foundations

Article

Extended Newton–Kantorovich Theorem for Solving Nonlinear Equations

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

1 Department of Mathematics, University of Houston, Houston, TX 77204, USA; samundra.regmi@untdallas.edu
2 Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA
3 Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, Karnataka 575025, India; sgeorge@nitk.edu.in
4 Department of Computing and Technology, Cameron University, Lawton, OK 73505, USA; christopher.argyros@cameron.edu
* Correspondence: iargyros@cameron.edu

Abstract: The Newton–Kantorovich theorem for solving Banach space-valued equations is a very important tool in nonlinear functional analysis. Several versions of this theorem have been given by Adley, Argyros, Ciarlet, Ezquerro, Kantorovich, Potra, Proinov, Wang, et al. This result, e.g., establishes the existence and uniqueness of the solution. Moreover, the Newton sequence converges to the solution under certain conditions of the initial data. However, the convergence region in all of these approaches is small in general; the error bounds on the distances involved are pessimistic, and information about the location of the solutions appears improvable. The novelty of our study lies in the fact that, motivated by optimization concerns, we address all of these. In particular, we introduce a technique that extends the convergence region; provides weaker sufficient semi-local convergence criteria; offers tighter error bounds on the distances involved and more precise information on the location of the solution. These advantages are achieved without additional conditions. This technique can be used to extend other iterative methods along the same lines. Numerical experiments illustrate the theoretical results.

Keywords: Traub's method; Banach space; Convergence criterion

MSC: 49M15; 65J15; 65G99

Citation: Regmi, S.; Argyros, I.K.; George, S.; Argyros, C.I. Extended Newton–Kantorovich Theorem for Solving Nonlinear Equations. *Foundations* **2022**, *2*, 504–511. https://doi.org/10.3390/foundations2020033

Academic Editors: Martin Bohner and Dimplekumar N. Chalishajar

Received: 23 March 2022
Accepted: 2 June 2022
Published: 8 June 2022

1. Introduction

A plethora of applications from applied sciences can be converted to solving nonlinear equation

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

where operator $G \in C^1(D, B_2)$, $D \subset B_1$ is an open set and B_1, B_2 are Banach spaces [1]. It is desirable to find the solution $x^* \in D$ of equation $G(x) = 0$ in closed form, but this goal can be achieved only in rare cases. This is the explanation for why most solution techniques for equation $G(x) = 0$ are iterative. These techniques generate a sequence converging to x^* [1–6]. The most widely used method is Newton's (NM) defined by

$$x_0 \in D,\ x_{m+1} = x_m - G'(x_m)^{-1}G(x_m),\ \text{for } \forall m = 0, 1, 2, \cdots. \tag{2}$$

Numerous researchers and practitioners have provided a convergence analysis for NM in Banach space starting from the so-called Newton–Kantorovich theorem [1]. However, the semilocal convergence criterion (see (18)) can easily be violated despite the fact that NM may be convergent to x^* (see also the numerical work). This criterion has mainly been used by other investigators (see Adly et al. [7], Argyros et al. [2–5], Ciarlet et al. [8], Equerro et al. [6], Potra [9], Proinov [10] and the references therein).

In this article, the Kantorovich criterion is weaker, but no new conditions are employed. Moreover, the error distances on $\|x^* - x_n\|$ and the information on the uniqueness ball are improved. This is due to the usage of parameters at least as small as those in the aforementioned works. We extended the recent results in [8], which in turn simplified earlier proofs. Our technique involves the determination of a region D_0 of D, where the iterates x_m also remain. However, in D_0, the Lipschitz constants involved are at least as tight as the corresponding ones depending on D. This modification helps us provide a finer semi-local convergence analysis for NM. The advantages were mentioned in the introduction.

Section 2 includes the semi-local convergence of NM. The examples appear in Section 3. Concluding remarks can be found in Section 4.

2. Convergence Analysis

The notation $U(x_0, \rho)$ is used for the open ball of radius $\rho > 0$ and the center at a point $x_0 \in D$. Moreover, the ball $U[x_0, \rho]$ is the closure of the open ball. Furthermore, $L(B_1, B_2)$ is the space of continuous operators that are linear.

