ELSEVIER

Journal of Computational and Applied Mathematics 79 (1997) 131–145

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

A new semilocal convergence theorem for Newton's method

José M. Gutiérrez *

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

Received 21 March 1995; revised 3 December 1996

Abstract

A new semilocal convergence theorem for Newton's method is established for solving a nonlinear equation $F(x)=0$, defined in Banach spaces. It is assumed that the operator F is twice Fréchet differentiable, and F'' satisfies a Lipschitz type condition. Results on uniqueness of solution and error estimates are also given. Finally, these results are compared with those that use Kantorovich conditions.

Keywords: Nonlinear equations in Banach spaces; Newton's method; Convergence theorem; Error estimates; Majorizing sequences

AMS classification: 47H10, 65J15

1. Introduction

Let X, Y be Banach spaces and $F:\Omega\subseteq X\to Y$ be a nonlinear twice Fréchet differentiable operator in an open convex domain $\Omega_0\subseteq\Omega$. Let us assume that $F'(x_0)^{-1}\in\mathscr{L}(Y,X)$ exists at some $x_0\in\Omega_0$, where $\mathscr{L}(Y,X)$ is the set of bounded linear operators from Y into X.

Newton's method for solving the equation

$$F(x)=0 \tag{1.1}$$

is defined, starting from x_0, as follows:

$$x_{n+1}=x_n-F'(x_n)^{-1}F(x_n),\quad n\geqslant 0, \tag{1.2}$$

provided that $F'(x_n)^{-1}\in\mathscr{L}(Y,X)$ exists at each step.

* E-mail: jmguti@siur.unirioja.es.

PII S0377-0427(96)00169-0

Most of the authors study the convergence of the sequence (1.2) towards a solution of (1.1) under the conditions of the Kantorovich theorem [6–8], or closely related ones [9, 11, 13]. In these results it is assumed that the second Fréchet derivative F'' is continuous and bounded in Ω_0, or the weaker assumption of the Lipschitz continuity of F' in Ω_0.

Huang [5] has recently obtained a new convergence theorem for Newton's method, assuming that F'' satisfies a Lipschitz type condition. This new result is an alternative to Newton–Kantorovich theorem and it can be used in situations where this theorem fails, as we see in some examples. When both theorem fulfill, we compare them in order to obtain the best results on existence and uniqueness of solution for (1.1).

In this paper we assume that F satisfies the condition

$$\|F'(x_0)^{-1}(F''(x) - F''(x_0))\| \leqslant k\|x - x_0\|, \quad x \in \Omega_0. \tag{1.3}$$

We introduce the linear operator $L_F(x) : X \to X$, formally defined by

$$L_F(x) = F'(x)^{-1}F''(x)F'(x)^{-1}F(x). \tag{1.4}$$

This operator and its properties were studied in [3]. Notice that in the scalar case, if f is a convex function, $L_f(t)$ is a punctual measure of the convexity of f, called degree of logarithmic convexity (see [4]) and is defined as follows:

$$L_f(t) = \frac{f(t)f''(t)}{f'(t)^2}.$$

Assuming (1.3), and using the linear operator $L_F(x)$, we give a convergence theorem for Newton's method, in the same way as for Huang's result. We obtain a cubic polynomial which majorizes F and we establish results on convergence and error estimates for (1.2), as well as uniqueness of the solution for (1.1).

Next, we extend these results to another situation which includes, as particular cases, condition (1.3) or the hypothesis of Hölder continuity for F'' (see [1]).

2. Main results

One of the techniques to prove the convergence of a sequence $\{z_n\}$ in a Banach space is the use of a majorizing sequence [12], i.e., a real nonnegative sequence $\{s_n\}$ which satisfies

$$\|z_{n+1} - z_n\| \leqslant s_{n+1} - s_n, \quad n \geqslant 0.$$

Note that the convergence of $\{s_n\}$ implies the convergence of $\{z_n\}$.

In this paper, we present a new method for finding majorizing sequences for Newton's method, by using the linear operator $L_F(x)$ and its connection with this method. We can write the sequence (1.2) in the form $x_{n+1} = G(x_n)$, where

$$G(x) = x - F'(x)^{-1}F(x). \tag{2.1}$$

If G is a differentiable operator at x, it was shown [3] that $G'(x) = L_F(x)$. By applying [8, Theorem XVIII.1.1] we obtain the next result.

