Eur. J. Math. Anal. 3 (2023) 15
doi: 10.28924/ada/ma.3.15

Developments on the Convergence Analysis of Newton-Kantorovich Method for Solving Nonlinear Equations

Samundra Regmi[1], Ioannis K. Argyros[2,*], Santhosh George[3], Michael I. Argyros[4]

[1]*Department of Mathematics, University of Houston, Houston, TX 77204, USA*
sregmi5@uh.edu

[2]*Department of Computing and Mathematical Sciences, Cameron University, Lawton, OK 73505, USA*
iargyros@cameron.edu

[3]*Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, India-575 025*
sgeorge@nitk.edu.in

[4]*Department of Computer Science, University of Oklahoma, Norman, 73019, OK, USA*
michael.i.argyros-1@ou.edu

[*]*Correspondence: iargyros@cameron.edu*

Abstract. Developments are presented for the semi-local convergence of Newton's method to solve Banach space-valued nonlinear equations. By utilizing a new methodology, we provide a finer convergence analysis with no additional conditions than in earlier results. In particular, this is done by introducing the center-Lipschitz condition by which we construct a stricter domain than the original domain of the operator. Then, the Lipschitz constants in the new domain are at least as small as the original constants leading to weaker sufficient convergence criteria, tighter error bounds on the error distances involved, and a piece of better information on the location of the solution. These benefits are obtained under the same computational cost since in practice the computation of the original constants requires the computation of the new constants as special cases. The same benefits are obtained if the Lipschitz conditions are replaced by Hölder conditions or even more general $\omega-$continuity conditions. This methodology can be applied to other methods using such as the Secant, Stirling's Newton-like, and other methods along the same lines. Numerical examples indicate that the new results can be utilized to solve nonlinear equations, but not earlier ones.

1. Introduction

Consider the problem of finding a solution $x_* \in \Omega$ of the equation

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

Received: 26 Jan 2023.

Key words and phrases. Newton-Kantorovich method; Convergence; Banach space.

where $F : \Omega \longrightarrow \mathcal{E}_2$ is a continuously differentiable operator in the Fréchet-sense, $\mathcal{E}_1, \mathcal{E}_2$ are Banach spaces and $\Omega \subset \mathcal{E}_1$ is an open set.

The solution x_* in closed form is desirable. But this is possible only in special cases. So, most solution methods for (1.1) are iterative methods. The convergence regions for these methods are small in general, so their applicability is reduced. The error bounds are also pessimistic (in general).

Among the iterative methods, the most famous one is Newton's method (NM) defined for $n = 0, 1, 2, \ldots$ by

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

Kantorovich provided the semi-local convergence analysis of NM utilizing the contraction mapping principle attributed to Banach. In particular, he presented two different proofs using majorant functions or recurrence relations [15]. His so-called Newton-Kantorovich Theorem is that no assumption on the solution is made and at the same time, the existence of the solution x_* is established. Numerous researchers used this theorem in applications and also as a theoretical tool [1–16]. But the convergence criteria may not hold although NM may converge. Motivated by these concerns and optimization considerations we present new results that not only extend the convergence region but also provide more precise error estimates and better knowledge of the location of the solution. The novelty of the article is that these benefits require no additional conditions. This is how the usage of NM is extended. The technique used can be applied to extend other iterative methods along the same lines.

2. Convergence Analysis

Let $\alpha > 0$, $\lambda \geq 0$ and $x_0 \in \Omega$ be such that $\|F'(x_0)^{-1}\| \leq \alpha$, $\|F'(x_0)^{-1}F(x_0)\| \leq \lambda$ and $F'(x_0)^{-1} \in L(\mathcal{E}_2, \mathcal{E}_1)$, the space of bounded linear operators from $\mathcal{E}_2$ to $\mathcal{E}_1$. By $B(x, b)$, $B[x, b]$ we denote the open and closed balls in $\mathcal{E}_1$, respectively with center $x \in \mathcal{E}_1$ and of radius $b > 0$.

Some Lipschitz-type conditions are needed.

