Applied Mathematics Letters 26 (2013) 566–570

Contents lists available at SciVerse ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

On the local convergence of Newton's method under generalized conditions of Kantorovich

J.A. Ezquerro *, D. González, M.A. Hernández

University of La Rioja, Department of Mathematics and Computation, C/ Luis de Ulloa s/n, 26004 Logroño, Spain

ARTICLE INFO

Article history:
Received 11 September 2012
Received in revised form 19 December 2012
Accepted 20 December 2012

Keywords:
Newton's method
Local convergence
Order of convergence

ABSTRACT

Following an idea similar to that given by Dennis and Schnabel (1996) in [2], we prove a local convergence result for Newton's method under generalized conditions of Kantorovich type.

1. Introduction

Let X and Y be two Banach spaces and $F : \Omega \subseteq X \to Y$ an operator defined on a non-empty open convex domain Ω of X with values in Y. The best-known iteration for solving the equation

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

is Newton's method, which is defined as follows:

$$\begin{cases} x_0 \in \Omega, \\ x_n = x_{n-1} - [F'(x_{n-1})]^{-1}F(x_{n-1}), \quad n \in \mathbb{N}. \end{cases} \tag{2}$$

The first study of Newton's method in Banach spaces was given by L.V. Kantorovich [1], who obtained a first result on the semilocal convergence, which is known as the Newton–Kantorovich theorem. Kantorovich proved the theorem under the following conditions on the operator F and the starting point x_0:

(S_1) There exists $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$ for some $x_0 \in \Omega$, $\|\Gamma_0\| \le \beta$ and $\|\Gamma_0 F(x_0)\| \le \eta$, where $\mathcal{L}(Y, X)$ is the set of bounded linear operators from Y to X.
(S_2) $\|F''(x)\| \le M$ for $x \in \Omega$.
(S_3) $M\beta\eta \le \frac{1}{2}$ and $B\left(x_0, \frac{1-\sqrt{1-2M\beta\eta}}{M\beta}\right) \subset \Omega$.

While the semilocal convergence results require conditions on the operator F (see (S_2)) and the starting point x_0 (see (S_1)), the local convergence results require conditions on the operator F and a solution x^* of Eq. (1). An interesting local result, given by Dennis and Schnabel in [2], for Newton's method requires the following conditions:

(L_1) Let x^* be a solution of Eq. (1) such that the operator $[F'(x^*)]^{-1}$ exists, $B(x^*, r) \subset \Omega$ and $\|[F'(x^*)]^{-1}\| \le \gamma$, with $r, \gamma > 0$.
(L_2) $\|F''(x)\| \le M$ for $x \in \Omega$.

* Corresponding author.
E-mail addresses: jezquer@unirioja.es (J.A. Ezquerro), daniel.gonzalez@unirioja.es (D. González), mahernan@unirioja.es (M.A. Hernández).

0893-9659/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aml.2012.12.012

Dennis and Schnabel prove, under (L_1) and (L_2), that for any starting point in $B(x^*, \varepsilon)$, where $\varepsilon = \min\{r, R\}$ and $R = \frac{1}{2\gamma M}$, Newton's method is convergent. The local results provide what we call a ball of convergence, $B(x^*, \varepsilon)$. From the value ε, this ball of convergence gives information about the accessibility of the solution x^* of the equation to solve by the iterative method considered to approximate x^*.

Rall in [3] and Rheinboldt in [4] give results similar to that given by Dennis and Schnabel, but with different radii for the balls of convergence. On the other hand, Argyros in [5] generalizes the previous conditions by modifying condition (L_2) by means of a Lipschitz condition for the second derivative of F.

In this work, we analyse the local convergence of Newton's method under generalized conditions. In particular, we generalize the local convergence conditions given by Dennis and Schnabel in (L_1) and (L_2). Moreover, we observe that the new local convergence conditions required do not reduce the accessibility of the solution when it is approximated by Newton's method. Furthermore, in some cases, this accessibility improves that given by Dennis and Schnabel under conditions (L_1) and (L_2).

