Proc. Estonian Acad. Sci. Phys. Math., 1999, **48**, 2, 146–151
https://doi.org/10.3176/phys.math.1999.2.08

SOME APPROXIMATE METHODS FOR SOLVING NONLINEAR ILL-POSED PROBLEMS

Otu VAARMANN

Tallinn Technical University, Ehitajate tee 5, 12618 Tallinn, Estonia; vaarmann@ioc.ee

Received 25 September 1998, in revised form 18 January 1999

Abstract. The problem of the approximate solution of the equation $F(x) = 0$, where F is a Frechet-differentiable operator from a Hilbert space H_1 into another Hilbert space H_2 such that the range of F' is not necessarily closed, is discussed. Under certain conditions on a test function and the required solution, a few convergence results are obtained for iterative regularization methods based on the Gauss–Newton method. To get a more realistic impression of convergence properties of the methods under consideration, their approximate variants are studied.

Key words: nonlinear operator equation, ill-posedness, iterative regularization methods, approximate methods.

1. INTRODUCTION

A fundamental problem occurring in scientific computing is that of solving the nonlinear equation

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

where F is a differentiable operator from one abstract space into another. In what follows we suppose that F is Frechet-differentiable and it is acting between the Hilbert spaces H_1 and H_2. Frequently the problem (1) is ill-posed, e.g., we cannot assume the existence of $(F')^{-1}$ or its boundedness. Ill-posedness is more seriously expressed in instability of the solution process and a straightforward implementation of standard methods may produce unrealistic results. This challenges us to generate special methods, in particular, which are stable and/or give us a generalized solution.

One way to find a generalized solution is to seek it in the least squares sense, i.e. minimizing the functional

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

or

$$\Phi(x) = \frac{1}{2} \| WF(x) \|^2, \tag{3}$$

where W is a weighting operator and may, in general, depend on the iteration number [1].

2. METHODS

A popular method for minimizing the functional (2) is the Gauss–Newton method

$$x_{k+1} = x_k - [F'^*(x_k)F'(x_k)]^{-1} F'^*(x_k)F'(x_k),\ k = 0, 1...,\tag{4}$$

which is equivalent to minimizing the linearized functional

$$g_k(h) = \frac{1}{2} \| F(x_k) + F'(x_k)h \|^2 \tag{5}$$

at every iteration point x_k for the correction term h.

In general, the inverse $(F'^*F)^{-1}$ does not always exist. To overcome this difficulty, we can replace (5) by the Tikhonov functional

$$\Psi_k(h) = \frac{1}{2}(\| F(x_k) + F'(x_k)h \|^2 + \alpha \| \bar{x}_k - u_0 \|^2), \tag{6}$$

where $\alpha > 0$, $\bar{x}_k = x_k + h$, and u_0 is an element of H_1, the so-called test function. However, following the idea of iterative regularization [2,3], we suppose that α is a sequence $\{\alpha_k\}$ of positive numbers which have to be properly chosen.

Since

$$\nabla\Psi_k(h) = F'^*(x_k)[F(x_k) + F'(x_k)h] + \alpha_k h + \alpha_k(x_k - u_0)$$

and

$$\nabla^2\Psi_k(h) = F'^*(x_k)F'(x_k) + \alpha_k I,$$

where I is the identity mapping and F'^* denotes the dual mapping of F', on the basis of the equation

$$\nabla\Psi_k(0) + \nabla^2\Psi_k(h) = 0$$

we obtain a one-parameter iteratively regularized Gauss–Newton method

$$x_{k+1} = x_k - M_k{}^{-1}[F'^*(x_k)F(x_k) + \alpha_k(x_k - u_0)], \tag{7}$$

where $M_k = B_k + \alpha_k I, B_k = B(x_k) = F'^*(x_k)F'(x_k)$, and $h_k = x_{k+1} - x_k$. So far the convergence of (7) is insufficiently investigated. Although methods of Gauss–Newton type allow us to discard the non-monotonicity condition, these still impose rather rigid requirements upon the choice of a test function and the a priori condition for the required solution.

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

$$x_{k+1} = x_k - D_k[F'^*(x_k)F(x_k) + \alpha_k(x_k - u_0)], \tag{8}$$

where D_k is an approximation to $M_k{}^{-1}$ satisfying the condition $\| I - D_k(B_k + \alpha_k I) \| \leq \mu_k \leq \mu < 1$.

