Journal of Computational and Applied Mathematics 235 (2011) 1515–1522

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Local convergence of Newton's method under majorant condition

O.P. Ferreira *

IME/UFG, CP-131, CEP 74001-970-Goiânia, GO, Brazil

ARTICLE INFO

Article history:
Received 13 May 2009
Received in revised form 26 August 2010

Keywords:
Newton's method
Majorant condition
Local convergence
Banach space

ABSTRACT

A local convergence analysis of Newton's method for solving nonlinear equations, under a majorant condition, is presented in this paper. Without assuming convexity of the derivative of the majorant function, which relaxes the Lipschitz condition on the operator under consideration, convergence, the biggest range for uniqueness of the solution, the optimal convergence radius and results on the convergence rate are established. Besides, two special cases of the general theory are presented as applications.

1. Introduction

Newton's method and its variant are powerful tools for solving nonlinear equations in real or complex Banach space. In the past few years, a couple of papers have dealt with the issue of local and semi-local convergence analysis of Newton's method and its variants by relaxing the assumption of Lipschitz continuity of the derivative of the function, which define the nonlinear equation under consideration; see [1–10].

In [4,9], under a majorant condition and generalized Lipschitz condition, respectively, local convergence, quadratic rate and estimate of the best possible convergence radius of Newton's method as well as uniqueness of the solution for the nonlinear equation in question were established. In the analysis presented in [4], *convexity* of the derivative of the scalar majorant function was assumed and in [9] the *nondecrement* of the positive integrable function which defines the generalized Lipschitz condition was assumed. These assumptions seem to be actually natural in the local analysis of Newton's method. The convergence, uniqueness, superlinear rate and estimate of the best possible convergence radius will be established in this paper without assuming the convexity of the derivative of the majorant function or that the function which defines the generalized Lipschitz condition is nondecreasing. In particular, this analysis shows that the convexity of the derivative of the majorant function or that the function which defines the generalized Lipschitz condition is nondecreasing is needed only to obtain quadratic convergence rate of the sequence generated by Newton's Method. Also, as in [4], the analysis presented provides a clear relationship between the majorant function with the nonlinear operator under consideration. Besides improving the convergence theory this analysis permits us to obtain two new important special cases, namely, [7,10] (see also, [4,9]) as applications. It is worth pointing out that the majorant condition used here is equivalent to Wang's condition (see [10]) that the derivative of the majorant function is always convex.

The organization of the paper is as follows. In Section 1.1, some notations and one basic result used in the paper are presented. In Section 2, the main result is stated and in Section 2.1 some properties of the majorant function are established and the main relationships between the majorant function and the nonlinear operator used in the paper are presented. In Section 2.2, the uniqueness of the solution and the optimal convergence radius are obtained. In Section 2.3 the main result is proved and two applications of this result are given in Section 3. Some final remarks are given in Section 4.

* Tel.: +55 62 3521 1424; fax: +55 62 3521 1180.
E-mail address: orizon@mat.ufg.br.

0377-0427/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2010.08.038

1.1. Notation and auxiliary results

The following notations and results are used throughout our presentation. Let X, Y be Banach spaces. The open and closed balls at x are denoted, respectively, by

$$B(x,\delta) = \{y \in X;\ \|x-y\| < \delta\} \quad \text{and} \quad B[x,\delta] = \{y \in X;\ \|x-y\| \leqslant \delta\}.$$

Let $\Omega \subseteq X$ be an open set. The Fréchet derivative of $F : \Omega \to Y$ is the linear map $F'(x) : X \to Y$.

Lemma 1 (*Banach's Lemma*). *Let $B : X \to X$ be a bounded linear operator. If $I : X \to X$ is the identity operator and $\|B - I\| < 1$, then B is invertible and $\|B^{-1}\| \le 1/(1 - \|B - I\|)$.*

2. Local analysis for Newton's method

Our goal is to state and prove a local theorem for Newton's method, which generalizes Theorem 2.1 of [4]. First, we will prove some results regarding the scalar majorant function, which relaxes the Lipschitz condition. Then we will establish the main relationships between the majorant function and the nonlinear function. We will also prove the uniqueness of the solution in a suitable region and the optimal ball of convergence. Finally, we will show the well definedness of Newton's method and convergence, also results on the convergence rates will be given. The statement of the theorem is:

