MATHEMATICA, Tome 55 (78), N° 2, 2013, pp. 103–111

EXPANDING THE APPLICABILITY OF A NEWTON-LAVRENTIEV REGULARIZATION METHOD FOR ILL-POSED PROBLEMS

IOANNIS K. ARGYROS and SANTHOSH GEORGE

Abstract. We present a semilocal convergence analysis for a simplified Newton-Lavrentiev regularization method for solving ill-posed problems in a Hilbert space setting. We use a center-Lipschitz instead of a Lipschitz condition in our convergence analysis. This way we obtain: weaker convergence criteria, tighter error bounds and more precise information on the location of the solution than in earlier studies (such as [13]).

MSC 2010. 65H10, 65J20, 65J15, 65G99, 47H99, 49M15.

Key words. Newton-Lavrentiev regularization method, ill-posed problem, Hilbert space, semilocal convergence, Lipschitz condition.

1. INTRODUCTION

Let X and H be Hilbert spaces. Let $U(x,R)$ and $\overline{U(x,R)}$ stand, respectively, for the open and closed ball in X with center x and radius $R > 0$. Let also $L(X,H)$ be the space of all bounded linear operators from X into H.

In this study we are concerned with the problem of approximating a solution of the equation

$$AF(x) = y, \tag{1.1}$$

where $A : H \to H$ is a positive self-adjoint operator with its range $R(A)$ not closed in H and $F : D(F) \subseteq X \to H$.

Many problems from computational sciences and other disciplines can be brought in a form similar to equation (1.1) using mathematical modelling, e.g., [1], [4], [5], [6], [17], [18]. The solutions of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative. The study of the convergence of iterative procedures is usually based on two types of methods: on semi-local, respectively, on local convergence analysis. While the semi-local convergence analysis is based on the information around an initial point and gives conditions ensuring the convergence of the iterative procedure, the local one is based on the information around a solution and aims to find estimates of the radii of convergence balls.

Since $R(A)$ is not closed, the equation (1.1) is ill-posed in the sense that small perturbations of the data y can lead to large deviations of the "solution". In this case regularization techniques are needed to obtain stable approximate solutions for (1.1), e.g., [5]–[7], [8], [10], [16]. A solution x^* of (1.1) is called an x_0-*minimum norm solution* of (1.1) if

$$\|x^* - x_0\| = \min\{\|x - x_0\| : AF(x) = y,\ x \in D(F)\}. \tag{1.2}$$

We are interested in a solution $x^* \in D(F)$ of equation (1.1) which satisfies

$$\|F(x^*) - F(x_0)\| = \min\{\|F(x) - F(x_0)\| : AF(x) = y,\ x \in D(F)\}, \tag{1.3}$$

instead of (1.2). We also assume that $y^\delta \in H$ are the available noisy data with

$$\|y - y^\delta\| \leq \delta. \tag{1.4}$$

George and Nair used in [13] the Newton-Lavrentiev regularization method (NLRM) defined for each $n \in \{1, 2, \dots\}$, a fixed $\alpha > 0$ and $\delta > 0$ by

$$x^\delta_{n,\alpha} = x^\delta_{n-1,\alpha} - F'(x_0)^{-1}[(F(x^\delta_{n-1,\alpha}) - z^\delta_\alpha], \tag{1.5}$$

where $z^\delta_\alpha = F(x_0) + (K^*K + \alpha I)^{-1}(K^*y^\delta - AF(x_0))$ and $x^\delta_{0,\alpha} = x_0$ is an initial point to generate a sequence $\{x^\delta_{n,\alpha}\}$ for obtaining approximate solutions x^δ_α of the equation

$$F(x) = z^\delta_\alpha. \tag{1.6}$$

The regularization parameter α is chosen following the adaptive parameter selection procedure due to Pereverzev and Schock [15].

(NLRM) can be written in the more condensed form

$$x^\delta_{n,\alpha} = G^\delta_\alpha(x^\delta_{n-1,\alpha}),\ n \in \{1, 2, \dots\}, \tag{1.7}$$

where $G^\delta_\alpha(x) = x - F'(x_0)^{-1}(F(x) - z^\delta_\alpha)$. A semilocal convergence analysis was given by George and Nair in [13] under the conditions:

(C1) There exist $x_0 \in X,\ b > 0$ and $w > 0$ such that $F'(x_0)^{-1} \in L(H, X)$, $\|F'(x_0)\| \leq b$ and $\|F(\hat{x}) - F(x_0)\| \leq w$.