Definition 2.1. *Operator F' is center-Lipschitz continuous about x_0 on Ω if there exists $L_0 > 0$ such that for all $u \in \Omega$*

$$\|F'(u) - F'(x_0)\| \leq L_0\|u - x_0\|. \tag{2.1}$$

Set

$$\Omega_0 = B(x_0, \frac{1}{\alpha L_0}) \cap \Omega. \tag{2.2}$$

Definition 2.2. *Operator F' is $1-$Restricted Lipschitz continuous on Ω_0 if there exists $L > 0$ such that*

$$\|F'(u) - F'(v)\| \leq L\|u - v\| \tag{2.3}$$

for all $u \in \Omega_0$, $v = u - F'(u)^{-1}F(u) \in \Omega_0$.

Definition 2.3. *Operator F' is $2-$Restricted Lipschitz continuous on Ω_0 if there exists $L_1 > 0$ such that for all $u, v \in \Omega_0$*

$$\|F'(u) - F'(v)\| \leq L_1 \|u - v\|. \tag{2.4}$$

Definition 2.4. *Operator F' is Lipschitz continuous on Ω if there exists $L_2 > 0$ such that for all $u, v \in \Omega$*

$$\|F'(u) - F'(v)\| \leq L_2 \|u - v\|. \tag{2.5}$$

Definition 2.5. *Assume:*

$$\lambda \alpha L_0 < 1 \tag{2.6}$$

and

$$\Omega_1 = B(x_1, \frac{1}{\alpha L_0} - \|x_1 - x_0\|) \subset \Omega \tag{2.7}$$

Then, operator is $3-$ restricted Lipschitz continuous on Ω_1 if there exists a constant $K > 0$ such that for all $u \in \Omega_1$

$$\|F'(u) - F'(v)\| \leq K \|u - v\| \tag{2.8}$$

for $v = u - F'(u)^{-1} F(u) \in \Omega_1$.

REMARK 2.6. *By the definition of sets Ω_0 and Ω_1, we get*

$$\Omega_0 \subseteq \Omega, \tag{2.9}$$

and

$$\Omega_1 \subseteq \Omega_0. \tag{2.10}$$

Indeed, if $y \in \Omega_1$, then we obtain

$$\begin{aligned} \|y - x_1\| &\leq \frac{1}{\alpha L_0} - \|x_1 - x_0\| \Rightarrow \|y - x_1\| + \|x_1 - x_0\| \leq \frac{1}{\alpha L_0} \\ &\Rightarrow \|y - x_0\| \leq \frac{1}{\alpha L_0} \Rightarrow y \in \Omega_0 \Rightarrow \Omega_1 \subseteq \Omega_0. \end{aligned}$$

It follows by these definitions, (2.9) and (2.10) that if the best constants are chosen in the Definitions 2.1-2.5, then

$$L \leq L_1 \leq L_2, \tag{2.11}$$

$$L_0 \leq L_2, \tag{2.12}$$

and

$$K \leq L. \tag{2.13}$$

Hence, parameter K can replace results on Newton's using the constants L, L_1 and L_2. Notice also that $L_0 = L_0(F', \Omega)$, $L = L(F', \Omega_0), L_1 = L_1(F', \Omega_0)$, $L_2 = L_2(F', \Omega)$ and $K = K(F, \Omega_0, \Omega_1)$. Examples, where (2.9)-(2.13) are strict can be found in the Numerical Section.

Notice that the computation of the constant L_2 requires the computation of the other constants as special cases. Hence, no additional effort is needed to compute them. Moreover, they all depend on the initial data (x_0, F, Ω).

It is also worth noticing that under (2.1) we obtain

$$\|F'(u)^{-1}\| \leq \frac{\alpha}{1-\alpha L_0\|u-x_0\|}. \tag{2.14}$$

This is a tighter estimate than using the stronger (2.5) to get

$$\|F'(u)^{-1}\| \leq \frac{\alpha}{1-\alpha L_2\|u-x_0\|}. \tag{2.15}$$

We assume from now on that

$$L_0 \leq K. \tag{2.16}$$

But if $K < L_0$ then, the following results hold with L_0 replacing K.