Theorem 2. *Let X, Y be Banach spaces, $\Omega \subseteq X$ be an open set and $F : \Omega \to Y$ be a continuously differentiable function. Let $x_* \in \Omega$, $R > 0$ and $\kappa := \sup\{t \in [0,R) : B(x_*,t) \subset \Omega\}$. Suppose that $F(x_*) = 0$, $F'(x_*)$ is invertible and there exists an $f : [0, R) \to \mathbb{R}$ continuously differentiable such that*

$$\left\|F'(x_*)^{-1}\left[F'(x) - F'(x_* + \tau(x - x_*))\right]\right\| \le f'(\|x - x_*\|) - f'(\tau\|x - x_*\|), \tag{1}$$

for all $\tau \in [0,1]$, $x \in B(x_*,\kappa)$ and

(h1) $f(0) = 0$ *and* $f'(0) = -1$;
(h2) f' *is strictly increasing.*

Let $\nu := \sup\{t \in [0,R) : f'(t) < 0\}$, $\rho := \sup\{\delta \in (0,\nu) : [f(t)/f'(t) - t]/t < 1,\ t \in (0,\delta)\}$ and

$$r := \min\{\kappa, \rho\}.$$

Then the sequences with starting points $x_0 \in B(x_*,r)/\{x_*\}$ and $t_0 = \|x_0 - x_*\|$, respectively, namely

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad t_{k+1} = |t_k - f(t_k)/f'(t_k)|, \quad k = 0, 1, \ldots, \tag{2}$$

are well defined; $\{t_k\}$ is strictly decreasing, is contained in $(0,r)$ and converges to 0 and $\{x_k\}$ is contained in $B(x_*,r)$ and converges to the point x_* which is the unique zero of F in $B(x_*,\sigma)$, where $\sigma := \sup\{t \in (0,\kappa) : f(t) < 0\}$ and there hold:

$$\lim_{k\to\infty} [\|x_{k+1} - x_*\|/\|x_k - x_*\|] = 0, \qquad \lim_{k\to\infty} [t_{k+1}/t_k] = 0. \tag{3}$$

Moreover, if $f(\rho)/(\rho f'(\rho) - 1) = 1$ and $\rho < \kappa$ then $r = \rho$ is the best possible convergence radius.
If, additionally, given $0 \le p \le 1$

(h3) *the function $(0,\nu) \ni t \mapsto [f(t)/f'(t) - t]/t^{p+1}$ is strictly increasing,*

then the sequence $\{t_{k+1}/t_k^{p+1}\}$ is strictly decreasing and there holds

$$\|x_{k+1} - x_*\| \le \left[t_{k+1}/t_k^{p+1}\right]\|x_k - x_*\|^{p+1}, \quad k = 0, 1, \ldots. \tag{4}$$

Remark 1. The first equation in (3) means that $\{x_k\}$ converges superlinearly to x_*. Moreover, because the sequence $\{t_{k+1}/t_k^{p+1}\}$ is strictly decreasing $t_{k+1}/t_k^{p+1} \le t_1/t_0^{p+1}$, for $k = 0, 1, \ldots$. So, the inequality in (4) implies $\|x_{k+1} - x_*\| \le [t_1/t_0^{p+1}]\|x_k - x_*\|^{p+1}$, for $k = 0, 1, \ldots$. As a consequence, if $p = 0$ then $\|x_k - x_*\| \le t_0[t_1/t_0]^k$ for $k = 0, 1, \ldots$ and if $0 < p \le 1$ then

$$\|x_k - x_*\| \le t_0\,(t_1/t_0)^{[(p+1)^k - 1]/p}, \quad k = 0, 1, \ldots.$$

Example 1. The following continuously differentiable functions satisfy (h1), (h2) and (h3):

(i) $f : [0, +\infty) \to \mathbb{R}$ such that $f(t) = t^{1+p} - t$;
(ii) $f : [0, +\infty) \to \mathbb{R}$ such that $f(t) = e^{-t} + t^2 - 1$.

Letting $0 < p < 1$, the derivative of first function as well as that of the second is not convex.

Similarly to the proof of Proposition 2.6 in [4], always f has derivative f' convex, we can prove that (h3) holds with $p = 1$. In this case, Newton's sequence converges with quadratic rate. Indeed, the next example shows that the convexity of f' was necessary in [4] to obtain quadratic convergence rate.

Example 2. Let $g : \mathbb{R} \to \mathbb{R}$ be given by $g(t) = t^{5/3} - t$. Note that $g(0) = 0, g'(0) = -1$ and letting $p = 2/3$ in Example 1 the function f is a majorant function to g. Newton's method applied to g with starting point t_0 "near" 0 generates the following sequence:

$$t_{k+1} = \left(2\, t_k^{5/3}\right) \Big/ \left(5\, t_k^{2/3} - 3\right), \quad k = 0, 1, \ldots .$$

Theorem 2 implies that the sequence $\{t_k\}$ converges to 0 with superlinear rate. It easy to see that $\{t_k\}$ does not converge to 0 with quadratic rate. So, in particular, it follows from [4] that there is no majorant function having convex derivative for the function g.

From now on, we assume that the hypotheses of Theorem 2 hold, with the exception of (h3) which will be considered to hold only when explicitly stated.

2.1. Preliminary results

In this section, we will prove all the statements in Theorem 2 regarding the sequence $\{t_k\}$ associated to the majorant function. The main relationships between the majorant function and the nonlinear operator as well as the results in Theorem 2 related to the uniqueness of the solution and the optimal convergence radius will be also established.

