Mathematical Models and Methods in Applied Sciences

CONTINUOUS ANALOG OF THE GAUSS-NEWTON METHOD

RUBEN G. AIRAPETYAN *, ALEXANDER G. RAMM † and ALEXANDRA B. SMIRNOVA ‡
Department of Mathematics, Kansas State University,
Manhattan, Kansas 66506-2602, U.S.A.

A Continuous Analog of discrete Gauss-Newton Method (CAGNM) for numerical solution of nonlinear problems is suggested. In order to avoid the ill-posed inversion of the Fréchet derivative operator some regularization function is introduced. For the CAGNM a convergence theorem is proved. The proposed method is illustrated by a numerical example in which a nonlinear inverse problem of gravimetry is considered. Based on the results of the numerical experiments practical recommendations for the choice of the regularization function are given.
Keywords: Continuous Gauss-Newton method; iterative scheme; Fréchet derivative; regularization.

1. Introduction

Let H_1 and H_2 be real Hilbert spaces and $\varphi : H_1 \to H_2$ a nonlinear operator. Let us consider the equation:

$$\varphi(x) = 0. \tag{1.1}$$

We assume that the following condition on φ holds.

Condition A: Problem (1.1) has a solution $\hat{x}$, not necessarily unique.

In the well-known Newton's method ([5]) one constructs a sequence $\{x_n\}$ for $n = 0, 1, \ldots$ which converges to a solution (in general non unique) of equation (1.1). The first term x_0 is an initial approximation point and the other terms are constructed by means of the following iterative process:

$$x_{k+1} = x_k - \varphi'(x_k)^{-1}\varphi(x_k), \tag{1.2}$$

where $\varphi'(x)$ is the Fréchet derivative of the operator φ. Recall that $\varphi'(x)$ is a linear operator from H_1 to H_2. The usual necessary condition for the realization of the Newton method is the bounded invertibility of $\varphi'(x_k)$, that is, the existence

*E-mail: airapet@math.ksu.edu
†E-mail: ramm@math.ksu.edu
‡E-mail: smirn@math.ksu.edu

of a bounded linear operator $[\varphi'(x_k)]^{-1}$ for all k. Actually in order to provide the convergence of Newton iterations one needs bounded invertibility of φ' in a ball $B(\hat{x}, R) := \{x : x \in H_1, ||x - \hat{x}|| \le R\}$. However this condition does not hold in many important applications. In order to avoid this restriction several modifications of the Newton method have been developed. In this paper we consider the Gauss-Newton procedure for equation (1.1) (see e.g. ([5])):

$$x_{k+1} = x_k - [\varphi'^*(x_k)\varphi'(x_k)]^{-1}\varphi'^*(x_k)\varphi(x_k), \quad x_0 = x_0. \tag{1.3}$$

If the operator $\varphi'^*(x)\varphi'(x)$ is not boundedly invertible one needs some regularization procedure. In order to construct such a procedure one can introduce a sequence of positive numbers α_k, $\alpha_k \to 0$, and replace iterative method (1.3) by the following one ([5,2]):

$$x_{k+1} = x_k - [\varphi'^*(x_k)\varphi'(x_k) + \alpha_k I]^{-1}[\varphi'^*(x_k)\varphi(x_k) + \alpha_k(x_k - x_0)], \tag{1.4}$$

where I is the identity operator.

The methods constructed above can be also considered as discrete analogs of some continuous methods (called sometimes continuation methods). In ([3]) the following Cauchy problem has been considered as a continuous analog of (1.2):

$$\dot{x}(t) = -\varphi'(x(t))\varphi(x(t)), \quad x(0) = x_0, \quad \dot{x}(t) := \frac{dx}{dt}. \tag{1.5}$$

A solution to problem (1.1) can be obtained as limit of the function $x(t)$ for $t \to \infty$. If one solves this Cauchy problem by means of Euler's method with a stepsize $\tau = 1$ one gets (1.2) with $x_k = x(k)$. Continuous analogs of iterative methods have several advantages over the discrete ones. Convergence theorems for continuous methods usually can be obtained easier. If a convergence theorem is proved for a continuous method, that is, for the Cauchy problem for a differential equation, for instance (1.5), one can construct various finite difference schemes for the solution of this Cauchy problem. These difference schemes give discrete methods for the solution of equation (1.1). For instance the methods of Euler and Runge-Kutta can be used. More detailed information about the applications and modifications of continuous Newton methods can be found in ([3,8,1]).

The aim of this paper is to construct a continuous analog of iterative scheme (1.4), to prove a convergence theorem for this continuous analog of (1.4), and to test the method numerically by applying it to a practically interesting nonlinear inverse problem of gravimetry.

