J. Nonlinear Funct. Anal. 2016 (2016), Article ID 24

MATHRES
Journal of Nonlinear Functional Analysis
Available online at http://jnfa.mathres.org

ON NEWTON'S METHOD FOR SEMISMOOTH EQUATIONS

I. K. ARGYROS[1], D. GONZÁLEZ[2,*]

[1]Cameron University, Department of Mathematical Sciences, Lawton, OK, USA

[2]Escuela de Ciencias Físicas y Matemáticas, Universidad de Las Américas, Quito, Ecuador

Abstract. We present local and semilocal convergence results for Newton's method in order to approximate solutions of semismooth equations. The local convergence results are given under weaker conditions than in earlier studies resulting to a larger convergence ball and a smaller radio of convergence. In the semilocal convergence case weaker conditions than before conditions are employed to show the convergence of Newton's method. Numerical examples illustrating the advantages of our approach are also presented in this study.

Keywords. Newton's method; Local–semilocal convergence; Semismooth function; Convergence ball.

1. Introduction

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

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

where F is a continuous mapping from a subset D of $\mathbb{R}^n$ into $\mathbb{R}^n$.

Many problems in Computational Sciences and other disciplines can be brought in form $F(x) = 0$, using Mathematical Modelling [3, 4, 8, 9, 10, 12, 15, 18, 19]. In particular, a large number of problems in Applied Mathematics and also in Engineering are solved by finding the solutions of certain equations of the form $F(x) = 0$. For example, dynamic systems are

*Corresponding author.

E-mail addresses: ioannisa@cameron.edu (I. K. Argyros), daniel.gonzalez.sanchez@udla.edu.ec (D. González)

Accepted February 18, 2016

mathematically modelled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time invariant system is driven by the equation $\dot{x} = Q(x)$ (for some suitable operator Q), where x is the state. Then, the equilibrium states are determined by solving on equation like $F(x) = 0$. Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequence converges to an optimal solution of the problem at hand. Except in special cases, the solutions of these equations cannot be found in closed form. That is why most commonly used solution methods for these equations are iterative. In particular, the practice of Numerical Analysis for finding such solutions is essentially connected to variants of Newton's method.

Newton's method is defined by

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k) \quad \text{for each} \quad k = 0,1,2\ldots, \tag{2}$$

where x_0 is an initial point and F is a continuously Fréchet differentiable function on D, i. e., F is a smooth function.

The study about convergence matter of iterative methods is usually centered on two types: semi-local and local convergence analysis. The semi-local convergence analysis is based on the information around an initial point, to give convergence conditions guaranteeing the convergence of the iterative process; while the local one is, based on the information around a solution to find estimates of the radii of the convergence balls.

There is a plethora on local as well as semilocal convergence results on Newton's method (2) under various Lipschitz type conditions on F'. We refer the reader to [1]-[19] and the references therein for this type of results.

However, in many interesting applications F is not a smooth function [6]-[9]. In particular, we are interested in the case when F is a semismooth function. We present local as well as semilocal convergence results under weaker conditions than in earlier studies such as [1]-[19]. In the case of local convergence, our convergence ball is larger and the ratio of convergence smaller than before [1]-[19]. These advantages are also obtained under weaker hypotheses. This type of improved convergence results are important in Computational Mathematics, since this way we have a wider choice of initial guess and we compute less iterates in order to obtain a desired error tolerance.

The paper is organized as follows: semilocal and local convergence analysis of Newton method is given in Section 2. Finally, the numerical examples illustrating the theoretical results are given in the concluding Section 3.

2. Main results

2. 1. Semilocal convergence

Theorem 2.1.1. *Suppose that the operator $F : D \subseteq X \to Y$ is Frchet-differentiable on some subset D of a Banach space X with values in a Banach space Y.*

$$F'(x)^{-1} \in L(Y,X) \quad \textit{for each} \quad x \in D; \tag{3}$$

$$\|F'(x_0)^{-1}F(x_0)\| \leq \eta \quad \textit{for some} \quad x_0 \in D; \tag{4}$$

$$\|F'(y)^{-1}(F(y) - F(x) - F'(y)(y-x))\| \leq K\|y-x\|^{1+\gamma_0} \tag{5}$$

for each $x, y \in D$ and some $\gamma_0, K \geq 0$;

$$\|F'(y)^{-1}(F'(y) - F'(x))(y-x)\| \leq M\|y-x\|^{1+\gamma} \tag{6}$$

for each $x, y \in D$ and some $\gamma, M \geq 0$;