2.1.1. The scalar sequence

In this section, we will prove the statements in Theorem 2 involving $\{t_k\}$. First, we will prove that the constants κ, ν, ρ and σ are positive. We begin by proving that κ, ν and σ are positive.

Proposition 3. *The constants κ, ν and σ are positive and $t - f(t)/f'(t) < 0$, for all $t \in (0, \nu)$.*

Proof. Since Ω is open and $x_* \in \Omega$, we can immediately conclude that $\kappa > 0$. As f' is continuous in 0 with $f'(0) = -1$, there exists $\delta > 0$ such that $f'(t) < 0$ for all $t \in (0, \delta)$. So, $\nu > 0$. Now, because $f(0) = 0$ and $f'(0) = -1$, there exists $\delta > 0$ such that $f(t) < 0$ for all $t \in (0, \delta)$. Hence $\sigma > 0$.

It remains to show that $t - f(t)/f'(t) < 0$, for all $t \in (0, \nu)$. Since f' is strictly increasing, f is strictly convex. So, $0 = f(0) > f(t) - tf'(t)$, for $t \in (0, R)$. If $t \in (0, \nu)$ then $f'(t) < 0$, which, combined with the last inequality yields the desired inequality. □

According to (h2) and definition of ν, we have $f'(t) < 0$ for all $t \in [0, \nu)$. Therefore, Newton's iteration map for f is well defined in $[0, \nu)$. Let us call it n_f:

$$\begin{array}{llll} n_f : [0, \nu) & \to & (-\infty, 0] \\ t & \mapsto & t - f(t)/f'(t). \end{array} \tag{5}$$

Proposition 4. $\lim_{t\to 0} |n_f(t)|/t = 0$. *As a consequence, $\rho > 0$ and $|n_f(t)| < t$ for all $t \in (0, \rho)$.*

Proof. Using definition (5), Proposition 3, $f(0) = 0$, and definition of ν, a simple algebraic manipulation gives

$$\frac{|n_f(t)|}{t} = [f(t)/f'(t) - t]/t = \frac{1}{f'(t)} \frac{f(t) - f(0)}{t - 0} - 1, \quad t \in (0, \nu). \tag{6}$$

Because $f'(0) = -1 \neq 0$ the first statement follows by taking the limit in (6), as t goes to 0.

Since $\lim_{t\to 0} |n_f(t)|/t = 0$, the first equality in (6) implies that there exists $\delta > 0$ such that

$$0 < [f(t)/f'(t) - t]/t < 1, \quad t \in (0, \delta).$$

So, we conclude that ρ is positive. Therefore, the first equality in (6) together with the definition of ρ implies that $|n_f(t)|/t = [f(t)/f'(t) - t]/t < 1$, for all $t \in (0, \rho)$, as required. □

Using (5), it easy to see that the sequence $\{t_k\}$ is equivalently defined as

$$t_0 = \|x_0 - x_*\|, \qquad t_{k+1} = |n_f(t_k)|, \quad k = 0, 1, \ldots . \tag{7}$$

Corollary 5. *The sequence $\{t_k\}$ is well defined, is strictly decreasing and is contained in $(0, \rho)$. Moreover, $\{t_k\}$ converges to 0 with superlinear rate, i.e., $\lim_{k\to\infty} t_{k+1}/t_k = 0$. If, additionally, (h3) holds then the sequence $\{t_{k+1}/t_k^{p+1}\}$ is strictly decreasing.*

Proof. Since $0 < t_0 = \|x_0 - x_*\| < r \le \rho$, using Proposition 4 and (7) it is simple to conclude that $\{t_k\}$ is well defined, is strictly decreasing and is contained in $(0, \rho)$. So, we have proved the first statement of the corollary.

Because $\{t_k\} \subset (0, \rho)$ is strictly decreasing it converges. So, $\lim_{k\to\infty} t_k = t_*$ with $0 \le t_* < \rho$ which together with (7) implies $0 \le t_* = |n_f(t_*)|$. But, if $t_* \ne 0$ then Proposition 4 implies $|n_f(t_*)| < t_*$, hence $t_* = 0$. Now, $\lim_{k\to\infty} t_k = 0$. Thus, the definition of $\{t_k\}$ in (7) and the first statement in Proposition 4 imply that $\lim_{k\to\infty} t_{k+1}/t_k = \lim_{k\to\infty} |n_f(t_k)|/t_k = 0$ and the second statement is proved.

Since $\{t_k\}$ is strictly decreasing, the last statement is an immediate consequence of (h3). □

2.1.2. Relationship between the majorant function and the nonlinear operator

In this section, we will present the main relationships between the majorant function f and the nonlinear operator F.

Lemma 6. *If $\|x - x_*\| < \min\{\kappa, \nu\}$, then $F'(x)$ is invertible and*

$$\|F'(x)^{-1}F'(x_*)\| \leqslant 1/|f'(\|x - x_*\|)|.$$

In particular, F' is invertible in $B(x_*, r)$.