The paper is organized as follows. In section 2 a continuous analog of method (1.4) is described and a convergence theorem for this method is formulated. In section 3 this convergence theorem is proved. In section 4 an inverse gravimetry problem is considered and the proposed method is numerically tested. In our numerical experiments comparison of different regularization functions is done. Based on the results of the numerical experiments some recommendations are given for the choice of the regularization function.

2. Continuous Gauss-Newton Method and Convergence Theorem

In order to describe convergence rates we introduce the following

Definition 2.1. A positive function $\alpha(t) \in C^1[0,\infty)$ is said to be a convergence rate function if $\alpha(t)$ decreases monotonically to zero as $t \to \infty$, $\alpha(t) \in C^1[0,\infty)$ and $\ln \alpha(t)$ is concave, that is, $\dot{\alpha}(t)/\alpha(t)$ is monotonically increasing.

Remark 2.2. The number $\alpha(0)$ can be chosen sufficiently large and simultaneously the number $|\dot{\alpha}(0)/\alpha(0)|$ can be sufficiently small. Here and below the over dot denotes the derivative with respect to time $\dot{x} := dx/dt$. For example, one can choose $\alpha(t) = b/(t+a)$, where a and b are positive constants such that a and b/a are sufficiently large.

A continuous analog of iterative process (1.4) is the following Cauchy problem:

$$\dot{x}(t) = -[\varphi'^*(x(t))\varphi'(x(t)) + \alpha(t)I]^{-1}[\varphi'^*(x(t))\varphi(x(t)) + \alpha(t)(x(t) - x_0)], \quad (2.1)$$

$$x(0) = x_0.$$

Denote by $\mathrm{Ran}(L)$ the range of the linear operator L. The convergence of the continuous analog of Gauss-Newton method (CAGNM) is established by the following theorem, in which (and throughout this paper) the norms $||\varphi'(x)||$ and $||\varphi''(x)||$ are the norms of linear and bilinear operators from H_1 to H_2 and from $H_1 \times H_1$ to H_2 respectively.

Theorem 2.3. Let $\alpha(t)$ be a convergence rate function. Assume that there exists a positive number R for which Condition A and the following conditions hold:

(i) The Fréchet derivatives $\varphi'(x)$ and $\varphi''(x)$ exist in the ball $B(\hat{x}, R)$ and satisfy the following inequalities:

$$||\varphi'(x)|| \le N_1, \quad ||\varphi''(x)|| \le N_2 \quad \forall x \in B(\hat{x}, R), \quad (2.2)$$

where

$$\frac{\alpha(0)}{N_1 N_2}\left(1 - 2N_1N_2||v|| + \frac{\dot{\alpha}(0)}{\alpha(0)}\right) \le R.$$

(ii)

$$x_0 \in B(\hat{x}, R) \cap [\hat{x} + \mathrm{Ran}(\varphi'^*(\hat{x})\varphi'(\hat{x}))]. \quad (2.3)$$

(iii) For some v, such that $\hat{x} - x_0 = \varphi'^*(\hat{x})\varphi'(\hat{x})v$, the following inequalities hold:

$$1 - 2N_1N_2||v|| + \frac{\dot{\alpha}(0)}{\alpha(0)} > 0, \quad (2.4)$$

$$\left(1 - 2N_1N_2||v|| + \frac{\dot{\alpha}(0)}{\alpha(0)}\right)^2 - 2N_1N_2||v|| > 0. \quad (2.5)$$

Then the following conclusions hold:

(i) The solution $x = x(t)$ of problem (2.1) exists, and $x(t) \in B(\hat{x}, R)$ for $t \in [0, \infty)$,

(ii) $||x(t) - \hat{x}|| = O(\alpha(t))$ for $t \to \infty$.

Remark 2.4. Condition (ii) in Theorem 2.3 gives some restriction on the choice of an initial approximation point. It is not easy to verify this condition algorithmically. However some kind of this condition is necessary if one works with equation (1.1) with the operator $\varphi'^*(x)\varphi'(x)$, which is not boundedly invertible. If the operator $\varphi'^*\varphi'$ is injective but is not boundedly invertible, then the image of the linear selfadjoint operator $\varphi'^*\varphi'$ is dense in $B(\hat{x}, R)$ and consequently the set of the suitable initial approximation points satisfying condition (ii) is also dense in $B(\hat{x}, R)$. As our numerical results show (see section 4) the proposed method is practically efficient.