Three crucial types of Lipschitz continuity of G' are introduced, so we can connect them to each other. We suppose from now on that there exists a point $x_0 \in D$ such that $F'(x_0)^{-1} \in L(B_2, B_1)$.

Definition 1. *Operator $H \in C^1(D, B_2)$ is center-Lipschitz continuous on D if there exists parameter $a_0 > 0$ such that*

$$\|H'(x_0)^{-1}(H'(v) - H'(x_0))\| \leq a_0\|v - x_0\| \tag{3}$$

for all $v \in D$.

Set

$$D_0 = U(x_0, \frac{1}{a_0}) \cap D. \tag{4}$$

Definition 2. *Operator $H \in C^1(D, B_2)$ is restricted center-Lipschitz continuous on D_0 if there exists $a > 0$ such that*

$$\|H'(x_0)^{-1}(H'(v_1) - H'(v_2))\| \leq a(\|v_1 - v_2\|) \tag{5}$$

for all $v_1, v_2 \in D_0$.

Definition 3. *Operator $H \in C^1(D, B_2)$ is Lipschitz continuous on D if there exists parameter $a_1 > 0$ such that*

$$\|H'(x_0)^{-1}(H'(w_1) - H'(w_2))\| \leq a_1\|w_1 - w_2\| \tag{6}$$

for all $w_1, w_2 \in D$.

Remark 1. *It follows from these definitions that*

$$a_0 \leq a_1 \tag{7}$$

and

$$a \leq a_1 \tag{8}$$

since

$$D_0 \subseteq D. \tag{9}$$

Suppose that there exists $b > 0$ such that

$$\|G'(x_0)^{-1}\| \leq b. \tag{10}$$

The following estimates were used in the convergence analysis [8] for $G' = H'$ or $G' = H'(x_0)^{-1}H'$:

$$\|G'(z)^{-1}\| \leq \frac{b}{1 - b\bar{a}_1\|z - x_0\|} \text{ for all } z \in U(x_0, \frac{1}{b\bar{a}_1}) \subset D. \tag{11}$$

If

$$\|G'(w_1) - G'(w_2)\| \leq \bar{a}_1\|w_1 - w_2\| \text{ for} \,\forall\, w_1, w_2 \in D \tag{12}$$

or the affine invariant form

$$\|G'(z)^{-1}G'(x_0)\| \leq \frac{1}{1 - a_1\|z - x_0\|} \text{ for } \forall\, z \in U(x_0, \frac{1}{a_1}) \subset D. \tag{13}$$

These estimates are based on (6), but instead we use (3), which is weaker and more precise than (6), to obtain them, instead of using (11) and (13).

$$\|G'(z)^{-1}\| \leq \frac{b}{1 - b\bar{a}_0\|z - x_0\|} \text{ for} \,\forall\, z \in U(x_0, \frac{1}{b\bar{a}_0}) \subset D. \tag{14}$$

If

$$\|G'(w) - G'(x_0)\| \leq \bar{a}_0\|w - x_0\| \text{ for} \,\forall\, w \in D \tag{15}$$

and

$$\|G'(z)^{-1}G'(x_0)\| \leq \frac{1}{1 - a_0\|z - x_0\|} \text{ for} \,\forall\, z \in U(x_0, \frac{1}{a_0}) \subset D, \tag{16}$$

respectively. Notice that parameters a_0, a, and a_1 depend on set D_0, D and operator G as follows: $a_0 = a_0(D, G)$, $a_1 = a_1(D, G)$, whereas $a = a(D_0, G)$. The computation of the Lipschitz constant a_1 induces that of a_0 and a as special cases. Hence, no additional computational effort is required under our technique. Suppose that

$$a_0 \leq a. \tag{17}$$

If $a_0 > a$, then the results that follow hold with a_0 replacing a. Based on the above, we can use the tighter constant a, instead of a_1, in the proofs of previous results. In particular, we present extensions of the Theorems 3–5 in [8], respectively. The proofs are omitted as identical to the corresponding ones in [8] provided that the stated modifications are made.