Based on the above we present two extended theorems on Newton's method.

An important role is played in the convergence of NM by the majorizing sequence $\{s_n\}$ defined by

$$\begin{aligned} s_0 &= 0, s_{n+1}-s_n = -\frac{p(s_n)}{p_0'(s_n)} \\ &= \frac{\alpha K(s_n-s_{n-1})^2}{1-L_0\alpha s_n}, \\ p(s) &= \frac{K}{2}s^2 - \frac{s}{\alpha} + \frac{\lambda}{\alpha}, \\ p_0(s) &= \frac{L_0}{2}s^2 - \frac{s}{\alpha\frac{\lambda}{\alpha}}. \end{aligned}$$

THEOREM 2.7. *(Extended Newton-Kantorovich Theorem [1,2,10,12,13,15,16]) Under conditions (2.1), (2.6)-(2.8) further suppose $B(x_0, s_*) \subset \Omega$,*

$$H = K\alpha\lambda \leq \frac{1}{2}. \tag{2.17}$$

Then, Newton's method (1.2) initiated at $x_0 \in \Omega$ generates a sequence $\{x_n\}$ such that:$\{x_n\} \subseteq B(x_0, s_*)$, $\lim_{n\longrightarrow\infty} x_n = x_* \in B[x_0, s_*]$.

$$\|x_{n+1}-x_n\| \leq s_{n+1}-s_n \tag{2.18}$$

$$\|x_*-x_n\| \leq s_*-s_n, \tag{2.19}$$

where, $\lim_{n\longrightarrow\infty} s_n = s_* = \frac{1-\sqrt{1-2H}}{K\alpha}$ and $s_{**} = \frac{1+\sqrt{1-2H}}{K\alpha}$. Moreover, the following items hold for $\tau = \frac{s_*}{s_{**}}$

$$s_* - s_n = \begin{cases} \frac{(s_{**}-s_*)\tau^{2^n}}{1-\tau^{2^n}}, & \text{if } s_* < s_{**} \\ \frac{1}{2^n}s_*, & \text{if } s_* = s_{**}. \end{cases}$$

Furthermore, the element x_* is the unique solution of equation $F(x) = 0$ in $B[x_0, \bar{s}]$, where $\bar{s} = \frac{2}{L_0\alpha} - s_*$ if $L_0\alpha s_* < 2$.

Proof. Simply replace L_2 by K and use (2.14) instead of (2.15) in the proof of the version of Newton-Kantorovich Theorem given in [10] (see also [3–9, 14–16].

□

REMARK 2.8. *(i)If $K = L_2$, the result of Theorem 2.7 reduces to one in the Newton-Kantorovich Theorem where*

$$H_K = L_2\alpha\lambda \leq \frac{1}{2}, \tag{2.20}$$

$$\begin{aligned} t_0 &= 0, t_{n+1} - t_n = -\frac{\bar{p}(t_n)}{\bar{p}'(t_n)} \\ &= \frac{\alpha L_2(t_n - t_{n-1})^2}{1 - L_2\alpha t_n}, \\ \bar{p}(s) &= \frac{L_2}{2}s^2 - \frac{s}{\alpha} + \frac{\lambda}{\alpha}, \end{aligned}$$

and $\lim_{n\longrightarrow\infty} t_n = t_* = \frac{1-\sqrt{1-2K_K}}{L_2\alpha}$ *and* $t_{**} = \frac{1+\sqrt{1-2K_K}}{L_2\alpha}$, $\bar{\bar{s}} = \frac{2}{L_2\alpha} - t_*$, $\mu = \frac{t_*}{t_{**}}$,

$$t_* - t_n = \begin{cases} \frac{(t_{**}-t_*)\mu^{2^n}}{1-\mu^{2^n}}, & \text{if } t_* < t_{**} \\ \frac{1}{2^n}t_*, & \text{if } t_* = t_{**}. \end{cases}$$