3. Proof of Theorem 2.3

The main part of the proof is to show that the solution to problem (2.1) does not leave the ball $B(\hat{x}, R)$ (Lemma 3.3). In order to prove it, let us assume that there exists such a point $t_1 \in [0, \infty)$ that $x(t)$ intersects the boundary of $B(\hat{x}, R)$ for the first time at $t = t_1$. Hence $x(t)$ belongs to the interior of the $B(\hat{x}, R)$ for $t \in [0, t_1)$ and $||x(t_1) - \hat{x}|| = R$. Let us introduce an auxiliary function

$$w(t) := ||x(t) - \hat{x}||/\alpha(t). \tag{3.1}$$

First, in Lemma 3.1, we derive a nonlinear differential inequality for $w(t)$. From this differential inequality we get the estimate which shows that for all $t \in [0, t_1]$ the points of the integral curve of problem (2.1) belong to the interior of the ball $B(\hat{x}, R)$. This contradiction proves that the integral curve of the solution does not leave the above ball, and consequently problem (2.1) has the global solution for $t \in [0, \infty)$. Also we show the boundedness of the $w(t)$ and this implies, by formula (3.1), strong convergence of $x(t)$ to $\hat{x}$ for $t \to \infty$.

Lemma 3.1. If the assumptions of Theorem 2.3 hold then the differential inequality

$$\frac{dw}{dt} \leq C_1 w^2 - C_2 w + C_3, \tag{3.2}$$

is valid for $t \in [0, t_1]$, where

$$C_1 = \frac{N_1 N_2}{2}, \quad C_2 = 1 - 2N_1 N_2 ||v|| + \frac{\dot{\alpha}(0)}{\alpha(0)}, \quad C_3 = ||v||. \tag{3.3}$$

Proof. The Gâteaux derivative $\varphi''(x, \xi_1, \xi_2)$ is a bilinear operator such that

$\varphi'(x+\xi_1)\xi_2 - \varphi'(x)\xi_2 := \varphi''(x, \xi_1, \xi_2) + \eta\xi_2$, and $||\eta|| \cdot ||\xi_1||^{-1} \to 0$ for $\xi_1 \to 0, ||\xi_1|| > 0$.

Let us define operators $K, G : B(\hat{x}, R) \times H_1 \times H_1 \to H_2$ by the formulas:

$$K(x, \xi_1, \xi_2) = \int_0^1 \int_0^1 \varphi''(x + st\xi_1, \xi_1, \xi_2) t dt ds \tag{3.4}$$

and

$$G(x, \xi_1, \xi_2) = \int_0^1 \varphi''(x + t\xi_1, \xi_1, \xi_2) dt. \tag{3.5}$$

Then from (2.2) we get

$$|K(x, \xi_1, \xi_2)| \leq \frac{N_2}{2} ||\xi_1|| \cdot ||\xi_2||, \quad |G(x, \xi_1, \xi_2)| \leq N_2 ||\xi_1|| \cdot ||\xi_2||. \tag{3.6}$$

The following formulas will be used:

$$\varphi(\hat{x}) - \varphi(x) = \varphi'(x)(\hat{x} - x) + K(x, \hat{x} - x, \hat{x} - x), \tag{3.7}$$

and

$$(\varphi'(\hat{x}) - \varphi'(x))\xi = G(x, \hat{x} - x, \xi), \tag{3.8}$$

where K and G are defined by (3.4) and (3.5) respectively. Let us derive formulas (3.7) and (3.8). One has

$$\varphi(\hat{x}) - \varphi(x) = \int_0^1 \frac{d}{dt} \varphi(x + t(\hat{x} - x)) dt = \int_0^1 \varphi'(x + t(\hat{x} - x))(\hat{x} - x) dt$$

$$= \varphi'(x)(\hat{x} - x) + \int_0^1 [\varphi'(x + t(\hat{x} - x)) - \varphi'(x)](\hat{x} - x) dt$$

$$= \varphi'(x)(\hat{x} - x) + \int_0^1 t dt \int_0^1 ds [\varphi''(x + st(\hat{x} - x))](\hat{x} - x)(\hat{x} - x)$$

and

$$[\varphi'(\hat{x}) - \varphi'(x)]\xi = \int_0^1 \frac{d}{dt} \varphi'(x + t(\hat{x} - x))\xi dt = \int_0^1 \varphi''(x + t(\hat{x} - x), \hat{x} - x, \xi) dt.$$

Since $\hat{x}$ solves (1.1), one can rewrite equation (2.1) as

$$\frac{dx}{dt} = -[\varphi'^*(x)\varphi'(x) + \alpha I]^{-1}[\varphi'^*(x)(\varphi(x) - \varphi(\hat{x})) + \alpha(x - x_0)].$$