Theorem 1. *Suppose the following exist:*

(i) *The point $x_0 \in D$ such that $G'(x_0)$ is onto and one-to-one $G'(x_0)^{-1} \in L(B_2, B_1)$ and (2) holds.*

(ii) *Constants $\bar{a}_0, b, d$ such that $0 < q_0 := \bar{a}_0 bd \leq \frac{1}{2}$, $U(x_0, \rho) \subset D$,*

$$\|G'(x_0)^{-1}G(x_0)\| \leq d,$$

(2) and (15) hold and $\rho_0 = \frac{1}{b\bar{a}_0}$. Then, the following assertions hold

(1) *$G'(z) \in L(B_1, B_2)$ is onto and one-to-one $G'(z)^{-1} \in L(B_2, B_1)$ for all $z \in U(x_0, \rho_0)$; sequence $\{x_m\} \subset U(x_0, \rho)$, $\lim_{m \longrightarrow \infty} x_m = x^* \in U[x_0, \rho], G(x^*) = 0$, so that*

$$\|x_m - x^*\| \leq \frac{\rho_0}{2^m}\left(\frac{\rho_-}{\rho_0}\right)^{2^m}, \; if \; q_0 < \frac{1}{2}$$

or

$$\|x_m - x^*\| \leq \frac{\rho_0}{2^m},\ if\ q_0 = \frac{1}{2},$$

where $\rho_- = \frac{1-\sqrt{1-2q_0}}{\bar{a}_0 b} < \rho_0$.

(2) *Additionally, if* $q_0 < \frac{1}{2}$, $U(x_0, \rho_+) \subset D$, *and* (15) *hold, then the only solution where of equation* $G(x) = 0$ *is* $U(x_0, \rho_+)$ *is* $x^* \in U[x_0, \rho_+]$, *where* $\rho_+ = \frac{1+\sqrt{1-2q_0}}{\bar{a}_0 b} < \rho_0$.
If $q_0 = \frac{1}{2}$, *and* $U[x_0, \rho_0] \subset D$, *then the only solution of equation* $G(x) = 0$ *in* $U(x_0, \rho_0)$ *is* $x^* \in U[x_0, \rho_0]$.

(3) *Scalar sequence* $\{s_m\}$ *defined by*

$$s_0 = 0,\ s_{m+1} - s_m = s_m - \frac{p(s_m)}{p'(s_m)} = \frac{\bar{a}_0 b (s_m - s_{m-1})^2}{2(1 - \bar{a}_0 b s_m)}$$

is such that

$$\|x_{m+1} - x_m\| \leq s_{m+1} - s_m$$

$$\|x_{m+1} - x^*\| \leq \rho_- - s_{m+1} = \frac{\bar{a}_0 b (\rho_- - s_m) 62}{2(1 - \bar{a}_0 b s_m)},$$

where $p(t) = \frac{\bar{a}_0 b t^2 - 2t + 2d}{2}$. *Notice that sequence* $\{s_m\}$ *is majorizing for* $\{x_m\}$, *is nondecreasing and* $\lim_{m \longrightarrow \infty} s_m = \rho_-$.

Proof. Simply replace a_1 by $\bar{a}_0$ in the proof of Theorem 3 in [8]. □

Theorem 2. *Suppose the following exist:*

(i) *The point* $x_0 \in D$ *such that* $G \in C^1(D, B_2)$ *is onto and one-to-one,*

(ii) *Constants* d *and* a *such that* $0 < d \leq \frac{a}{2}$, $U(x_0, a) \subset D$,

$$\|G'(x_0)^{-1} G(x_0)\| \leq d,$$

and

$$\|G'(x_0)^{-1}(G'(v) - G'(w))\| \leq \frac{1}{a}\|v - w\|$$

for $\forall v, w \in D_0$.
Then, the following assertions hold

(1) $G'(z) \in L(B_1, B_2)$ *is onto and one-to-one,* $G'(z)^{-1} \in L(B_2, B_1)$ *for all* $z \in U(x_0, a)$; *sequence* $\{x_m\} \subset U(x_0, \rho_-)$, $\lim_{m \longrightarrow \infty} x_m = x^* \in U[x_0, \rho_-], G(x^*) = 0$, *where* $\rho_- = a(1 - \sqrt{1 - \frac{2d}{a}}) \leq a$. *Moreover, the following error bounds hold*