Lemma 2.1. *With the previous notations, let us assume that a real function f satisfies*
(i) $\|F'(x_0)^{-1}F(x_0)\| \leqslant -f(t_0)/f'(t_0)$,
(ii) $\|L_F(x)\| \leqslant L_f(t)$, *for* $\|x-x_0\| \leqslant t-t_0$.
Then, the Newton sequence

$$t_{n+1} = t_n - f(t_n)/f'(t_n),$$

starting at t_0, *is a majorizing sequence of* (1.2), *i.e.*,

$$\|x_{n+1}-x_n\| \leqslant t_{n+1}-t_n, \quad n \geqslant 0.$$

In what follows, we write $\Gamma_0 = F'(x_0)^{-1}$, and assume that F satisfies (1.3) and, besides,

$$\|\Gamma_0 F(x_0)\| \leqslant a, \tag{2.2}$$

$$\|\Gamma_0 F''(x_0)\| \leqslant b. \tag{2.3}$$

Let us denote

$$B(x_0; r) = \{x \in X;\ \|x-x_0\| < r\}$$

and

$$\overline{B(x_0; r)} = \{x \in X;\ \|x-x_0\| \leqslant r\}.$$

Consider the cubic polynomial defined by

$$p(t) = a - t + \frac{b}{2}t^2 + \frac{k}{6}t^3, \tag{2.4}$$

where k, a and b are given as above. Next, we obtain a result on the existence of the linear operator $L_F(x)$ and some properties of the polynomial $p(t)$.

Lemma 2.2. *Let*

$$m = \frac{2}{b + \sqrt{b^2 + 2k}}. \tag{2.5}$$

Then, the linear operator $L_F(x)$ *is well defined for* $x \in B(x_0; m)$ *and satisfies*

$$\|L_F(x)\| \leqslant \frac{(b + k\|x-x_0\|)\|\Gamma_0 F(x)\|}{[1 - \frac{1}{2}k\|x-x_0\|^2 - b\|x-x_0\|]^2}.$$

Proof. Notice that

$$\int_{x_0}^{x} \Gamma_0[F''(y) - F''(x_0)]\,\mathrm{d}y = \Gamma_0 F'(x) - I - \Gamma_0 F''(x_0)(x-x_0).$$

Therefore, from (1.3) it is deduced that

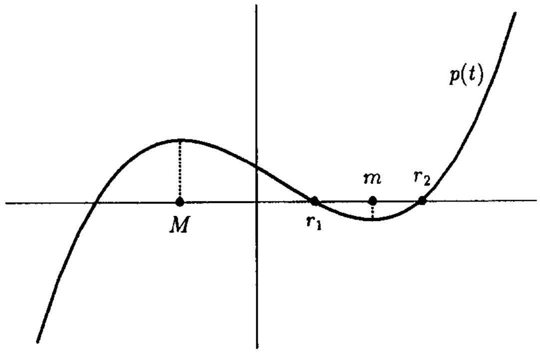

Fig. 1. Polynomial defined in (2.4).

$$\|\Gamma_0 F'(x) - I - \Gamma_0 F''(x_0)(x - x_0)\| = \left\| \int_{x_0}^{x} \Gamma_0 [F''(y) - F''(x_0)]\,\mathrm{d}y \right\| \leqslant \frac{k\|x - x_0\|^2}{2}.$$

So, if $x \in B(x_0; m)$, we have

$$\|\Gamma_0 F'(x) - I\| \leqslant \tfrac{1}{2} k\|x - x_0\|^2 + b\|x - x_0\| < 1.$$

Then, by Banach's theorem on existence of inverse operators (see [8]), the operator $\Gamma_0 F'(x)$ has a continous inverse in $B(x_0; m)$, and

$$\|[\Gamma_0 F'(x)]^{-1}\| \leqslant \frac{1}{1 - (\frac{1}{2}k\|x - x_0\|^2 + b\|x - x_0\|)}.$$

Taking into account (1.3) and (2.3), we obtain

$$\|\Gamma_0 F''(x)\| \leqslant b + k\|x - x_0\|,$$

and so, the result follows from (1.4). □

Notice that the polynomial $p(t)$ has a maximum at

$$t = M = -\frac{b + \sqrt{b^2 + 2k}}{k} < 0,$$

and a minimum at $t = m > 0$, where m is given by (2.5) (see Fig. 1). A necessary and sufficient condition for p to have positive roots is

$$p(m) = p\left(\frac{2}{b + \sqrt{b^2 + 2k}}\right) \leqslant 0. \tag{2.6}$$

Each one of the following conditions are equivalent to (2.6):

$$6ab^3 + 9a^2k^2 + 18abk \leqslant 3b^2 + 8k$$

or

$$3ak^2 + 3bk + b^3 \leqslant [b^2 + 2k]^{3/2}. \tag{2.7}$$

Theorem 2.3. *Let F be the operator defined in* (1.1) *satisfying* (1.3), (2.2) *and* (2.3). *Let p be the polynomial defined in* (2.4). *Assume that $B(x_0; m) \subseteq \Omega_0$. If* (2.6) *holds, then p has two positive roots, r_1, r_2 $(r_1 \leqslant r_2)$ and the sequence $\{x_n\}$ defined by* (1.2) *converges to x^*, solution of* (1.1) *in $\overline{B(x_0; r_1)}$. If $r_1 < r_2$ the solution is unique in $B(x_0; r_2)$. If $r_1 = r_2$ the solution is unique in $\overline{B(x_0; r_1)}$.*

Proof. Under the previous assumptions we prove that

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n, \quad n \geqslant 0$$

where $\{t_n\}$ is the Newton sequence to solve $p(t) = 0$, starting at $t_0 = 0$.