Proof. Neither assumption (h3) nor the fact that the derivative of the majorant function is convex is necessary for proving this lemma. The proof follows the same pattern of Lemma 2.9 of [4]. □

Newton's iteration at a point happens to be a zero of the linearization of F at such a point. So, we study the linearization error at a point in Ω

$$E_F(x, y) := F(y) - \left[F(x) + F'(x)(y - x)\right], \quad y, x \in \Omega. \tag{8}$$

We will bound this error by the error in the linearization on the majorant function f

$$e_f(t, u) := f(u) - \left[f(t) + f'(t)(u - t)\right], \quad t, u \in [0, R). \tag{9}$$

Lemma 7. *If $\|x_* - x\| < \kappa$, then $\|F'(x_*)^{-1}E_F(x, x_*)\| \le e_f(\|x - x_*\|, 0)$.*

Proof. Neither assumption (h3) nor the fact that the derivative of the majorant function is convex is necessary for proving this lemma. The proof follows the same pattern of Lemma 2.10 of [4]. □

Lemma 6 guarantees, in particular, that F' is invertible in $B(x_*, r)$ and consequently, Newton's iteration map is well defined. Let us call N_F, Newton's iteration map for F in that region:

$$\begin{aligned} N_F : B(x_*, r) &\to Y \\ x &\mapsto x - F'(x)^{-1}F(x). \end{aligned} \tag{10}$$

Now, we establish an important relationship between Newton's iteration maps n_f and N_F. As a consequence, we obtain that $B(x_*, r)$ is invariant under N_F. This result will be important to assert the definition of Newton's method.

Lemma 8. *If $\|x - x_*\| < r$ then $\|N_F(x) - x_*\| \le |n_f(\|x - x_*\|)|$. As a consequence,*

$$N_F(B(x_*, r)) \subset B(x_*, r).$$

Proof. Since $F(x_*) = 0$, the inequality is trivial for $x = x_*$. Now assume that $0 < \|x - x_*\| \le t$. Lemma 6 implies that $F'(x)$ is invertible. Thus, because $F(x_*) = 0$, direct manipulation yields

$$x_* - N_F(x) = -F'(x)^{-1}\left[F(x_*) - F(x) - F'(x)(x_* - x)\right] = -F'(x)^{-1}E_F(x, x_*).$$

Using the above equation, Lemmas 6 and 7, it is easy to conclude that

$$\|x_* - N_F(x)\| \le \| - F'(x)^{-1}F'(x_*)\|\|F'(x_*)^{-1}E_F(x, x_*)\| \le e_f(\|x - x_*\|, 0)/|f'(\|x - x_*\|)|.$$

On the other hand, taking into account that $f(0) = 0$, the definitions of e_f and n_f imply that

$$e_f(\|x - x_*\|, 0)/|f'(\|x - x_*\|)| = |n_f(\|x - x_*\|)|.$$

So, the first statement follows by combining the above two expressions.

Take $x \in B(x_*, r)$. Since $\|x - x_*\| < r$ and $r \le \rho$, the first part together with the second part of Proposition 4 implies that $\|N_F(x) - x_*\| \le |n_f(\|x - x_*\|)| < \|x - x_*\|$ and the last result follows. □

Lemma 9. *If (h3) holds and $\|x - x_*\| \le t < r$ then $\|N_F(x) - x_*\| \le [|n_f(t)|/t^{p+1}]\|x - x_*\|^{p+1}$.*

Proof. The inequality is trivial for $x = x_*$. If $0 < \|x - x_*\| \le t$ then assumption (h3) and (5) give $|n_f(\|x - x_*\|)|/\|x - x_*\|^{p+1} \le |n_f(t)|/t^{p+1}$. So, using Lemma 8 the statement follows. □

2.2. Uniqueness and optimal convergence radius

In this section we will obtain the uniqueness of the solution and the optimal convergence radius.

Lemma 10. *The point x_* is the unique zero of F in $B(x_*, \sigma)$.*

Proof. Neither assumption (h3) nor the fact that the derivative of the majorant function is convex is necessary for proving this lemma. The proof follows the same pattern of Lemma 2.13 of [4]. □

Lemma 11. *If $f(\rho)/(\rho f'(\rho)) - 1 = 1$ and $\rho < \kappa$, then $r = \rho$ is the optimal convergence radius.*

Proof. The proof follows the same pattern of Lemma 2.15 of [4]. □

2.3. Newton's sequence

In this section, we will prove the statements in Theorem 2 involving Newton's sequence $\{x_k\}$. First, note that the first equation in (2) together with (10) implies that the sequence $\{x_k\}$ satisfies

$$x_{k+1} = N_F(x_k), \quad k = 0, 1, \ldots, \tag{11}$$

which is indeed an equivalent definition of this sequence.