(C2) $F : U(x_0, R) \subseteq X \to H$, for some $R > 0$, is Fréchet differentiable and there exists a constant $L > 0$ such that for each $x, y \in U(x_0, R)$ the following Lipschitz condition holds

$$\|F'(x) - F'(y)\| \leq L\|x - y\|.$$

(C3) For $\alpha > 0$ and $\delta > 0$ set

$$h := 2b^2L\left(\frac{\delta}{\alpha} + w\right) \tag{1.8}$$

and

$$q := 1 - \sqrt{1 - h}. \tag{1.9}$$

Suppose that

$$\frac{\delta}{\alpha} + w < \frac{1}{2b}\min\left\{R, \frac{1}{bL}\right\}. \tag{1.10}$$

We shall refer to (C1)–(C2) as the (C) conditions. Note that (1.10) implies

$$h < 1 \tag{1.11}$$

and

$$r = \frac{1-\sqrt{1-h}}{bL} < R. \tag{1.12}$$

In the present paper we present a semilocal convergence analysis for (NLRM) with the following advantages over the one given in [13]:

(A1) less expensive computational cost for the Lipschitz constants,
(A2) weaker convergence conditions,
(A3) tighter error estimates on $\|x^{\delta}_{n,\alpha} - x^{\delta}_{\alpha}\|$ and $\|x_{n,\alpha} - x_{\alpha}\|$,
(A4) more precise information on the location of the solution.

We shall refer to (A1)–(A4) as the (A) advantages. Let us explain how these advantages are obtained. Note that in view of (C2) we get:

(C2)' There exists $L_0 > 0$ such that center Lipschitz condition

$$\|F'(x) - F'(x_0)\| \leq L_0\|x - x_0\|.$$

holds for each $x \in U(x_0, R)$

Consider also:

(C3)' Put

$$h_0 := 2b^2 L_0 \left(\frac{\delta}{\alpha} + w\right) \tag{1.13}$$

and

$$q_0 := 1 - \sqrt{1 - h_0}. \tag{1.14}$$

Suppose that

$$\frac{\delta}{\alpha} + w < \frac{1}{2b} \min\left\{R, \frac{1}{bL_0}\right\}. \tag{1.15}$$

Note that (1.15) yields

$$h_0 < 1 \tag{1.16}$$

and

$$r_0 = \frac{1-\sqrt{1-h_0}}{bL_0} < R. \tag{1.17}$$

We shall refer to (C1), (C2)' and (C3)' as the (C_0) conditions.

Note that the inequality

$$L_0 \leq L \tag{1.18}$$

holds in general, and that $\frac{L}{L_0}$ can be arbitrarily large, see [1], [2]–[4].

Our semilocal convergence analysis is based on the (C_0) conditions. Observe that

$$h < 1 \Longrightarrow h_0 < 1, \tag{1.19}$$

$$\frac{h_0}{h} \to 0 \text{ as } \frac{L_0}{L} \to 0, \tag{1.20}$$

and

$$r_0 \leq r. \tag{1.21}$$

If $L_0 < L$, then strict inequality will hold in (1.21), too. The estimate (1.20) shows that the applicability of these methods can be expanded infinitely many times when compared to the results in [13]. The computation of L_0 is less expensive than that of L. Finally, there are problems for which (C2)' holds but not (C2) (see the examples at the end of this study).

The proof of the result in [13] involving (C2) can be given by simply using (C2)'. Based on this crucial observation, the results in [13] can be given using the (C_0) conditions instead of (C). That is why in Section 2 we simply present the results without any proofs. We refer the reader to [13] for these proofs. Finally, in Section 3 we provide examples where (C2) does not hold but (C2)' holds.

2. CONVERGENCE ANALYSIS

We present the semilocal result for (NLRM).