First notice that x_1 is defined and besides

$$\|x_1 - x_0\| = \|\Gamma_0 F(x_0)\| \leqslant a = t_1 - t_0 < r_1 \leqslant m.$$

So, $x_1 \in B(x_0; m)$ and because of Lemma 2.2, the linear operator $L_F(x_1)$ exists. Then, we can define x_2 by means of

$$x_2 - x_1 = \int_{x_0}^{x_1} L_F(x)\,\mathrm{d}x. \tag{2.8}$$

For $x \in [x_0, x_1]$, we have $x = x_0 + s(x_1 - x_0)$, where $0 \leqslant s \leqslant 1$. By Taylor's formula [11],

$$\Gamma_0 F(x) = \Gamma_0 F(x_0) + (x - x_0) + \tfrac{1}{2}\Gamma_0 F''(x_0)(x - x_0)^2 + \int_{x_0}^{x} \Gamma_0[F''(y) - F''(x_0)](x - y)\,\mathrm{d}y.$$

As $x - x_0 = s(x_1 - x_0) = -s\Gamma_0 F(x_0)$, then

$$\|\Gamma_0 F(x)\| \leqslant (1 - s)\|\Gamma_0 F(x_0)\| + \tfrac{1}{2}\|\Gamma_0 F''(x_0)\|\|(x - x_0)^2\| + \left\|\int_{x_0}^{x} \Gamma_0[F''(y) - F''(x_0)](x - y)\,\mathrm{d}y\right\|.$$

Taking into account (1.3), (2.2), and writing $sa = t$, we obtain

$$\|\Gamma_0 F(x)\| \leqslant (1 - s)a + \tfrac{1}{2}bs^2a^2 + \tfrac{1}{6}ka^3s^3 = a - t + \tfrac{1}{2}bt^2 + \tfrac{1}{6}kt^3 = p(t). \tag{2.9}$$

From Lemma 2.2 we deduce

$$\|L_F(x)\| \leqslant \frac{(b + kt)(a - t + \frac{1}{2}bt^2 + \frac{1}{6}kt^3)}{[1 - \frac{1}{2}kt^2 - bt]^2} = L_p(t),$$

and, by (2.8),

$$\|x_2 - x_1\| = \left\|\int_{x_0}^{x} L_F(x)\,\mathrm{d}x\right\| \leqslant \int_{t_0}^{t_1} L_p(t)\,\mathrm{d}t = t_2 - t_1.$$

Using similar arguments, one can show that

$$\|x_3 - x_2\| \leqslant t_3 - t_2,$$

and so on.

The convergence of $\{t_n\}$ [9] implies that $\{x_n\}$ is a Cauchy sequence and therefore converges to a limit x^*. From (2.9) and the continuity of F, we have $F(x^*) = 0$.

Finally, to prove the uniqueness, suppose $r_1 < r_2$ and $\tilde{x}$ is another solution of (1.1) in $B(x_0; r_2)$. Then,

$$\|\tilde{x} - x_0\| \leqslant \rho(r_2 - t_0), \quad \text{with } 0 < \rho < 1.$$

Following Huang's technique [5], it can be shown that

$$\|\tilde{x} - x_n\| \leqslant \rho^{2^n}(r_2 - t_n), \quad n \geqslant 0.$$

If $r_1 = r_2$, and $\tilde{x} \in \overline{B(x_0; r_1)}$, we have

$$\|\tilde{x} - x_n\| \leqslant r_1 - t_n, \quad n \geqslant 0.$$

In both cases, $\|\tilde{x} - x_n\| \to 0$ when $n \to \infty$, and therefore $\tilde{x} = x^*$. □

Ostrowski [9] obtained an error expression for Newton's method applied to a quadratic polynomial, in terms of the polynomial roots. Using a similar technique, we establish the following result for the cubic polynomial (2.4).

Theorem 2.4. *Let p be the polynomial defined in (2.4), with a negative root, $-r_0$, and two positive roots, $r_1 \leqslant r_2$. Then, the Newton sequence*

$$t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}, \quad t_0 = 0, \tag{2.10}$$

converges to r_1. Moreover, if $r_1 < r_2$, we have for $n \geqslant 0$

$$(r_2 - r_1)\frac{\alpha^{2^n}}{r - \alpha^{2^n}} \leqslant r_1 - t_n \leqslant (r_2 - r_1)\frac{\theta^{2^n}}{R - \theta^{2^n}},$$

where

$$0 < r = \frac{r_0 - r_2}{r_0 - r_1} < 1, \quad 0 < R = 1 - \frac{r_2 - r_1}{r_0 + r_1} < 1,$$

$$0 < \alpha = r\frac{r_1}{r_2} < 1, \quad 0 < \theta = R\frac{r_1}{r_2} < 1.$$

If $r_1 = r_2$, then

$$r_1\left(\frac{r_0 - r_1}{2r_0 - r_1}\right)^n \leqslant r_1 - t_n \leqslant \frac{r_1}{2^n}.$$

Proof. Notice that p is a decreasing, convex function in $[0, m]$, $p(0) > 0 \geqslant p(m)$, and $r_1 \leqslant m$. Then, it is well known [9] that Newton sequence (2.10) converges to r_1. Moreover, $\{t_n\}$ is an increasing sequence.