During the last fifty years, a lot of authors have studied the convergence of Newton's method, both local and semilocal, by modifying the conditions required to the operator F, namely (S_2) or (L_2), see [6–9] and the references given there. In [10] we present a generalization of (L_2) that consists of considering the condition $\|F''(x)\| \le \omega(\|x\|)$, $x \in \Omega$, where $\omega : [0, +\infty) \to \mathbb{R}$ is a non-decreasing continuous function such that $\omega(0) \ge 0$. In this work, we present a generalization of the previous condition to high order derivatives of the operator F; in particular,

$$\|F^{(k)}(x)\| \le \omega(\|x\|), \quad x \in \Omega,\ k \ge 3, \tag{3}$$

where $\omega : [0, +\infty) \to \mathbb{R}$ is a non-decreasing continuous function such that $\omega(0) \ge 0$. It is clear that an interesting situation is given when (1) is a polynomial equation of degree k, since the operator $F^{(k)}(x)$ is such that $\|F^{(k)}(x)\| \le M$, $x \in \Omega$, and consequently $F^{(k)}(x)$ always satisfies condition (3). Even, for more general equations, by using Taylor's series, Eq. (1) can be approximated by polynomial equations.

In Section 2, we prove a new local convergence result for Newton's method. In Section 3, we present an example where we show that the new local convergence conditions do not restrict the accessibility of Newton's method when it is used to approximate a solution of a particular equation. We even find situations in which the accessibility is improved.

2. Local convergence and the order of convergence

Now, we obtain a new local convergence result for Newton's method when the operator F satisfies condition (3). For this, we follow an idea similar to that given by Dennis and Schnabel in [2].

Theorem 1. *Let $F : \Omega \subseteq X \to Y$ be a nonlinear operator that is k ($k \ge 3$) times continuously differentiable on a non-empty open convex domain Ω of a Banach space X with values in a Banach space Y. Let x^* be a solution of $F(x) = 0$ such that the operator $[F'(x^*)]^{-1}$ exists, $B(x^*, r) \subseteq \Omega$, $\|[F'(x^*)]^{-1}\| \le \gamma$ and $\|F^{(i)}(x^*)\| \le \alpha_i$ (for $i = 2, 3, \dots, k-1$) with $r, \gamma, \alpha_i > 0$. Suppose that condition (3) is satisfied and there exists the smallest positive root R of the equation*

$$\gamma\left(\sum_{i=1}^{k-2} \frac{\alpha_{i+1}}{i!} t^{i-1} + \frac{t^{k-2}}{(k-1)!}\omega(\|x^*\| + t)\right) t - \delta = 0, \tag{4}$$

where $\delta \in \left(0, \frac{k}{2k-1}\right)$. Then, there exists $\varepsilon > 0$ such that Newton's sequence $\{x_n\}$ is well-defined and converges to x^* for every $x_0 \in B(x^*, \varepsilon)$. Moreover,

$$\|x^* - x_n\| < \frac{\delta}{\varepsilon}\|x^* - x_{n-1}\|^2, \quad n \in \mathbb{N}. \tag{5}$$

Proof. Let $\varepsilon = \min\{r, R\}$. First, we prove, for all $x \in B(x^*, \varepsilon)$, that there exists $[F'(x)]^{-1}$ and $\|[F'(x)]^{-1}\| \le \frac{\gamma}{1-\delta}$. For this, we consider

