Some approximate Gauss–Newton-type methods for nonlinear ill-posed problems

Inga Kangro[a]*, Raul Kangro[b], and Otu Vaarmann[a]

[a] Institute of Cybernetics, Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia
[b] Institute of Mathematical Statistics, University of Tartu, Liivi 2, 50409 Tartu, Estonia

Abstract. This paper treats numerical methods for solving the nonlinear ill-posed equation $F(x) = 0$, where the operator F is a Fréchet differentiable operator from one Hilbert space into another Hilbert space. Two parametric approximate Gauss–Newton-type methods are developed, a local convergence theorem is proved under certain conditions on a test function and the required solution, and some computational aspects are discussed. The validity of the theoretical convergence rate estimates is illustrated by the numerical results of solving two sample problems, one in a finite-dimensional and the other in an infinite-dimensional Hilbert space.

Key words: nonlinear operator equation, ill-posedness, iterative regularization, approximate methods, Gauss–Newton-type methods.

1. INTRODUCTION

An important problem in scientific computing is that of solving a nonlinear equation

$$F(x) = 0, \tag{1}$$

where F is a differentiable operator from an abstract space into another. Further we shall suppose that F is Fréchet differentiable and it is acting from a Hilbert space H into another Hilbert space. Depending on the properties of the Fréchet derivative F', the problem under consideration may have a unique solution or infinitely many solutions or no solution in the classical sense and one has to seek for a generalized solution. One way to determine a generalized solution is to seek it in the least squares sense by minimizing the functional

$$f(x) = \frac{1}{2} \|F(x)\|^2. \tag{2}$$

Real-life problems are frequently ill-posed. For example, those arising in inverse problems are often ill-posed because they typically involve the estimation of certain quantities based on indirect measurements. Ill-posedness is an essential aspect of computation that is often seriously expressed in instability of the solution (the solution is unstable under data perturbations) and the usage of standard, non-regularized methods can produce uncontrollable and unacceptable error propagation. For solving ill-posed problems regularizing algorithms containing several variable parameters may be fruitful, because a proper selection of parameters

---

* Corresponding author, inga@ioc.ee

involved sometimes improves the convergence properties, reduces the amount of computation, and provides a wider choice of initial guesses.

Much work has been devoted to the study of the convergence of the iteratively regularized methods. Not long ago, many monographs, e.g. [1,7], and a number of papers [3,6,11,12,16,17], were devoted to this topic. In the works mentioned above, as a rule, problems with the inaccurate data are considered and the convergence of the Gauss–Newton method is proved under the assumption on source-like representation of the exact solution x^*:

$$x^* - u_0 = (F'^*(x^*)F'(x^*))^p v, \quad 0 < p \le 1,$$

holding for an element $v \in H$. The writing $F'^*(x)$ denotes the adjoint operator of $F'(x)$. In order to obtain simple estimates of the terms including approximations to the inverse operators, here we use the value $p = 1$. It is well known that without extra assumptions, usually source-type assumptions or some assumptions concerning the smoothness of the solution, one cannot get a specific rate of convergence.

In contrast to many other papers on ill-posed problems, we do not consider the question of choosing the level of regularization based on the known or estimated level of the noise in the data. In this paper the data are assumed to be exact and we consider the question of finding a solution of (1) in the least squares sense when $F'^*(x)F'(x)$ is not necessarily continuously invertible for all x in a sufficiently large neighbourhood of the exact solution x^*.

This paper is a continuation of papers [8,10,13,14] and it treats approximate Gauss–Newton-type methods for solving nonlinear ill-posed problems for which procedural and rounding errors are unavoidable. Frequently the use of finite-difference approximations to the derivatives gives rise to an inexact method. Approximations may also be regarded as a result of inevitable inaccuracy of computation or they are due to the strategy used for solving linear problems at each iteration, i.e. associated linear equations are solved intentionally approximately by taking finitely many steps of an iterative procedure or the inverse operator appearing in a method is approximated by a formula that is less time-consuming to compute. The purpose of this paper is to discuss basic concepts of the approach and therefore the emphasis is on the theoretical background rather than the practical solution.

2. METHODS AND CONVERGENCE THEOREM

A large number of real-life problems are solved by finding the solution of certain equation or by minimizing a functional under consideration. Here for finding stationary points of the functional (2) or, equivalently, for solving the nonlinear normal equation

$$F'^*(x)F(x) = 0$$

a family of iterative methods based on the notion of regularization is developed. Recall that here $F'^*(x)$ denotes the adjoint operator of $F'(x)$.

