Advances and Applications in Mathematical Sciences
Volume 22, Issue 9, July 2023, Pages 1977-1989

EXTENDED NEWTON'S METHOD ON BANACH SPACE WITH A CONVERGENCE STRUCTURE

IOANNIS K. ARGYROS[1*], JINNY ANN JOHN[2], JAYAKUMAR JAYARAMAN[2] and SAMUNDRA REGMI[3]

[1]Department of Computing
and Mathematical Sciences
Cameron University, Lawton, 73505, OK, USA

[2]Department of Mathematics
Puducherry Technological University
Pondicherry, India 605014
E-mail: jinny3@pec.edu
jjayakumar@ptuniv.edu.in

[3]Department of Mathematics
University of Houston
Houston, TX, USA
E-mail: sregmi5@uh.edu

Abstract

A unifying semi-local convergence theory for Newton's method is extended on Banach space with a convergence structure. The new theory is obtained without additional conditions and is leading to an enlarged convergence domain, improved estimates on the error distances as well as more precise knowledge of the whereabouts of the solution of the equation about a neighbourhood of the initial point. An application of the theory is included along with an example.

1. Introduction

Let Q be an operator defined on a Banach space B with values in B. Let us also define the operator $F : B \to B$ by

2020 Mathematics Subject Classification: 65J15, 49H17.
Keywords: Banach space, convergence structure, Newton's method, semi-local convergence.
*Corresponding author; E-mail: iargyros@cameron.edu
Received July 12, 2022; Accepted December 14, 2022

$$F(x) = MQ(x_0 + x), \tag{1.1}$$

where x_0 is an initial point. We shall determine a solution x^* of the equation

$$F(x) = 0, \tag{1.2}$$

where M is an approximation to $Q'(x_0)^{-1}$.

Next, in order to make the article as self contained as possible we reintroduce some concepts that can also be found with more details in [12]. Motivated by optimization considerations we present a new semi-local convergence analysis for Newton's method with the following benefits:

(1) Extended convergence domain.

(2) Tighter error estimates on the distances involved.

(3) A more precise information on the uniqueness ball containing the solution x^*.

This is obtained using more precise majorants than in [12] and under the same computational cost, since the new functions are tighter and special cases of the ones in [12]. Relevant research on a real Banach space can be found in [1-11, 13-23].

2. Preliminaries

Definition 2.1. A Banach space B with Convergence Structure (CS) is a triplet (B, U, W) with the following properties:

(i) The pair $(B, \|\cdot\|)$ is a Banach space (real).

(ii) The triplet $(U, K, \|\cdot\|_U)$ is a real Banach space which is ordered by the convex and closed cone K. Moreover, the norm $\|\cdot\|_U$ is assumed to be monotone on K.

(iii) W is a convex closed cone in $B \times U$ such that $\{0\} \times K \subset W \subset B \times K$.

(iv) The mapping $|\cdot| : H \to K$ is well defined for each $x \in H = \{x \in B \mid \exists v \in K : (x, v) \in W\}$ by

$$|x| = \inf\{v \in K \mid (x, v) \in W\}.$$

(v) For each $x \in H : \| x \| \leq \| \, | x | \, \|_U$.

Recall [12] that the set H satisfies $H + H \subset H$ and for each $p > 0 : pH \subset H$. Moreover, the set

$$S(\alpha) = \{x \in B \mid (x, \alpha) \in W\}$$

defines a "generalized" neighbourhood of zero.

Let $B = \mathbb{R}^j$ be equipped with the maximum norm. Then, the following examples constitute the motivation for the Definition 2.1

(a) $U = \mathbb{R}, W = \{(x, v) \in \mathbb{R}^j \times \mathbb{R}, \| x \|_\infty \leq v\}$.

(b) $U = \mathbb{R}^j, W = \{(x, v) \in \mathbb{R}^j \times \mathbb{R}^j, | x | \leq v\}$, the component wise absolute value.

(c) $U = \mathbb{R}^j, W = \{(x, v) \in \mathbb{R}^j \times \mathbb{R}^j, | x | \leq v\}$.