$$\|x_m - x^*\| \leq \frac{a}{2^m}\left(\frac{\rho_-}{\rho_0}\right)^{2^m},\ if\ d < \frac{a}{2}$$

or

$$\|x_m - x^*\| \leq \frac{\rho_0}{2^m},\ if\ d = \frac{a}{2}.$$

(2) *If* $0 < d < \frac{a}{2}$, $U(x_0, \rho_+) \subset D$, *and*

$$\|G'(x_0)^{-1}(G'(v) - G'(w))\| \leq \frac{1}{a}\|v - w\|$$

for all $v, w \in U(x_0, \rho_+)$, where $\rho_+ = a(1 + \sqrt{1 - \frac{2d}{3}})$. Then, $x^* \in U[x_0, \rho_-]$ is the only solution of the equation $G(x) = 0$ in $U(x_0, \rho_+)$. If $d = \frac{a}{2}$ and $U[x_0, a] \subset D$, then $x^* \in U[x_0, a]$ is the only solution of the equation $G(x) = 0$ in $U[x_0, a]$.

Proof. Simply use a for $\frac{1}{r}$ in the proof of Theorem 4 in [8]. □

Theorem 3. *Suppose the following exist:*

(i) The point $x_0 \in D$ such that $G'(x_0) \in L(B_1, B_2)$ is onto and one-to-one,

(ii) Constants $a_1 > 0$ such that $U[x_0, a_1] \subset D$,

$$\|G'(x_0)^{-1}G(x_0)\| \leq \frac{a_1}{2},$$

and

$$\|G'(x_0)^{-1}(G'(v) - G'(w))\| \leq \frac{1}{a_1}\|v - w\|$$

for all $v, w \in U(x_0, a_1)$.

Then, $G'(z) \in L(B_1, B_2)$ is onto and one-to-one, $G'(z)^{-1} \in L(B_2, B_1)$ for all $z \in U(x_0, a_1)$; sequence $\{x_m\} \subset U(x_0, a_1)$, $\lim_{m \longrightarrow \infty} x_m = x^* \in U[x_0, a_1], G(x^*) = 0$. Moreover, the following error bounds hold

$$\|x_m - x^*\| \leq \frac{a_1}{2^m}.$$

Furthermore, the point $x^* \in U[x_0, a_1]$ is the only solution of equation $G(x) = 0$ in $U[x_0, a_1]$.

Proof. Simply use a_1 for $\frac{1}{r}$ in the proof of Theorem 5 in [8]. □

Remark 2. *(a) Our results clearly reduce to the corresponding results in [8] if all constants a are equal to each other.*

Concerning the comparison to the earlier works in [1,6,7,9–11], it is noted that the following Kantorovich criterion is used

$$da_1 \leq \frac{1}{2}. \tag{18}$$

However, in the present study,

$$da \leq \frac{1}{2} \tag{19}$$

is used instead and

$$da_1 \leq \frac{1}{2} \Rightarrow da \leq \frac{1}{2}. \tag{20}$$

Hence, the Kantorovich criterion is weakened. It is important to notice that no new conditions are used, and parameter a is a specialization of a_1. Moreover, the computation of parameter a_1 requires that of parameter a. (b) Based on Remark 1, the advantages of our technique have been justified (see also the numerical work).

Finally, the existence of the solution x^* ball can be further extended, though not necessarily under all the conditions of these three results. This is shown next for Theorem 1. The same can be done for the other two results.

Proposition 1. *Suppose*

(i) *There exists a point $h^* \in U(x_0, r_0) \subseteq D$ for some $r_0 > 0$ solving equation $G(x) = 0$ which is simple.*

(ii) *There exist $x_0 \in D$, $b > 0$, $\bar{a}_0 > 0$ such that $G'(x_0)^{-1} \in L(B_2, B_1)$, $\|G'(x_0)^{-1}\| \leq b$ and*

$$\|G'(u) - G'(x_0)\| \leq \bar{a}_0 \|u - x_0\|$$

for $\forall u \in D$.

