INFORMATICA, 1993, Vol.4, No.1-2, 188-198

ON THE KANTOROVICH HYPHOTHESIS FOR NEWTON'S METHOD

Dana PETCU

University of Timişoara
B-dul Vasile Pârvan 4, 1900 Timişoara, Romania

Abstract. The article is dedicated to Newton's method for solving nonlinear equation systems. The Kantorovich convergence theorem assumes that the derivative of the system function is Lipschitz continuous. Our purpose is to provide error estimates in the case of a Hölder continuous derivative.

Key words: numerical analysis, Kantorovich's theorems, Newton method, nonlinear equations, error estimate.

1. Introduction. The problem we discuss is to find solutions of the systems of equation

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

where $F: D \subset R^n \to R^n$ is a given operator. The Newton's iterates are

$$x^{k+1} = x^k - F'(x^k)^{-1}F(x^k), \quad k = 0, 1, \ldots, \tag{2}$$

where F' denotes the G-derivative. The use of the Newton's iteration is considered as a standard numerical solutions. The convergence analysis has a long history. In 1948 Kantorovich gave a famous convergence theorem. A later proof, is due to Ortega and Rheinboldt (1970):

Newton-Kantorovich Theorem. *Assume that F: $D \subset R^n \to R^n$ is F-differentiable on a convex set $D_0 \subset D$ and that*

$$\|F'(x) - F'(y)\| \leqslant \gamma\|x - y\|, \quad \forall x, y \in D_0. \tag{3}$$

Suppose that there exists an $x^0 \in D_0$ such that $\|F'(x^0)^{-1}\| \leqslant \beta$ and $\alpha = \beta\gamma\mu \leqslant 1/2$, where $\mu \geqslant \|F'(x^0)^{-1}F(x^0)\|$. Set

$$t^* = (\beta\gamma)^{-1}\left[1-(1-2\alpha)^{1/2}\right], \quad t^{**} = (\beta\gamma)^{-1}\left[1+(1-2\alpha)^{1/2}\right],$$

and assume that $\overline{S}(x^0,t^*) \subset D_0$. Then the Newton iterates are well-defined, remain in $\overline{S}(x^0,t^*)$, and converge to a solution x^* of $F(x) = 0$ which is unique in $S(x^0,t^{**}) \cap D_0$. Moreover, the error estimate

$$\|x^* - x^k\| \leqslant (\beta\gamma 2^k)^{-1}(2\alpha)^{2^k}, \quad k = 0,1,\dots \tag{4}$$

holds.

We note that the error estimates (also in later versions) suppose the condition of Lipschitz continuity.

Counter example. For the following system

$$\begin{cases} x_1 + \sqrt{x_2^3} = 4, \\ \sqrt{x_1^3} + x_2 = 8 \end{cases} \tag{5}$$

the Lipschitz inequality (3) doesn't holds for the Euclidean norm.

Proof. If $x = (4, n^{-2})$, $y = (4,0)$, then (3) holds if $\gamma \geqslant 3/2n$, which is impossible for all $n \geqslant 0$.

Note. Instead of (3), we have only a Hölder condition:

$$\|F'(x) - F'(y)\| \leqslant \gamma\|x-y\|^p, \quad \forall x,y \in D_0, \quad 0 \leqslant p \leqslant 1 \tag{6}$$

with $p = 1/2$, $\gamma = 3/2$.

In the case of only Hölder continuous derivative, *Newton Attraction theorem* (Ortega and Rheinboldt, 1970) states that, if x^* is a point of attraction of the iteration, then we have superlinear convergence.

We did not find something about the error estimate of the iteration applied to a system like (5). There are several references (Anselsone and Moore, 1966; Dennis, 1969; Gragg and Tapia, 1974) about some generalizations of Kantorovich theorem, however, which assume some conditions for the second derivative.

The purpose of this paper is to find formulas of the error estimate for Newton method in the case a Hölder continuous derivative.