Case (a) is the classical convergence analysis in Banach space. Case (b) is used for component wise analysis and error estimates. Case (c) is used for the monotone convergence analysis.

We consider monotonicity in the space $B \times U$. Let $(x_m, v_m) \in W^{\mathbb{R}}$ be an increasing sequence.

Then, we have

$$(x_m, v_m) \leq (x_{m+i}, v_{m+i}) \Rightarrow 0 \leq (x_{m+i} - x_m, v_{m+i} - v_m).$$

If $v_n \to v$, then we get $0 \leq (x_{m+i} - x_m, v - v_m)$. Consequently, by the condition (v) of Definition 2.1

$$\| x_{m+i} - x_m \| \leq \| v - v_m \|_U \to 0 \text{ as } n \to +\infty.$$

Thus, the sequence $\{x_m\}$ is complete. We use the sequence $v_m = a_0 - a_m$, where the sequence $\{a_m\} \in K^{\mathbb{R}}$ is decreasing to find the estimate

$$0 \leq (x_{m+i} - x_m, a_m - a_{m+i}) \leq (x_{m+i} - x_m, a_m).$$

Then, if $x_m \to x^*$, this leads to the estimate $| x^* - x_m | = a_m$.

The notation $\mathscr{L}(B^m)$ stands for the space of multilinear, symmetric and bounded operators $M : B^m \to B$ on the Banach space B. Then, for an ordered Banach space U, we define

$$L_+(U^m) = \{L \in \mathscr{L}(U^m) \mid 0 \leq x_i \Rightarrow 0 \leq L(x_1, x_2, ..., x_m)\}.$$

A mapping $L \in C^1(U_M \to U)$ on an open subset U_M of an ordered Banach space U is defined to be order convex on the interval $[s_1, s_2] \subset U_M$ if for $s_3, s_4 \in [s_1, s_2]$

$$s_3 \leq s_4 \Rightarrow L'(s_4) - L'(s_3) \in L_+(U).$$

Definition 2.2. The set of bounds for an operator $M \in \mathscr{L}(B^m)$ is defined by

$$B(M) = \{L \in M_+(U^m) \mid (x_i, v_i) \in H \Rightarrow [M(x_1, x_2, \dots, x_m),$$

$$L(v_1, v_2, \dots, v_m)] \in W\}.$$

The following auxiliary result is needed.

Lemma 2.3. *Let* $M[0, 1] \to \mathscr{L}(B^m)$ *and* $L : [0, 1] \to L_+(U^m)$ *be continuous mappings*. Then, for each $t \in [0, 1]$

$$L(t) \in B(M(t)) \Rightarrow \int_0^1 L(t)dt \in B(\int_0^1 M(t)dt).$$

This result will be used for the remainder of Taylor's Theorem. Let $P : B_1 \to B_1$ be a mapping on a subset B_1 of a normed space. Then, $P^m(x)$ stands for the m-fold application of P. In case of convergence

$$P^\infty(x) = \lim_{m \to \infty} P^m(x).$$

The definition of the right inverse follows:

Definition 2.4. Let $M \in \mathscr{L}(B)$ and $w \in B$ be an arbitrary point. Then, we define

$$M^* w = z \Leftrightarrow z \in P^{\infty}(0),\ P(x) = (I - M)x + w \Leftrightarrow z = \sum_{m=0}^{+\infty} (I - M)^m w$$

provided this limit exists.

Under the preceding notation, the Newton's method for generating a sequence $\{x_m\}$ approximating a solution x^* of equation (1.1) is defined for $x_0 = 0$ and each $m = 0, 1, 2, \ldots$ by

$$x_{m+1} = x_m + F'(x_m)^*(-F(x_m)), \tag{2.1}$$

where $F'(x_m)^*$ follows the notation as in Definition 2.4.

3. Convergence

Some notions of generalized Lipschitz continuity are developed and compared to.