Proposition 12. *The sequence $\{x_k\}$ is well defined, is contained in $B(x_*, r)$ and converges to the point x_* the unique zero of F in $B(x_*, \sigma)$ and there hold:*

$$\lim_{k\to\infty} [\|x_{k+1} - x_*\|/\|x_k - x_*\|] = 0. \tag{12}$$

If, additionally, (h3) holds then the sequences $\{x_k\}$ and $\{t_k\}$ satisfy

$$\|x_{k+1} - x_*\| \le \left[t_{k+1}/t_k^{p+1}\right]\|x_k - x_*\|^{p+1}, \quad k = 0, 1, \ldots. \tag{13}$$

Proof. As $x_0 \in B(x_*, r)$ and $r \le \nu$, combining (11), inclusion $N_F(B(x_*, r)) \subset B(x_*, r)$ in Lemmas 8 and 6, it is easy to conclude that $\{x_k\}$ is well defined and remains in $B(x_*, r)$.

We are going to prove that $\{x_k\}$ converges towards x_*. Since $\|x_k - x_*\| < r \le \rho$, for $k = 0, 1, \ldots,$ we obtain from (11), Lemma 8 and Proposition 4 that

$$\|x_{k+1} - x_*\| = \|N_F(x_k) - x_*\| \le |n_f(\|x_k - x_*\|)| < \|x_k - x_*\|, \quad k = 0, 1, \ldots. \tag{14}$$

So, $\{\|x_k - x_*\|\}$ is strictly decreasing and convergent. Let $\ell_* = \lim_{k\to\infty} \|x_k - x_*\|$. Because $\{\|x_k - x_*\|\}$ rests in $(0, \rho)$ and is strictly decreasing we have $0 \le \ell_* < \rho$. Thus, the continuity of n_f in $[0, \rho)$ and (14) implies $0 \le \ell_* = |n_f(\ell_*)|$ and from Proposition 4 we have $\ell_* = 0$. Therefore, the convergence of $\{x_k\}$ to x_* is proved. The uniqueness was proved in Lemma 10.

For proving the equality in (12) note that equation (14) implies

$$[\|x_{k+1} - x_*\|/\|x_k - x_*\|] \le \left[|n_f(\|x_k - x_*\|)|/\|x_k - x_*\|\right], \quad k = 0, 1, \ldots.$$

Since $\lim_{k\to\infty} \|x_k - x_*\| = 0$ the desired equality follows from the first statement in Proposition 4.

Now we will show (13). First, we will prove by induction that the sequences $\{t_k\}$ and $\{x_k\}$ defined, respectively, in (11) and (7) satisfy

$$\|x_k - x_*\| \le t_k, \quad k = 0, 1, \ldots. \tag{15}$$

Because $t_0 = \|x_0 - x_*\|$, the above inequality holds for $k = 0$. Now, assume that $\|x_k - x_*\| \le t_k$. Using (11), Lemma 9, the induction assumption and (7) we obtain that

$$\|x_{k+1} - x_*\| = \|N_F(x_k) - x_*\| \le \frac{|n_f(t_k)|}{t_k^{p+1}} \|x_k - x_*\|^{p+1} \le |n_f(t_k)| = t_{k+1},$$

and the proof by induction is complete. Therefore, it easy to see that the desired inequality follows by combining (11), (15), Lemma 9 and (7). □

The proof of Theorem 2 follows from Corollary 5, Lemmas 10 and 11 and Proposition 12.

3. Special cases

In this section, we will present two special cases of Theorem 2.

3.1. Convergence result under Hölder-like condition

In this section we will present the convergence theorem for Newton's method under an affine invariant Hölder-like condition which has appeared in [7,10].

Theorem 13. *Let X, Y be Banach spaces, $\Omega \subseteq X$ be an open set and $F : \Omega \to Y$ be a continuously differentiable function. Let $x_* \in \Omega$ and $\kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}$. Suppose that $F(x_*) = 0$, $F'(x_*)$ is invertible and there exists a constant $K > 0$ and $0 < p \le 1$ such that*

$$\left\| F'(x_*)^{-1} \left[F'(x) - F'(x_* + \tau(x - x_*)) \right] \right\| \le K(1 - \tau^p)\|x - x_*\|^p, \quad x \in B(x_*, \kappa), \tau \in [0, 1]. \tag{16}$$

Let $r = \min\{\kappa, [(p + 1)/((2p + 1)K)]^{1/p}\}$. Then, the sequences with starting points $x_0 \in B(x_*, r)/\{x_*\}$ and $t_0 = \|x_0 - x_*\|$, respectively,

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad t_{k+1} = \frac{K\, p\, t_k^{p+1}}{(p + 1)[1 - K\, t_k^p]}, \quad k = 0, 1, \ldots,$$

are well defined, $\{t_k\}$ is strictly decreasing, is contained in $(0, r)$ and converges to 0 and $\{x_k\}$ is contained in $B(x_*, r)$, converges to x_* which is the unique zero of F in $B(x_*, [(p + 1)/K]^{1/p})$ and there holds

$$\|x_{k+1} - x_*\| \le \frac{K\, p}{(p + 1)[1 - K\, t_k^p]}\, \|x_k - x_*\|^{p+1}, \quad k = 0, 1, \ldots.$$

Moreover, if $[(p + 1)/((2p + 1)K)]^{1/p} < \kappa$ then $r = [(p + 1)/((2p + 1)K)]^{1/p}$ is the best possible convergence radius.