From the condition (iii) of Theorem 2.3 and from (3.7) we get

$$\frac{dx}{dt} = -[\varphi'^*(x)\varphi'(x) + \alpha I]^{-1}[-\varphi'^*(x)(\varphi'(x)(\hat{x} - x) + K(x, \hat{x} - x, \hat{x} - x)) +$$

$$\alpha(x-\hat{x})+\alpha\varphi'^*(\hat{x})\varphi'(\hat{x})v],$$

and therefore

$$\frac{dx}{dt}=-(x-\hat{x})-[\varphi'^*(x)\varphi'(x)+\alpha I]^{-1}[-\varphi'^*(x)K(x,\hat{x}-x,\hat{x}-x)+\alpha\varphi'^*(x)\varphi'(x)v+$$

$$\alpha(\varphi'^*(\hat{x})\varphi'(\hat{x})-\varphi'^*(x)\varphi'(x))v].$$

Since $\hat{x}$ does not depend on t, it follows from (3.8) that

$$\frac{d(x(t)-\hat{x})}{dt}=\frac{dx}{dt}=-(x-\hat{x})-[\varphi'^*(x)\varphi'(x)+\alpha I]^{-1}[-\varphi'^*(x)K(x,\hat{x}-x,\hat{x}-x)$$

$$+\alpha\varphi'^*(x)\varphi'(x)v+\alpha(\varphi'^*(\hat{x})-\varphi'^*(x))\varphi'(\hat{x})v+\alpha\varphi'^*(x)G(x,\hat{x}-x,v)].$$

Now let us derive an inequality for $\frac{d}{dt}||x-\hat{x}||^2$. One has

$$\frac{d}{dt}||x-\hat{x}||^2=-2||x-\hat{x}||^2+2([\varphi'^*(x)\varphi'(x)+\alpha I]^{-1}[\varphi'^*(x)K(x,\hat{x}-x,\hat{x}-x)],x-\hat{x})-$$

$$2\alpha([\varphi'^*(x)\varphi'(x)+\alpha I]^{-1}\varphi'^*(x)\varphi'(x)v,x-\hat{x})+2\alpha(\varphi'(\hat{x})v,G(x,\hat{x}-x,[\varphi'^*(x)\varphi'(x)$$

$$+\alpha I]^{-1}(x-\hat{x})))+2\alpha([\varphi'^*(x)\varphi'(x)+\alpha I]^{-1}\varphi'^*(x)G(x,\hat{x}-x,v),x-\hat{x}).$$

Since the operator $\varphi'^*(x)\varphi'(x)$ is selfadjoint and nonnegative we have the following spectral representation:

$$[\varphi'^*(x)\varphi'(x)+\alpha I]^{-1}\varphi'^*(x)\varphi'(x)=\int_0^\infty\frac{\lambda}{\lambda+\alpha}dE_\lambda,$$

where E_λ is the resolution of the identity of the selfadjoint operator $\varphi'^*(x)\varphi'(x)$. Since $0\le\lambda/(\lambda+\alpha)\le 1$ for $\alpha>0$ and $\lambda\ge 0$, it follows that

$$||[\varphi'^*(x)\varphi'(x)+\alpha(t)I]^{-1}\varphi'^*(x)\varphi'(x)||\le 1. \tag{3.9}$$

Also one has the following estimate:

$$||[\varphi'^*(x)\varphi'(x)+\alpha(t)I]^{-1}||\le 1/\alpha(t). \tag{3.10}$$

From (3.9),(3.10) and (3.6) one gets the following differential inequality for $A(t):=||x(t)-\hat{x}||$:

$$\dot{A}\le -A+\frac{N_1N_2}{2\alpha}A^2+\alpha||v||+2N_1N_2||v||A.$$

In order to finish the proof of the Lemma 3.1, we derive from the last inequality the inequality for $w(t)$ by taking into account that $\dot{\alpha}(t)/\alpha(t)$ is monotonically increasing function. □

The following lemma is a simple corollary of the more general results established in ([6]).

Lemma 3.2. Let $f(t,u)$ be a continuous function on $[0,T]\times(-\infty,+\infty)$ such that the Cauchy problem

$$\dot{u} = f(t,u(t)), \quad u(0) = u_0 \tag{3.11}$$

is uniquely solvable on $[0,T]$ and $v(t)$ be a differentiable function defined on $[0,T]$ and satisfies the conditions

$$\dot{v} \le f(t,v(t)), \quad t \in [0,T], \quad v(0) = v_0. \tag{3.12}$$

If $v_0 \le u_0$ then

$$v(t) \le u(t) \text{ for } t \in [0,T].$$

It follows from inequality (2.5) that $c = \sqrt{C_2^2 - 4C_1C_3} > 0$ for constants C_1, C_2, C_3 defined in (3.3). Let u_1 and u_2 be correspondingly the smaller and the larger roots of the equation $C_1u^2 - C_2u + C_3 = 0$. For u_0 satisfying the inequality $u_2 < u_0 < C_2/(2C_1)$ the solution of the Cauchy problem