THEOREM 2.1. *Suppose that the (C_0) conditions hold. Then the sequence $\{x^\delta_{n,\alpha}\}$, generated by (NLRM), is well-defined, remains in $U(x_0, r_0)$ for each $n \in \{0, 1, 2, \dots\}$ and converges to a unique solution $x^\delta_\alpha \in U(x_0, r_0)$ of the equation*

$$F(x) = z^\delta_\alpha \tag{2.22}$$

such that

$$\|x^\delta_{n,\alpha} - x^\delta_\alpha\| \leq \frac{q_0^n r_0}{1 - q_0}, \tag{2.23}$$

where

$$q_0 = 1 - \sqrt{1 - h_0}. \tag{2.24}$$

If in addition $bL_0R < 1$, then the following inequalities hold

$$\|x^\delta_\alpha - \hat{x}\| \leq \frac{b}{1 - bL_0R}\|z^\delta_\alpha - F(\hat{x})\|$$

and

$$\|x^\delta_{n,\alpha} - \hat{x}\| \leq \frac{q_0^n r_0}{1 - q_0} + \frac{b}{1 - bL_0R}\|z^\delta_\alpha - F(\hat{x})\|.$$

Moreover, if $F(\hat{x}) - F(x_0) \in \overline{R(A)}$, then, for $z_\alpha = F(x_0) + (A + \alpha I)^{-1}(y - F(x_0))$, the following assertion holds

$$\|F(\hat{x}) - z_\alpha\| \to 0 \ \text{ as } \ \alpha \to 0.$$

Concerning error bounds under source conditions we present the following result.

THEOREM 2.2. *Suppose that the* (C_0) *conditions hold, that* $bL_0R < 1$ *and that there exists a continuous, strictly increasing function* $\varphi : (0, a] \to (0, \infty)$ *with* $a > \|A\|$ *satisfying:*

- $\lim \lambda \to 0 \varphi(\lambda) = 0$,
- $\sup_{\lambda \geq 0} \dfrac{\alpha\varphi(\lambda)}{\lambda + \alpha} \leq \varphi(\alpha)$, *for all* $\alpha \in (0, a]$,
- *there exists* $v \in X$ *with* $\|v\| \leq 1$ *such that* $F(\hat{x}) = \varphi(K^*K)v$.

Then the following inequality holds for each $n \in \{1, 2, \dots\}$

$$\|\hat{x} - x^\delta_{n,\alpha}\| \leq \frac{b}{1 - bL_0R}(\varphi(\alpha) + \frac{\delta}{\sqrt{\alpha}}) + \frac{q_0^n r_0}{1 - q_0}.$$

Let $\psi(\lambda) := \lambda\sqrt{\varphi^{-1}(\lambda)}$, *where* $0 < \lambda \leq \|A\|$, *and*

$$\alpha_\delta := \varphi^{-1}(\psi^{-1}(\delta)).$$

If $\psi(\alpha_\delta) + w < \min\{R, \frac{1}{bL_0}\}$ *and* $r_\delta := \min\{q_0^n \leq \frac{\delta}{\alpha_\delta}\}$, *then*

$$\|x^\delta_{n_\delta,\alpha_\delta} - \hat{x}\| = O(\psi^{-1}(\delta)).$$

Concerning an adaptive selection of the parameter, which does not involve even the regularization method in an explicit sense, we have the following result (see [12] and [13] for more details).

THEOREM 2.3. *Suppose that*

$$2bw < R < \min\left\{2b(1 + w), \frac{1}{bL_0}\right\}, \quad \mu > \frac{2b}{R - 2bw}.$$

Define α_j, *for* $j \in \{0, 1, 2, \dots, N\}$, *and* k, *respectively, by* $\alpha_0 := \delta, \alpha_j := \mu^j\delta$, *for* $j \in \{1, 2, \dots, N\}$, *and*

$$k := \max\{i : \|z^\delta_{\alpha_i} - z^\delta_{\alpha_j}\| \leq 4\mu^{-j}, i = 0, 1, \dots, i\}.$$