Proof. It is immediate to prove that F, x_* and $f : [0, \kappa) \to \mathbb{R}$, defined by $f(t) = Kt^{p+1}/(p + 1) - t$, satisfy the inequality (1) and the conditions (h1), (h2) and (h3) in Theorem 2. In this case, it is easy to see that ρ and ν, as defined in Theorem 2, satisfy

$$\rho = [(p + 1)/((2p + 1)K)]^{1/p} \le \nu = [1/K]^{1/p},$$

and, as a consequence, $r = \min\{\kappa, [(p+1)/((2p+1)K)]^{1/p}\}$. Moreover, $f(\rho)/(\rho f'(\rho)) - 1 = 1$, $f(0) = f([(p+1)/K]^{1/p}) = 0$ and $f(t) < 0$ for all $t \in (0, [(p + 1)/K]^{1/p})$. Therefore, the result follows by invoking Theorem 2. □

Remark 2. Since Theorem 13 is a special case of Theorem 2 it follows from Remark 1 that

$$\|x_k - x_*\| \le \left[\frac{K\, p\, \|x_0 - x_*\|^p}{(p + 1)[1 - K\, \|x_0 - x_*\|^p]} \right]^{[(p+1)^k - 1]/p} \|x_0 - x_*\|, \quad k = 0, 1, \ldots.$$

Remark 3. If $F : \Omega \to Y$ satisfies the Lipschitz condition $\|F'(x) - F'(y)\| \le L\|x - y\|$, for all $x, y \in \Omega$, where $L > 0$, then it also satisfies condition (16) with $p = 1$ and $K = L\|F'(x_*)^{-1}\|$. In this case, the best possible convergence radius for Newton's method is $r = 2/(3L\|F'(x_*)^{-1}\|)$; see [11,12]. We point out that the convergence radius of affine invariant theorems are insensitive to invertible linear transformation of F. On the other hand, theorems with the Lipschitz condition are sensitive; see [13]. For more details about affine invariant theorems on Newton's method see [14] (see also [15]).

3.2. Convergence result under generalized Lipschitz condition

In this section, we will present a local convergence theorem on Newton's method under a generalized Lipschitz condition due to X. Wang, which appeared in [10] (see also [9]). It is worth point out that the result in this section does not assume that the function which defines the generalized Lipschitz condition is nondecreasing.

Theorem 14. *Let X, Y be Banach spaces, $\Omega \subseteq X$ be an open set and $F : \Omega \to Y$ be a continuously differentiable function. Let $x_* \in \Omega$ and $\kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}$. Suppose that $F(x_*) = 0$, $F'(x_*)$ is invertible and there exists a positive integrable function $L : [0, R) \to \mathbb{R}$ such that*

$$\left\| F'(x_*)^{-1} \left[F'(x) - F'(x_* + \tau(x - x_*)) \right] \right\| \le \int_{\tau\|x-x_*\|}^{\|x-x_*\|} L(u)\mathrm{d}u, \tag{17}$$

for all $\tau \in [0, 1]$, $x \in B(x_*, \kappa)$. Let $\bar{\nu} > 0$ be the constant defined by

$$\bar{\nu} := \sup\left\{ t \in [0, R) : \int_0^t L(u)\mathrm{d}u - 1 < 0 \right\},$$

and let $\bar{\rho} > 0$ *and* $\bar{r} > 0$ *be the constants defined by*

$$\bar{\rho} := \sup \left\{ t \in (0, \delta) : \int_0^t L(u)u\mathrm{d}u \Big/ \left[t \left(1 - \int_0^t L(u)\mathrm{d}u \right) \right] < 1, t \in (0, \delta) \right\}, \quad \bar{r} = \min \{\kappa, \bar{\rho}\}.$$

Then, the sequences with starting points $x_0 \in B(x_*, \bar{r})/\{x_*\}$ *and* $t_0 = \|x_0 - x_*\|$*, respectively,*