$$\begin{aligned}
\|I - [F'(x^*)]^{-1}F'(x)\| &\le \|[F'(x^*)]^{-1}\| \left\| \int_0^1 F''\Big(x + \tau(x^* - x)\Big)\, d\tau\, (x^* - x) \right\| \\
&\le \|[F'(x^*)]^{-1}\| \Bigg[\left\| \int_0^1 \left(\sum_{i=2}^{k-1} \frac{(\tau - 1)^{i-2}}{(i-2)!} F^{(i)}(x^*)(x^* - x)^{i-2}\, d\tau \right)(x^* - x) \right\| \\
&\quad + \frac{1}{(k-3)!} \int_0^1 \int_0^1 \left\| F^{(k)}\left(x^* + s(\tau - 1)(x^* - x)\right) \right\| (1 - s)^{k-3}(1 - \tau)^{k-2} \|x^* - x\|^{k-1}\, ds\, d\tau \Bigg] \\
&\le \|[F'(x^*)]^{-1}\| \Bigg[\sum_{i=2}^{k-1} \frac{1}{(i-1)!} \alpha_i \|x^* - x\|^{i-1}
\end{aligned}$$

$$+ \frac{1}{(k-3)!}\omega(\|x^*\|+\varepsilon)\left(\int_0^1 (1-s)^{k-3}\,ds\right)\left(\int_0^1 (1-\tau)^{k-2}\,d\tau\right)\|x^*-x\|^{k-1}\Bigg]$$

$$\leq \gamma\left[\sum_{i=1}^{k-2}\frac{1}{i!}\alpha_{i+1}\varepsilon^{i-1} + \frac{1}{(k-1)!}\omega(\|x^*\|+\varepsilon)\varepsilon^{k-2}\right]\varepsilon,$$

since

$$F''\left(x+\tau(x^*-x)\right) = \sum_{i=2}^{k-1}\frac{(\tau-1)^{i-2}}{(i-2)!}F^{(i)}(x^*)(x^*-x)^{i-2} + \frac{1}{(k-3)!}\int_0^1 F^{(k)}\left(x^*+s(\tau-1)(x^*-x)\right)(1-s)^{k-3}(\tau-1)^{k-2}(x^*-x)^{k-2}\,ds,$$

$$\left\|\int_0^1\left(\sum_{i=2}^{k-1}\frac{(\tau-1)^{i-2}}{(i-2)!}F^{(i)}(x^*)(x^*-x)^{i-2}\right)d\tau\,(x^*-x)\right\| \leq \sum_{i=2}^{k-1}\frac{1}{(i-1)!}\alpha_i\|x^*-x\|^{i-1},$$

$$\left\|F^{(k)}\left(x^*+s(\tau-1)(x^*-x)\right)\right\| \leq \omega(\|x^*\|+\varepsilon).$$

As $\varepsilon \leq R$ and R is a solution of (4), then

$$\|I-[F'(x^*)]^{-1}F'(x)\| \leq \gamma\left[\sum_{i=1}^{k-2}\frac{\alpha_{i+1}}{i!}R^{i-1} + \frac{R^{k-2}}{(k-1)!}\omega(\|x^*\|+R)\right]R = \delta < 1. \quad (6)$$

Now, by the Banach lemma on invertible operators, the operator $[F'(x)]^{-1}$ exists and $\|[F'(x)]^{-1}\| < \frac{1}{1-\delta}\|[F'(x^*)]^{-1}\| \leq \frac{\gamma}{1-\delta}$.

As $x_0 \in B(x^*,\varepsilon)$, then the operator $\Gamma_0 = [F'(x_0)]^{-1}$ exists, $\|\Gamma_0\| \leq \frac{\gamma}{1-\delta}$ and x_1 is well-defined. Moreover,

$$\begin{aligned} x_1 - x^* &= x_0 - \Gamma_0 F(x_0) - x^* \\ &= \Gamma_0\int_0^1\left(\sum_{i=2}^{k-1}\frac{1}{(i-2)!}F^{(i)}(x^*)(t-1)^{i-2}(x^*-x_0)^{i-2}\right. \\ &\quad \left. + \frac{1}{(k-3)!}\int_0^1 F^{(k)}\left(x^*+s(t-1)(x^*-x_0)\right)(1-s)^{k-3}(t-1)^{k-2}(x^*-x_0)^{k-2}\,ds\right)(1-t)(x^*-x_0)^2\,dt, \end{aligned}$$

