Electronic Journal: Southwest Journal of Pure and Applied Mathematics
Internet: http://rattler.cameron.edu/swjpam.html
ISSN 1083-0464
Vol. No. 01, 1995, pp. 30-36.
Submitted: May 12, 1995. Published: October 16, 1995.

RESULTS ON CONTROLLING THE RESIDUALS OF PERTURBED NEWTON-LIKE METHODS ON BANACH SPACES WITH A CONVERGENCE STRUCTURE

Ioannis K. Argyros

Abstract. In this study, we provide sufficient conditions for controlling the residuals of perturbed Newton-like methods, in such a way that the convergence of such methods is ensured. In order to achieve this, we apply a theorem due to Kantorovich on a setting of a regular space with a convergence structure.

(1991) A.M.S. (MOS) Subject Classification Codes. 65J15, 65B05, 47D15, 47H17, 49D15.

Key Words and Phrases. Newton-like methods, residuals.

I. INTRODUCTION

In this study, we are concerned with approximating a solution x^* of the nonlinear operator equation

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

where F is a Frechet-differentiable operator defined on a convex subset D of a Banach space X with values in X, and Q is a non-differentiable nonlinear operator with the same domain and values in X.

We generate a sequence $\{x_n\}$ $(n \geq 0)$ using the perturbed Newton-like method scheme given by

$$x_{n+1} = x_n + \delta_n \quad (n \geq 0) \tag{2}$$

where the correction δ_n satisfies

$$A(x_n)\delta_n = -(F(x_n) + Q(x_n)) + r_n \quad (n \geq 0) \tag{3}$$

for a suitable $r_n \in X$ $(n \geq 0)$ called residual. The importance of studying perturbed Newton-like methods comes from the fact that variants of Newton's method can be considered as procedures of this type. In fact, the approximations (2) and (3) characterize any iterative process in which the corrections are taken as approximate solutions of the Newton equations. This happens when approximation (2) is solved by any iterative method or when therein the derivative is replaced by a suitable approximation. In [4] we provided sufficient conditions for the convergence of iteration (2) to a solution x^* of equation (1), by assuming that X is a Banach space with a convergence structure (to be precised later).

Here we derive sufficient conditions for controlling the residuals r_n in such a way that the convergence of the sequence $\{x_n\}$ $(n \geq 0)$ to a solution of equation (1) is ensured.

We also refer the reader to [5], [6], [8] and the references there for relevant work, which however is valid on a Banach space X without a convergence structure. The advantages of working on a Banach space with a convergence structure have been explained in some detail in [4], [6], [7].

Department of Mathematics, Cameron University, Lawton, OK 73505, U.S.A.
E-mail Address: ioannisa@cua.cameron.edu

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

II. PRELIMINARIES

We will need the definitions:

Definition 1. The triple (X, V, E) is a Banach space with a convergence structure if

(C_1) $X, \|\cdot\|)$ is a real Banach space;

(C_2) $(V, C, \| \|_v)$ is a real Banach space which is partially ordered by the closed convex cone C; the norm $\| \|_v$ is assumed to be monotone on C;

(C_3) E is a closed convex cone in $X \times V$ satisfying $\{0\} \times C \subseteq E \subseteq X \times C$;

(C_4) the operator $| \,|: D \to C$ is well defined:

$$|x| = \inf\{q \in C | (x, q) \in E\}$$

for

$$x \in D = \{x \in X | \exists q \in C : (x, q) \in E\};$$

and

(C_5) for all $x \in D$ $\|x\| \le \||x|\|_v$.

The set

$$U(a) = \{x \in X | (x, a) \in E\}$$

defines a sort of generalized neighborhood of zero.