We shall consider the case where the range of $F'(x)$ is not necessarily closed and the product operator $F'^*(x)F'(x)$ may not have the bounded inverse operator for every x. The Gauss–Newton method

$$
\begin{aligned}
x_{k+1} &= x_k - [B(x_k)]^{-1} F'^*(x_k) F(x_k), \\
B(x_k) &= F'^*(x_k) F'(x_k), \quad k = 0, 1, \ldots,
\end{aligned}
$$

is equivalent to minimizing the functional

$$g_k(h) = \frac{1}{2} \left\| F(x_k) + F'(x_k)h \right\|^2$$

at every iteration point x_k for the correction term h. In order to cope with the possibility of having a non-invertible or ill-conditioned $F'(x_k)$ at some iteration step it is desirable to develop algorithms based on the Tikhonov functional

$$\Phi_k(h) = \frac{1}{2} \left(\left\| F(x_k) + F'(x_k)h \right\|^2 + \alpha \left\| x_k + h - u_0 \right\|^2 \right),$$

where $\alpha > 0$ and u_0 is an element of H, the so-called test function.

Following the idea of iterative regularization [2,5,15,16], we assume that α in the Tikhonov functional is changed at each iteration step according to a sequence $\{\alpha_k\}$ of properly chosen positive numbers. From the stationarity condition for the Tikhonov functional we obtain a one-parameter iterative Gauss–Newton method

$$x_{k+1} = x_k - M_k^{-1} \left[F'^*(x_k)F(x_k) + \alpha_k(x_k - u_0) \right],$$

where $M_k = B(x_k) + \alpha_k I$ and I denotes the identity mapping. As $B(x)$ is positive semidefinite for every $x \in H$, the operator M_k is clearly invertible for every k.

In this paper we shall study a class of two-parameter regularized Gauss–Newton-type methods

$$x_{k+1} = x_k - \varepsilon_k D_k \left[F'^*(x_k)F(x_k) + \alpha_k(x_k - u_0) \right], \tag{3}$$

where $0 < \varepsilon_k \le 2$ and D_k is an approximation to M_k^{-1} satisfying the condition

$$\| I - D_k M_k \| \le \mu_k \le \mu_0 < 1, \quad k = 1, 2, \ldots.$$

The motivation for introducing the approximate inverse D_k of M_k is the fact that often it is not possible or reasonable to apply M_k^{-1} to an element $x \in H$ exactly and the result is found by solving the equation $M_k y = x$ approximately by a numerical method.

To get convergence results for the method (3), let us suppose that the operator F is Fréchet differentiable with Lipschitz continuous derivative in the region $S \subset H$ under consideration with

$$\left\| F'(x) \right\| \le K, \quad \left\| F'(x) - F'(y) \right\| \le L \|x - y\|, \quad K, L > 0, \ x, y \in S \tag{4}$$

and that equation (1) has a solution x^* with the representation

$$x^* - u_0 = F'^*(x^*)F'(x^*) v \tag{5}$$

holding for an element $v \in H$ and a test function u_0.

We shall show that for an appropriate choice of α_k the sequence $\{x_k\}$ defined by (3) converges to the solution x^* of (1) provided the parameters ε_k satisfy the condition

$$|1 - \varepsilon_k| \le c_\varepsilon < 1, \tag{6}$$

and that the parameters μ_k and $\|v\|$ are sufficiently small, i.e.

$$\mu_k \le \mu_0 < \frac{1 - c_\varepsilon}{1 + c_\varepsilon}, \quad k = 1, 2, \ldots \tag{7}$$

and

$$\sqrt{2}(1 + c_\varepsilon)(KL\|v\|)^{1/2}(1 + \mu_0) + c_\varepsilon + (1 + c_\varepsilon)\mu_0 + 2KL(1 + c_\varepsilon)(1 + \mu_0)\|v\| = q < 1. \tag{8}$$

In order to make the statement and the proof of the convergence result shorter, let us introduce some additional constants:

$$a_1 = \frac{1}{2}(1 + c_\varepsilon)KL(1 + \mu_0), \tag{9}$$

$$a_2 = c_\varepsilon + (1 + c_\varepsilon)\mu_0 + 2KL(1 + c_\varepsilon)(1 + \mu_0)\|v\|, \tag{10}$$

$$a_3 = (1 + c_\varepsilon)(1 + \mu_0)\|v\|. \tag{11}$$

Note that

$$q = 2\sqrt{a_1 a_3} + a_2.$$

Now we are ready to state the convergence theorem.