$$0 \leq \alpha := K\eta^{\gamma_0} + M\eta^{\gamma} < 1 \tag{7}$$

and for

$$r = \frac{\eta}{1-\alpha}, \tag{8}$$

$$\overline{U}(x_0, r) \subseteq D. \tag{9}$$

Then, sequence $\{x_n\}$ generated by Newton's method is well defined, remains in $\overline{U}(x_0,r)$ for each $n=0,1,2,\ldots$ and converges to a solution $x^* \in \overline{U}(x_0,r)$ of equation $F(x)=0$. Moreover, the following estimates hold:

$$\|x_{n+1}-x_n\| \le \alpha \|x_n - x_{n-1}\| \quad \text{for each} \quad n=1,2,\ldots \tag{10}$$

and

$$\|x_n - x^*\| \le \frac{\alpha^n \eta}{1-\alpha} \quad \text{for each} \quad n=0,1,2,\ldots \tag{11}$$

Proof. It follows from (4), (7), (8) and Newton's method for $n=0$ that

$$\|x_1-x_0\| = \|F'(x_0)^{-1}F(x_0)\| \le \eta \le \frac{\eta}{1-\alpha} = r.$$

Hence, $x_1 \in \overline{U}(x_0,r)$. Using Newton's method for $n=1$, we get the approximation

$$\begin{aligned} F(x_1) &= F(x_1)-F(x_0)-F'(x_0)(x_1-x_0) \\ &= [F(x_1)-F(x_0)-F'(x_1)(x_1-x_0)] \\ &\quad +(F'(x_1)-F'(x_0))(x_1-x_0). \end{aligned} \tag{12}$$

Then, since $x_1 \in D$ we have that $F'(x_1)^{-1} \in L(Y,X)$. In view of (4), (5), (6), (7) and (12) we get that

$$\begin{aligned} \|x_2-x_1\| &= \|F'(x_1)^{-1}F(x_1)\| \\ &\le \|F'(x_1)^{-1}(F(x_1)-F(x_0)-F'(x_1)(x_1-x_0))\| \\ &\quad +\|F'(x_1)^{-1}(F'(x_1)-F'(x_0))(x_1-x_0)\| \\ &\le K\|x_1-x_0\|^{1+\gamma_0}+M\|x_1-x_0\|^{1+\gamma} \\ &= (K\|x_1-x_0\|^{\gamma_0}+M\|x_1-x_0\|^{\gamma})\|x_1-x_0\| \\ &\le (K\eta^{\gamma_0}+M\eta^{\gamma})\|x_1-x_0\| = \alpha\|x_1-x_0\| \end{aligned} \tag{13}$$

and

$$\begin{aligned} \|x_2-x_0\| &\le \|x_2-x_1\|+\|x_1-x_0\| \le \alpha\|x_1-x_0\|+\|x_1-x_0\| \\ &= \frac{1-\alpha^2}{1-\alpha}\|x_1-x_0\| \le \frac{1-\alpha^2}{1-\alpha}\eta \le r, \end{aligned} \tag{14}$$

which show that (10) holds for $n = 1$ and $x_2 \in \overline{U}(x_0, r)$.Let us assume that (10) holds for all $m \le n$ and $x_m \in \overline{U}(x_0, r)$. Then, by simply using x_{m-1}, x_m in place of x_0, x_1 in (12)–(14) we get that

$$\|x_{m+1} - x_m\| \le \alpha \|x_m - x_{m-1}\| \tag{15}$$

and

$$\|x_{m+1} - x_0\| \le \frac{1-\alpha^{m+1}}{1-\alpha}\|x_1 - x_0\| \le \frac{1-\alpha^{m+1}}{1-\alpha}\eta \le r, \tag{16}$$

which complete the induction for (10) and $x_{m+1} \in \overline{U}(x_0, r)$. It follows that sequence $\{x_n\}$ is complete and as such it converges to some $x^* \in \overline{U}(x_0, r)$ (since $\overline{U}(x_0, r)$ is a closed set). By letting $m \to +\infty$ in the estimate

$$\|F'(x_m)^{-1}F(x_m)\| = \|x_{m+1} - x_m\| \le \alpha^{m+1}\eta$$

and since $F'(x_m)^{-1} \in L(Y, X)$, we get that $F(x^*) = 0$. We also have that

$$\begin{aligned} \|x_{n+m} - x_n\| &\le \|x_{n+m} - x_{n+m-1}\| + \|x_{n+m-1} - x_{n+m-2}\| + \cdots + \|x_{n+1} - x_n\| \\ &\le (\alpha^{n+m-1} + \alpha^{n+m-2} + \cdots + \alpha^n)\|x_1 - x_0\| \\ &\le \alpha^n \frac{1-\alpha^m}{1-\alpha}\|x_1 - x_0\| \le \alpha^n \frac{1-\alpha^m}{1-\alpha}\|x_1 - x_0\|. \end{aligned} \tag{17}$$