Definition 2. An operator $L \in C^1(V_1 \to V)$ defined on an open subset V_1 of an ordered Banach space V is order convex on $[a, b] \subseteq V_1$ if

$$c, d \in [a, b], \; c \le d \Rightarrow L'(d) - L'(c) \in L_+(V),$$

where for $m \ge 0$

$$L_+(V^m) = \{L \in L(V^m) | 0 \le x_i \Rightarrow 0 \le L(x_1, x_2, \dots, x_m)\}$$

and $L(V^m)$ denotes the space of m-linear, symmetric, bounded operators on V.

Definition 3. The set of bounds for an operator $H \in L(X^m)$ is defined to be

$$B(H) = \{L \in L_+(V^m) | (x_i, q_i) \in E \Rightarrow (H(x_1, \dots, x_m), L(q_1, \dots, q_m)) \in E\}.$$

Definition 4. A partially ordered topological space V is called regular if every order bounded increasing sequence has a limit.

We will need the following proposition due to Kantorovich [3], [5].

Proposition. Let V be a regular partially ordered topological space (POTL-space) and let x, y be two points of V such that $x \le y$. If $H : [x, y] \to V$ is a continuous isotone operator having the property that $x \le Hx$ and $y \ge Hy$, then there exists a point $z \in [x, y]$ such that $z = Hz$.

III. CONVERGENCE ANALYSIS

We will need the following basic result:

Lemma. Let V be a regular partially ordered topological space, an operator $L \in C^1(V_1 \to V)$ with $[0, a] \subseteq V_1 \subseteq V$ for some $a \in V$, an operator $M \in C(V_1 \to V)$, operators $R < I$, $B > I$, T, $K \in L_+(V)$, a point $c \in V$ with $c > 0$ and a point $p \in [0, a]$.

Assume:

(a) The equation

$$\begin{aligned} g(q) = BT\big[&L(p+q) - L(p) - L'(p)q + M(p+q) - M(p) + K(p+q) - K(p)\big] \\ &- (I - R)q + c = 0 \end{aligned} \tag{4}$$

has solutions in the interval $[0, a]$ and denote by q^* the least of them.

(b) Let $G \in L_+(V)$ be given and $R_+ \in L_+(V)$, $c_+, p_+ \in V$ be such that the following conditions are satisfied:

$$R_+ < \min\{(G-2I)BTL'(p) + BTGL'(p_+) + G(R-I) + I, I\} = \alpha \tag{5}$$

$$\begin{aligned} 0 < c_+ \leq\ & BTG(L(p_+) + M(p_+) + K(p_+)) + GBT(L(p) + M(p) + K(p)) \\ & -2BT(L(p) + M(p) + K(p)) + (R_+ + G - I - BTGL'(p_+))c = \beta \end{aligned} \tag{6}$$

and

$$0 \leq p_+ \leq p + c, \tag{7}$$

where α and β are functions of the operators and points involved.

(c) The following estimate is true

$$M(p) \leq M(p+q) \tag{8}$$

for all $p, q \in [0, a]$.

Then the equation

$$\begin{aligned} g_+(q) =\ & BGT\big[L(p_+ + q) - L(p_+) - L'(p_+)q + M(p_+ + q) - M(p_+) + K(p_+ + q) - K(p_+)\big] \\ & - (I - R_+)q + c_+ = 0 \end{aligned} \tag{9}$$

has nonnegative solutions and the least of them, denoted by q_+^* lies in the interval $[c_+, q^* - c]$.

Proof. Using the hypothses $g(q^*) = 0$ and $R < I$ we deduce from (4) that $c \leq q^*$. We will show that