$$\dot{u} = C_1u^2 - C_2u + C_3, \quad u(0) = u_0. \tag{3.13}$$

is given by the formula:

$$\left|\frac{u-u_2}{u-u_1}\right| = \frac{u_2-u_0}{u_0-u_1}e^{ct}.$$

Let us show that $u(t)$ is defined for all $t \in [0,\infty)$. Indeed, $u_0 \in (u_1,u_2)$, so for sufficiently small t one has

$$u = u_1 + \frac{u_2-u_1}{\frac{u_2-u_0}{u_0-u_1}e^{ct}+1}. \tag{3.14}$$

Thus $u_1 < u(t) < u(0) < u_2$. This means that $u(t)$ does not leave the interval (u_1,u_2) for all $t \in [0,\infty)$ and $u(t)$ is well defined for all $t \in [0,\infty)$.

From condition (i) of Theorem 2.3 one obtains the following estimate:

$$w(0) = \frac{||x_0-\hat{x}||}{\alpha(0)} < \frac{1-2N_1N_2||v|| + \frac{\dot{\alpha}(0)}{\alpha(0)}}{N_1N_2} = \frac{C_2}{2C_1}.$$

Therefore from Lemmas 3.1 and 3.2 it follows that

$$\frac{||x(t_1)-\hat{x}||}{\alpha(t_1)} \le u_1 + \frac{u_2-u_1}{\frac{u_2-u_0}{u_0-u_1}e^{ct_1}+1} < u_0 < \frac{C_2}{2C_1}. \tag{3.15}$$

Thus

$$||x(t_1)-\hat{x}|| < \frac{C_2}{2C_1}\alpha(t_1) < \frac{C_2}{2C_1}\alpha(0) \le R.$$

This contradicts the assumption $||x(t_1)-\hat{x}|| = R$. So the following lemma is proved.

Lemma 3.3. If the assumptions of Theorem 2.3 hold and for an arbitrary positive T the solution of the problem (2.1) exists on the interval $[0, T]$, then the integral curve of the solution of (2.1) lies in the interior of the ball $B(\hat{x}, R)$ for all t from the interval $[0, T]$.

Now let us show that there exists the unique solution of (2.1) on $[0, \infty)$provided that x_0 satisfies conditions ii) and iii) of Theorem 2.3. The Cauchy problem (2.1) is equivalent to the integral equation

$$x(t) = x_0 + \int_0^t F(s, x(s))ds, \tag{3.16}$$

where

$$F(s, x(s)) := -[\varphi'^*(x(s))\varphi'(x(s)) + \alpha(s)I]^{-1}[\varphi'^*(x(s))\varphi(x(s)) + \alpha(s)(x(s) - x_0)].$$

Let us fix an arbitrary large positive number T and use the successive approximation method to solve equation (3.16) on $[0, T]$:

$$x_{n+1}(t) = x_0 + \int_0^t F(s, x_n(s))ds, \quad x_0(t) = x_0, \quad t \in [0, T]. \tag{3.17}$$

Since $\varphi'(x)$ and $\varphi''(x)$ are assumed to be bounded in $B(\hat{x}, R)$, see (2.2), and $\alpha(t)$ is positive on $[0, T]$, for every $t \in [0, T]$ the function $F(t, x)$ has bounded Fréchet derivative with respect to x in $B(\hat{x}, R)$. So one has:

$$||F(t, x_1) - F(t, x_2)|| \le K(T)||x_1 - x_2||$$

for all $t \in [0, T]$ and x_1, x_2 belong to $B(\hat{x}, R)$. Thus, one easily gets the estimate

$$||x_{n+1}(s) - x_n(s)|| \le ||F(x_0)||K^n(T)\frac{T^n}{n!}$$

valid on the maximal subinterval $[0, T_1] = \{t : t \in [0, T] \text{ and } x(t) \in B(\hat{x}, R)\}$. Therefore iterative process (3.17) converges uniformly and determines the unique solution of equation (3.16) on $[0, T_1]$. If $T_1 < T$, it follows from the maximality of the subinterval $[0, T_1]$ that $x(T_1)$ is a boundary point of $B(\hat{x}, R)$. But this contradicts to Lemma 3.3. So the solution of the problem (2.1) exists and belongs to the interior of the ball $B(\hat{x}, R)$ on every interval $[0, T]$ and consequently on $[0, \infty)$.