By letting $m \to +\infty$ in (17) we obtain (11).

Remark 2.1.2.

a) Condition (6) does not necessarily imply that F' is Lipschitz and according to (12) cannot be avoided if we want to show convergence. Let $\gamma = 0$, e. g., then, by (6) F' is not Hölder or necessarily Lipschitz.

b) If we use

$$\|F'(y)^{-1}\| \le K_1, \tag{18}$$

$$\|F(y) - F(x) - F'(y)(y - x)\| \le K_2\|y - x\|^{1+\gamma_0}, \tag{19}$$

$$\|(F'(y) - F'(x))(y - x)\| \le M_1\|y - x\|^{1+\gamma}. \tag{20}$$

Then due to the estimates

$$\begin{aligned}\|F'(y)^{-1}(F(y)-F(x)-F'(y)(y-x))\| &\leq \|F'(y)^{-1}\|\|F(y)-F(x)\\ &\quad -F'(y)(y-x)\|\\ &\leq K_1K_2\|y-x\|^{1+\gamma_0}\end{aligned}$$

and

$$\begin{aligned}\|F'(y)^{-1}(F'(y)-F'(x))(y-x)\| &\leq \|F'(y)^{-1}\|\|(F'(y)-F'(x))(y-x)\|^{1+\gamma}\\ &\leq K_1M_1\|y-x\|^{1+\gamma}, \qquad (21)\end{aligned}$$

we can set $K=K_1K_2$, $M=K_1M_1$. If (5), (6) are replaced by (18)–(20), then, the conclusions of Theorem 2.1.1 hold in this stronger though setting.

c) Notice that due the estimate

$$\|F'(y)^{-1}(F'(y)-F'(x))(y-x)\| \leq \|F'(y)^{-1}(F'(y)-F'(x))\|\|y-x\|$$

M in (6) can be chosen to be an upper bound on $\|F'(y)^{-1}(F'(y)-F'(x))\|$. That is

$$\|F'(y)^{-1}(F'(y)-F'(x))\| \leq M,$$

which must be less than one.

Notice also that $\|F'(y)-F'(x)\| \leq M_2$, holds, if e. g. F' is continuous. Then, for $\gamma=0$, we can choose $M=K_1M_2$.

d) If we drop γ_0,γ from (5), (6) and define α by $\alpha=K+M$, then the same proof leads to even weaker results. Indeed, suppose that D is convex and only (6) holds. In this case we choose $K=\frac{M}{2}$, since in this case

$$\left\|F'(y)^{-1}\int_0^1(F'(x+\theta(y-x))-F'(y))\,d\theta\right\| \leq K.$$

2.2. Local convergence

Theorem 2.2.1. *Suppose that $F: D\subseteq X\to Y$ is a Frchet differentiable operator; there exists $x^*\in D$ such that $F(x^*)=0$; $F'(x)^{-1}\in L(Y,X)$ for each $x\in D$;*

$$\|F'(y)^{-1}(F(y)-F(x^*)-F'(y)(y-x^*))\| \leq \lambda\|y-x^*\|^{1+b} \qquad (22)$$

for each $x,y \in D$, some $\lambda > 0$ and $b \geq 0$ and for

$$R = \min\left\{\frac{1}{\lambda}, \frac{1}{\lambda^b}\right\}, \tag{23}$$

$$\overline{U}(x_0,R) \subseteq D. \tag{24}$$

Then, sequence $\{x_n\}$ generated by Newton's method converges to x^* provided that $x_0 \in U(x^*,R)$. Moreover, the following estimates hold for each $n = 0,1,2,\ldots$

$$\|x_{n+1} - x^*\| \leq \lambda \|x_n - x^*\|^{1+b} \leq \|x_0 - x^*\| < R \tag{25}$$

and

$$\|x_{n+1} - x^*\| \leq \lambda^{-\frac{1}{b}} (\lambda \|x_0 - x^*\|)^{(1+b)^{n+1}} \tag{26}$$