$$g_+(q^* - c) \leq 0. \tag{10}$$

From equation (9) and using (6), we obtain in turn

$$\begin{aligned} g_+(q^* - c) \leq\ & BGT\big[L(p + c + q^* - c) - L(p_+) - L'(p_+)(q^* - c) \\ & + M(p + c + q^* - c) - M(p_+) + K(p + c + q^* - c) - K(p_+)\big] - (I - R_+)(q^* - c) + c_+ \\ =\ & g(q^*) - BT\big[L(p + q^*) - L(p) - L'(p)q^* + M(p + q^*) - M(p) + K(p + q^*) - K(p)\big] \\ & + BTG\big[L(p + q^*) - L(p_+) - L'(p_+)(q^* - c) + M(p + q^*) - M(p_+) + K(p + q^*) - K(p_+)\big] \\ & + (I - R)q^* - c - (I - R_+)(q^* - c) + c_+ \\ =\ & \big[BTL'(p) - BTGL'(p_+) + (I - R) - (I - R_+)\big]q^* \\ & + (G - I)\big\{(I - R)q^* - c - BT\big[L(p) + M(p) + K(p) + L'(p)q^*\big]\big\} \\ & + BT\big[L(p) + M(p) + K(p) - GL(p_+) - GM(p_+)\big] \\ & + BTGL'(p_+)c + (I - R_+)c + c_+ - c \\ =\ & \big[(2I - G)BTL'(p) - BTGL'(p_+) + R_+ + G(I - R) - I\big]q^* \\ & + 2BT(L(p) + M(p) + K(p)) - BTG(L(p_+) + M(p_+) + K(p_+)) \\ & - GBT(L(p) + M(p) + K(p)) + (BTGL'(p_+) - G + I - R_+)c + c_+ \leq 0 \end{aligned}$$

since (5) and (6) are satisfied.

Moreover from (9) for $q = c_+$, we obtain

$$g_+(c_+) \geq 0. \tag{11}$$

By inequalities (10), (11), the fact that g is continuous and isotone on $[c_+, q^* - c]$, and since V is a regular partially ordered topological space, from the proposition, we deduce that there exists a point q_+^* with

$$g_+(q_+^*) = 0 \tag{12}$$

and

$$c_+ \leq q_+^* \leq q^* - c. \tag{13}$$

We can assume that q_+^* denotes the least of the solutions of equation (9).

That completes the proof of the lemma.

The following result is a consequence of the above lemma.

Theorem 1. Let $\{c_n\} \in V$, $\{T_n\}$, $\{R_n\}$, $\{G_n\} \in L_+(V)$ $(n \geq 0)$ be sequences and V as in the above lemma. Assume:

(a) There exists a sequence $\{p_n\} \in [0, a] \subseteq V_1 \subseteq V$ for some $a \in V$ with $p_0 = 0$ and

$$p_{n+1} \leq \sum_{j=0,1,\ldots,n} c_j \text{ for } n \geq 0. \tag{14}$$

(b) $R_0 < I$ and the function

$$\begin{aligned} g_0(q) = {} & BT_0\big[L(p_0+q) - L(p_0) - L'(p_0)q + M(p_0+q) - M(p_0) + K(p_0+q) - K(p_0)\big] \\ & - (I - R_0)q + c_0 = 0 \end{aligned} \tag{15}$$

has roots on $[0, a]$, where B, L, M, K are as in the above lemma. Denote by q_0^* the least of them.

(c) The following conditions are satisfied for all $n \geq 0$

$$R_{n+1} < \alpha_{n+1}, \tag{16}$$

$$0 < c_{n+1} \leq \beta_{n+1} \tag{17}$$

and

$$0 \leq p_{n+1} \leq p_n + c_n. \tag{18}$$

(d) The linear operators T_n are boundedly invertible for all $n \geq 0$, and set $G_n = T_{n+1}T_n^{-1}$ $(n \geq 0)$.

(e) Condition (8) is satisfied.