(iii) *There exists $r \geq r_0$ such that*

$$\frac{b\bar{a}_0}{2}(r_0 + r) < 1.$$

Set $\Omega = U[x_0, r] \cap D$. Then, the only solution of equation $G(x) = 0$ in the region Ω is h^*.

Proof. Let $\bar{h} \in \Omega$ be a solution of equation $G(x) = 0$. Define $Q = \int_0^1 G'(h^* + \theta(\bar{h} - h^*))d\theta$. By employing (ii),

$$\begin{aligned} \|G'(x_0)^{-1}\|\|G'(x_0) - Q\| &\leq b\bar{a}_0 \int_0^1 ((1-\theta)\|\bar{h} - x_0\| + \theta\|h^* - x_0\|)d\theta \\ &\leq \frac{\bar{a}_0 b}{2}(r_0 + r) < 1. \end{aligned}$$

Then, the estimate, $\bar{h} = h^*$ is implied as a consequence of the Banach result on the inverse of linear operators and $Q(\bar{h} - h^*) = Q(\bar{h}) - Q(h^*) = 0$. □

The set Ω is clearly larger than the ones given in all three preceding results, since in the former the uniqueness depends on a or a_1.

Remark 3. *The developed technique can be applied to extend the applicability of NM under weaker Lipschitz conditions inaugurated by Wang Xinghva along the same lines* [12,13].

3. Numerical Experiments

Two examples show that new criteria can be verified to solve equations, but the ones in [8] (or [1–7,9–11,14,15]) cannot.

Example 1. *Consider $B_1 = B_2 = C[0,1]$. The max-norm is used. Set $D = U(x_0, 3)$. Let operator G on D be given by*

$$G(\beta)(\gamma) = \beta(\gamma) - \alpha(\gamma) - \int_0^1 T(\gamma, s)\delta^3(s)ds \tag{21}$$

$\gamma \in [0,1]$, $\beta \in C[0,1]$, where $\alpha \in C[0,1]$ is given and T is the Green's kernel function

$$T(\gamma, \zeta) = \begin{cases} (1-\gamma)\zeta, & \text{if } \zeta \leq \gamma \\ \gamma(1-\zeta), & \text{if } \gamma \leq \zeta \end{cases}. \tag{22}$$

Definition (21) and (22) are used to show that the derivative G', according to Fréchet, is given by

$$[G'(\delta)(\beta)](\gamma) = \beta(\gamma) - 3\int_0^1 T(\gamma, \zeta)\delta^2(s)\alpha(s)ds, \tag{23}$$

$\gamma \in [0,1]$, $\beta \in C[0,1]$. Let $\alpha(\gamma) = x_0(\gamma) = 1$. Then, by (21)–(23), we obtain $G'(x_0)^{-1} \in L(B_2, B_1)$,

$$\|I - G'(x_0)\| < \frac{3}{8}, \ \|G'(x_0)^{-1}\| \leq \frac{8}{5}, \ d = \frac{1}{5}$$

$\bar{a}_0 = \frac{12}{5}$, $a_1 = \frac{18}{5}$, and

$$D_0 = U(1,3) \cap U(1, \frac{5}{12}) = U(1, \frac{5}{12}).$$

Consequently, it follows $a = \frac{3}{2}$, so $a_0 < a_1$ and $a < a_1$. Then, the sufficient convergence criterion (18) is not satisfied, since

$$da_1 = 1.152 > \frac{1}{2}$$

holds. It follows that no assurance for convergence can be provided by these results in the aforementioned references. However, our criterion (19) holds, since

$$da = 0.48 < \frac{1}{2}.$$

Example 2. *Define scalar function*

$$G(t) = v_0 t + v_1 + v_2 \sin e^{v_3 t},$$

for $t_0 = 0$, where v_j, $j = 0, 1, 2, 3$ are real parameters. It follows by this definition that if v_3 is large enough and v_2 is small enough, the fraction $\frac{a_0}{a_1}$ is small (enough), i.e., $\frac{a_0}{a_1} \longrightarrow 0$.