To obtain the error expression, we can write the polynomial (2.4) in the form

$$p(t)=\tfrac{1}{6}k(r_1-t)(r_2-t)(r_0+t).$$

Comparing the polynomial coefficients of t^2, we have $r_1+r_2\leqslant r_0$. We also have

$$p'(t)=-\tfrac{1}{6}k[(r_2-t)(r_0+t)+(r_1-t)(r_0+t)-(r_1-t)(r_2-t)].$$

Let us denote $a_n=r_1-t_n$, $b_n=r_2-t_n$ and $c_n=r_0+t_n$. Then, by (2.10)

$$\begin{aligned}a_{n+1}&=r_1-t_{n+1}=r_1-t_n+\frac{p(t_n)}{p'(t_n)}\\&=a_n-\frac{a_nb_nc_n}{b_nc_n+a_nc_n-a_nb_n}=\frac{a_n^2(c_n-b_n)}{b_nc_n+a_nc_n-a_nb_n}.\end{aligned}$$

In a similar way, we obtain

$$b_{n+1}=\frac{b_n^2(c_n-a_n)}{b_nc_n+a_nc_n-a_nb_n}.$$

So, we have

$$\frac{a_{n+1}}{b_{n+1}}=\left(\frac{a_n}{b_n}\right)^2\frac{c_n-b_n}{c_n-a_n}.$$

Notice that

$$\frac{c_n-b_n}{c_n-a_n}=\frac{(r_0-r_2)+2t_n}{(r_0-r_1)+2t_n}$$

and

$$r=\frac{r_0-r_2}{r_0-r_1}\leqslant\frac{(r_0-r_2)+2t_n}{(r_0-r_1)+2t_n}\leqslant\frac{r_0-r_2+2r_1}{r_0+r_1}=R. \tag{2.11}$$

That means that

$$\frac{a_{n+1}}{b_{n+1}}\leqslant R\left(\frac{a_n}{b_n}\right)^2\leqslant R^3\left(\frac{a_{n-1}}{b_{n-1}}\right)^4\leqslant\cdots\leqslant R^{2^{n+1}-1}\left(\frac{a_0}{b_0}\right)^{2^{n+1}}=\frac{1}{R}\left(R\frac{r_1}{r_2}\right)^{2^{n+1}}.$$

Taking into account that $b_n=r_2-r_1+a_n$, it follows that

$$a_n\left(1-\frac{\theta^{2^n}}{R}\right)\leqslant\frac{r_2-r_1}{R}\theta^{2^n},$$

and

$$r_1-t_n\leqslant(r_2-r_1)\frac{\theta^{2^n}}{R-\theta^{2^n}}.$$

For the lower estimate, we obtain from (2.11)

$$\frac{a_{n+1}}{b_{n+1}}\geqslant r\left(\frac{a_n}{b_n}\right)^2\geqslant r^3\left(\frac{a_{n-1}}{b_{n-1}}\right)^4\geqslant\cdots\geqslant r^{2^{n+1}-1}\left(\frac{a_0}{b_0}\right)^{2^{n+1}}=\frac{1}{r}\left(r\frac{r_1}{r_2}\right)^{2^{n+1}}=\frac{1}{r}\alpha^{2^{n+1}}.$$

Therefore

$$a_n\left(1-\frac{\alpha^{2^n}}{r}\right) \geqslant \frac{r_2-r_1}{r}\alpha^{2^n},$$

and hence

$$(r_2-r_1)\frac{\alpha^{2^n}}{r-\alpha^{2^n}} \leqslant r_1 - t_n.$$

Finally, if $r_1 = r_2$, then

$$a_{n+1} = a_n\frac{c_n-a_n}{2c_n-a_n}.$$

Since

$$\frac{r_0-r_1}{2r_0-r_1} \leqslant \frac{c_n-a_n}{2c_n-a_n} \leqslant \frac{1}{2},$$

we obtain the result. □

We extend the result obtained in the Theorem 2.3 to a more general situation. Assume, instead of (1.3), that F satisfies

$$\|\Gamma_0[F''(x)-F''(x_0)]\| \leqslant k\|x-x_0\|^p, \quad k>0, \ p \geqslant 0, \ x \in \Omega_0. \tag{2.12}$$

Observe that for $p=0$, we have $\|\Gamma_0 F''(x)\| \leqslant k + \|\Gamma_0 F''(x_0)\| = k'$, and we are in the situation of the Kantorovich theorem [8, Theorem XVIII.1.6]. When $p=1$, we have (1.3). If $p \in (0,1)$, it is said that F'' is Hölder continous on Ω_0.