Then, the equation

$$\begin{aligned} g_n(q) = {} & BG_nT_n\big[L(p_n+q) - L(p_n) - L'(p_n)q + M(p_n+q) - M(p_n) + K(p_n+q) - K(p_n)\big] \\ & - (I - R_n)q + c_n = 0 \end{aligned} \tag{19}$$

has solutions in $[0, a]$ for every $n \geq 0$ and denoting by q_n^* the least of them, we have

$$\sum_{j=n,\ldots,\infty} c_j \leq q_n^* \quad (n \geq 0). \tag{20}$$

Proof. Let us assume that for some nonnegative integer n, $I - R_n > 0$, $g_n(q)$ has roots on $[0, a]$ and denote by q_n^* the least of them. We use induction on n. We also observe that this is true by hypothesis (b) for $n = 0$. Using (14), (16), (17), (18), the lemma, and setting $c = c_n$, $c_+ = c_{n+1}$, $R = R_n$, $R_+ = R_{n+1}$, $G = G_n$ we deduce that q_{n+1}^* exists, and

$$c_{n+1} \leq q_{n+1}^* \leq q_n^* - c_n. \tag{21}$$

The induction is now complete and (20) follows immediately from (21).

That completes the proof of the theorem. From now on we assume that X is a Banach space with a convergence structure in the sense of [4].

The following result is an immediate consequence of Theorem 1.

Theorem 2. Assume:

(a) the hypotheses of Theorem 1 are satisfied;

(b) there exists a sequence $\{x_n\}$ $(n \geq 0)$ in a Banach space X with a convergence structure such that $|x_{n+1} - x_n| \leq c_n$. Then

(i) the sequence $\{x_n\}$ $(n \geq 0)$ converges to some point x^*;

(ii) moreover the following error bounds are true

$$|x^* - x_n| \leq q_n^* \tag{22}$$

and

$$|x^* - x_{n+1}| \leq q_n^* - c_n \text{ for all } n \geq 0. \tag{23}$$

We can introduce the main result:

Theorem 3. Let X be a Banach space with convergence structure (X, V, E) with $V = (V, C, \| \|_v)$, an operator $F \in C^1(D \to X)$ with $D \subseteq X$, an operator $Q \in C(D \to X)$, an operator $A(x) \in C(X \to D)$, an operator $L \in C^1(V_1 \to V)$ with $V_1 \subseteq V$, an operator $M \in C(V_1 \to V)$, an operator $K \in L_+(V)$, and a point $a \in C$ such that the following conditions are satisfied:

(a) the inclusions $U(a) \subseteq D$ and $[0, a] \subseteq V_1$ are true;

(b) L is order-convex on $[0, a]$ and satisfies

$$K + L'(|x| + |y|) - L'(|x|) \in B(A(x) - F'(x + y)) \tag{24}$$

for all $x, y \in U(a)$ with $|x| + |y| \leq a$;

(c) M satisfies the condition

$$0 \leq (Q(x) - Q(x + y), M(|x| + |y|) - M(|x|) \in E, \quad M(0) = 0 \tag{25}$$

for all $x, y \in U(a)$ with $|x| + |y| \leq a$;

(d) for the sequences $\{c_n\}$, $\{T_n\}$, $\{R_n\}$, $\{G_n\}$, $\{p_n\}$ $(n \geq 0)$ the hypotheses (14), (b), (16), (18) and (d) of Theorem 1 are satisfied;

(e) the following conditions are also satisfied

$$|\delta_n| \leq c_n \leq T_n| - (F(x_n) + Q(x_n))| \leq \gamma_n \leq \beta_n \quad (n \geq 1), \tag{26}$$

$$| - r_n| \leq T_n^{-1} R_n c_n \tag{27}$$

where

$$\gamma_n = T_n \big[L(p_n + c_n) - L(p_n) - L'(p_n)c_n + M(p_n + c_n) - M(p_n) + K(p_n + c_n) - K(p_n)\big] + R_n c_n \quad (n \geq 1). \tag{28}$$

Then

(i) the sequence $\{x_n\}$ $(n \geq 0)$ generated by

$$x_{n+1} = x_n + \delta_n \text{ with } x_0 = 0$$

remains in $\overline{U}(x_0, t_0^*)$ and converges to a solution x^* of equation $F(x) = 0$;

(ii) moreover, the error estimates (22) and (23) are true where q_n^* is the least root in $[0, a]$ of the function $g_n(q)$ defined in (19), with $p_n = \|x_n - x_0\|$ $(n \geq 0)$.