Definition 3.1. The operator L_0 is order-convex and generalized center-Lipschitz on the domain $[0, \alpha^*]$ for some α^* and for each $x \in S(\alpha^*)$ with $|x| \leq \alpha^*$

$$L_0'(|x|) - L_0'(|0|) \in T(F'(0) - F'(x)).$$

Suppose that $L_0(\alpha_0)\alpha_0 \leq I$, for some $\alpha_0 \leq \alpha^*$.

Definition 3.2. The operator L is order-convex and restricted generalized Lipschitz on the domain $[0, \alpha_0]$ and for each $x, y \in S(\alpha_0)$ with $|x| + |y| \leq \alpha_0$

$$L'(|x| + |y|) - L'(|x|) \in T(F'(x) - F'(x + y)).$$

Denote by α the smallest point in $[0, \alpha_0)$ satisfying this condition.

Definition 3.3. The operator L_1 is order convex and generalized

Lipschitz on the domain $[0, \alpha^*]$ and for each $x, y \in S(\alpha^*)$ with $|x| + |y| \leq \alpha^*$

$$L_1'(|x| + |y|) - L_1'(|x|) \in T(F'(x) - F'(x + y)).$$

Denote by α_1 the smallest point in $(0, \alpha^*]$ satisfying this condition.

Remark 3.4. It follows by these definitions that

$$L_0' \leq L_1'$$

$$\text{and } L' \leq L_1',$$

$$\text{since } [0, \alpha_0] \subset [0, \alpha^*] \text{ and } [0, \alpha] \subset [0, \alpha^*].$$

Notice that $L_0 = L_0(\alpha, B, U, E)$, $L_1 = L_1(\alpha, B, U, E)$ but $L = L(\alpha_0, \alpha, B, U, E)$. That is L_0 used to define L. In view of the above, tighter L can replace L_1 in the results in [12].

We assume that

$$L_0' \leq L'.$$

Otherwise, the results that follow hold for L_2 replacing L, where L_2 is the maximum operator on $[0, \alpha]$.

Next, we present the semi-local convergence of the Newton sequence $\{x_n\}$ which extends the corresponding Theorem 5 in [12].

Theorem 3.5. *Assume there exists a Banach space B with $CS(B, U, V)$ with $U = (U, K, \|\cdot\|_U)$, a mapping $F \in C^1(B_F \to B)$ with $B_F \subset B$, a mapping $L_0 \in C^1(U_{L_0} \to U)$ with $U_{L_0} \subset U$, such that there exist α_0, α^* with $\alpha_0 \leq \alpha^*$ such that each $x \in S(\alpha^*)$ with $|x| \leq \alpha^*$*

$$L_0'(|x|) - L_0'(|0|) \in B(F'(0) - F'(x))$$

$$\text{and } L_0'(\alpha_0) \leq I.$$

Moreover, assume

(i) $S(\alpha_0) \subset B_F$ *and* $[0, \alpha_0] \subset U_1$.

(ii) *The operator* L *is order convex on* $[0, \alpha_0]$ *and satisfies for* $x, y \in S(\alpha_0)$ *with*

$$| x | + | y | \leq \alpha_0$$

$$L'(| x | + | y |) - L'(| x |) \in B(F'(x) - F'(x + y)).$$

(iii) $L'(0) \in B(I - F'(0))$ *and* $(-F(0), L(0)) \in W$.

(iv) $L(\alpha_0) \leq \alpha_0$.

(v) $L'(\alpha_0)^m \alpha_0 \to 0$ *as* $n \to +\infty$.

Then, the following assertions hold:

(1) *The Newton sequence* $\{x_m\}$ *is well defined and converges to the unique solution* $x^* \in S(\alpha_0)$ *of the equation* $F(x) = 0$.

(2) *Define the sequence* $\{d_m\}$ *for* $d_0 = 0$ *and each* $n = 0, 1, 2, \ldots$ *by*

$$c_m = \| x_{m+1} - x_m \|,$$

$$d_{m+1} = L(d_m) + L'(| x_m |)c_m.$$

Then, for $b = L^{\infty}(0)$ *being the smallest fixed point of* L *in* $[0, \alpha_0]$, *the sequence* $(x_m, d_m) \in (B \times U)^R$ *is well defined, belongs in* W^R, *is monotone and satisfies* $d_m \leq b$.