Theorem 1. *Let F be a Fréchet differentiable operator with Lipschitz continuous derivative and the relations (4)–(8) hold. If the regularization parameter α_k is chosen according to*

$$\alpha_k = c_\beta \beta^k, \ q \leq \beta < 1, \tag{12}$$

where the coefficient c_β satisfies the inequality

$$c_\beta \geq \frac{\|x_0 - x^*\|}{\delta} \tag{13}$$

with

$$\delta = \frac{\beta - a_2 + \sqrt{(\beta - a_2)^2 - 4a_1 a_3}}{2a_1},$$

the sequence $\{x_k\}$ defined by (3) converges to the solution x^* with

$$\|x_k - x^*\| \leq c_\beta \delta \beta^k, \quad k = 1, 2, \ldots.$$

Remark 1. If the test function u_0 is not equal to x^*, parameter δ is well-defined since for $\beta \geq q$ we have

$$(\beta - a_2)^2 \geq (q - a_2)^2 = 4a_1 a_3.$$

Moreover, since

$$\delta \geq \frac{q - a_2}{2a_1} = \sqrt{\frac{a_3}{a_1}} = \frac{\sqrt{2\|v\|}}{\sqrt{KL}},$$

the condition (13) is satisfied for all $\beta \in [q, 1)$ if $c_\beta \geq \frac{\sqrt{KL}}{\sqrt{2\|v\|}}\|x_0 - x^*\|$. If $u_0 = x^*$, we can choose the parameters α_k according to the conditions of the theorem only for $\beta > q$.

Proof. Denote $h_k = x_k - x^*$ and $\tau_k = \|x_k - x^*\|$. Note that according to (13), we have $\tau_0 \leq \delta \alpha_0$. Assume that $\tau_k \leq \delta \alpha_k$ for some $k \geq 0$ and show that then also

$$\tau_{k+1} \leq \delta \alpha_{k+1}.$$

Since, according to assumption (4), we have

$$\|F(x_k) - F(x^*) - F'(x_k)h_k\| = \left\| \int_0^1 [F'(x^* + th_k) - F'(x_k)]h_k \, dt \right\|$$

$$\leq \int_0^1 (1-t)L\|h_k\|^2 \, dt = \frac{L}{2}\|h_k\|^2,$$

we can write

$$F'^*(x_k)[F(x_k) - F(x^*)] = F'^*(x_k)F'(x_k)(x_k - x^*) + G,$$

where

$$\|G\| \leq \frac{KL}{2}\|x_k - x^*\|^2. \tag{14}$$

Taking into account that $F(x^*) = 0$ and the relation (3), we have

$$
\begin{aligned}
x_{k+1} - x^* &= x_k - x^* - \varepsilon_k D_k \left[F'^*(x_k) F(x_k) + \alpha_k (x_k - u_0) \right] \\
&= x_k - x^* - \varepsilon_k D_k \left\{ F'^*(x_k) \left[F(x_k) - F(x^*) \right] + \alpha_k (x_k - x^*) + \alpha_k (x^* - u_0) \right\} \\
&= x_k - x^* - \varepsilon_k D_k \left[M_k (x_k - x^*) + G + \alpha_k (x^* - u_0) \right] \\
&= x_k - x^* - \left[M_k^{-1} + (\varepsilon_k - 1) M_k^{-1} + \varepsilon_k (D_k - M_k^{-1}) \right] M_k (x_k - x^*) - \varepsilon_k D_k G - \varepsilon_k \alpha_k D_k (x^* - u_0) \\
&= - \left[(\varepsilon_k - 1)(x_k - x^*) + \varepsilon_k (D_k M_k - I)(x_k - x^*) + \varepsilon_k D_k G + \varepsilon_k \alpha_k D_k (x^* - u_0) \right].
\end{aligned}
$$

On account of the relation (5) we obtain

$$
\begin{aligned}
x_{k+1} - x^* = \ &- \left[\varepsilon_k - 1 + \varepsilon_k (D_k M_k - I) \right] (x_k - x^*) \\
&- \varepsilon_k D_k M_k M_k^{-1} G - \varepsilon_k \alpha_k D_k M_k M_k^{-1} F'^*(x_k) F'(x_k) v \\
&- \varepsilon_k \alpha_k D_k M_k M_k^{-1} \left[F'^*(x^*) F'(x^*) - F'^*(x_k) F'(x_k) \right] v.
\end{aligned}
$$

Clearly,

$$
M_k^{-1} F'^*(x_k) F'(x_k) v = M_k^{-1} \left[F'^*(x_k) F'(x_k) + \alpha_k I - \alpha_k I \right] v = \left(I - \alpha_k M_k^{-1} \right) v.
$$

Let $\langle \cdot, \cdot \rangle$ denote the inner product in H. Due to the relation $\left\| M_k^{-1} \right\| \leq \frac{1}{\alpha_k}$ and positivity of M_k^{-1} we have that

$$
0 \leq \langle \alpha_k M_k^{-1} x, x \rangle \leq \|x\|^2
$$

and thus

$$
\left\| I - \alpha_k M_k^{-1} \right\| = \sup_{\|x\| \leq 1} \left| \langle (I - \alpha_k M_k^{-1}) x, x \rangle \right| \leq 1.
$$