To finish the proof of Theorem 2.3 it is sufficient to note that equation (3.15) implies estimate

$$||x(t) - \hat{x}|| \le \frac{C_2}{2C_1}\alpha(t)$$

for all $t \in [0, \infty)$. □

4. Numerical Results

To test numerically the method described above, we chose the inverse gravimetry problem ([7]). The goal of the numerical test is to illustrate the choice of the regularization function $\alpha(t)$ and to compare two methods of solving the Cauchy problem (2.1): the Euler method, which corresponds to the iterative scheme (1.4), and the Runge-Kutta method.

Let the sources of a gravitational field with a constant density ρ be distributed in the domain

$$D = \{-l \leq t \leq l, \quad -H \leq z \leq -H + x(t)\},$$

where $x(t)$ is an interface between two media, l and H are parameters of the domain. The potential V of such a field is given by the double integral:

$$V(t,z) = \frac{1}{2\pi} \int_D \int \rho \ln \frac{1}{\sqrt{(t-s)^2 + (z-\tau)^2}} dS =$$

$$= -\frac{\rho}{4\pi} \int_{-l}^{l} ds \int_{-H}^{-H+x(s)} \ln[(t-s)^2 + (z-\tau)^2] d\tau.$$

For the z - component of the gravitational field one has

$$-\frac{\partial V(t,z)}{\partial z} = -\frac{\rho}{4\pi} \int_{-l}^{l} ds \int_{-H}^{-H+x(s)} \frac{\partial}{\partial \tau} \ln[t-s)^2 + (x-\tau)^2] d\tau =$$

$$= \frac{\rho}{4\pi} \int_{-l}^{l} \ln \frac{(t-s)^2 + (z+H)^2}{(t-s)^2 + (z+H-x(s))^2} ds.$$

In particular, on the surface $z = 0$ we obtain the following nonlinear operator equation

$$\varphi(x) \equiv \frac{\rho}{4\pi} \int_{-l}^{l} K(t,s,x(s)) ds - y(t) = 0, \tag{4.1}$$

where

$$K(t,s,x(s)) = \ln \frac{(t-s)^2 + H^2}{(t-s)^2 + (H-x(s))^2}.$$

The gravity strength anomaly $y(t) = -\frac{\partial V(t,0)}{\partial z}$ is given and the interface between two media (with and without the sources of a gravimetry field) $x(s)$ is to be determined.

Let φ act between the pair of Hilbert spaces H_1 and H_2. Assume that $H_1 = H^1[-l,l]$ or $L_2[-l,l]$ and $H_2 = L_2[-l,l]$. The Fréchet derivative of this operator is the following one

$$\varphi'(x)h = \int_{-l}^{l} \frac{2(H-x(s))h(s)}{(t-s)^2 + (H-x(s))^2} ds. \tag{4.2}$$

For any fixed $x \in \{x \in L_2[-l,l], x \leq H-\varepsilon, \varepsilon > 0\}$ the kernel

$$K'_x(t,s,x(s)) \equiv \frac{2(H-x(s))}{(t-s)^2+(H-x(s))^2}$$

is a square integrable function on $[-l,l]\times[-l,l]$, therefore $\varphi'(x)$ in (4.2) is a compact linear operator in $L_2[-l,l]$. This means that the operators $\varphi'(x)$ and $\varphi'^*(x)\varphi'(x)$ are not boundedly invertible. So, one can not use classical iterative schemes such as Newton or Gauss - Newton in the case of equation (4.1).

We solve the Cauchy problem (2.1) for φ given by (4.1) with some regularization function $\alpha(t)$. The problem is numerically solved by means of two finite difference methods, namely, Euler's method

$$x_{k+1} = x_k + \tau F(t_k, x_k), \quad x_0 = x(0) \tag{4.3}$$

and the Runge - Kutta method

$$x_{k+\frac{1}{2}} = x_k + \frac{\tau}{2} F(t_k, x_k), \tag{4.4}$$

$$x_{k+1} = x_k + \tau F(t_{k+\frac{1}{2}}, x_{k+\frac{1}{2}}), \quad x_0 = x(0).$$

Here

$$F(t_k,x_k) \equiv -[\varphi'^*(x_k)\varphi'(x_k)+\alpha(t_k)I]^{-1}(\varphi'^*(x_k)\varphi(x_k)+\alpha(t_k)(x_k-x_0))$$