Then the following inequality holds

$$\|F(\hat{x}) - z^\delta_{\alpha_k}\| \leq 6\mu\psi^{-1}(\delta),$$

where $\psi(t) := t\varphi(t)$, *for* $0 < t < \|A\|$. *Furthermore, if*

$$h_{0,k} := 2b^2L_0(w + \frac{1}{\mu^k}) < 1$$

and if $n_{0,k} := \min\{n : q^n_{0,k} \leq \frac{1}{\mu^k}\}$ *with* $q_{0,k} = 1 - \sqrt{1 - h_{0,k}}$, *then the following equality holds*

$$\|x^\delta_{n_{0,k},\alpha_k} - \hat{x}\| = O(\psi^{-1}(\delta)).$$

ALGORITHM 2.4. For $i, j \in \{0, 1, \dots, N\}$ we have

$$z^\delta_{\alpha_i} - z^\delta_{\alpha_j} = (\alpha_j - \alpha_i)(A + \alpha_jI)^{-1}(A + \alpha_iI)^{-1}(y^\delta - AF(x_0)).$$

Hence the adaptive algorithm associated with the choice of the parameter specified in Theorem 2.3 involves the following steps:

Part I:

- $i = 1$.
- Solve for $w_i : (A + \alpha_i I)w_i = y^\delta - AF(x_0)$.
- Solve for $z_{ij} : (A + \alpha_i I)z_{ij} = (\alpha_j - \alpha_i)w_i, \ \ j \leq i$.
- If $\|z_{ij}\| > 4\mu^{-j}$, then take $k = i - 1$.
- Otherwise, repeat with $i + 1$ in place of i.

Part II:

- $n = 1$.
- If $q_{0,k}^n \leq \mu^{-k}$, then take $n_{0,k} = n$.
- Otherwise, repeat with $n + 1$ in place of n.

Part III:

- Solve for $u_{j-1} : F'(x_0)u_{j-1} = F(x^\delta_{j-1,\alpha_k}) - z^\delta_{\alpha_k}$.
- $x^\delta_{j,\alpha_k} := x^\delta_{j-1,\alpha_k} - u_{j-1}, \ j = 1, 2, \ldots, n_{0,k}$.

REMARK 2.5. If $L_0 = L$ the results reduce to the ones in [13]. Otherwise, i.e., if $L_0 < L$, these results constitute an improvement with advantages (A) as stated in the introduction of this paper. Note also that in this case $q_0 < q$ and the choice of the parameters α, δ and of the functions φ and ψ are tighter. Moreover, the choice of x_0 plays the role of a selection criterion, see [11]. Finally note that the algorithm is also tighter, which in practice leads to fewer steps to achieve a desired error tolerance $\epsilon > 0$.

3. EXAMPLES

In this section we first present two examples where (C2) is not satisfied but (C2)' is satisfied.

EXAMPLE 3.1. Let $X = Y = \mathbb{R}$, $D = [0, \infty)$, $x_0 = 1$ and define the function F on D by

$$F(x) = \frac{x^{1+\frac{1}{i}}}{1+\frac{1}{i}} + c_1 x + c_2, \tag{3.25}$$

where c_1, c_2 are real parameters and $i > 2$ is an integer. Then $F'(x) = x^{1/i} + c_1$ is not Lipschitz on D. However, the central Lipschitz condition (C2)' holds for $L_0 = 1$. Indeed, we have

$$\begin{aligned} \|F'(x) - F'(x_0)\| &= |x^{1/i} - x_0^{1/i}| = \frac{|x - x_0|}{x_0^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}} \\ &\leq L_0 |x - x_0|. \end{aligned}$$

EXAMPLE 3.2. We consider the integral equations

$$u(s) = f(s) + \lambda \int_a^b G(s,t)u(t)^{1+1/n} \mathrm{d}t, \ \ n \in \mathbb{N}, \tag{3.26}$$

where f is a given continuous function satisfying $f(s) > 0, s \in [a, b]$, λ is a real number, and the kernel G is continuous and positive in $[a, b] \times [a, b]$.

For example, when $G(s, t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$\begin{cases} u'' = \lambda u^{1+1/n} \\ u(a) = f(a),\ u(b) = f(b). \end{cases}$$

These type of problems have been considered in [1], [2], [14].