Therefore we can estimate

$$
\left\| M_k^{-1} F'^*(x_k) F'(x_k) v \right\| \leq \|v\|.
$$

Note that from (4) it follows that the inequality

$$
\left\| F'^*(x^*) F'(x^*) - F'^*(x_k) F'(x_k) \right\| \leq 2KL \|x^* - x_k\|, \quad k = 1, 2, \ldots \tag{15}
$$

is valid. Making use of (15) and the inequalities

$$
\varepsilon_k \leq 1 + c_\varepsilon, \quad \|D_k M_k\| = \|I + D_k M_k - I\| \leq 1 + \mu_0,
$$

we obtain

$$
\begin{aligned}
\tau_{k+1} \ &\leq \ \frac{\varepsilon_k KL(1 + \mu_0)}{2\alpha_k} \tau_k^2 + \left(|\varepsilon_k - 1| + \varepsilon_k \mu_k \right) \tau_k + \varepsilon_k \alpha_k (1 + \mu_0) \|v\| + 2KL \varepsilon_k (1 + \mu_0) \tau_k \|v\| \\
&\leq \ (1 + c_\varepsilon) \left[\frac{KL(1 + \mu_0)}{2\alpha_k} \tau_k^2 + \alpha_k (1 + \mu_0) \|v\| \right] + \left(c_\varepsilon + (1 + c_\varepsilon) \mu_0 + 2KL(1 + c_\varepsilon)(1 + \mu_0) \|v\| \right) \tau_k.
\end{aligned}
$$

Using the assumption $\tau_k \leq \delta \alpha_k$ and the notations (9)–(11), we get

$$
\tau_{k+1} \leq (a_1 \delta^2 + a_2 \delta + a_3) \alpha_k.
$$

Since δ is a solution of the equation

$$a_1 z^2 + (a_2 - \beta)z + a_3 = 0,$$

we have

$$a_1 \delta^2 + a_2 \delta + a_3 = \beta \delta,$$

hence

$$\tau_{k+1} \leq \delta \beta \alpha_k = \delta \alpha_{k+1}.$$

Thus, according to the principle of mathematical induction, we have proved that

$$\tau_k \leq \delta \alpha_k \; \forall k.$$

The statement of the theorem is proved.

3. NUMERICAL EXPERIMENTS

In order to illustrate the accuracy of our theoretical convergence result, we consider two test problems.

Our first test problem is solving the algebraic system of equations

$$F(x) = 0,$$

where the function F is the Powell singular function (see [9])

$$F(x) = \begin{pmatrix} x_1 + 10x_2 \\ \sqrt{5}(x_3 - x_4) \\ (x_2 - 2x_3)^2 \\ \sqrt{10}(x_1 - x_4)^2 \end{pmatrix}.$$

The exact solution is $x^* = (0,0,0,0)^T$ and the Jacobian F', given by

$$F'(x) = \begin{pmatrix} 1 & 10 & 0 & 0 \\ 0 & 0 & \sqrt{5} & -\sqrt{5} \\ 0 & 2(x_2 - 2x_3) & -4(x_2 - 2x_3) & 0 \\ 2\sqrt{10}(x_1 - x_4) & 0 & 0 & -2\sqrt{10}(x_1 - x_4) \end{pmatrix}$$

is singular at x^*. Note that for such problems also unmodified Newton's method has a linear convergence rate (see [4]). For the initial approximation we took $x_0 = (\frac{3}{\sqrt{11}}, -\frac{1}{\sqrt{11}}, 0, \frac{1}{\sqrt{11}})^T$. Since

$$F'^*(x^*)F'(x^*) = \begin{pmatrix} 1 & 10 & 0 & 0 \\ 10 & 100 & 0 & 0 \\ 0 & 0 & 5 & -5 \\ 0 & 0 & -5 & 5 \end{pmatrix},$$

the assumption (5) requires us to choose u_0 in the form $u_0 = (a, 10a, b, -b)$ for some real numbers a and b. As for this problem the coefficients K and L in (4) are not small (by direct computations it is possible to show that K can be taken to be 13 in the unit sphere and L is $\sqrt{160}$), the norm of the vector v in (5) should be quite small for the local convergence result to apply. We used $a = 0.001$ and $b = 0.001$ in our computations; this corresponds to $\|v\| \approx 1.7 \times 10^{-4}$.