(3) *If* $S_m(W) = L(| x_m | + v) - L(| x_m |) - L'(| x_m |)v$, *define*

$$R_m(v) = (I - L'(| x_m |))^* S_m(v) + c_m.$$

Then, if S_m *is monotone on the interval* $I_m = [0, \alpha_0 - | x_m |]$ *and there exists* $v_m \in K$ *with* $| x_m | + v_m \leq \alpha_0$ *and*

$$S_m(v_m) + L'(| x_m |)(v_m - c_m) \leq v_m - c_m,$$

then $R_m : [0, v_m] \to [0, v_m]$ *is well defined and monotone. A choice for* v_m *is* $\alpha_0 - d_n$.

(4) *If* $v \in I_m$ *satisfies* $R_m(v) \leq v$, *then*

$$c_m \leq R_m(v) = g \leq v \text{ and } R_{m+1}(g - c_m) \leq g - c_m.$$

(5) *If any solution* $v \in I_m$ *of* $R_m(v) = v$, *then*

$$| x^* - x_m | \leq R_m^{\infty}(0) \leq v.$$

Proof. It follows from the proof of Theorem 5 in [12], with $[0, \alpha_0]$, L, replacing $[0, \alpha_0]$, L_1 respectively.

Notice that L_0 is only used to define L.

Remark 3.6. If $L_1 = L$, then the results of Theorem 3.5 reduce to the ones in Theorem 5 in [12]. Otherwise, the new results extend the ones in [12] with advantages already stated in the introduction.

Theorem 3.7. *Assume the conditions* (i)-(iii) *of Theorem* 3.5 *and one more condition* (iv). *There exists* $p \in (0, 1)$ *such that* $L(x_0) \leq p\alpha_0$.

Then, *there exists* $\overline{\alpha} \in [0, p\alpha_0]$ *satisfying the conditions of Theorem* 3.5. *Moreover*, *the solution* $x^* \in S(\overline{\alpha})$ *of the equation* $F(x) = 0$ *is unique in* $S(\alpha_0)$.

Remarks similar to Remark 3.6 can follow for the Theorem 3.7. This theorem extends Theorem 6 in [12]. The monotonicity case given in Theorem 13 in [12] can also be immediately extended if $B = U$, $H = C^2$ and $|\cdot| = I$.

Next, applications of the theory follow.

4. Applications

Case 1. The Banach space with a real norm $\| \cdot \|$.

We assume that the derivative F' exists and $F'(0) = I$, the identity operator on B. Moreover, assume there exists a monotone mapping $h_0 : [0, \alpha^*] \to \mathbb{R}$ such that for each $x \in S(\alpha^*)$

$$\| F'(x) - F'(0) \| \leq h_0(\| x \|) \| x \|. \tag{4.1}$$

Then, define the operator L_0 such that

$$L_0'(v) = \int_0^v h_0(t)dt. \tag{4.2}$$

In the interesting case, when the monotone mapping h_0 is constant, then

$$L_0'(v) = h_0 v. \tag{4.3}$$

It follows by $h_0(t) \leq h_0(\alpha)$ that the condition

$$L_0'(v) < I \tag{4.4}$$

holds if

$$\alpha_0 = \frac{1}{h_0} < \alpha^*. \tag{4.5}$$

Assume that there exists a monotone mapping $h : [0, \alpha_0] \to \mathbb{R}$ such that for each $x \in S(\alpha_0)$

$$\| F''(x) \| \leq h(\| x \|). \tag{4.6}$$

Define the operator L on the interval $[0, \alpha_0]$

$$L(v) = \| F(0) \| + \int_0^v ds \int_0^s h(t)dt. \tag{4.7}$$

In the case when the mapping h is a constant then for $h(t) \leq h(\alpha)$, the condition (4.7) holds if