To get convergence results for (7) and (8), let us suppose that the operator F is twice Frechet-differentiable in the region under discussion, with

$$\| F'(x) \| \leq K_1, \quad \| F''(x) \| \leq K_2, \quad K_1, K_2 > 0, \tag{9}$$

the solution x^* of (1) has the representation

$$x^* - u_0 = F'^*(x^*)F'(x^*)v \tag{10}$$

holding for an element $v \in H_1$ and a test function u_0, and the inequality

$$\| F'^*(x^*)F'(x^*) - F'^*(x_k)F'(x_k) \| \leq C_0 \| x^* - x_k \| \tag{11}$$

is valid.

It can be shown that the sequence $\{x_k\}$ defined by (8) converges to a solution x^* of (1) provided $\| v \|$ and μ_k are sufficiently small, i.e.

$$\mu_k \leq \mu_0 \leq \| v \|,$$

$$[2K_1K_2(1 + \mu_0)^2 \| v \|]^{1/2} + (C_0\mu_0 + C_0 + 1) \| v \| = q < 1 \tag{12}$$

and the parameter α_k is chosen as

$$\alpha_k = \left[\frac{K_1K_2}{2 \| v \|}\right]^{1/2} \tau_k, \tag{13}$$

where $\| x_k - x^* \| \leq \tau_k$.

Theorem. *Let the operator F be twice Frechet-differentiable and the relations* (9)–(13) *hold. Then the sequence $\{x_k\}$ defined by* (8) *converges to x^*, with*

$$\| x_k - x^* \| \leq q^k \tau_0,$$

where $\| x_0 - x^* \| \leq \tau_0$.

Proof. According to the Taylor formula,

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

with

$$\| G \| \leq \frac{K_1 K_2}{2} \| x_k - x^* \|^2. \tag{14}$$

Taking into account that $F(x^*) = 0$, on the basis of (8) and (14) we have

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

Owing to the relation

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

we have

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

Making use of the inequalities

$$\mu_k \leq \| v \|, \quad \| {M_k}^{-1} \| \leq \frac{1}{\alpha_k}, \quad \| M_k \| \leq {K_1}^2 + \alpha_0,$$

$$\| D_k M_k \| = \| I + D_k M_k - I \| \leq 1 + \mu_k,$$

we obtain

$$\begin{aligned} &\| x_{k+1} - x^* \| \\ &\leq \frac{K_1 K_2 (1 + \mu_k)}{2\alpha_k} \| x_k - x^* \|^2 + \mu_k \| x_k - x^* \| + (1 + \mu_k)\alpha_k \| v \| \\ &\quad + C_0(1 + \mu_k) \| x^* - x_k \| \| v \| \end{aligned}$$

$$= \frac{K_1 K_2 (1+\mu_k)}{2\alpha_k} {\tau_k}^2 + (1+\mu_k)\alpha_k \parallel v \parallel$$

$$+ (1+\mu_k)[C_0\mu_k + C_0 + 1] \parallel x_k - x^* \parallel \parallel v \parallel$$

$$\leq \frac{a_k \tau_k}{2[a_k/2b_k \parallel v \parallel]^{1/2}} + \left(\frac{a_k}{2b_k \parallel v \parallel}\right)^{1/2} b_k \parallel v \parallel \tau_k$$

$$+ [C_0\mu_k + C_0 + 1] \parallel v \parallel \tau_k,$$

where $a_k = K_1 K_2 (1+\mu_k)$ and $b_k = 1+\mu_k$.

It is not hard to prove that

$$\frac{a_k}{2[a_k/2b_k \parallel v \parallel]^{1/2}} + \frac{{a_k}^{1/2} b_k \parallel v \parallel}{(2b_k \parallel v \parallel)^{1/2}} = (2a_k b_k \parallel v \parallel)^{1/2}.$$

Thus $\parallel x_{k+1} - x^* \parallel \leq q\tau_k = \tau_{k+1}$ and $\tau_k = q^k \tau_0$.