First we present the results obtained for method (3) with the exact inverse $D_k = M_k^{-1}$. We considered three different constant step sizes ε_k (namely 0.7, 1.0, and 1.3) and also the method with ε_k chosen randomly

in the interval $[0.7, 1.3]$. According to Remark 1, we can use $c_\beta = 690$ for all values of β. Thus the regularization parameter was chosen in the form $\alpha_k = 690\beta^k$ for $\beta \in \{0.9, 0.7, 0.5, 0.3\}$. The results corresponding to 30 iterations are presented in Table 1. Here τ_k is the Euclidean distance between x^* and x_k. We can see the observed reductions in errors during the last four iterations, the average observed reduction coefficient based on the last 10 iterations, and the final error. The last line of the table corresponds to the standard Gauss–Newton method ($\varepsilon_k = 1$, $\beta = 0$). According to our result, the theoretical overall reduction rate of the error is guaranteed to be at least β for $\beta \geq 0.684$ if $\varepsilon_k = \varepsilon \in \{0.7, 1.3\}$ and for $\beta \geq 0.295$ if $\varepsilon_k = \varepsilon = 1$. We see that the numerical results confirm the theoretical estimates. It is not surprising that for the concrete problem we may get a higher convergence rate by choosing a value of β that is smaller than the ones allowed by the theorem but the last line for the standard Gauss–Newton method shows clearly that a too small value may cause slower convergence speeds.

In order to test our theoretical results in the case of approximate inversion of M_k, we used the method with $D_0 = M_0^{-1}$ and

$$D_k = D_{k-1} \sum_{i=0}^{3} (I - M_k D_{k-1})^i, \ k = 1, 2, \ldots.$$

The computations were done for $\beta \in \{0.9, 0.8, 0.7, 0.6, 0.5\}$ and it was verified that the accuracy of approximation of M_k^{-1} was not smaller than $\mu_0 = 0.1$. The corresponding results are presented in Table 2. From the convergence theorem it follows that in the case $\mu_0 = 0.1$ the linear convergence rate β is guaranteed for $\beta \geq 0.43$ if $\varepsilon_k = 1$ and for $\beta \geq 0.86$ if $\varepsilon_k = \varepsilon \in \{0.7, 1.3\}$. We see that the statement of the theorem is confirmed.

For our second test problem we consider the equation

$$\sin\left(\int_0^t x(s)\,ds\right) = t, \ t \in [0, 0.5], \tag{16}$$

with the exact solution

$$x^*(t) = \frac{1}{\sqrt{1 - t^2}}, \ t \in [0, 0.5].$$

Table 1. Numerical results for the exact D_k

ε	β	$\dfrac{\tau_{27}}{\tau_{26}}$	$\dfrac{\tau_{28}}{\tau_{27}}$	$\dfrac{\tau_{29}}{\tau_{28}}$	$\dfrac{\tau_{30}}{\tau_{29}}$	$\left(\dfrac{\tau_{30}}{\tau_{20}}\right)^{\frac{1}{10}}$	τ_{30}
0.7	0.9	0.936	0.934	0.933	0.932	0.939	2.6E-03
0.7	0.7	0.701	0.701	0.701	0.7	0.704	3.3E-06
0.7	0.5	0.5	0.5	0.5	0.5	0.5	1.9E-10
0.7	0.3	0.309	0.309	0.309	0.308	0.31	5.7E-16
1.0	0.9	0.935	0.934	0.932	0.931	0.938	2.5E-03
1.0	0.7	0.701	0.701	0.7	0.7	0.703	2.7E-06
1.0	0.5	0.5	0.5	0.5	0.5	0.5	1.1E-10
1.0	0.3	0.3	0.3	0.301	0.301	0.3	2.5E-17
1.3	0.9	0.935	0.933	0.932	0.931	0.937	2.4E-03
1.3	0.7	0.701	0.701	0.7	0.7	0.702	2.4E-06
1.3	0.5	0.5	0.5	0.5	0.5	0.5	9.0E-11
1.3	0.3	0.283	0.319	0.283	0.321	0.3	2.0E-16
Random	0.9	0.922	0.947	0.931	0.921	0.939	2.5E-03
Random	0.7	0.703	0.57	0.788	0.591	0.696	2.3E-06
Random	0.5	0.648	0.433	0.569	0.35	0.486	9.3E-11
Random	0.3	0.207	0.213	0.478	0.326	0.303	2.9E-17
Standard GN	0.0	0.5	0.5	0.5	0.5	0.5	7.1E-07