$$\frac{1}{2} h^2 \alpha^2 + \| F(0) \| \leq \alpha. \tag{4.8}$$

This can happen if

$$q = 2\| F(0) \| h \leq 1 \tag{4.9}$$

and

$$\alpha \leq \alpha^*. \tag{4.10}$$

Assume that there exists a monotone mapping $h_1 : [0, \alpha^*] \to \mathbb{R}$ such that for each $x \in S(\alpha^*)$

$$\| F''(x) \| \leq h(\| x \|). \tag{4.11}$$

Define the operator L_1 on the interval $[0, \alpha^*]$ by

$$L_1(v) = \| F(0) \| + \int_0^v ds \int_0^1 h_1(t) dt. \tag{4.12}$$

This time we have that if the mapping h_1 is constant

$$L_1(v) = \frac{1}{2} h_1^2 v^2 + \| F(0) \|. \tag{4.13}$$

Then, the corresponding Kantorovich type condition [1-3, 8] is

$$q_1 = 2 \| F(0) \| h_1 \leq 1 \tag{4.14}$$

and

$$\alpha_1 \leq \alpha^*. \tag{4.15}$$

The equation (4.14) is the celebrated Kantorovich convergence condition for Newton's method [[8], Ch.18].

Notice that

$$h_0 \leq h_1 \tag{4.16}$$

and

$$h \leq h_1 \tag{4.17}$$

thus

$$h_1 \leq 1 \Rightarrow h = 1 \tag{4.18}$$

but not necessarily vice versa unless if $h = h_1$.

Moreover, α and α_1 can be given in explicit form as

$$\alpha = \frac{2 \| F(0) \|}{1 + \sqrt{1 - q}} \tag{4.19}$$

and

$$\alpha_1 = \frac{2 \| F(0) \|}{1 + \sqrt{1 - q_1}} \tag{4.20}$$

provided that (4.9) and (4.14) hold. Then, in view of (4.17), (4.19) and (4.20), we deduce that

$$\alpha \leq \alpha_1. \tag{4.21}$$

Hence, the convergence criterion given in Theorem 5 in [12] is weakened without additional conditions, since the computation of L_1 requires that of L_0 and L as special cases.

In the next example, we find explicitly the mappings and parameters introduced in Case 1.

Example 4.1. Let $B = \mathbb{R}^3$ and $\alpha^* = 1$. Then, for $w = (w_1, w_2, w_3)^{tr}$ and some $\gamma \in \mathbb{R}$ define the operator F on B by

$$F(w) = (e^{w_1} + \gamma - 1 + w_2 + w_3, \frac{e-1}{2} w_2^2 + w_2 + w_3, w_2 + w_3). \tag{4.22}$$

Then, the derivatives F' and F'' are given by

$$F'(w) = \begin{bmatrix} e^{w_1} & 0 & 0 \\ 0 & (e-1)w_2 + 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.23}$$

and

$$F''(w) = \left[\begin{array}{ccc|ccc|ccc} e^{w_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & (e-1) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{4.24}$$

It follows from (4.22) and (4.23) that $F(0) = (\gamma, 0, 0)^{tr}$ and $F'(0) = I$. Then, we get

$$h_0(v) = (e-1)v,$$

$$\alpha_0 = \frac{1}{e-1},$$

$$h(v) = ve^{\frac{1}{\alpha_0}} = ve^{\frac{1}{e-1}},$$

$$h_1(v) = ve,$$

$$q = 2|\gamma|e^{\alpha_0},$$

$$q_1 = 2|\gamma|e,$$

$$\alpha = \frac{1 - \sqrt{1 - 2|\gamma| e^{\alpha_0}}}{e^{\alpha_0}}$$

$$\text{and } \alpha_1 = \frac{1 - \sqrt{1 - 2|\gamma| e}}{e}.$$

Specifying $\gamma = \frac{3}{8(e-1)}$, we get

$$q = 0.58390407 < 1$$

but

$$q_1 = 1.18648253 > 1.$$

Hence, the previous Kantorovich conditions (4.14) in [12] is not fulfilled. Hence, there is no guarantee that the Newton's method converges to the solution. But the new condition (4.9) is fulfilled.