Equations of the form (3.26) generalize the equations

$$u(s) = \int_a^b G(s, t) u(t)^n \mathrm{d}t \tag{3.27}$$

studied in [1], [2], [14]. Instead of (3.26) we can try to solve the equation $F(u) = 0$, where

$$F : \Omega \to C[a, b],\ \text{with } \Omega = \{u \in C[a, b] : u(s) \geq 0,\ s \in [a, b]\},$$

and

$$F(u)(s) = u(s) - f(s) - \lambda \int_a^b G(s, t) u(t)^{1+1/n} \mathrm{d}t.$$

The norm we consider in this case is the max-norm.

The derivative F' is given by

$$F'(u)v(s) = v(s) - \lambda \left(1 + \frac{1}{n}\right) \int_a^b G(s, t) u(t)^{1/n} v(t) \mathrm{d}t,\ v \in \Omega.$$

First of all, we notice that F' does not satisfy any Lipschitz-type condition in Ω. Let us consider, for instance, $[a, b] = [0, 1]$, $G(s, t) = 1$ and $y(t) = 0$. Then $F'(y)v(s) = v(s)$ and

$$\|F'(x) - F'(y)\| = |\lambda| \left(1 + \frac{1}{n}\right) \int_a^b x(t)^{1/n} \mathrm{d}t.$$

Assuming that F' is a Lipschitz function, then

$$\|F'(x) - F'(y)\| \leq L_1 \|x - y\|,$$

or, equivalently,

$$\int_0^1 x(t)^{1/n} \mathrm{d}t \leq L_2 \max_{x \in [0,1]} x(s),\ \text{for all } x \in \Omega, \tag{3.28}$$

and for a constant L_2. But this is not true. Consider, for example, the functions

$$x_j(t) = \frac{t}{j},\ \ j \geq 1,\ t \in [0, 1].$$

If these are substituted into (3.28), we obtain

$$\frac{1}{j^{1/n}(1 + 1/n)} \leq \frac{L_2}{j} \Leftrightarrow j^{1-1/n} \leq L_2(1 + 1/n),\ \ \forall j \geq 1.$$

This inequality is not true when $j \to \infty$. Therefore condition (3.28) is not satisfied in this case. However, condition (C2)' holds true. To show this, let $x_0(t) = f(t)$ and $\gamma = \min_{s\in[a,b]} f(s),\ \alpha > 0$. For $v \in \Omega$ we then have

$$\begin{aligned}\|[F'(x) \quad &- \quad F'(x_0)]v\| \\ &= |\lambda|\left(1+\frac{1}{n}\right)\max_{s\in[a,b]}\left|\int_a^b G(s,t)(x(t)^{1/n} - f(t)^{1/n})v(t)\mathrm{d}t\right| \\ &\le |\lambda|(1+\frac{1}{n})\max_{s\in[a,b]} G_n(s,t),\end{aligned}$$

where $G_n(s,t) = \frac{G(s,t)|x(t)-f(t)|}{x(t)^{(n-1)/n}+x(t)^{(n-2)/n}f(t)^{1/n}+\cdots+f(t)^{(n-1)/n}}\|v\|$.

Hence

$$\begin{aligned}\|[F'(x) - F'(x_0)]v\| &= \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}\max_{s\in[a,b]}\int_a^b G(s,t)\mathrm{d}t\|x - x_0\| \\ &\le L_0\|x - x_0\|,\end{aligned}$$

where $L_0 = \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}N$ and $N = \max_{s\in[a,b]}\int_a^b G(s,t)\mathrm{d}t$. Thus condition (C2)' holds true for sufficiently small λ.

In the last example we compare the "h" conditions.

Example 3.3. Let $X = Y = \mathbb{R}$, $D = [0,\infty)$, $x_0 = 1$, $A = F$, $D(F) = \overline{U(x_0, 1-p)}$ for $p \in (0, \frac{1}{2})$, $y = 0$, $y^\delta = 0$, $\delta = 0$ and define the function F on D by

$$F(x) = x^3 - p.$$

Then, using (3.28) and (C3), (C3)', we have that $b = \frac{1}{3}$, $\eta = \frac{1}{3}(1-p)$ and $L_0 = 3(3-p) < L = 6(2-p)$. Then we have that $h = 2bL\eta > 1$ for all $p \in (0, \frac{1}{2})$. Hence there is no guarantee that the method converges to $\hat{x} = \sqrt[3]{p}$. However, our new condition $4h_0 = 2bL_0\eta \le 1$ holds for all $p \in [0.418861170, 0.5)$. So the new result can be applied for $p \in [0.418861170, 0.5)$, but not the old ones.