Table 2. Numerical results for an approximate D_k

ε	β	$\frac{\tau_{27}}{\tau_{26}}$	$\frac{\tau_{28}}{\tau_{27}}$	$\frac{\tau_{29}}{\tau_{28}}$	$\frac{\tau_{30}}{\tau_{29}}$	$\left(\frac{\tau_{30}}{\tau_{20}}\right)^{\frac{1}{10}}$	τ_{30}
0.7	0.9	0.936	0.934	0.933	0.932	0.939	2.6E-03
0.7	0.8	0.82	0.817	0.814	0.812	0.826	1.5E-04
0.7	0.7	0.701	0.701	0.701	0.7	0.704	3.3E-06
0.7	0.6	0.6	0.6	0.6	0.6	0.6	3.7E-08
0.7	0.5	0.5	0.5	0.5	0.5	0.5	1.9E-10
1.0	0.9	0.935	0.934	0.932	0.931	0.938	2.5E-03
1.0	0.8	0.818	0.815	0.813	0.811	0.824	1.4E-04
1.0	0.7	0.701	0.701	0.7	0.7	0.703	2.7E-06
1.0	0.6	0.6	0.6	0.6	0.6	0.6	2.6E-08
1.0	0.5	0.5	0.5	0.5	0.5	0.5	1.1E-10
1.3	0.9	0.935	0.933	0.932	0.931	0.937	2.4E-03
1.3	0.8	0.817	0.815	0.812	0.81	0.823	1.3E-04
1.3	0.7	0.701	0.701	0.7	0.7	0.702	2.4E-06
1.3	0.6	0.6	0.6	0.6	0.6	0.6	2.3E-08
1.3	0.5	0.5	0.5	0.5	0.5	0.5	9.0E-11

We view it as an operator equation in the space $L^2(0,0.5)$ of the square integrable functions. By denoting

$$Tu(t) = \int_0^t u(s)\,ds, \quad g(t) = t,$$

the former equation can be viewed as an operator equation of the form

$$F(x) = \sin(Tx) - g = 0, \ x \in L^2(0,0.5).$$

Clearly the Fréchet derivative is

$$[F'(x)z](t) = \cos(Tx(t))Tz(t), \quad z \in L^2(0,0.5),$$

and its adjoint operator is

$$[F'^*(x)z](t) = \int_t^{0.5} \cos(Tx(s))z(s)\,ds, \quad z \in L^2(0,0.5).$$

By direct calculation we get that the composition $F'^*(x)F'(x)$ is the integral operator

$$[F'^*(x)F'(x)z](t) = \kappa_x(t)\int_0^t z(s)\,ds + \int_t^{0.5} \kappa_x(s)z(s)\,ds, \ t \in [0,0.5],$$

where the function κ_x depends on x and is defined by

$$\kappa_x(t) = \int_t^{0.5} \cos^2(Tx(s))ds.$$

Note that for any $x,z \in L^2(0,0.5)$ the function $[F'^*(x)F'(x)z](t)$ is 0 at $t = 0.5$ and its (one-sided) derivative is 0 at $t = 0$, so the condition

$$x^* - u_0 = F'^*(x^*)F'(x^*)v$$

for some $v \in L^2(0,0.5)$ requires us to choose u_0 so that it is continuously differentiable and satisfies $u_0(0.5) = \frac{2}{\sqrt{3}}$, $u_0'(0) = 0$. It can be shown that any function u_0 that satisfies the given conditions and has a second derivative in $L^2(0,0.5)$ is an acceptable choice. We use the function

$$u_0(t) = \frac{2}{\sqrt{3}}\frac{1+t^2}{1+0.5^2}$$

for our test function. The corresponding function v is

$$v(t) = \left(\frac{x^{*\prime}(t) - u_0'(t)}{\kappa_{x^*}'(t)} \right)'$$

and by direct computation we get $\|v\| \approx 0.49$. It is relatively easy to check that the conditions (5) are satisfied with $K = \frac{1}{\sqrt{8}}$, $L = \frac{1}{4}$, so from Remark 1 we get that any coefficient $c_\beta \geq 0.23$ can be used for defining α_k in Theorem 1. We are going to use $c_\beta = 1$.

We start our iteration process by choosing $x_0(t) = 0$, $t \in [0, 0.5]$. In order to compute approximately

$$z = M_k^{-1} \left[F'^*(x_k) F(x_k) + \alpha_k (x_k - u_0) \right],$$

we solve the integral equation

$$\alpha_k z + F'^*(x_k) F'(x_k) z = F'^*(x_k) F(x_k) + \alpha_k (x_k - u_0)$$

numerically as follows. We fix the number of intervals $n = n(k)$ and consider the uniform grid

$$t_i = ih, \ h = \frac{1}{2n}, \ i = 0, 1, \ldots, n.$$

Denote $\eta_i = \frac{t_{i-1} + t_i}{2}$ and determine a piecewise constant approximate solution with the values ξ_i on the subintervals $(t_{i-1}, t_i]$, $i = 1, \ldots, n$, to the integral equation by requiring

$$\alpha_k \xi_i + h \, \kappa_{x_k}(\eta_i) \sum_{j=1}^{i} \xi_j + h \sum_{j=i+1}^{n} \kappa_{x_k}(\eta_j) \xi_j = b_i, \ i = 1, \ldots, n,$$