More precisely, we try to generalize the Kantorovich and Mysovskii theorems:

Newton-Mysovskii Theorem. (Ortega and Rheinboldt, 1970). *Suppose that $F: D \subset R^n \to R^n$ is F-differentiable on a convex set $D_0 \subset D$ and that for each $x \in D_0$, $F'(x)$ is nonsingular and satisfies the Lipschitz condition (3) and $\|F'(x)^{-1}\| \leqslant \beta$, $\forall x \in D_0$. If $x^0 \in D_0$ is such that $\|F'(x^0)^{-1}F(x^0)\| \leqslant \mu$ and $\alpha = \beta\gamma\mu/2 < 1$, as well as $\overline{S}(x^0, r_0) \subset D_0$, where*

$$r_0 = \mu \sum_{j=0}^{\infty} \alpha^{2^j - 1}, \tag{7}$$

then the Newton iterates (2) remain in $\overline{S}(x^0, r_0)$ and converge to a solution x^* of $F(x) = 0$. Moreover, the following estimate hold

$$\|x^* - x^k\| \leqslant \varepsilon_k \|x^k - x^{k-1}\|^2, \quad k = 1, 2, \ldots \tag{8}$$

where

$$\varepsilon_k = \frac{\alpha}{\mu} \sum_{j=0}^{\infty} \left(\alpha^{2^k}\right)^{2^j - 1} \leqslant \alpha\left[\mu(1 - \alpha^{2^k})\right]^{-1}, \quad k = 1, 2, \ldots. \tag{9}$$

The results are stated in Theorem 1 and 2. For the case of Theorem 1 the proof follows the way of the above theorem, but in the case of Ortega's proof for Kantorovich theorem the analogy method do not work.

In the case $p = 1$ the given formulas provide the classic estimates.

2. Main results. We propose the following generalization of the Newton-Mysovskii theorem:

Theorem 1. *Suppose that $F: D \subset R^n \to R^n$ is F-differentiable on a convex set $D_0 \subset D$ and that for each $x \in D_0$, $F'(x)$ is nonsingular and satisfies the Hölder condition (6) with $p \leqslant 1$ and $\|F'(x)^{-1}\| \leqslant \beta$, $\forall x \in D_0$. If $x^0 \in D_0$ is such that $\|F'(x^0)^{-1}F(x^0)\| \leqslant \mu$ and $\alpha = \frac{1}{p+1}\beta\gamma\mu^p < 1$, as well as $\overline{S}(x^0, r_0) \subset D_0$, where*

$$r_0 = \mu \sum_{j=0}^{\infty} \alpha^{((p+1)^j - 1)/p}, \tag{10}$$

then the Newton iterates (2) remain in $\overline{S}(x^0, r_0)$ and converge to a solution x^* of $F(x) = 0$. Moreover, the following estimate hold

$$\|x^* - x^k\| \leqslant \varepsilon_k \|x^k - x^{k-1}\|^{p+1}, \quad k = 1, 2, \ldots, \tag{11}$$

where

$$\begin{aligned} \varepsilon_k &= \frac{\alpha}{\mu^p} \sum_{j=0}^{\infty} \left(\alpha^{(p+1)^k/p}\right)^{((p+1)^j - 1)/p} \\ &\leqslant \alpha\mu^{-p}\left(1 - \alpha^{(p+1)^k/p}\right)^{-1}, \quad k = 1, 2, \ldots. \end{aligned} \tag{12}$$

For the Kantorovich theorem, we find a partial generalization:

Theorem 2. *Assume that $F: D \subset R^n \to R^n$ is F-differentiable on a convex set $D_0 \subset D$ and the inequality (6) holds for $p \leqslant 1$. Suppose that there exists an $x^0 \in D_0$ such that $\|F'(x^0)^{-1}\| \leqslant \beta$ and $\alpha = \beta\gamma\mu^p \leqslant \frac{p}{p+1}$, where $\mu \geqslant \|F'(x^0)^{-1}F(x^0)\|$ and $\overline{S}(x^0, (\beta\gamma)^{-1/p}) \subset D_0$. Then the Newton iterates (2) are well-defined, remain in $\overline{S}(x^0, (\beta\gamma)^{-1/p})$, and converge to a solution x^* of $F(x) = 0$. Moreover, the error estimate*

$$\|x^* - x^k\| \leqslant \frac{1}{p}\left(\frac{p}{\beta\gamma}\right)^{1/p}(p+1)^{-k}\left[\frac{(p+1)^p\alpha}{p}\right]^{(p+1)^k/p}, \quad k = 0, 1, \ldots \tag{13}$$

holds.

We apply the last theorem to our counter example.