and an equal grid size $\tau > 0$ defines the node points,

$$t_k = k\tau, \quad k = 0, 1, ..$$

For the successful realization of Continuous Gauss-Newton method an appropriate regularization function should be chosen. At the beginning of the process values of $\alpha(t)$ should not be very small for the operator $\varphi'^*(x(t))\varphi'(x(t)) + \alpha(t)I$ to be stably invertible and at the same time $\alpha(t)$ should tend to zero sufficiently fast to ensure convergence of the function $x(t)$ to the solution of problem (4.1). In the numerical experiments the functions $\alpha_0/(\beta+t)^m$, $\alpha_0 e^{-\beta t}$ and $\alpha_0 2^{-\beta t}$ were used. The experiments have shown that for all the considered functions the numerical solution is evaluated with appropriate accuracy for sufficiently large range of the parameters m, α_0 and β. The following tables illustrate the dependence of the accuracy of the numerical results on parameters for the following data $l = 1, \quad H = 2, \quad \rho = 1, \quad x_0 = 1$. For the numerical tests the function $y(t)$ in (4.1) was chosen as the solution of the direct problem for the model function $x_{mod}(t) = (1-t^2)^2$. The integral in (4.1) was calculated by Simpson's formula with the number of node points equals 201 and with a step size equal to 0.01.

In the tables below Δ_E and Δ_R are the absolute errors, σ_E and σ_R are the discrepancies $||\varphi(x(t))||$, see (4.1), of the Euler and the Runge-Kutta methods respectively. The first table shows the dependence of the absolute errors and the discrepancies $\varphi(x)$ on the regularization function.

Table 1.

$\alpha_0 = 0.1, \quad \tau = 0.1$					
$\alpha(t)$	N	Δ_E	σ_E	Δ_R	σ_R
$\alpha_0(1+t)^{-2}$	38	0.26	$3.75 \cdot 10^{-2}$	0.27	$4.09 \cdot 10^{-2}$
$\alpha_0(1+t)^{-4}$	25	0.25	0.25	0.11	0.11
$\alpha_0(1+t)^{-6}$	43	0.12	0.12	$1.75 \cdot 10^{-2}$	$2.03 \cdot 10^{-2}$
$\alpha_0(1+t)^{-8}$	61	$5.30 \cdot 10^{-2}$	$2.87 \cdot 10^{-3}$	$5.33 \cdot 10^{-2}$	$3.58 \cdot 10^{-3}$
$\alpha_0(1+t)^{-10}$	79	$2.25 \cdot 10^{-2}$	$4.90 \cdot 10^{-4}$	$2.23 \cdot 10^{-2}$	$6.22 \cdot 10^{-4}$
$\alpha_0 2^{-3.5t}$	127	$1.07 \cdot 10^{-2}$	$1.14 \cdot 10^{-5}$	$1.08 \cdot 10^{-2}$	$3.03 \cdot 10^{-5}$
$\alpha_0 e^{-3.5t}$	85	$1.08 \cdot 10^{-2}$	$2.84 \cdot 10^{-4}$	$1.08 \cdot 10^{-2}$	$3.60 \cdot 10^{-4}$

Then it is assumed that the type of a function $\alpha(t)$ is chosen and the dependence of the absolute errors and the discrepancies on parameters β (Tables 2 and 3) and α_0 (Table 4) is analyzed.

Table 2.

$\alpha(t) = \alpha_0 e^{-\beta t}, \quad \alpha_0 = 0.1, \quad \tau = 0.1$					
β	N	Δ_E	σ_E	Δ_R	σ_R
1	29	0.26	$7.60 \cdot 10^{-2}$	0.26	$9.10 \cdot 10^{-2}$
2	156	$1.09 \cdot 10^{-2}$	$4.58 \cdot 10^{-6}$	$1.10 \cdot 10^{-2}$	$1.52 \cdot 10^{-5}$
3	100	$1.06 \cdot 10^{-2}$	$5.78 \cdot 10^{-5}$	$1.07 \cdot 10^{-2}$	$9.30 \cdot 10^{-5}$
4	73	$1.15 \cdot 10^{-2}$	$7.54 \cdot 10^{-4}$	$1.14 \cdot 10^{-2}$	$1.10 \cdot 10^{-3}$
5	56	$1.52 \cdot 10^{-2}$	$4.50 \cdot 10^{-3}$	$1.50 \cdot 10^{-2}$	$6.12 \cdot 10^{-3}$
6	45	$2.08 \cdot 10^{-2}$	$1.40 \cdot 10^{-2}$	$2.17 \cdot 10^{-2}$	$1.84 \cdot 10^{-2}$
7	37	$3.90 \cdot 10^{-2}$	$3.06 \cdot 10^{-2}$	$3.89 \cdot 10^{-2}$	$3.80 \cdot 10^{-2}$
8	31	0.11	$5.82 \cdot 10^{-2}$	$8.59 \cdot 10^{-2}$	$6.52 \cdot 10^{-2}$
9	26	0.18	0.10	0.15	0.11
10	23	0.24	0.14	0.20	0.16

Table 3.