since

$$F''\left(x_0+t(x^*-x_0)\right) = \sum_{i=2}^{k-1}\frac{(t-1)^{i-2}}{(i-2)!}F^{(i)}(x^*)(x^*-x_0)^{i-2} + \frac{1}{(k-3)!}\int_0^1 F^{(k)}\left(x^*+s(t-1)(x^*-x_0)\right)(1-s)^{k-3}(t-2)^{k-2}(x^*-x_0)^{k-2}\,ds.$$

Thus,

$$\begin{aligned} \|x^*-x_1\| &\leq \frac{\gamma}{1-\delta}\left[\sum_{i=2}^{k-1}\frac{i-1}{i!}\alpha_i\|x^*-x_0\|^{i-1} + \frac{k-1}{k!}\omega(\|x^*\|+\varepsilon)\|x^*-x_0\|^{k-1}\right]\|x^*-x_0\| \\ &\leq \frac{\gamma}{1-\delta}\left[\sum_{i=1}^{k-2}\frac{i}{(i+1)!}\alpha_{i+1}R^i + \frac{k-1}{k!}\omega(\|x^*\|+R)R^{k-1}\right]\|x^*-x_0\| \leq \frac{\delta(k-1)}{(1-\delta)k}\|x^*-x_0\|, \end{aligned}$$

since

$$\sum_{i=2}^{k-1}\frac{i-1}{i!}\alpha_i\|x^*-x_0\|^{i-1} \leq \sum_{i=1}^{k-2}\frac{i}{(i+1)!}\alpha_{i+1}R^i,$$

$$\frac{k-1}{k!}\omega(\|x^*\|+\varepsilon)\|x^*-x_0\|^{k-1} \leq \frac{k-1}{k!}\omega(\|x^*\|+R)R^{k-1},$$

$$\sum_{i=1}^{k-2}\frac{i}{(i+1)!}\alpha_{i+1}R^i \leq \frac{k-1}{k}\sum_{i=1}^{k-2}\frac{\alpha_{i+1}}{i!}R^i.$$

Following now an inductive argument, for all $n \in \mathbb{N}$, we have

$$\|x^* - x_n\| \leq \frac{\gamma}{1-\delta}\left(\sum_{i=2}^{k-1}\frac{i-1}{i!}\alpha_i R^{i-2} + \frac{k-1}{k!}\omega(\|x^*\|+R)R^{k-1}\right)\|x^* - x_{n-1}\|^2$$
$$\leq \frac{\delta(k-1)}{(1-\delta)k}\|x^* - x_{n-1}\|,$$

and then $\|x^* - x_n\| \leq \left(\frac{\delta(k-1)}{(1-\delta)k}\right)^n \|x^* - x_0\|$, so $\lim_{n\to+\infty} x_n = x^*$.

On the other hand, (5) follows from

$$\gamma\left(\sum_{i=1}^{k-2}\frac{\alpha_{i+1}}{i!}R^{i-1} + \frac{R^{k-2}}{(k-1)!}\omega(\|x^*\|+R)\right) = \frac{\delta}{R} < \frac{\delta}{\varepsilon}.$$

The proof is complete. □

Note that the higher value of δ is, the greater the root R of Eq. (4) is, since ω is a non-decreasing continuous function.

Remark 2. From (5), it follows that Newton's method has Q-order of convergence at least 2 [11]. Moreover, if $\delta\varepsilon < 1$, then

$$\|x_n - x^*\| < \frac{\delta}{\varepsilon}\|x_{n-1} - x^*\|^2 \leq \left(\frac{\delta}{\varepsilon}\right)^{1+2+\cdots+2^{n-1}}\|x_0 - x^*\|^{2^n} < \left(\sqrt{\delta\varepsilon}\right)^{2^n}\sqrt{\frac{\varepsilon}{\delta}},$$

and consequently, Newton's method has R-order of convergence at least 2 [11].