Proof. We have that $x_1 \in U(x^*,R)$ by the choide of R. Then, using the estimate

$$x_1 - x^* = -F'(x_1)^{-1}[F(x_1) - F(x^*) - F'(x_1)(x_1 - x^*)] \tag{27}$$

and (22), we get that

$$\begin{aligned} \|x_1 - x^*\| &= \|-F'(x_1)^{-1}[F(x_1) - F(x^*) - F'(x_1)(x_1 - x^*)]\| \\ &\leq \lambda \|x_1 - x^*\|^{1+b} \leq \|x_1 - x^*\| < R \end{aligned} \tag{28}$$

by the choice of R, which show (25) and (26) for $n = 0$. Suppose that (25) and (26) hold for each $k \leq n$. Then, we have that

$$x_{k+1} - x^* = -F'(x_k)^{-1}[F(x_k) - F(x^*) - F'(x_k)(x_k - x^*)]. \tag{29}$$

By (22) and (29) we get that

$$\begin{aligned} \|x_{k+1} - x^*\| &= \|F'(x_k)^{-1}[F(x_k) - F(x^*) - F'(x_k)(x_k - x^*)]\| \\ &\leq \lambda \|x_k - x^*\|^{1+b} \leq \lambda \left(\lambda \|x_{k-1} - x^*\|^{1+b}\right)^{1+b} \\ &\leq \lambda \lambda^{1+b} \|x_{k-1} - x^*\|^{1+b^2} \leq \lambda^{1+(1+b)+\cdots+(1+b)^k} \|x_0 - x^*\|^{(1+b)^{k+1}} \\ &= \lambda^{\frac{(1+b)^{k+1}+\cdots+(1+b)+1}{(1+b)-1}} \|x_0 - x^*\|^{(1+b)^{k+1}} \\ &= \lambda^{-\frac{1}{b}} (\lambda \|x_0 - x^*\|)^{k+1}, \end{aligned} \tag{30}$$

which show (25), (26) for all n and that $\lim_{k\to+\infty} x_k = x^*$.

Remark 2.2.2.

(a) Condition (22) certainly holds if replaced by the stronger

$$\|F'(y)^{-1}(F(y)-F(x)-F'(y)(y-x))\| \leq \lambda_1 \|y-x\|^{1+b} \tag{31}$$

for each $x,y \in D$. Note that $\lambda \leq \lambda_1$.

(b) Notice also that a condition of the type (6) (or (4)) is not needed for the local convergence.

3. Examples

In this section, we present an example for the local convergence case.

Example 3.1. *Let $X = Y = \mathbb{R}$, $D = U(0,1)$ and define function F on D by*

$$F(x) = e^x - 1. \tag{32}$$

Then, we have that $x^* = 0$. Using (32) we get in turn that

$$\begin{aligned} F'(y)^{-1}(F(y)-F(x^*)-F'(y)(y-x^*)) &= 1-y-\left(1-y+\frac{y^2}{2!}-\frac{y^3}{3!}+\frac{y^4}{4!}-\cdots\right) \\ &= \left(\frac{1}{2!}-\frac{y}{3!}+\frac{y^2}{4!}-\frac{y^3}{5!}+\cdots\right)y^2. \end{aligned}$$

So

$$\begin{aligned} \|F'(y)^{-1}\left(F(y)-F(x^*)-F'(y)(y-x^*)\right)\| &= \left|\frac{1}{2!}-\frac{y}{3!}+\frac{y^2}{4!}-\frac{y^3}{5!}+\cdots\right||y|^2 \\ &\leq \left(\frac{1}{2!}+\frac{|y|}{3!}+\frac{|y|^2}{4!}+\frac{|y|^3}{5!}+\cdots\right)|y|^2 \\ &\leq \left(\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}+\cdots\right)|y|^2 \\ &= (e-2)|y|^2. \end{aligned}$$

Hence, we can choose $\lambda = e-2$ *and* $b=1$. *Moreover, we have that*

$$\begin{aligned} F'(y)^{-1}\left(F(y)-F(x)-F'(y)(y-x)\right) &= e^{-y}(e^y-1-e^x+1-e^y(y-x)) \\ &= 1-(y-x)-e^{x-y} \\ &= 1+x-y-\left(1+(x-y)+\frac{(x-y)^2}{2!}+\frac{(x-y)^3}{3}+\cdots\right) \\ &= \left(\frac{1}{2!}+\frac{x-y}{3!}+\frac{(x-y)^2}{4!}+\cdots\right)(x-y)^2. \end{aligned}$$