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad t_{k+1} = \int_0^{t_k} L(u)u\mathrm{d}u \Big/ \left(1 - \int_0^{t_k} L(u)\mathrm{d}u \right), \quad k = 0, 1, \ldots,$$

are well defined, $\{t_k\}$ *is strictly decreasing, is contained in* $(0, \bar{r})$ *and converges to* 0*,* $\{x_k\}$ *is contained in* $B(x_*, \bar{r})$*, converges to* x_* *which is the unique zero of* F *in* $B(x_*, \bar{\sigma})$*, where*

$$\bar{\sigma} := \sup \left\{ t \in (0, \kappa) : \int_0^t L(u)(t-u)\mathrm{d}u - t < 0 \right\}.$$

and there hold: $\lim_{k\to\infty} t_{k+1}/t_k = 0$ *and* $\lim_{k\to\infty}[\|x_{k+1} - x_*\|/\|x_k - x_*\|] = 0$*. Moreover, if*

$$\int_0^{\bar{\rho}} L(u)u\mathrm{d}u \Big/ \left[\bar{\rho} \left(1 - \int_0^{\bar{\rho}} L(u)\mathrm{d}u \right) \right] = 1,$$

and $\bar{\rho} < \kappa$ *then* $\bar{r} = \bar{\rho}$ *is the best possible convergence radius.*
If, additionally, given $0 \le p \le 1$

(h) *the function* $(0, \nu) \ni t \mapsto t^{1-p}L(t)$ *is nondecreasing,*

then the sequence $\{t_{k+1}/t_k^{p+1}\}$ *is strictly decreasing and there holds*

$$\|x_{k+1} - x_*\| \le \left[t_{k+1}/t_k^{p+1} \right] \|x_k - x_*\|^{p+1}, \quad k = 0, 1, \ldots. \tag{18}$$

Proof. Let $\bar{f} : [0, \kappa) \to \mathbb{R}$ be a differentiable function defined by

$$\bar{f}(t) = \int_0^t L(u)(t-u)\mathrm{d}u - t. \tag{19}$$

Note that the derivative of the function f is given by

$$\bar{f}'(t) = \int_0^t L(u)\mathrm{d}u - 1.$$

Because L is integrable $\bar{f}'$ is continuous (in fact $\bar{f}'$ is absolutely continuous). So, it is easy to see that (17) becomes (1) with $f' = \bar{f}'$. Moreover, because L is positive the function $f = \bar{f}$ satisfies conditions (h1) and (h2) in Theorem 2. Direct algebraic manipulation yields

$$\frac{1}{t^{p+1}} \left[\frac{\bar{f}(t)}{\bar{f}'(t)} - t \right] = \left[\frac{1}{t^{p+1}} \int_0^t L(u)u\mathrm{d}u \right] \frac{1}{|\bar{f}'(t)|}.$$

If assumption (h) holds then Lemma 2.2 of [10] implies that the first term on the right hand side of the above equation is nondecreasing in $(0, \nu)$. Now, since $1/|\bar{f}'|$ is strictly increasing in $(0, \nu)$ the above equality implies that (h3) in Theorem 2, with $f = \bar{f}$, also holds. Therefore, the result follows from Theorem 2 with $f = \bar{f}$, $\nu = \bar{\nu}$, $\rho = \bar{\rho}$, $r = \bar{r}$ and $\sigma = \bar{\sigma}$. □

Remark 4. Since Theorem 14 is a special case of Theorem 2 it follows from Remark 1 that if $p = 0$ then $\|x_k - x_*\| \le q^k\|x_0 - x_*\|$, for $k = 0, 1, \ldots$ and if $0 < p \le 1$ then

$$\|x_k - x_*\| \le q^{[(p+1)^k - 1]/p}\|x_0 - x_*\|, \quad k = 0, 1, \ldots,$$

where

$$q = \int_0^{\|x_0 - x_*\|} L(u)u\mathrm{d}u \Big/ \left[\|x_0 - x_*\| \left(1 - \int_0^{\|x_0 - x_*\|} L(u)\mathrm{d}u \right) \right].$$

Remark 5. It was shown in [9] that if L is positive and nondecreasing then the sequence generated by Newton's method converges with quadratic rate. From Theorem 14 we conclude that the assumption on the nondecrement of L is needed only to obtain the quadratic convergence rate of Newton's sequence. This result was also obtained in [10].

Finally, we observe that if the positive integrable function $L : [0, R) \to \mathbb{R}$ is nondecreasing then the strictly increasing function $f' : [0, R) \to \mathbb{R}$, defined by

$$f'(t) = \int_0^t L(u)\mathrm{d}u - 1,$$

is convex. In this case, is not hard to prove that inequalities (1) and (17) are equivalents. On the other hand, if f' is strictly increasing and not necessarily convex then inequalities (1) and (17) are not equivalent. Because there exist functions strictly increasing, continuous, with derivative zero almost everywhere, see [16] (see also [17]). Note that these functions are not absolutely continuous, so they cannot be represented by an integral.

4. Final remarks

Theorem 14 has many interesting special cases, including Smale's theorem on Newton's method (see [18]) for analytical functions; see [10]. Theorem 2 has Nesterov–Nemirovskii's theorem on Newton's method (see [19]) for self-concordant functions as a special case; see [4].