In any case, and following the proof of the Theorem 2.3, we obtain a majorizing sequence for (1.2), by using Newton's method for the real equation $f(t)=0$, where

$$f(t) = a - t + \frac{b}{2}t^2 + \frac{k}{(p+1)(p+2)}t^{p+2}, \quad t \geqslant 0, \tag{2.13}$$

with k, a and b as before.

Note. The equation $f'(t)=0$ has only one positive solution. This we call m. Moreover, m is a minimum of f. Therefore, $f(m) \leqslant 0$ is a necessary and sufficient condition for the existence of positive solutions of $f(t)=0$. Let us denote these solutions r_1 and r_2 ($r_1 \leqslant r_2$). So, we can write

$$f(t) = (r_1-t)(r_2-t)g(t) \quad \text{with} \quad g(r_1) \neq 0 \neq g(r_2). \tag{2.14}$$

Observe that f is a decreasing, convex function in $[0,m]$, and $f(0) > 0 \geqslant f(m)$. These conditions are sufficient for the convergence of the sequence

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)}, \quad t_0 = 0, \tag{2.15}$$

to r_1. Moreover, $\{t_n\}$ is an increasing sequence.

Repeating the proofs of Lemma 2.2 and Theorem 2.3, we obtain the next results.

Lemma 2.5. *Let m be the minimum of the function f given by* (2.13). *Then, the linear operator $L_F(x)$ is defined for $x \in B(x_0; m)$ and*

$$\|L_F(x)\| \leqslant \frac{(b + k\|x - x_0\|^p)\|\Gamma_0 F(x)\|}{[1 - k/(p+1)\|x - x_0\|^{p+1} - b\|x - x_0\|]^2}.$$

Theorem 2.6. *Let F be as in Theorem 2.3, with* (2.12) *instead of* (1.3). *Suppose that f is given by* (2.13) *and has two positive roots. Then, the Newton sequence* (1.2) *converges to a solution x^* of* (1.1) *in $\overline{B(x_0; r_1)}$, where r_1 is the smallest root of* (2.13). *Besides, the solution is unique in $B(x_0; r_2)$, where $r_2 \neq r_1$ is the another root of* (2.13). *If $r_1 = r_2$ the solution is unique in $\overline{B(x_0; r_1)}$.*

Next, we obtain error estimates for the sequence (2.15), when f is given by (2.13).

Theorem 2.7. *The sequence $\{t_n\}$ defined in* (2.15) *is an increasing convergent sequence to r_1. Moreover, if $r_1 < r_2$, we have*

$$\frac{(r_2 - r_1)}{r - \alpha^{2^n}}\alpha^{2^n} \leqslant r_1 - t_n \leqslant \frac{(r_2 - r_1)}{R - \theta^{2^n}}\theta^{2^n},$$

where

$$r = \min_{t\in[0,r_1]} H(t), \qquad R = \max_{t\in[0,r_1]} H(t), \quad \theta = R\frac{r_1}{r_2}, \qquad \alpha = r\frac{r_1}{r_2},$$

and

$$H(t) = \frac{g(t) - (r_2 - t)g'(t)}{g(t) - (r_1 - t)g'(t)}.$$

If $r_1 = r_2$, we have

$$\hat{r}^n r_1 \leqslant r_1 - t_n \leqslant \widehat{R}^n r_1,$$

where

$$\hat{r} = \min_{t\in[0,r_1]} \widehat{H}(t), \qquad \widehat{R} = \max_{t\in[0,r_1]} \widehat{H}(t),$$

$$\widehat{H}(t) = \frac{g(t) - (r_1 - t)g'(t)}{2g(t) - (r_1 - t)g'(t)}.$$

Proof. We have already stated that $\{t_n\}$ is an increasing convergent sequence to r_1. Taking into account (2.14),

$$f'(t) = -(r_2 - t)g(t) - (r_1 - t)g(t) + (r_1 - t)(r_2 - t)g'(t).$$

Let $a_n = r_1 - t_n$, $b_n = r_2 - t_n$. Following the proof of Theorem 2.4, we have

$$a_{n+1} = a_n^2\frac{b_n g'(t_n) - g(t_n)}{f'(t_n)}, \qquad b_{n+1} = b_n^2\frac{a_n g'(t_n) - g(t_n)}{f'(t_n)}.$$

Then,

$$r\left(\frac{a_n}{b_n}\right)^2 \leqslant \frac{a_{n+1}}{b_{n+1}} = \left(\frac{a_n}{b_n}\right)^2 \frac{g(t_n)-b_n g'(t_n)}{g(t_n)-a_n g'(t_n)} \leqslant R\left(\frac{a_n}{b_n}\right)^2,$$

and therefore,

$$\frac{1}{r}\left(r\frac{r_1}{r_2}\right)^{2^n} \leqslant \frac{a_n}{b_n} \leqslant \frac{1}{R}\left(R\frac{r_1}{r_2}\right)^{2^n}.$$

Taking into account that $b_n = r_2 - r_1 + a_n$, the first part follows.