Example. In the case of system (5), the error estimate is

$$\|x^* - x^k\| \leqslant \frac{2}{9}\|F'(x^0)^{-1}\|^{-2}\left(\frac{2}{3}\right)^{(\frac{3}{2})^k + k}$$

for an x^0 such that

$$\|F'(x^0)^{-1}\|_2^3 \|F(x^0)\|_2 \leqslant \frac{4}{81}$$

holds.

3. Proofs. We need the following earlier results of Ortega and Rheinboldt (1970).

The mean-value theorem for functions with Hölder continuous F-derivative. *Let $F: D \subset R^n \to R^n$ be continuously differentiable on a convex set $D_0 \subset D$ and suppose that, for constants $\gamma \geqslant 0$ and $p \geqslant 0$, F' satisfies (6). Then, for any $x, y \in D_0$,*

$$\|Fy - Fx - F'(x)(y-x)\| \leqslant \frac{\gamma}{p+1}\|y-x\|^{p+1}, \tag{14}$$

Perturbation Lemma. *Let $A, C \in L(R^n)$ and assume that A is invertible, with $\|A^{-1}\| \leqslant \beta$. If $\|A - C\| \leqslant \gamma$ and $\beta\gamma < 1$, then C is also invertible, and $\|C^{-1}\| \leqslant \beta/(1-\gamma\beta)$.*

DEFINITION. Let $\{x^k\}$ be any sequence in R^n. Then a sequence $\{t^k\}$ for which

$$\|x^{k+1} - x^k\| \leqslant t_{k+1} - t_k, \quad k = 0, 1, \ldots$$

holds is a majorizing sequence for $\{x^k\}$.

Note that any majorizing sequence is necessarily monotonically increasing.

Majorizing sequences arise as solutions of certain nonlinear difference equations. The idea is given in the following two lemma.

Lemma 1. *For $G: D \subset R^n \to R^n$, suppose that there exists an monotone function $\phi: [0, \infty) \to [0, \infty)$ such that on some set $D_0 \subset D$*

$$\|G^2(x) - G(x)\| \leqslant \phi(\|G(x) - x\|), \quad \forall x, G(x) \in D_0, \tag{15}$$

assume, further, that for some $x^0 \in D_0$ the iterates $x^k = G^k(x^0)$, $k = 1, 2, \ldots$ remain in D_0 and the sequence t_k defined by

$$t_{k+1} = t_k + \phi(t_k - t_{k-1}), \quad t_0 = 0, \quad t_1 \geqslant \|G(x^0) - x^0\|, \quad k = 1, 2, \ldots$$

converges to $t^* < \infty$. Then $\lim_{k\to\infty} x^k = x^*$ exists and the estimate

$$\|x^* - x^k\| \leqslant t^* - t_k, \quad k = 0, 1, \ldots \tag{16}$$

holds. Moreover, if $x^* \in D$ and G is continuous at x^*, then $x^* = Gx^*$.

If an inequality of the form

$$\|G^2(x) - G(x)\| \leqslant \phi(\|G(x) - x)\|, \ \|G(x) - x^0\|, \|x - x^0\|), \quad \forall x, \ G(x) \in D_0 \tag{17}$$

holds, Lemma 1 extends to a more general setting:

Lemma 2. *Let $G: D \subset R^n \to R^n$ and $\phi: J_1 \times J_2 \times J_3 \subset R^3 \to [0, \infty)$, where each J_i is an interval of the form $[0, \alpha]$, $[0, \alpha)$, or $[0, \infty)$ and ϕ is monotone in each variable. Suppose that there is a set $D_0 \subset D$ and an $x^0 \in D_0$ such that (17) holds whenever $x, G(x) \in D_0$, and that with $t_0 = 0$, $t_1 \geqslant \|x^0 - G(x^0)\|$ the solution of the difference equation*

$$t_{k+1} - t_k = \phi(t_k - t_{k-1}, t_k, t_{k-1}), \quad k = 1, 2, \ldots$$

exists and converges to $t^* < \infty$. Finally, assume either that $\overline{S}(x^0, t^*) \subset D_0$ or that $S(x^0, t^*) \subset D_0$ and $t_k < t^*$ for all $k \geqslant 0$. Then the iterates $x^{k+1} = G(x^k)$, $k = 0, 1, \ldots$ are well-defined, lie in $\overline{S}(x^0, t^*)$, converge to some $x^* \in \overline{S}(x^0, t^*)$, and satisfy (16).