Proof. Let us assume that $x_n, x_{n+1} \in U(x_0, q_0^*)$, where the existence of q_0^* is guaranteed from hypotheses (d). We note that $|\delta_0| \leq c_0$. Using the approximation

$$-(F(x_{n+1}) + Q(x_{n+1})) = (F(x_n) - F(x_{n+1}) + A(x_n)(x_{n+1} - x_n)) + (Q(x_n) - Q(x_{n+1})) - r_n, \tag{29}$$

(24), (25), (27) and setting $p_n = |x_n - x_0|$, we obtain in turn

$$\begin{aligned}
&| - (F(x_{n+1}) + Q(x_{n+1}))| \leq |F(x_n) - F(x_{n+1}) + A(x_n)(x_{n+1} - x_n)| + |Q(x_n) - Q(x_{n+1})| + | - r_n| \\
&\leq \int_0^1 \big[L'(p_n + t|x_{n+1} - x_n|) - L'(p_n) + K\big]|x_{n+1} - x_n|\, dt + M(p_n + |x_{n+1} - x_n|) - M(p_n) + | - r_n| \\
&\leq L(p_n + c_n) - L(p_n) - L'(p_n)c_n + Kc_n + M(p_n + c_n) - M(p_n) + T_n^{-1} R_n c_n.
\end{aligned}$$

Hence, by (26) we get

$$c_{n+1} \leq T_{n+1}| - (F(x_{n+1}) + Q(x_{n+1}))| \leq \gamma_n \leq \beta_n \quad (n \geq 1),$$

which shows (17).

It can easily be seen that by using induction on n, the hypotheses of Theorem 2 are satisfied. Hence, by (24) and (27) the iteration $\{x_n\}$ $(n \geq 0)$ remains in $U(x_0, q_0^*)$ and converges to x^* so that (22) and (23) are satisfied. Moreover, from the estimate

$$| - (F(x_n) + Q(x_n))| \leq |A(x_n) - F'(x_n)|c_n + |F'(x_n)|c_n + | - r_n|,$$

(24), (27), the continuity of F, F', A, T_n, R_n and $c_n \to 0$ as $n \to \infty$, we deduce that

$$F(x^*) + Q(x^*) = 0.$$

That completes the proof of the theorem.

IV. APPLICATIONS

We will complete this study with an example that shows how to choose L, M, in practical applications. For simplicity, we assume that $A(x) = F'(x)$ for all $x \in D$.

Example. We discuss the case of a real Banach space with norm $\|\ \|$. Assume that $F'(0) = I$ and there exists a monotone operator

$$f : [0, a] \to IR$$

such that

$$\|F''(x)\| \le f(\|x\|) \text{ for all } x \in U(a) \tag{30}$$

and a continuous nondecreasing function ℓ on $[0, r]$, $r \le a$ such that

$$\|Q(x) - Q(y)\| \le \ell(r)\|x - y\| \tag{31}$$

for all $x, y \in U(\frac{r}{2})$.

We showed in [3], (see also [9]) that (18) implies that

$$\|Q(x + e) - Q(x)\| \le h(r + \|e\|) - h(r), \quad x \in U(a), \quad \|e\| \le a - r, \tag{32}$$

where

$$h(r) = \int_0^r \ell(t)\, dt. \tag{33}$$

Conversely, it is not hard to see that we may assume, without loss of generality, that the function h and all functions $h(r+t) - h(r)$ are monotone in r. Hence, we may assume that $h(r)$ is convex and hence differentiable from the right. Then, as in [9], we show that (32) implies (31) and $\ell(r) = h'(r + 0)$. Hence, we can set

$$L_1(q) = L(q) = \|F(0) + Q(0)\| + \int_0^q ds \int_0^s f(t)\, dt \tag{34}$$

and

$$M(q) = \int_0^q \ell(t)\, dt. \tag{35}$$

Define the functions f_1, f_2, f_3 on $[0, a]$ by

$$f_1(q) = L(q) - q, \ f_2(q) = f_1(q) + h(q) \text{ and } f_3(q) = \int_0^q f(t)\, dt.$$