So,

$$\begin{aligned} \|F'(y)^{-1}\left(F(y)-F(x)-F'(y)(y-x)\right)\| &= \left|\frac{1}{2!}+\frac{x-y}{3!}+\frac{(x-y)^2}{4!}+\cdots\right||y-x|^2 \\ &\leq \left(\frac{1}{2!}+\frac{|x-y|}{3!}+\frac{|x-y|^2}{4!}+\cdots\right)|y-x|^2 \\ &\leq \left(\frac{1}{2!}+\frac{2}{3!}+\frac{2^2}{4!}+\cdots\right)|y-x|^2 \\ &\leq (e^2-3)|y-x|^2. \end{aligned}$$

Hence, we can choose $\lambda_1 = e^2-3$ *and* $b=1$. *Therefore, we obtain*

$$R_1 = \frac{1}{\lambda_1} = 0.227839421 < 1.392211191 = \frac{1}{\lambda} = R.$$

That is, we deduce that the new convergence ball is larger and the radio of convergence smaller than the old convergence ball and the old radio of convergence of Newton's method is guaranteed by Theorem 2.1.1 provided that $x_0 \in U(x_0,R)$.

REFERENCES

[1] S. Amat, S. Busquier and M. Negra, Adaptive approximation of nonlinear operators, Numer. Funct. Anal. Optim. 25 (2004), 397–405.

[2] S. Amat, S. Busquier, D. El Kebir and J. Molina, A fast Chebyshev's method for quadratic equations, Appl. Math. Comput. 148 (2004), 461-474.

[3] I. K. Argyros, Convergence and Applications of Newton–type iterations, Springer, (2008).

[4] I. K. Argyros and S. Hilout, Computational methods in nonlinear analysis, World Scientific Publications Company. New Jersey, (2013).

[5] I. K. Argyros and S. Hilout, On an improved convergence analysis of Newton's method, Appl. Math. Comput. 225 (2013), 372–386.

[6] E. Bierstone and P. D. Milman, Semianalitic and subanalitic sets, IHES. Publ. Math. 67 (1988), 5–42.

[7] J. Bolte, A. Daniilidis and S. A. Lewis, Tame mapping are semismooth, Math. Program. 117 (2009), 5–19.

[8] F. H. Clarke, Optimization and Nonsmooth Analysis, Society for Industrial and Applied Mathematics (1990).

[9] J. P. Dedieu, Penalty functions in subanalytic optimization, Optimization 26 (1992), 27–32.

[10] J. E. Jr. Dennis and R. B. Schnabel, Numerical methods of unconstrained optimization and nonlinear equations, Prentice–Hall, Englewood Cliffs, N. J. (1983).

[11] P. Deuflhard, Newton methods for nonlinear problems: Affine invariance and adaptive algorithms, Berlin, Springer–Verlag, (2004).

[12] J. A. Ezquerro, J. M. Gutiérrez, M. A. Hernández and M. A. Salanova, Solving nonlineal integral equations arising in radiative transfer, Numer. Funct. Anal. Optim. 20 (1999), 661–673.

[13] J. A. Ezquerro, J. M. Gutiérrez, M. A. Hernández and M. A. Salanova, Chebyshev–like methods and quadratic equations, Revue d'Analyse Numérique et de Théorie de l'Approximation, 28 (1) (2000), 23–35.

[14] M. A. Hernández, A note on Halley's method, Numer. Math. 59 (1991), 273–276.

[15] M. A. Hernández and N. Romero, On a characterization of some Newton–like methods of R–order at least three, J. Comput. Appl. Math. 183 (2005), 53–66.

[16] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, SIAM J. Control Optim. 15 (1997), 957–972.

[17] L. Qi and and J. Sun, A nonsmooth version of Newton's method, Math. Program. 58 (1993), 353–367.

[18] R. T. Rockafellar, Favorable classes of Lipschitz–continuous functions in subgraient optimization, in E. Nurminski. Nondifferentiable Optimization, Pergamon Press, New York, (1982) 125–143.

[19] L. Van Den Dries and C. Miller, Geometric categories and o-minimal structures, Duke Math. J. 84 (1996), 497–540.