Proof of Theorem 1. We define $Gx = x - F'(x)^{-1}F(x)$. Then

$$\|G^2(x) - G(x)\| \leqslant \beta\|F(G(x)) - F(x) - F'(x)(G(x) - x)\| \leqslant \frac{1}{p+1}\beta\gamma\|G(x) - x\|^{p+1},$$

and the associate difference equation is the following:

$$t_{k+1} - t_k = \frac{1}{p}\beta\gamma(t_k - t_{k-1})^{p+1}, \quad t_0 = 0, \quad t_1 = \mu, \quad k = 1, 2, \ldots.$$

We show by induction that

$$t_{k+1} - t_k \leqslant \mu\alpha^{((p+1)^k - 1)/p}, \quad k = 0, 1, \ldots \tag{18}$$

holds. For $k = 0$ that is correct because $t_1 - t_0 = \mu \leqslant \mu$. If it holds for $k = j - 1$, then

$$t_{j+1} - t_j \leqslant \frac{1}{p+1}\beta\gamma\left(\mu\alpha^{((p+1)^{j-1} - 1)/p}\right)^{p+1} = \mu\alpha^{\frac{(p+1)^j - 1}{p}}.$$

Inequality (18) provides

$$t^* = \lim_{k\to\infty} t_k = \lim_{k\to\infty} \sum_{j=0}^{k-1}(t_{j+1} - t_j) \leqslant \mu \sum_{j=0}^{\infty} \alpha^{\frac{(p+1)^j-1}{p}} = r_0.$$

Now the convergence statement follows from Lemma 1.

In order to obtain an error estimate, we define $\alpha_0 = \beta\gamma/(p+1)$ and from

$$\begin{aligned}\|x^{k+1} - x^k\| &\leqslant \alpha_0\|x^k - x^{k-1}\|^{p+1}\\ &\leqslant \alpha_0(t_k - t_{k-1})^{p+1} \leqslant \alpha_0\mu^{p+1}\alpha^{((p+1)^k-(p+1))/p}\end{aligned}$$

results

$$\begin{aligned}\|x^{k+m} - x^k\| &\leqslant \sum_{j=k}^{k+m-1} \|x^{j+1} - x^j\| \leqslant \sum_{j=1}^{m} \alpha_0^{((p+1)^j-1)/p}\|x^k - x^{k-1}\|^{(p+1)^j}\\ &\leqslant \|x^k - x^{k-1}\|^{p+1} \sum_{j=1}^{\infty} \alpha_0^{((p+1)^j-1)/p}(t_k - t_{k-1})^{((p+1)^j-(p+1))/p}.\end{aligned}$$

Since

$$\begin{aligned}&\sum_{j=1}^{\infty}\alpha_0^{((p+1)^j-1)/p}(t_k - t_{k-1})^{((p+1)^j-(p+1))/p}\\ &\quad\leqslant \sum_{j=1}^{\infty}\alpha_0^{((p+1)^j-1)/p}\left(\frac{\alpha^{(p+1)^{k-1}/p}}{\alpha_0}\right)^{((p+1)^j-(p+1))/p}\\ &\quad= \alpha_0\sum_{j=0}^{\infty}\alpha^{(p+1)^k((p+1)^j-1)/p^2} = \varepsilon_k < \alpha\mu^{-p}\left[1 - \alpha^{(p+1)^k/p}\right]^{-1},\end{aligned}$$

the proposed inequality holds.

Proof of Theorem 2. Set $D_1 = S(x^0, (\beta\gamma)^{-1/p}) \cap D_0$. Then for $x \in D_1$ we have

$$\|F'(x) - F'(x^0)\| \leqslant \gamma\|x - x^0\|^p < \frac{1}{\beta} \leqslant \frac{1}{\|F'(x^0)^{-1}\|}.$$

Thus, the perturbation lemma shows that $F'(x)$ is nonsingular for all $x \in D_1$, and that

$$\|F'(x)^{-1}\| \leqslant \frac{\beta}{1 - \beta\gamma\|x - x^0\|^p}, \quad \forall x \in D_1,$$

so that $G(x) = x - F'(x)^{-1}F(x)$ is well defined on D_1.