If $r_1 = r_2$, then

$$a_{n+1} = r_1 - t_{n+1} = a_n \frac{g(t_n)-a_n g'(t_n)}{2g(t_n)-a_n g'(t_n)},$$

and the second part also holds. □

3. Illustrative examples and concluding remarks

We have studied the Newton sequence (1.2) under different assumptions from those of the Kantorovich theorem. Now, we analyze both conditions in two different ways: accessibility of solution and results on existence and uniqueness of solution.

The Kantorovich theorem assumes that F satisfies

$$\|F'(x_0)^{-1}(F'(x)-F'(y))\| \leqslant c\|x-y\|, \quad x,y \in \Omega_0, \tag{3.1}$$

and $ac \leqslant \frac{1}{2}$, where a is given by (2.2). Notice that (3.1) is slightly weaker than the original Kantorovich assumption (see [8, 13]).

In regard to the accessibility of solution, Theorem 2.3 and the Kantorovich theorem are not comparable. The following examples show situations where the Kantorovich assumptions fail and the Theorem 2.3 fulfills or vice versa.

Example 3.1 (*Huang* [5]). Let $X = [-1,1]$, $Y = \mathbb{R}$, $x_0 = 0$ and $f: X \to Y$ the polynomial

$$f(x) = \tfrac{1}{6}x^3 + \tfrac{1}{6}x^2 - \tfrac{5}{6}x + \tfrac{1}{3}.$$

In this case, $c = \frac{8}{5}$ and $a = \frac{2}{5}$. Then $ac = 16/25 > \frac{1}{2}$. Therefore, Kantorovich condition fails and we cannot guarantee the convergence of the Newton sequence starting from x_0.

However, under the assumptions of the Theorem 2.3, and with the same notation, we have $a = \frac{2}{5}$, $b = \frac{2}{5}$ and $k = \frac{6}{5}$. Then,

$$3ak^2 + 3bk + b^3 = 3\frac{36}{25}\frac{2}{5} + 3\frac{6}{5}\frac{2}{5} + \frac{8}{125} = \frac{404}{125},$$

$$(b^2+2k)^{3/2} = \left(\frac{4}{25} + 2\frac{6}{5}\right)^{3/2} = \left(\frac{8}{5}\right)^3 = \frac{512}{125}.$$

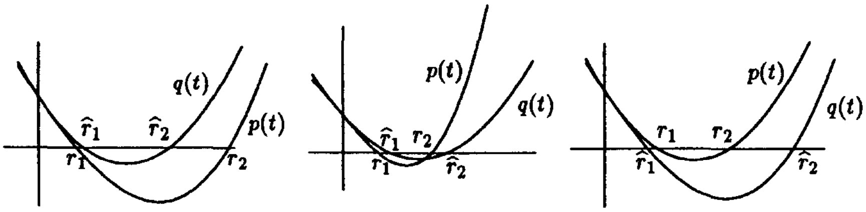

Fig. 2. Location of the roots of p and q.

Therefore, condition (2.7) holds, and consequently f satisfies the hypothesis of the Theorem 2.3. Hence, the Newton's sequence converges starting from x_0.

Example 3.2. Let $X = Y = \mathbb{R}$, $x_0 = 0$ and $f : X \to Y$ the function

$$f(x) = \sin x - 5x - 8.$$

In this case, $a = 2$, $b = 0$ and $c = k = \frac{1}{4}$. Then $ac = \frac{1}{2}$ and the hypothesis of the Kantorovich theorem holds. However, the polynomial (2.4) appearing in the Theorem 2.3,

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a = \frac{1}{24}t^3 - t + 2,$$

has not got positive roots. Consequently, we cannot use the Theorem 2.3 in order to prove the convergence of the Newton's sequence converges starting from x_0.

Sometimes, the convergence of (1.2) can be established using the Kantorovich theorem or the Theorem 2.3 indistinctly. Then we wonder which result gives us more accurate information on the solutions of (1.1). Under the assumptions of the Theorem 2.3 we locate the solutions of (1.1) in terms of the roots of the polynomial (2.4) (error estimates are given in Theorem 2.4). Under Kantorovich assumptions the information is obtained from the quadratic polynomial

$$q(t) = \tfrac{1}{2}ct^2 - t + a. \tag{3.2}$$

Let us denote by $\hat{r}_1, \hat{r}_2$ $(\hat{r}_1 \leqslant \hat{r}_2)$ the roots of q. Then

$$p(\hat{r}_j) = \tfrac{1}{2}\hat{r}_j^2\left(\tfrac{1}{3}k\hat{r}_j - (c - b)\right), \quad j = 1, 2.$$

Observe that

$$p(\hat{r}_1) \leqslant 0 \Leftrightarrow k(1 - \sqrt{1 - 2ac}) \leqslant 3c(c - b),$$

$$p(\hat{r}_2) \leqslant 0 \Leftrightarrow k(1 + \sqrt{1 - 2ac}) \leqslant 3c(c - b).$$

Our goal now is to get the smallest region where the solution is located and the biggest one where this solution is unique. We distinguish three situations (see Fig. 2):

Case 1: $k(1 + \sqrt{1 - 2ac}) \leqslant 3c(c - b)$. Then $r_1 \leqslant \hat{r}_1$, $\hat{r}_2 \leqslant r_2$ and, consequently, the solution x^* is located in $\overline{B(x_0; r_1)}$ and is unique in $B(x_0; r_2)$.