Then, in view of estimates (2.11)-(2.13) we have

$$H_K \leq \frac{1}{2} \Rightarrow H \leq \frac{1}{2}, \tag{2.21}$$

$$s_* \leq t_*, \bar{\bar{s}} \leq \bar{s}, \tag{2.22}$$

$$0 \leq s_{n+1} - s_n \leq t_{n+1} - t_n \tag{2.23}$$

and

$$0 \leq s_* - s_n \leq t_* - t_n. \tag{2.24}$$

Estimates (2.21)-(2.24) justify the advantages (A) as stated in the introduction.
(ii)A more careful look at the proof shows that tighter sequence $\{r_n\}$ defined by

$$\begin{aligned} r_0 &= 0, r_1 = \lambda, \; r_2 = r_1 + \frac{\alpha L_0(r_1 - r_0)^2}{2(1 - L_0\alpha r_1)}, \\ r_{n+2} &= r_{n+1} + \frac{K\alpha(r_{n+1} - r_n)^2}{2(1 - L_0\alpha r_{n+1})}, \end{aligned}$$

also majorizes sequence $\{x_n\}$. The sufficient convergence criterion for this sequence is given by

$$H_A = \bar{K}\alpha\lambda \leq \frac{1}{2}, \tag{2.25}$$

where $\bar{K} = \frac{1}{8}(4L_0 + \sqrt{KL_0 + 8L_0^2} + \sqrt{L_0K})$. *This criterion was given by us in [4] for* $K = L - 2$. *Notice that*

$$H \leq \frac{1}{2} \Rightarrow H_A \leq \frac{1}{2}. \tag{2.26}$$

Hence, if (2.25) and $\{r_n\}$ *replace (2.17) and* $\{s_n\}$ *the conclusions of Theorem 2.7 hold with these changes too.*

(iii)Suppose that there exist $a > 0, b > 0$ *such that*

$$\|F'(x_0 + \theta(x_1 - x_0)) - F'(x_0)\| \leq \tau a \|x_1 - x_0\| \tag{2.27}$$

and

$$\|F'(x_1) - F'(x_0)\| \leq b\|x_1 - x_0\| \tag{2.28}$$

for all $\tau \in [0, 1]$. *Then, it was shown in [5] that sequence* $\{q_n\}$ *defined by*

$$\begin{aligned} q_0 &= 0, q_1 = \lambda, \; q_2 = q_1 + \frac{\alpha a(q_1 - q_0)^2}{2(1 - b\alpha q_1)}, \\ q_{n+2} &= q_{n+1} + \frac{K\alpha(q_{n+1} - q_n)^2}{2(1 - L_0\alpha q_{n+1})} \end{aligned}$$

is also majorizing for sequence $\{x_n\}$. *The convergence criterion for sequence* $\{q_n\}$ *is given by*

$$H_{AA} = \frac{\lambda}{2c} \leq \frac{1}{2}, \tag{2.29}$$

where

$$p_1(s) = (Ka + 2dL_0(a - 2b))s^2 + 4p(L_0 + b)s - 4d,$$

$$d = \frac{2K}{K + \sqrt{K^2 + 8L_0K}},$$

and

$$c = \begin{cases} \frac{1}{L_0 + b}, & Ka + 2dL_0(a - 2b) = 0 \\ \text{positive root of } p_1, & Ka + 2dL_0(a - 2b) > 0 \\ \text{smaller positive root of } p_1, & Ka + 2dL_0(a - 2b) < 0. \end{cases}$$

Notice that $b \leq a \leq L_0$. *Hence,* $\{q_n\}$ *is a tighter majorizing sequence than* $\{r_n\}$.