Let us talk about some computational aspects of Newton's method for solving the nonlinear equation

$$F(x) = 0, \tag{20}$$

where $F : \Omega \to \mathbb{R}^n$ is a continuously differentiable function and $\Omega \subseteq \mathbb{R}^n$ is an open set. Note that the first equality in (2) is equivalent to

$$x_{k+1} = x_k + S_k, \qquad F'(x_k)S_k = -F(x_k) \quad k = 0, 1, \ldots. \tag{21}$$

Since the solution of the linear systems in (21) for large systems is computationally expensive, namely, at each iteration the derivative at x_k must be computed and stored. Besides, the solution of the linear system in (21) is required. To circumvent these drawbacks, Dembo, Eisenstat and Steihaug introduced in [20] the inexact Newton's method. The inexact Newton's methods for solving nonlinear equation (20) is any method which, given an initial point x_0, generates a sequence $\{x_k\}$ as follows:

$$x_{k+1} = x_k + S_k, \qquad F'(x_k)S_k = -F(x_k) + r_k, \quad \|r_k\| \le \theta_k \|F(x_k)\|, \ k = 0, 1, \ldots,$$

for a suitable forcing sequence $\{\theta_k\}$, which is used to control the level of accuracy. Therefore, solutions of practical problems are obtained by computational implementations of the inexact Newton-like methods. The analysis of these methods under majorant condition will be done in the future.

Acknowledgements

The author was supported by CNPq Grants 302618/2005-8 and 475647/2006-8, PRONEX–Optimization (FAPERJ/CNPq) and FUNAPE/UFG.

References

[1] J. Chen, The convergence analysis of inexact GaussNewton methods for nonlinear problems, Comput. Optim. Appl. 40 (1) (2008) 97–118.
[2] J. Chen, W. Li, Convergence behaviour of inexact Newton methods under weak Lipschitz condition, J. Comput. Appl. Math. 191 (1) (2006) 143–164.
[3] C. Li, W.P. Shen, Local convergence of inexact methods under the Holder condition, J. Comput. Appl. Math. 222 (2008) 544–560.
[4] O.P. Ferreira, Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle, IMA J. Numer. Anal. 29 (3) (2009) 746–759.
[5] O.P. Ferreira, M.L.N. Gonçalves, Local convergence analysis of inexact Newton-like methods under majorant condition, Comput. Optim. Appl., 2009, in press (doi:10.1007/s10589-009-9249-2).
[6] O.P. Ferreira, B.F. Svaiter, Kantorovich's majorants principle for Newton's method, Comput. Optim. Appl. 42 (2009) 213–229.
[7] Z. Huang, The convergence ball of Newton's method and the uniqueness ball of equations's under Hölder-type continuos derivatives, Comput. Math. Appl. 47 (2004) 247–251.
[8] P.D. Proinov, General local convergence theory for a class of iterative processes and its applications to Newton's method, J. Complexity 25 (1) (2009) 8–62.
[9] X. Wang, Convergence of Newton's method and uniqueness of the solution of equation in Banach space, IMA J. Numer. Anal. 20 (2000) 123–134.
[10] X. Wang, C. Li, Convergence of Newton's method and Uniqueness of the Solution of Equations in Banach Spaces II, Acta Math. Sin. (Engl. Ser.) 19 (2) (2003) 405–412.
[11] L.B. Rall, A note on the convergence of Newton's method, SIAM J. Numer. Anal. 11 (1) (1974) 34–36.
[12] W.C. Rheinboldt, An adaptive continuation process for solving systems of nonlinear equations, Banach Center Publ. 3 (1978) 129–142. Papers, Fifth Semester, Stefan Banach Internat. Math. Center, Warsaw, 1975.
[13] J.F. Traub, H. Wozniakowski, Convergence and complexity of Newton iteration for operator equation, J. Assoc. Comput. Mach. 26 (2) (1979) 250–258.
[14] P. Deuflhard, Newton methods for nonlinear problems. Affine invariance and adaptive algorithms, in: Series Computational Mathematics, Vol. 35, Springer, Heidelberg, 2004.
[15] P. Deuflhard, G. Heindl, Affine invariant convergence for Newtons method and extensions to related methods, SIAM J. Numer. Anal. 16 (1) (1979) 1–10.
[16] L. Takács, An increasing continuous singular function, Amer. Math. Monthly 85 (1) (1978) 35–37.
[17] H. Okamoto, M. Wunsch, A geometric construction of continuous, strictly increasing singular functions, Proc. Japan Acad. Ser. A Math. Sci. 83 (7) (2007) 114–118.
[18] S. Smale, Newton method estimates from data at one point, in: R. Ewing, K. Gross, C. Martin (Eds.), The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, Springer-Verlag, New York, 1986, pp. 185–196.
[19] Y. Nesterov, A. Nemirovskii, Interior-point polynomial algorithms in convex programming, in: SIAM Studies in Applied Mathematics, Vol. 13, Philadelphia, 1994.
[20] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (1982) 400–408.