If $p \in (0,1)$ and we choose, say $p = 0.7$, then we get that

$$L_0 = 2.3 < L = 2.6,\ \eta = 0.1,\ h = 0.52,\ h_0 = 0.46.$$

Hence the old and the new hypotheses are satisfied by

$$q_0 = 0.265153077 < q = 0.307179677.$$

That is the ratio of convergence is tighter with the new approach.

REFERENCES

[1] Argyros, I.K., *Convergence and Application of Newton-type Iterations*, Springer, 2008.

[2] Argyros, I.K., *A Semilocal convergence for directional Newton methods*, Math. Comp., **80** (2011), 327–343.

[3] Argyros, I.K. and Hilout, S., *Weaker conditions for the convergence of Newton's method*, J. Complexity, **28** (2012), 364–387.

[4] Argyros, I.K., Cho, Y.J. and Hilout, S., *Numerical methods for equations and its applications*, CRC Press, Taylor and Francis, New York, 2012.

[5] Binder, A., Engl, H.W. and Vessela, S., *Some inverse problems for a nonlinear parabolic equation connected with continuous casting of steel: stability estimate and regularization*, Numer. Funct. Anal. Optim., **11** (1990), 643–671.

[6] Engl, H.W., Hanke, M. and Neubauer, A., *Tikhonov regularization of nonlinear differential equations*, in *Inverse Methods in Action* (P.C. Sabatier, ed.), Springer-Verlag, New York, 1990, 92–105.

[7] Engl, H.W., Hanke, M. and Neubauer, A., *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1993.

[8] Kirsch, A., *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1996.

[9] Mair, B.A., *Tikhonov regularization for finitely and infinitely smoothing operators*, SIAM J. Math. Anal., **25** (1994), 135–147.

[10] Raus, T., *On the discrepancy principle for the solution of ill-posed problems*, Acta Comment. Univ. Tartu. Math., **672** (1984), 16–26.

[11] Engl, H.W., Kunisch, K. and Neubauer, A., *Convergence rates for Tikhonov regularization of nonlinear ill-posed problems*, Inverse Problems, **5** (1989), 523–540.

[12] George, S. and Kunhanandan, M., *An iterative regularization method for Ill-posed Hammerstein type operator equation*, Inverse Ill-posed Probl. Ser., **17** (2009), 831–844.

[13] George, S. and Nair, M.T., *A modified Newton-Lavrentiev regularization for nonlinear ill-posed Hammerstein-Type operator equation*, J. Complexity, **24** (2008), 228–240.

[14] Kantorovich, L.V. and Akilov, G.P., *Functional Analysis in Normed Spaces*, Pergamon Press, New York, 1964.

[15] Pereverzev, S. and Schock, E., *On the adaptive selection of the parameter in regularization of ill-posed problems*, SIAM. J. Numer. Anal., **43** (2005), 2060–2076.

[16] Scherzer, O., *A parameter choice for Tikhonov regularization for solving nonlinear inverse problems leading to optimal rates*, Appl. Math., **38** (1993), 479–487.

[17] Scherzer, O., *The use of Tikhonov regularization in the identification of electrical conductivities from overdetermined problems*, Inverse Problems, **5** (1989), 227–238.

[18] Scherzer, O., Engl, H.W. and Anderssen, R.S., *Parameter identification from boundary measurements in parabolic equation arising from geophysics*, Nonlinear Anal., **20** (1993), 127–156.

Received October 15, 2012
Accepted March 9, 2014

Cameron University
Department of Mathematical Sciences
Lawton, OK 73505, USA
E-mail: ioannisa@cameron.edu

National Institute of Technology Karnataka
Department of Mathematical and Computational Sciences
India-757 025
E-mail: sgeorge@nitk.ac.in