Criterion (2.29) was given by us in [4] for $K = L_2$. *Therefore (2.29) and* $\{q_n\}$ *can also replace (2.17) and* $\{s_n\}$ *in Theorem 2.7.*

(iv) It follows from the definition of sequence $\{s_n\}$ *that if*

$$L_0\alpha s_n < 1. \tag{2.30}$$

Then, sequence $\{s_n\}$ *is such that* $0 \leq s_n \leq s_{n+1}$ *and* $\lim_{n\longrightarrow\infty} s_n = s_* \leq \frac{1}{L_0\alpha}$. *Hence, weaker than all conditions (2.30) can be used in Theorem 2.7.*

3. Examples

We test the convergence criteria.

EXAMPLE 3.1. *Defined the real function f on $\Omega = B[x_0, 1-\delta]$, $x_0 = 1$, $\delta \in (0, \frac{1}{2})$ by*

$$f(s) = s^3 - \delta.$$

Then, the definitions are satisfied for $\lambda = \frac{1-\delta}{3}$, $\alpha = \frac{1}{3}$, $L_0 = 3(3-\delta), L_2 = 6(2-\delta)$, $L_1 = 6(1+\frac{1}{3-\delta}), x_1 = \frac{2+\delta}{3}$, $L = \frac{5(\frac{4-\delta}{3-\delta})^3+\delta}{3(\frac{4-\delta}{3-\delta})^2}$, $a = b = \delta + 5$,

$K = \frac{5h^3+\delta}{3h^2}$, and $h = \frac{\delta+2}{3} + \frac{3-(1-\delta)(3-\delta)}{3(1-\delta)}$.

Denote by M_1, M_2, M_3, M_4 the set of values $\delta \in (0, \frac{1}{2})$ for which (2.20), (2.17), (2.25) and (2.29) are satisfied, respectively. Then, by solving these inequalities for δ, we get $M_1 = \emptyset$, $M_2 = (0.0751, 0.5)$, $M_3 = (0.1320, 0.5)$ and $M_4 = (0.3967, 0.5)$.

Notice in particular that the Newton-Kantorovich criterion (2.20) [1, 9–15] cannot assure convergence of NM since $M_1 = \emptyset$.

A second example is provided to show that our conditions can be used to solve equations in cases where the ones in [1, 2, 10, 12, 13] cannot.

EXAMPLE 3.2. *Consider $\mathcal{E}_1 = \mathcal{E}_2 = C[0,1]$ with the norm-max. Set $\Omega = B(x_0, 3)$. Define, Hammerstein-type integral operator M on Ω by*

$$M(z)(w) = z(w) - y(w) - \int_0^1 T(w,t)v^3(t)dt, \tag{3.1}$$

$w \in [0,1]$, $z \in C[0,1]$, where $y \in C[0,1]$ is fixed and T is a Green's Kernel defined by

$$T(w,u) = \begin{cases} (1-w)u, & if\ u \le w \\ w(1-u), & if\ w \le u. \end{cases} \tag{3.2}$$

Then, the derivative M' according to Fréchet is defined by

$$[M'(v)(z)](w) = z(w) - 3\int_0^1 T(w,u)v^2(t)z(t)dt, \tag{3.3}$$

$w \in [0,1]$, $z \in C[0,1]$. Let $y(w) = x_0(w) = 1$. Then, using (3.1)-(3.3), we obtain $M'(x_0)^{-1} \in L(\mathcal{E}_2, \mathcal{E}_1)$, $\|I - M'(x_0)\| < \frac{3}{8}$, $\|M'(x_0)^{-1}\| \le \frac{8}{5} := \alpha$, $\lambda = \frac{1}{5}$, $L_0 = \frac{12}{5}$, $L_2 = \frac{18}{5}$, and $\Omega_0 = B(1,3) \cap B(1, \frac{5}{12}) = B(1, \frac{5}{12})$, so $L_1 = \frac{3}{2}$, and $L_0 < L_2$, $L_1 < L_2$. Set $K = L = L_1$. Then, the old sufficient convergence criterion is not satisfied, since $\alpha\lambda L_2 = \frac{1}{5}\frac{8}{5}\frac{18}{5} = \frac{144}{125} > \frac{1}{2}$ holds. Therefore, there is no guarantee that Newton's method (1.2) converges to x_* under the conditions of the aforementioned references. But our condition hold, since $dba = \frac{1}{5}\frac{8}{5}\frac{3}{2} = \frac{24}{50} < \frac{1}{2}$. Therefore, the conclusions of our Theorem 2.7 follow.