Case 2: $k(1 - \sqrt{1 - 2ac}) \leqslant 3c(c - b) < k(1 + \sqrt{1 - 2ac})$. In this situation $r_1 \leqslant \hat{r}_1$, $r_2 \leqslant \hat{r}_2$. Then the solution x^* belongs to $\overline{B(x_0; r_1)}$ and is the only one in $B(x_0; \hat{r}_2)$.

Case 3: $3c(c-b) \leqslant k(1-\sqrt{1-2ac})$. Now we have $\widehat{r}_1 \leqslant r_1$, $r_2 \leqslant \widehat{r}_2$, thus x^* is located in $\overline{B(x_0;\widehat{r}_1)}$ and is unique in $B(x_0;\widehat{r}_2)$.

In the cases 1 and 3 we get the best information from Theorem 2.3 and the Kantorovich theorem respectively. But in the case 2, the best information is obtained by mixing both results.

Example 3.3. Let $X=[0,1]\times[0,1]$, $Y=\mathbb{R}^2$, $(x_0,y_0)=(0,0)$ and $F: X \to Y$ given by

$$F(x,y)=\left(\frac{x^3}{24}+\frac{y^2}{4}-x+\frac{1}{3}, \frac{y^3}{8}+\frac{3x^2}{4}-3y+1\right).$$

We consider the max-norm in $\mathbb{R}^2$. For a bilinear operator B on X defined by the following calculation scheme:

$$B(x,y)=(x_1,x_2)\begin{pmatrix} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ \hline b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22}\end{pmatrix}\begin{pmatrix} y_1 \\ y_2\end{pmatrix}$$

$$=\begin{pmatrix} b_1^{11}x_1+b_1^{21}x_2 & b_1^{12}x_1+b_1^{22}x_2 \\ b_2^{11}x_1+b_2^{21}x_2 & b_2^{12}x_1+b_2^{22}x_2\end{pmatrix}\begin{pmatrix} y_1 \\ y_2\end{pmatrix}$$

$$=\begin{pmatrix} b_1^{11}x_1y_1+b_1^{21}x_2y_1+b_1^{12}x_1y_2+b_1^{22}x_2y_2 \\ b_2^{11}x_1y_1+b_2^{21}x_2y_1+b_2^{12}x_1y_2+b_2^{22}x_2y_2\end{pmatrix},$$

$$x=(x_1,x_2)\in X, \quad y=(y_1,y_2)\in X,$$

we consider the norm (see [2, 10])

$$\|B\| = \sup_{\|x\|=1} \max_i \sum_{j=1}^{2} \left|\sum_{k=1}^{2} b_i^{jk} x_k\right|.$$

Then

$$\Gamma_0 F(0,0)=\begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{3}\end{pmatrix}\begin{pmatrix} \frac{1}{3} \\ 1\end{pmatrix}=\begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{3}\end{pmatrix},$$

and $\|\Gamma_0 F(0,0)\| = \frac{1}{3} = a$.

On the other hand, if we compose a linear operator in $\mathbb{R}^2$

$$L=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}$$

with B we obtain a new bilinear operator, whose associated matrix is

$$LB=\begin{pmatrix} a_{11}b_1^{11}+a_{12}b_2^{11} & a_{11}b_1^{12}+a_{12}b_2^{12} \\ a_{11}b_1^{21}+a_{12}b_2^{21} & a_{11}b_1^{22}+a_{12}b_2^{22} \\ \hline a_{21}b_1^{11}+a_{22}b_2^{11} & a_{21}b_1^{12}+a_{22}b_2^{12} \\ a_{21}b_1^{21}+a_{22}b_2^{21} & a_{21}b_1^{22}+a_{22}b_2^{22}\end{pmatrix}.$$

As a particular case

$$\Gamma_0 F''(0,0) = \begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \\ \hline \frac{3}{2} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -\frac{1}{2} \\ \hline -\frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}$$

and $\|\Gamma_0 F''(0,0)\| = \frac{1}{2} = b$. Besides,

$$\Gamma_0[F''(x,y) - F''(0,0)] = \begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} x/4 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & 3y/4 \end{pmatrix} = \begin{pmatrix} -x/4 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & -y/4 \end{pmatrix}$$

and $\|\Gamma_0[F''(x,y) - F''(0,0)]\| \leqslant \frac{1}{4}\|(x,y)\|$. Hence $k = \frac{1}{4}$.

