extremal problems, *NAUKA Publishers,* Moscow, 1988.
19. J. H. Wilkinson, Algebraic eigenvalue problems, *NAUKA Publishers,* Moscow, 1970.
20. A. Sh. Zhuzhunashvili, V. V. Ivanov and Y. G. Tvalodze. Estimates of the total absolute error of a numerical solution of Volterra integral equations of the second kind, *UCD/CCM Report 102, University of Colorado at Denver, Center for Computational Mathematics,* Denver, Colorado, 1997.