$\alpha(t)=\alpha_0 e^{-\beta t}, \quad \alpha_0=0.1, \quad \tau=0.6$					
β	N	Δ_E	σ_E	Δ_R	σ_R
1	4	0.23	$5.35 \cdot 10^{-2}$	0.26	0.12
2	20	$4.13 \cdot 10^{-2}$	$1.00 \cdot 10^{-3}$	$1.09 \cdot 10^{-2}$	$6.11 \cdot 10^{-5}$
3	16	$8.90 \cdot 10^{-2}$	$5.04 \cdot 10^{-2}$	$1.05 \cdot 10^{-2}$	$1.73 \cdot 10^{-4}$
4	12	$1.46 \cdot 10^{-2}$	$1.29 \cdot 10^{-4}$	$1.25 \cdot 10^{-2}$	$1.79 \cdot 10^{-3}$
5	9	$1.76 \cdot 10^{-2}$	$7.14 \cdot 10^{-4}$	$1.84 \cdot 10^{-2}$	$1.08 \cdot 10^{-2}$
6	7	0.16	$3.85 \cdot 10^{-3}$	$2.98 \cdot 10^{-2}$	$3.22 \cdot 10^{-2}$
7	6	0.38	$2.58 \cdot 10^{-2}$	0.12	$5.61 \cdot 10^{-2}$
8	5	0.83	1.48	0.18	$8.60 \cdot 10^{-2}$
9	4	0.83	01.48	0.21	0.15
10	3	0.83	01.48	0.23	0.19

Table 4.

$\alpha(t)=\alpha_0 e^{-\beta t}, \quad \beta=3.5, \quad \tau=0.1$					
α_0	N	Δ_E	σ_E	Δ_R	σ_R
10^{-3}	71	$1.62 \cdot 10^{-2}$	$9.77 \cdot 10^{-4}$	$1.61 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$
10^{-2}	79	$1.33 \cdot 10^{-2}$	$4.55 \cdot 10^{-4}$	$2.29 \cdot 10^{-2}$	$2.90 \cdot 10^{-4}$
10^{-1}	85	$1.07 \cdot 10^{-2}$	$2.84 \cdot 10^{-4}$	$1.08 \cdot 10^{-2}$	$3.60 \cdot 10^{-4}$

Analyzing the results of the numerical experiments (a part of them is included in the Tables) one concludes the following:

(i) The Runge-Kutta method is more stable with respect to changes of the regularization function $\alpha(t)$ and the step size τ, than the Euler method;

(ii) for all the considered functions suitable parameters can be chosen, however in the case when $\alpha(t)=\alpha_0 e^{-\beta t}$ the accuracy with which the solution $x(t)$ is calculated is higher;

(iii) an appropriate range of values of the parameter α_0 is from 0.001 to 0.1, for larger values the accuracy is lower, and for smaller values the processes do not converge;

(iv) the range of appropriate values of β is large enough: from 2 to 6 for $\tau=0.6$ and from 2 to 7 for $\tau=0.1$.

Remark 4.1. The reason why result (i), formulated above, is emphasized can be understood if one remembers that problem (4.1) is ill-posed. If a problem is ill-posed then the usage of a higher-order accuracy difference scheme (or quadrature formula) may lead to less accurate results, as was observed in the literature (see e.g. ([4]), p. 155). On the other hand, if a problem is well-posed, then the usage of a higher-order accuracy scheme should lead to more accurate results.

Acknowledgments

The authors thank Dr. V.Protopopescu for useful remarks.

References

1. R.G. Airapetyan and I.V. Puzynin, *Newtonian iterative scheme with simultaneous iterations of inverse derivative, Comp. Phys. Comm.* **102** (1997) 97-108.
2. A.B. Bakushinskii, *Iterative methods for nonlinear operator equations without regularity. New approach, Dokl. Russian Acad. Sci.* **330** (1993) 282-284.
3. M.K. Gavurin, *Nonlinear functional equations and continuous analogies of iterative methods, Izv. Vuzov. Ser. Matematika.* **5** (1958) 18–31.
4. V.A.Morozov and A.I.Grebennikov, Methods of the Solution of Ill-Posed Problems. Alghorithmic Approach. (Moscow University Press, Moscow, 1992).
5. J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables (Academic Press, 1970).
6. J. Szarski, Differential Inequalities (PWN, Warszawa, 1967).
7. V.V. Vasin and A.L. Ageev, Ill-Posed Problems with a priori Information (Nauka. Ekaterinburg, 1993).
8. E.P. Zhidkov and I.V. Puzynin, *The solution of the boundary problems for second order nonlinear differential equations by means of the stabilization method, Soviet Math. Dokl.* **8** (1967) 614-616.