Notice also that

$$\alpha = 0.265330058 < \alpha^*.$$

Therefore, Newton's method converges to the solution $x^* = (0.1974082327, 0, 0)^{tr}$.

References

[1] I. K. Argyros, Unified convergence criteria for iterative Banach space valued methods with applications, Mathematics 9(16) (2021), 1942.

[2] I. K. Argyros, The Theory and Applications of Iteration Methods, 2nd edition, CRC Press/Taylor and Francis Publishing Group Inc., Boca Raton, Florida, USA, 2022.

[3] I. K. Argyros and S. Hilout, Weaker conditions for the convergence of Newton's method, Journal of Complexity 28(3) (2012), 364-387.

[4] X. Chen and T. Yamamoto, Convergence domains of certain iterative methods for solving nonlinear equations, Numerical Functional Analysis and Optimization 10(1-2) (1989), 37-48.

[5] E. Catinas, The inexact, inexact perturbed, and quasi-Newton methods are equivalent models, Mathematics of Computation 74(249) (2005), 291-301.

[6] P. Deuflhard, Newton Methods for Nonlinear Problems, Affine Invariance and Adaptive Algorithms, In Springer Series in Computational Mathematics, Springer, 2004.

[7] J. Ezquerro, J. Gutierrez, M. Hernandez, N. Romero and M. Rubio, The Newton method: from Newton to Kantorovich, Gac. R. Soc. Mat. Esp 13 (1) (2010), 53-76.

[8] L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, Pergamon Press, Oxford, 1964.

[9] M. A. Krasnoselskii, G. M. Vainikko, R. Zabreyko, Y. B. Ruticki and V. V. Stetsenko, Approximate Solution of Operator Equations, Springer Science and Business Media, 2012.

[10] P. Meyer, Newton's method in generalized Banach spaces, Numerical Functional Analysis and Optimization 9(3-4) (1987), 249-259.

[11] P. W. Meyer Das, Modifizierte Newton-Verfahren in verallgemeinerten Banach-Räumen, Numerische Mathematik 43(1) (1984), 91-104.

[12] P. W. Meyer, A unifying theorem on Newton's method, Numerical Functional Analysis and Optimization 13(5-6) (1992), 463-473.

[13] J. Ortega and W. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, SIAM, vol. 30 1970.

[14] F. A. Potra and V. Ptak, Sharp error bounds for Newton's process, Numerische Mathematik 34(1) (1980), 63-72.

[15] F.-A. Potra and V. Ptak, Nondiscrete induction and iterative processes, Pitman Advanced Publishing Program, vol. 10, 1984.

[16] P. D. Proinov, General local convergence theory for a class of iterative processes and its applications to Newton's method, Journal of Complexity 25(1) (2009), 38-62.

[17] P. D. Proinov, New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems, Journal of Complexity 26(1) (2010), 3-42.

[18] V. Ptak, Concerning the rate of convergence of Newton's process, Commentationes Mathematicae Universitatis Carolinae 16(4) (1975), 699-705.

[19] V. Ptak, The rate of convergence of Newton's process, Numerische Mathematik 25(3) (1975), 279-285.

[20] J. S. Vandergraft, Newton's method for convex operators in partially ordered spaces, SIAM Journal on Numerical Analysis 4(3) (1967), 406-432.

[21] T. Yamamoto, A unified derivation of several error bounds for Newton's process, Journal of Computational and Applied Mathematics 12 (1985), 179-191.

[22] T. Yamamoto, A convergence theorem for Newton-like methods in Banach spaces, Numerische Mathematik 51(5) (1987), 545-557.

[23] P. Zabrejko and P. Zlepko, On a generalization of the Newton-Kantorovich method for an equation with nondifferentiable operator, Ukr. Mat. Zhurn 34 (1982), 365-369.