3. An example

Next, we illustrate the previous result with the following example given in [2]. We choose the max-norm.

Let $F(x, y, z) = 0$ be a nonlinear system, where $F : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ and $F(x, y, z) = (x, y^2 + y, e^z - 1)$. It is obvious that $(0, 0, 0) = \bar{x}^*$ is a solution of the system.

From F, we deduce

$$F'(\bar{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2y+1 & 0 \\ 0 & 0 & e^z \end{pmatrix} \quad \text{and} \quad F'(\bar{x}^*) = \operatorname{diag}\{1, 1, 1\},$$

where $\bar{x} = (x, y, z)$. Hence, $[F'(\bar{x}^*)]^{-1} = \operatorname{diag}\{1, 1, 1\}$ and $\gamma = 1$. Moreover,

$$F''(\bar{x}) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e^z \end{array}\right),$$

and consequently, $\|F''(\bar{x}^*)\| \leq 2 = \alpha_2$. Furthermore, as

$$F'''(\bar{x}) = \left(\begin{array}{c|c|c|c|c|c|c|c|ccc} & & & & & & & & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & & 0 & 0 & e^z \end{array}\right),$$

then $\|F'''(\bar{x})\| \leq e^{\|\bar{x}\|}$, so $\omega(t) = e^t$ if $k = 3$ in Theorem 1. In this case, if we choose for example $\delta = \frac{3}{5} - 10^{-3}$, Eq. (4) is reduced to $\left(2 + \frac{t}{2}e^t\right)t - \left(\frac{3}{5} - 10^{-3}\right) = 0$, whose unique solution is $R = 0.274676\ldots$. Therefore, Newton's method is convergent from any starting point $\bar{x} \in B(\bar{x}^*, 0.274676\ldots)$.

From the above, it follows easily that $\|F^{(j)}(\bar{x}^*)\| = 1$ and $\|F^{(j)}(\bar{x})\| \leq e^{\|\bar{x}\|}$, for all $j \geq 3$, so $\alpha_j = 1$ and $\omega(t) = e^t$ for all $j \geq 3$. From the last data, we construct Table 1, where we show the different values of the radii R of the domains of starting points $B(\bar{x}^*, R)$ that are obtained for different values of k.

If we compare the results that appear in Table 1 with those given by Dennis and Schnabel in [2], we can emphasize three things. The first and most important is that our result is independent of the value r, since we can choose $\Omega = \mathbb{R}^3$, while Dennis and Schnabel cannot. The second is that if $r < \ln 2$, Dennis and Schnabel obtain $R = \frac{1}{4}$, so we improve the domain of starting points that Dennis and Schnabel obtain if $k = 3, 4, 5, 6, 7$. The third is that if $r \geq \ln 2$, Dennis and Schnabel obtain $R = \frac{1}{2\exp(r)}$, so we could improve the domain of starting points that Dennis and Schnabel obtain based on the values of k that can be taken, although they have to be calculated previously. Observe that the radius $R = \frac{1}{2\exp(r)}$ that Dennis and Schnabel obtain decreases exponentially depending on r, which does not occur with the values of R in our case, since they do not decrease exponentially, as we can see in Table 1. So, if for example we choose $r = \ln 5 = 1.609437\ldots$, Dennis and Schnabel obtain $R = 0.1$, and if we choose for example $k = 10000$, we obtain $R = 0.235051\ldots$, so we can conjecture that there will always be a sufficiently large k that allows us to improve the ball of convergence that Dennis and Schnabel obtain.