If $x, G(x) \in D_1$, it follows from the mean-value theorem that

$$\begin{aligned}\|G^2(x) - G(x)\| &= \|F'(G(x))^{-1}F(G(x))\| \\ &= \|F'(G(x))^{-1}[F(G(x)) - F(x) - F'(x)(G(x) - x)\| \\ &\leqslant \frac{1}{p+1}\beta\gamma\frac{\|G(x) - x\|^{p+1}}{1 - \beta\gamma\|G(x) - x^0\|^p} \\ &= \phi\left(\|G(x) - x\|, \|G(x) - x^0\|\right),\end{aligned}$$

where

$$\phi(s,t) = \frac{1}{p+1}\beta\gamma\frac{s^{p+1}}{1 - \beta\gamma t^p}.$$

We now apply Lemma 2 to the difference equation

$$t_{k+1} - t_k = \phi(t_k - t_{k-1}, t_k), \quad t_0 = 0, \quad t_1 = \mu, \quad k = 1, 2, \ldots.$$

We now show that $\{t_k\}$ is an increasing and bounded sequence, in fact that

$$t_{k+1}^p < \frac{1}{\beta\gamma}, \tag{19.1}$$

$$t_{k+1} > t_k, \tag{19.2}$$

$$t_{k+1} - t_k \leqslant \mu\alpha^k \tag{19.3}$$

holds for $k = 0, 1, \ldots$. We use the induction method. The inequalities clearly hold $k = 0$. We assume they hold for $k - 1$.

First we consider the inequality (19.1). It follows from (19.3) for $i \doteq 0(1)k - 1$ that

$$t_k \leqslant \mu + \ldots + \mu\alpha^{k-1} = \mu\frac{1 - \alpha^k}{1 - \alpha}. \tag{20}$$

From the recurrence relation for t_{k+1} we obtain the following equivalent form for (19.1):

$$\frac{(\beta\gamma)^{\frac{p+1}{p}}}{p+1}(t_k - t_{k-1})^{p+1} \leqslant (1 - \beta\gamma t_k^p)[1 - (\beta\gamma)^{1/p}t_k].$$

It is possible to consider the following more restricted inequality, if we consider the definition of α, (20), the condition that $p \leqslant 1$, and (19.3) for $k-1$:

$$\alpha^{(k-1)p+1}\alpha^{k-1+1/p} \leqslant \left[1-\alpha\left(\frac{1-\alpha^k}{1-\alpha}\right)^p\right]\left[1-\alpha^{1/p}\frac{1-\alpha^k}{1-\alpha}\right].$$

This follows from

$$\alpha^{k-1+1/p}+\alpha^{1/p}\frac{1-\alpha^k}{1-\alpha} \leqslant \left[\alpha^{(k-1)p+1}+\alpha\left(\frac{1-\alpha^k}{1-\alpha}\right)^p\right]^{1/p} \leqslant 1,$$

$$\alpha^p \leqslant \left(\frac{p}{p+1}\right)^p \leqslant \frac{1}{2},$$

$$\alpha^{(k-1)p+1}+\alpha\left(\frac{1-\alpha^k}{1-\alpha}\right)^p \leqslant \alpha[\alpha^{(k-1)p}+1+\ldots+\alpha^{(k-1)p}]$$

$$\leqslant \frac{1}{2}\left(1+\ldots+\frac{1}{2^{k-1}}\right)=1.$$

With this (19.1) is proved.

From (19.1) for k and (19.2) for $k-1$ we notice that

$$t_{k+1}-t_k=\frac{\beta\gamma}{p+1}\frac{(t_k-t_{k-1})^{p+1}}{1-\beta\gamma t_k^p}>0.$$

We consider now the inequality (19.3). From (19.3) for $k-1$ and (20), the inequality at k is satisfied if

$$f(k)=\frac{\alpha^{kp-p}}{p+1}+\alpha\left(\frac{1-\alpha^k}{1-\alpha}\right)^p \leqslant 1.$$

But this is true because $f'(k) \leqslant 0$ and therefore $f(k) \leqslant f(1)=1$, $\forall k \geqslant 1$.

The inequalities (19) prove that $t^* = \lim_{k\to\infty} t_k$ exists and $t^* \in \overline{S}(t_0,(\beta\gamma)^{-1/p})$.