The polynomial given by (2.4) is

$$p(t) = \tfrac{1}{3} - t + \tfrac{1}{4}t^2 + \tfrac{1}{24}t^3.$$

This polynomial has positive roots and therefore the Theorem 2.3 holds.

We also have

$$\Gamma_0 F''(x,y) = \begin{pmatrix} -x/4 & 0 \\ 0 & -1/2 \\ \hline -1/2 & 0 \\ 0 & -y/4 \end{pmatrix}$$

and then,

$$\|\Gamma_0 F''(x,y)\| \leqslant \tfrac{3}{4} = c, \qquad \forall (x,y) \in X = [0,1] \times [0,1].$$

So, we obtain $ac = \frac{1}{4} < \frac{1}{2}$, and Kantorovich theorem also holds. In this situation, the polynomial (3.2) is

$$q(t) = \tfrac{3}{8}t^2 - t + \tfrac{1}{3}.$$

Since

$$k(1 + \sqrt{1 - 2ac}) = \frac{1 + \sqrt{\frac{1}{2}}}{4} < 3c(c - b) = \frac{9}{16},$$

as in the case 1, we obtain better information from the polynomial (2.4). Actually, we know that the solution is located in $\overline{B(0; r_1)} = \overline{B(0; 0.3695)}$ and is the only one in $B(0; r_2) = B(0; 2.4533)$ instead

Table 1
Error comparisons

Iteration	$r_1 - t_n$	$\widehat{r}_1 - s_n$
0	0.3695850618081907	0.3905242917512699
1	0.0362517284748574	0.1067190958417936
2	0.0004701842187366	0.0163540286238108
3	0.0000000820038027	0.0000014159342765
4	0.0000000000000025	0.0000000000010632

of the regions obtained from the quadratic polynomial: $\overline{B(0;\widehat{r}_1)} = \overline{B(0;0.3905)}$ and $B(0;\widehat{r}_2) = B(0;2.2761)$.

Finally, we compare the error bounds we get from p and q. Let us denote (x^*, y^*) the solution of $F(x,y) = 0$, and $\{x_n\}$ the Newton sequence

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) - F'(x_n, y_n)^{-1}F(x_n, y_n), \quad (x_0, y_0) = (0,0).$$

Let r_1 and $\widehat{r}_1$ be the smallest positive roots of $p(t) = 0$ and $q(t) = 0$ respectively; $\{t_n\}$ and $\{s_n\}$ the Newton sequences

$$t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}, \qquad s_{n+1} = s_n - \frac{q(s_n)}{q'(s_n)}, \qquad s_0 = t_0 = 0.$$

We know that

$$\|(x^*, y^*) - (x_n, y_n)\| \leqslant r_1 - t_n$$

and

$$\|(x^*, y^*) - (x_n, y_n)\| \leqslant \widehat{r}_1 - s_n.$$

Both error estimates are compared in Table 1.

Acknowledgements

Supported in part by a grant of the University of La Rioja (ULR-96PYB25MHV).

References

[1] I.K. Argyros, Newton-like methods under mild differentiability conditions with error analysis, *Bull. Austral. Math. Soc.* **37** (1988) 131–147.

[2] I.K. Argyros, A Newton-like method for solving nonlinear equations in Banach space, *Acta Math. Hungh.* **27** (1992) 368–378.

[3] J.M. Gutiérrez, M.A. Hernández and M.A. Salanova, Accessibility of Solutions by Newton's Method, *Internat. J. Comput. Math.* **57** (1995) 239–247.

[4] M.A. Hernández, Newton–Raphson's method and convexity, *Zb. Rad. Prirod.-Mat. Fak. Ser. Mat.* **22**(1) (1992) 159–166.

[5] Z. Huang, A note on the Kantorovich theorem for Newton iteration, *J. Comput. Appl. Math.* **47** (1993) 211–217.
[6] L.V. Kantorovich, On Newton's method for functional equations, *Dokl. Akad. Nauk SSRR* **59** (1948) 1237–1240.
[7] L.V. Kantorovich, The majorant principle and Newton's method, *Dokl. Akad. Nauk SSRR* **76** (1951) 17–20.
[8] L.V. Kantorovich and G.P. Akilov, *Functional Analysis* (Pergamon, Oxford, 1982).
[9] A.M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces* (Academic Press, New York, 1943).
[10] L.B. Rall, Quadratic equations in Banach spaces, *Rend. Circ.. Mat. Palermo* (2) **10** (1961) 314–332.
[11] L.B. Rall, *Computational Solution of Nonlinear Operator Equations* (Robert E. Krieger, New York, 1979).
[12] W.C. Rheinboldt, A unified convergence theory for a class of iterative process, *SIAM J. Numer. Anal.* **5** (1968) 42–63.
[13] T. Yamamoto, A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions, *Numer. Math.* **49** (1986) 203–220.