Choose $B = T_0 = 1$, $R_0 = K = 0$ and $p_0 = 0$. Then by (15) we get

$$g_0(q) = f_2(q).$$

It can easily be seen that with the above choices of L and M conditions (24) and (25) are satisfied.

Suppose that the function g_0 has a unique zero q_0^* in $[0, a]$ and $g_0(a) \le 0$. It is then known [3], [9] that equation (1) admits a solution x^* in $U(q_0^*)$, this solution is unique in $U(a)$, and the iteration $\{x_n\}$ $(n \ge 0)$ given by (2) is well defined, remains in $U(q_0^*)$ for all $n \ge 0$ and converges to x^*. By applying the Banach lemma on invertible operators we can show that $A(x_n)$ is invertible and $\|A(x_n)^{-1}\| \le -f_1'(\|x_n\|)^{-1} = T_n$ $(n \ge 0)$.

Moreover, suppose that (30) is satisfied. Furthermore, assume that instead of conditions (16) and (17), the weaker condition (10) is satisfied. Using (30) and the approximation

$$r_n = \big[(F'(x_n) - F'(x_0)) + F'(x_0)\big]\delta_n + F(x_n) + Q(x_n)$$

we obtain

$$\|r_n\| \le \big(f_3(\|x_n\|) + 1\big)c_n + T_n^{-1}\gamma_n \quad (n \ge 1).$$

The above estimate provides us with a possible (but not the only) choice for the R_n's. Indeed, (27) will be true if the R_n's $(n \ge 0)$ can be chosen so that

$$(f_3(\|x_n\|) + 1)c_n + T_n^{-1}\gamma_n \le T_n^{-1} R_n c_n \quad (n \ge 1).$$

Using the above choices and Theorem 3, it finally also follows that estimates (22) and (23) are satisfied for all $n \ge 0$.

REFERENCES

[1] Argyros, I.K. The secant method in generalized Banach spaces, Appl. Math. Comp. 39, (1990), 111–121.

[2] ______________. Newton-like methods in generalized Banach spaces, Functiones et Approximatio Commentarii Mathematici, XXII, (1993), 13–20.

[3] Argyros, I.K. and Szidarovszky, F. The Theory and Applications of iteration methods, C.R.C. Press, Inc. Boca Raton Florida, U.S.A. 1993.

[4] Argyros, I.K. Perturbed Newton-like methods and nondifferentiable operator equations on Banach spaces with a convergence structure, Southwest Journal of Pure and Applied Mathematics (SWJPAM) Vol. 1, No. 1 (1995), 1–12.

[5] Kantorovich, L.V. and Akilov, G.P. Functional Analysis in normed spaces, Academic Press, New York, 1978.

[6] Meyer, P.W. Newton's method in generalized Banach spaces, Numer. Funct. Anal. Optimiz. 9 (3 and 4), (1987), 249–259.

[7] ______________. A unifying theorem for Newton's method, Numer. Funct. Anal. Optimiz. 13 (5 and 6), (1992), 463–473.

[8] Moret, I. A Kantorovich-type theorem for inexact Newton methods, Numer. Funct. Anal. and Optimiz. 10 (3 and 4), (1989), 351–365.

[9] Zabrejko, P.P. and Nguen, D.F. The majorant method in the theory of Newton-Kantorovich approximations and the Ptak error estimates, Numer. Funct. Anal. Optimiz. 9 (5 and 6), (1987), 671–684.