where

$$b_i = F_{ki} \cos(T x_k(\eta_i)) \frac{h}{2} + h \sum_{j=i+1}^{n} \left(F_{kj} \cos(T x_k(\eta_j)) \right) + \alpha_k (x_k(\eta_i) - u_0(\eta_i))$$

with

$$F_{ki} = \sin(T x_k(\eta_i)) - \eta_i, \ i = 1, \ldots, n.$$

This procedure corresponds to using a fully discretized collocation method and can be viewed as using an approximation D_k instead of M_k^{-1}. In order to keep the approximation error under control as α_k gets smaller, we have to increase the dimension n with k. We used $n(1) = 2$ and doubled the value of n at even values of k (so $n(2) = 4$, $n(3) = 4$, $n(4) = 8$ and so on). We used the method (3) with a constant step size $\varepsilon_k = \varepsilon$ for three different values of ε. The regularization parameter α_k was chosen to be of the form $\alpha_k = \beta^k$ for five different values of β. The numerical results are presented in Table 3.

Here $\tau_k = \|x_k - x^*\|_{L^2}$. In total, 25 iteration steps were carried through. The observed reductions in errors during the last four iterations, the average observed reduction coefficient based on the last 10 iterations and the final error are presented in the table. According to Theorem 1, the theoretical asymptotic reduction rate of the error is guaranteed to be at least β for sufficiently large $\beta < 1$ and the numerical results confirm the validity of the estimate: for $\beta \in \{0.9, 0.8, 0.7\}$ the observed reduction rate of the error is approaching the theoretical value. Our theorem does not cover the behaviour of the method if the parameter β is too small (less than q defined in (8), which is unfortunately unknown since the coefficient μ_0 is not known). From the numerical results we see that in the case of the test problem the convergence takes place also for relatively small values of β (namely for $\beta \in \{0.4, 0.2\}$) but the observed reduction rate is not related to the value of β any more and that a too small value of β does not improve the accuracy of the final result.

Table 3. Numerical results for equation (16)

ε	β	$\dfrac{\tau_{22}}{\tau_{21}}$	$\dfrac{\tau_{23}}{\tau_{22}}$	$\dfrac{\tau_{24}}{\tau_{23}}$	$\dfrac{\tau_{25}}{\tau_{24}}$	$\left(\dfrac{\tau_{25}}{\tau_{15}}\right)^{\frac{1}{10}}$	τ_{25}
0.7	0.9	0.954	0.951	0.949	0.946	0.958	1.7E-02
0.7	0.8	0.819	0.816	0.813	0.811	0.828	1.7E-03
0.7	0.7	0.707	0.708	0.706	0.707	0.71	7.5E-05
0.7	0.4	0.684	0.826	0.686	0.819	0.745	8.8E-06
0.7	0.2	0.648	0.773	0.648	0.772	0.709	1.3E-05
1.0	0.9	0.953	0.95	0.948	0.945	0.957	1.7E-02
1.0	0.8	0.817	0.814	0.812	0.81	0.825	1.5E-03
1.0	0.7	0.706	0.708	0.706	0.707	0.709	6.2E-05
1.0	0.4	0.651	0.864	0.656	0.858	0.737	7.2E-06
1.0	0.2	0.613	0.816	0.613	0.816	0.707	1.1E-05
1.3	0.9	0.952	0.95	0.947	0.944	0.956	1.6E-02
1.3	0.8	0.816	0.813	0.811	0.81	0.824	1.5E-03
1.3	0.7	0.706	0.708	0.706	0.707	0.708	5.6E-05
1.3	0.4	0.625	0.891	0.635	0.89	0.73	6.4E-06
1.3	0.2	0.574	0.872	0.574	0.872	0.707	9.6E-06

4. CONCLUDING REMARKS

In this paper we studied a family of two parametric iterative regularization methods based on the Gauss–Newton method for solving ill-posed problems in Hilbert spaces. It turned out that under suitable assumptions about the test function the family of methods has a geometric rate of convergence near the exact solution. The two numerical examples (one in a finite-, the other in an infinite-dimensional Hilbert space) confirm the validity of the obtained results. Of course, from the practical point of view there still remains the problem of choosing a good test function. Likewise, finding a proper regularization parameter α and, especially, a suitable compromise between the values of the parameters α and ε is a fragile problem. It is well known that mathematically efficient methods do not necessarily result in efficient computer programs. On the other hand, the worst case bounds of a method may be too pessimistic and in practice it may perform better than expected. Therefore the determination of the parameters on the basis of the heuristic considerations certainly deserves further investigation as well.