Remark 1. If the quantity μ_k tends to zero as $k \to \infty$, e.g., at the rate of $\mu_k = C_1 \parallel x_k - x^* \parallel$, $C_1 > 0$, then

$$\parallel x_{k+1} - x^* \parallel \leq \left[\frac{K_1 K_2 (1+\mu_k)}{2\alpha_k} + C_1\right] \parallel x_k - x^* \parallel^2 + (1+\mu_k)\alpha_k \parallel v \parallel$$

$$+ C_0 (1+\mu_k) \parallel x^* - x_k \parallel \parallel v \parallel.$$

If we now take

$$a_k = K_1 K_2 (1+\mu_k) + 2C_1\alpha_k, \ \ b_k = 1+\mu_k, \text{ and } \alpha_k = \left(\frac{a_k}{2b_k \parallel v \parallel}\right)^{1/2} \tau_k,$$

$$q_k = (2a_k b_k \parallel v \parallel)^{1/2} + C_0(1+\mu_k) \parallel v \parallel,$$

then it can be similarly shown that

$$q_k = (2K_1K_2 + 2C_1\alpha_k)^{1/2} + C_0(1+\mu_k) \parallel v \parallel \leq q$$

provided

$$\mu_0 = C_1 \parallel x_0 - x^* \parallel \leq \parallel v \parallel.$$

Remark 2. Sometimes discretization serves as a regularizer [4]. For the nonlinear equation

$$F(x) = y, \qquad (15)$$

with the compact operator F, a suitable precedent discretization can be obtained by means of projection methods, replacing Eq. (15) by the equation $Q_h F P_h x = Q_h y$, where P_h and Q_h are orthoprojectors [5]. To solve (15), frequently the Galerkin method is used which has a remarkable property: if the formally written Galerkin approximation converges, then the limit is necessarily an exact solution [6]. Some regularized Newton-type methods are studied in [7].

ACKNOWLEDGEMENT

Financial support from the Estonian Science Foundation (grant No. 2249) is greatly appreciated.

REFERENCES

1. Vaarmann, O. On solving ill-conditioned systems of nonlinear equations. *Proc. Estonian Acad. Sci. Phys. Math.*, 1994, **43**, 2, 49–63.
2. Bakushinskii, A. B. and Polyak, B. T. On solving variational inequalities. *DAN SSSR*, 1974, **219**, 1038–1041 (in Russian).
3. Bakushinskii, A. B. On the convergence of the iteratively regularized Gauss–Newton method. *Zh. Vychisl. Mat. Mat. Fiz.*, 1992, **32**, 1503–1509 (in Russian).
4. Vainikko, G. and Hämarik, U. Projection methods and self-regularization in ill-posed problems. *Soviet. Math.*, 1985, **29**, 1–20.
5. Smirnova, A. B. Discretization and iterative approximation of the solutions to nonlinear ill-posed problems. *Izv. Vyssh. Uchebn. Zaved. Mat.*, 1992, **8**, 73–79.
6. Domanskij, E. Sequential regularizability in the sense of Maslov for improperly posed problems. *Math. Notes*, 1994, **4**, 1036–1042.
7. Kaltenbacher, B. Some Newton-type methods to the regularization of nonlinear ill-posed problems. *Inverse Problems*, 1997, **13**, 729–753.

MÕNED MITTELINEAARSETE MITTEKORREKTSETE ÜLESANNETE APROKSIMATIIVSED LAHENDUSMEETODID

Otu VAARMANN

Paljud numbrilise analüüsi ülesanded taanduvad mittelineaarse võrrandi $F(x) = 0$ lahendamisele, kus F on nõutav arv kordi diferentseeruv operaator, mis toimib ühest abstraktsest ruumist teise. Tõsised raskused võrrandi lahendamisel tekivad siis, kui tuletisel F' ei eksisteeri ühtlaselt tõkestatud (pseudo-) pöördoperaatorit, s.t. meil on tegu oluliselt mittekorrektse ülesandega. Selliste ülesannete lahendamiseks on esitatud mõned aproksimatiivsed lahendusmeetodid, mis põhinevad iteratiivse regulariseerimise ja Gaussi–Newtoni idee kasutamisel. On tõestatud lokaalne koonduvusteoreem Hilberti ruumis ja tuletatud koonduvuskiiruse hinnangud ülalmainitud lahendusmeetoditele, kusjuures selles teoreemis ei nõuta pöördoperaatori ühtlast tõkestatust.