Table 1
Radii of the domains of convergence.

k	Eq. (4)	R
3	$t\left(\frac{t}{2}e^t+2\right)-\left(\frac{3}{5}-10^{-3}\right)=0$	0.274676 ...
4	$t\left(\frac{t^2}{6}e^t+\frac{t}{2}+2\right)-\left(\frac{4}{7}-10^{-3}\right)=0$	0.265549 ...
5	$t\left(\frac{t^3}{24}e^t+\frac{t^2}{6}+\frac{t}{2}+2\right)-\left(\frac{5}{9}-10^{-3}\right)=0$	0.258946 ...
6	$t\left(\frac{t^4}{120}e^t+\frac{t^3}{24}+\frac{t^2}{6}+\frac{t}{2}+2\right)-\left(\frac{6}{11}-10^{-3}\right)=0$	0.254559 ...
7	$t\left(\frac{t^5}{720}e^t+\frac{t^4}{120}+\frac{t^3}{24}+\frac{t^2}{6}+\frac{t}{2}+2\right)-\left(\frac{7}{13}-10^{-3}\right)=0$	0.251504 ...
8	$t\left(\frac{t^6}{5040}e^t+\frac{t^5}{720}+\frac{t^4}{120}+\frac{t^3}{24}+\frac{t^2}{6}+\frac{t}{2}+2\right)-\left(\frac{8}{15}-10^{-3}\right)=0$	0.249259 ...

Results corresponding to that of Dennis and Schnabel with the bigger radii $r < \frac{\sqrt{2}-1}{\sqrt{2}\gamma M}$ and $r < \frac{2}{3\gamma M}$ were proved under the same conditions by, respectively, Rall in [3] and Rheinboldt in [4]. But the same comments as were made concerning the result of Dennis and Schnabel remain valid for the results of Rall and Rheinboldt, since the radius of the ball of convergence depends on a bound for F'' or on the Lipschitz constant for F', which in turn depend on the value of r. In consequence, the same conclusions are obtained. On the other hand, Argyros obtains an interesting result in [5], by requiring a Lipschitz condition on F'', for which the same comments as were made above still apply.

Notice that the main advantage of our result is that it depends not on any bound calculated on the domain $B(x^*, r)$, but simply on the values of the successive derivatives of F at the solution x^*.

Acknowledgement

This work was supported in part by the project MTM2011-28636-C02-01 of the Spanish Ministry of Science and Innovation.

References

[1] L.V. Kantorovich, The majorant principle and Newton's method, Dokl. Akad. Nauk SSSR. 76 (1951) 17–20 (in Russian).
[2] J.E. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, Philadelphia, 1996.
[3] L.B. Rall, A note on the convergence of Newton's method, SIAM J. Numer. Anal. 11 (1974) 34–36.
[4] W.C. Rheinboldt, An adaptive continuation process for solving systems of nonlinear equations, Math. Models Numer. Methods, Banach Cent. Publ. 3 (1978) 129–142.
[5] I.K. Argyros, Comparing the radii of some balls appearing in connection to three local convergence theorems for Newton's method, Southwest J. Pure Appl. Math. 2 (1998) 24–28.
[6] I.K. Argyros, On the Newton–Kantorovich hypothesis for solving equations, J. Comput. Appl. Math. 169 (2) (2004) 315–332.
[7] F. Cianciaruso, E. de Pascale, Newton–Kantorovich approximations when the derivative is Hölderian: old and new results, Numer. Funct. Anal. Optim. 24 (7) (2003) 713–723.
[8] F. Cianciaruso, E. de Pascale, Estimates of majorizing sequences in the Newton–Kantorovich method: a further improvement, J. Math. Anal. Appl. 322 (2006) 329–335.
[9] J.A. Ezquerro, M.A. Hernández, Generalized differentiability conditions for Newton's method, IMA J. Numer. Anal. 22 (2) (2002) 187–205.
[10] J.A. Ezquerro, D. González, M.A. Hernández, Majorizing sequences for Newton's method from initial value problems, J. Comput. Appl. Math. 236 (2012) 2246–2258.
[11] F.A. Potra, V. Pták, Nondiscrete Induction and Iterative Processes, Pitman Publishing Limited, London, 1984.