ACKNOWLEDGEMENTS

REFERENCES

1. Bakushinsky, A. B. and Kokurin, M. Y. *Iterative Methods for Approximate Solution of Inverse Problems*. Springer, Dordrecht, 2004.
2. Bakushinskij, A. B. and Polyak, B. T. On the solution of variational inequalities. *Dokl. Akad. Nauk SSSR*, 1974, **219**, 1038–1041 (in Russian).
3. Chen, J. The convergence analysis of inexact Gauss–Newton methods for nonlinear problems. *Comput. Optim. Appl.*, 2008, **40**, 97–118.
4. Decker, D. W., Keller, H. B., and Kelley, C. T. Convergence rates for Newton's method at singular points. *SIAM J. Numer. Anal.*, 1983, **20**(2), 296–314.
5. Domanskii, E. Sequential regularizability in the sense of Maslov for improperly posed problems. *Math. Notes*, 1994, **56**, 1036–1042.

6. Kaltenbacher, B. and Neubauer, A. Convergence of projected iterative regularization methods for nonlinear problems with smooth solutions. *Inverse Probl.*, 2006, **22**, 1105–1119.

7. Kaltenbacher, B., Neubauer, A., and Scherzer, O. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Radon Series on Computational and Applied Mathematics 6. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

8. Kangro, I. and Vaarmann, O. Two-parameter regularized Gauss–Newton type methods for highly nonlinear least squares problems. In *The 20th International Conference EurOPT'2008, Selected Papers, Neringa, Lithuania, 2008: Continuous Optimization and Knowledge-Based Technologies*. VGTU Publishing House "Technica", Institute of Mathematics and Informatics, 2008, 160–164.

9. Moré, J. J., Garbow, B. S., and Hillström, K. E. Testing unconstrained optimization software. *ACM Trans. Math. Soft.*, 1987, **7**, 17–41.

10. Parts, I. and Vaarmann, O. Damped Gauss–Newton type methods for ill-posed optimization problems. In *International Conference on Operational Research SOBI2006, Tallinn, Estonia, 2006: Simulation and Optimisation in Business and Industry*. Technologija, Kaunas, Kaunas University of Technology, 2006, 119–123.

11. Qi-Nian, J. On the iteratively regularized Gauss–Newton method for solving nonlinear ill-posed problems. *Math. Computation*, 2000, **69**, 1603–1623.

12. Qinian, J. A convergence analysis of the iteratively regularized Gauss–Newton method under the Lipschitz condition. *Inverse Probl.*, 2008, **24**, Article ID 045002, 1–16.

13. Vaarmann, O. Solution of nonlinear least squares problems by Levenberg–Marquardt type methods. *Proc. Estonian Acad. Sci. Phys. Math.*, 1989, **38**, 146–153.

14. Vaarmann, O. On solving ill-conditioned systems of nonlinear equations. *Proc. Estonian Acad. Sci. Phys. Math.*, 1994, **43**, 49–63.

15. Vaarmann, O. Some approximate methods for ill-posed problems. *Proc. Estonian Acad. Sci. Phys. Math.*, 1999, **48**, 146–151.

16. Vasin, V. Iterative regularization techniques for ill-posed problems. *Russian Math. (Iz. VUZ)*, 1995, **11**, 69–84.

17. Vasin, V. V. and Mokrushin, A. A. Iterative process of Gauss–Newton type for ill-posed operator equations. *Dokl. Akad. Nauk.*, 2000, **371**, 35–37.

Mõned aproksimatiivsed Gauss-Newtoni tüüpi meetodid mittelineaarsete mittekorrektsete ülesannete lahendamiseks

Inga Kangro, Raul Kangro ja Otu Vaarmann

Paljud numbrilise analüüsi ülesanded taanduvad mittelineaarse võrrandi $F(x) = 0$ lahendamisele, kus F on nõutav arv kordi diferentseeruv operaator ühest Hilberti ruumist teise. Tõsised raskused selle võrrandi lahendamisel tekivad siis, kui mõne x korral otsitava lahendi x^* ümbruses ei eksisteeri operaatori F Fréchet' tuletisel $F'(x)$ ühtlaselt tõkestatud (pseudo)pöördoperaatorit, st meil on tegemist oluliselt mittekorrektse ülesandega. Selliste ülesannete ligikaudseks lahendamiseks on artiklis esitatud mõned lahendusmeetodid, mis põhinevad iteratiivse regulariseerimise ja Gauss-Newtoni idee kasutamisel. Teatud eeldustel testfunktsiooni ja otsitava lahendi kohta on tõestatud lokaalne koonduvusteoreem Hilberti ruumis ning tuletatud koonduvuskiiruse hinnangud eelmainitud lahendusmeetoditele. Samuti on esitatud arvutustulemused, mis on saadud kahe näiteülesande lahendamisel artiklis kirjeldatud meetodite abil.