4. Conclusions

A technique involving recurrent functions and restricted convergence domains was applied to extend the application of the Newton–Kantorovich theorem to solving nonlinear equations. The new results are better than earlier ones. Hence, they can replace them. No additional conditions are needed. The technique is a very general rendering that is useful to extend the usage of other iterative methods, such as the Secant, Stirling's, Kurchatov, Newton-type and others [4,6,11].

Author Contributions: Conceptualization, S.R., C.I.A., I.K.A. and S.G.; methodology, S.R., C.I.A., I.K.A. and S.G.; software, S.R., C.I.A., I.K.A. and S.G.; validation, S.R., C.I.A., I.K.A. and S.G.; formal analysis, S.R., C.I.A., I.K.A. and S.G.; investigation, S.R., C.I.A., I.K.A. and S.G.; resources, S.R., C.I.A., I.K.A. and S.G.; data curation, S.R., C.I.A., I.K.A. and S.G.; writing—original draft preparation, S.R., C.I.A., I.K.A. and S.G.; writing—review and editing, S.R., C.I.A., I.K.A. and S.G.; visualization, S.R., C.I.A., I.K.A. and S.G.; supervision, S.R., C.I.A., I.K.A. and S.G.; project administration, S.R., C.I.A., I.K.A. and S.G.; funding acquisition, S.R., C.I.A., I.K.A. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kantorovich, L.V.; Akilov, G.P. *Functional Analysis in Normed Spaces*; The Macmillan Co., Ltd.: New York, NY, USA, 1964.
2. Argyros, I.K. Unified Convergence Criteria for Iterative Banach Space Valued Methods with Applications. *Mathematics* **2021**, *9*, 1942. [CrossRef]
3. Argyros, I.K. *The Theory and Applications of Iteration Methods*, 2nd ed.; Engineering Series; Taylor and Francis Group, CRC Press: Boca Raton, FL, USA, 2022.
4. Argyros, I.K.; Magréñan, A.A. *A Contemporary Study of Iterative Procedures*; Elsevier: New York, NY, USA; Academic Press: Cambridge, MA, USA, 2018.
5. Argyros, I.K.; George, S. *Mathematical Modeling for the Solution of Equations and Systems of Equations with Applications*; Nova Publisher: Hauppauge, NY, USA, 2021; Volume-IV.
6. Ezquerro, J.A.; Hernez, M.A. *Newton's Procedure: An Updated Approach of Kantorovich's Theory*; Springer: Cham, Switzerland, 2018.

7. Adly, S.; Ngai, H.V.; Nguyen, V.V. Newton's procedure for solving generalized equations: Kantorovich's and Smale's approaches. *J. Math. Anal. Appl.* **2016**, *439*, 396–418. [CrossRef]
8. Ciarlet, P.G.; Madare, C. On the Newton-Kantorovich theorem. *Anal. Appl.* **2012**, *10*, 249–269. [CrossRef]
9. Potra, F.A.; Pták, V. *Nondiscrete Induction and Iterative Processes*; Research Notes in Mathematics 103; Pitman Advanced Publishing Program: Boston, MA, USA, 1984.
10. Proinov, P.D. New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems. *J. Complex.* **2010**, *26*, 3–42. [CrossRef]
11. Magréñan, A.A.; Gutiérrez, J.M. Real dynamics for damped Newton's procedure applied to cubic polynomials. *J. Comput. Appl. Math.* **2015**, *275*, 527–538. [CrossRef]
12. Xinghua, W. Convergence of Newton's method and uniqueness of the solution of equations in Banach space. *IMA J. Numer. Anal.* **2000**, *20*, 123134.
13. Xinghua, W. Convergence of Newton's method and inverse function theorem in Banach space. *Math. Comput.* **1999**, *68*, 169–186. [CrossRef]
14. Behl, R.; Maroju, P.; Martinez, E.; Singh, S. A study of the local convergence of a fifth order iterative procedure. *Indian J. Pure Appl. Math.* **2020**, *51*, 439–455.
15. Verma, R. *New Trends in Fractional Programming*; Nova Science Publisher: New York, NY, USA, 2019.