4. Conclusion

The technique of recurrent functions has been utilized to extend the sufficient conditions for convergence of NM for solving nonlinear equations. The new results are finer than the earlier ones. So, they can replace them. No additional conditions have been used. The technique is very general rendering useful to extend the usage of other iterative methods.

Declarations

The authors declare that there are no competing interests and that all authors contributed equally in conceptualization, methodology, formal analysis, and investigation. The original draft was prepared by I. K. Argyros and review and editing was done by S. Regmi, S. George, and M. Argyros.

References

[1] S. Adly, H.V. Ngai, V.V. Nguyen, Newton's methods for solving generalized equations: Kantorovich's and Smale's approaches, J. Math. Anal. Appl. 439 (2016) 396-418. https://doi.org/10.1016/j.jmaa.2016.02.047.

[2] S. Adly, R. Cibulka, H.V. Ngai, Newton's method for solving inclusions using set-valued approximations, SIAM J. Optim. 25 (2015) 159-184. https://doi.org/10.1137/130926730.

[3] I.K. Argyros, Unified convergence criteria for iterative Banach space valued methods with applications, Mathematics, 9 (2021) 1942. https://doi.org/10.3390/math9161942.

[4] I.K. Argyros, S. Hilout, Weaker conditions for the convergence of Newton's method, J. Complex. 28 (2012) 364-387. https://doi.org/10.1016/j.jco.2011.12.003.

[5] I.K. Argyros, S. Hilout, On an improved convergence analysis of Newton's method, Appl. Math. Comp. 225 (2013) 372-386; https://doi.org/10.1016/j.amc.2013.09.049.

[6] I.K. Argyros, A.A. Magréñan, A contemporary study of iterative procedures, Elsevier (Academic Press), New York, 2018. https://doi.org/10.1016/C2015-0-04301-5.

[7] I.K. Argyros, S. George, Mathematical modeling for the solution of equations and systems of equations with applications, Volume-IV, Nova Publisher, NY, 2021.

[8] R. Behl, P. Maroju, E. Martinez, S. Singh, A study of the local convergence of a fifth order iterative procedure, Indian J. Pure Appl. Math. 51 (2020) 439-455. https://doi.org/10.1007/s13226-020-0409-5.

[9] P.G. Ciarlet, C. Madare, On the Newton-Kantorovich theorem, Anal. Appl. 10 (2012) 249-269. https://doi.org/10.1142/S0219530512500121.

[10] R. Cibulka, A.L. Dontchev, J. Preininger, V. Veliov, T. Roubai, Kantorovich-type theorems for generalized equations, J. Convex Anal. 25 (2018) 459-486. http://hdl.handle.net/20.500.12708/144705.

[11] J.A. Ezquerro, M.A. Hernandez, Newton's procedure: An updated approach of Kantorovich's theory, Cham Switzerland, (2018). https://www.booksandcranniesva.com/book/9783319559759.

[12] L.V. Kantorovich, G.P. Akilov, Functional analysis in normed spaces, The Macmillan Co, New York, (1964).

[13] A.A. Magréñan, J.M. Gutiérrez, Real dynamics for damped Newton's procedure applied to cubic polynomials, J. Comp. Appl. Math. 275 (2015) 527–538; https://doi.org/10.1016/j.cam.2013.11.019.

[14] F.A. Potra, V. Pták, Nondiscrete induction and iterative processes, Research Notes in Mathematics 103, Pitman, Boston (1984). https://archive.org/details/nondiscreteinduc0000potr.

[15] P.D. Proinov, New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems, J. Complex. 26 (2010) 3-42. https://doi.org/10.1016/j.jco.2009.05.001.

[16] R. Verma, New trends in fractional programming, Nova Science Publisher, New York, USA, (2019). https://novapublishers.com/shop/new-trends-in-fractional-programming/.