Lemma 2 now ensures that all x^k are well-defined, remain in D_1, and converge to a $x^* \in D_1$. Moreover, G is continuous at x^* so that $x^* = G(x^*)$. This implies that $F(x^*)=0$.

In order to obtain the error estimate, we show first that

$$\frac{1}{1-\beta\gamma t_k^p} \leqslant \frac{1}{p}(p+1)^{kp}.$$

That result from (20), the boundary of α (from the hypothesis) and

$$\frac{p}{(p+1)^{kp}} + \frac{p}{p+1}\left[\frac{1-(p/(p+1))^k}{1-p/(p+1)}\right]^p$$
$$\leqslant f(k) = \frac{p}{(p+1)^{kp}} + \left(\frac{p}{p+1}\right)^{1-p}\left[1-\frac{1}{(p+1)^k}\right]^p \leqslant 1,$$

which is true because $f'(k) \geqslant 0$ and $f(k) \leqslant \lim_{k\to\infty} f(k) = [p/(p+1)]^{1-p} \leqslant 1$.

In addition

$$t_{k+1} - t_k \leqslant \left(\frac{p}{\beta\gamma}\right)^{1/p} \frac{1}{(p+1)^{k+1}}\left[\frac{(p+1)^p\alpha}{p}\right]^{(p+1)^k/p}, \quad k = 0, 1, \ldots$$

holds. This is clearly correct for $k = 0$ and, if it holds for $k-1$, then for k

$$t_{k+1} - t_k = \frac{\beta\gamma}{p+1}\frac{(t_k - t_{k-1})^{p+1}}{1-\beta\gamma t_k^p} \leqslant \frac{(p+1)^{kp-1}}{p}\beta\gamma(t_k - t_{k-1})^{p+1}$$
$$= (p+1)^{(p+1)^k-(k+1)}\left(\frac{\beta\gamma}{p}\right)^{((p+1)^k-1)/p}\left(\frac{\alpha}{\beta\gamma}\right)^{(p+1)^k/p}$$

Now, the error estimate on t_k follows from

$$t^* - t_k = \sum_{i=0}^{\infty}(t_{k+i+1} - t_{k+i})$$
$$\leqslant \left(\frac{p}{\beta\gamma}\right)^{1/p}\frac{1}{(p+1)^{k+1}}\left[\frac{(p+1)^p\alpha}{p}\right]^{(p+1)^k/p}$$
$$\times \sum_{i=0}^{\infty}\frac{1}{(p+1)^i}\left[\frac{(p+1)^p\alpha}{p}\right]^{(p+1)^i/p}.$$

Using $(p+1)^p\alpha/p \leqslant 1$, following from $\alpha \leqslant p/(p+1)$ and $(p+1)^{p-1} \leqslant 1$, we obtain

$$t^* - t_k \leqslant \left(\frac{p}{\beta\gamma}\right)^{1/p}\frac{1}{(p+1)^{k+1}}\left[\frac{(p+1)^p\alpha}{p}\right]^{(p+1)^k/p}\sum_{i=0}^{\infty}\frac{1}{(p+1)^i}.$$

The statement (13) follows from Lemma 2. With this Theorem 2 is proved.

Acknowledgement. The author is greatful to Dr. Georg Bader from University of Heidelberg for a number of suggestions that helped to improve this paper.

REFERENCES

Anselsone, P., and R.Moore (1966). An extension of the mention Kontorovich method for solving nonlinear equations with an applications to elastiaty. *J. Math. Anal. Appl.*, **13**, 476–501.

Denis, J.E. (1969). On the Kantorovich hypothesis for Menton's method. *SIAM J. Numer. Anal.*, **6**(3), 493–507.

Gragg, W.B., and R.A. Tapica (1974). Optimal error bounds for the Newton-Kantorovich theorem. *SIAM J. Numer. Anal.*, **11**, 10–13.

Ortega, J.M., and C.W. Rheinboldt (1970). *Iterative solution of Nonlinear Equations in Several Variables*. Academic Press, New York.

Received December 1992

D. Petcu is an assistant to the Departement of Informatics at the Mathematics Faculty of Timişoara University, Romania. At present, she is engaged in research on the doctoral thesis – about the numerical solutions for stiff differential equation systems.