Punjab University
Journal of Mathematics (ISSN 1016-2526)
Vol. 40 (2008) pp. 1-7

Local Convergence for Multistep Simplified Newton-like Methods

Ioannis K. Argyros
Department of Mathematical Sciences
Cameron University
Lawton, OK 73505, USA
E-mail: iargyros@cameron.edu.

Abstract. In this paper we provide a local convergence analysis for multistep Newton-like method (1.3) in order to approximate a solution of the nonlinear equation (1.1) in a Banach space setting. A refined and more flexible than before local [4]-[7] local convergence analysis of multistep simplified Newton-like methods for approximating solutions of nonlinear operator equations in Banach space is provided, by approximating not only the differentiable (see [4]-[7]) but also the non differentiable part (see also [1],[2]). A numerical example is used where our results compare favorably with earlier ones [4]-[7].

AMS (MOS) Subject Classification Codes: 65H10, 65J15, 47H17, 49M15.

Key Words: Local convergence, Banach space, radius of convergence, Fréchet derivative, Multi-step simplified Newton-like method.

1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution of equation

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

where f is a Fréchet-differentiable operator, g a continuous operator both defined on an open convex subset D of a Banach space X with values in a Banach space Y.

Newton-like (single step) method of the form

$$x^{n+1} = x^n - A(x^n)^{-1}F(x^n) \quad (n \geq 0 \tag{1.2}$$

has been used by several authors to approximate x^* [1]-[6]. With the exception of the works in [1]-[3] the authors take $A(x) \in L(X, Y)$ (the space of bounded linear operators from X into Y) to be a conscious approximation to the Fréchet-derivative $F'(x)$ of operator F. A survey of local and semilocal convergence results for method (1.2) can be found in [2].

However as already stated in [1], [3] there are several advantages (see Remark 3) if A is related not only to F' but also to the difference $g(x) - g(y)$. Here we extend these advantages (in the local convergence case) following some ideas in [5].

In order to compute each iterate in method (1.2) we solve the linear system $A(x^n)z = -F(x^n)$ and then set $x^{n+1} = x^n + z$ $(n \geq 0)$. The computation of $A(x^n)$ may be very expensive or impossible in general (for every $n \geq 0$). In practice we wish to use $A(x^n)$ instead of $A(x^{n-1}), ..., A(x^{n+m})$ to minimize the computational cost. That is why in [5] the multistep simplified Newton-like method was introduced for $x_0 \in D$ in the form:

$$\begin{aligned} x^{n,0} &= x^n \\ x^{n,i} &= x^{n,i-1} - A(x^n)^{-1}F(x^{n,i-1}), \quad i = 1, 2, ..., m \\ x^{n+1} &= x^{n,m} \quad (n \geq 0), \end{aligned} \tag{1.3}$$

where m is a natural number. Note that for $m = 1$ method (1.3) reduces to (1.2) which includes the so called simplified Newton-like method

$$x^{n+1} = x^n - A^{-1}F(x^n) \quad (n \geq 0), \tag{1.4}$$

with a constant linear operator A.

If $m = +\infty$ in (1.3) then the sequence $\{x^{0,i}\}$ also coincides with the one generated by (1.4) with $A = A(x^0)$. That is why in this study we assume m is finite. Local convergence results for method (1.3) were given in [5] for the interesting case $g \neq 0$ and $m > 1$. Here we show that under weaker hypotheses and the same computational cost the results in [5] can be improved (see more precisely Remark 3).

A numerical example is provided to justify the advantages of our approach over the ones in [5].

2. Local Convergence Analysis Of Simplified Newton-Like Method (1.3)

Suppose that equation (1.1) has a solution $x^* \in D$. We assume that there exists positive constants r_0, K, q, η and nonnegative constants c, e and an invertible linear operator L, such that for any

$$x, y \in U(x^*, r_0) = \{x \in X \,|\|x - x^*\| < r_0\} \subseteq D,$$

$A_1, A_2 \in L(Y, X)$, $A = A_1 + A_2$,

$$A(x)^{-1} \in L(X, Y)$$

such that

$$\begin{aligned} \left\|A(x)^{-1}L\right\| &\leq q, \\ \left\|A(x)^{-1}F(x)\right\| &\leq \eta, \\ \left\|L^{-1}(f'(x) - A_1(y))\right\| &\leq K\|x - y\| + c, \\ \left\|L^{-1}[g(x) - g(y) - A_2(x)(x - y)]\right\| &\leq e\|x - y\|. \end{aligned}$$

Define the scalar sequence $\{t_{n,i}\}$ by

$$t_{n,0} = 0, \; t_{n,i} = s_n(t_{n,i-1}), \; i = 1, ..., m+1, n \geq 0$$

where

$$\begin{aligned} s_n(t) &= q\left(\tfrac{K}{2}t + c + e\right)t + \eta_n, \\ \eta_0 &= \eta, \; \eta_n = t_{n-1,m+1} - t_{n-1,m} \; n \geq 1. \end{aligned}$$

Clearly $s_n(t)$ is an increasing function of $t \geq 0$. Therefore we have $t_{n,i} \leq t_{n,i+1}$. Further, define

$$t^* \geq \min(\max_n t_{n,m-1}, 2r_0),$$
$$b = q\left(\frac{Kt^*}{2} + c + e\right),$$

$$r_1 = \frac{2(1-b)}{qK},$$

and

$$a = \frac{qK}{2}.$$

We can state and show the local convergence theorem for Newton-like method (1.3).

Theorem 1. *Under the above assumptions, set $r^* = \min\{r_0, r_1\}$. If $b \in [0,1)$, then $U(x^*, r^*)$ is a convergence ball for (1.3). Moreover the following estimate holds for all $n \geq 0$:*

$$\left\|x^{n+1} - x^*\right\| \leq a(\|x^n - x^*\| + b)^m \|x^n - x^*\| \leq p^m \|x^n - x^*\|, \tag{2.5}$$

where,

$$p = a\left\|x^0 - x^*\right\| + b \in [0,1).$$

Proof. Let $x^0 \in U(x^*, r^*)$. Then we have

$$p < ar^* + b \leq ar_1 + b = 1$$

We shall prove the first inequality in (2.5) using induction on $k \geq 0$. We must show

$$\left\|x^{k,i} - x^{k,i-1}\right\| \leq t_{k,i} - t_{k,i-1} \;\; i = 1.....m \tag{2.6}$$

and

$$\left\|x^{k,i} - x^*\right\| \leq (a\left\|x^k - x^*\right\| + b)^i \left\|x^k - x^*\right\|, \; i = 1.....m \tag{2.7}$$

For $k = 0$, we have

$$\left\|x^{0,1} - x^{0,0}\right\| = \left\|x^{0,1} - x^0\right\| = \left\|A(x^0)^{-1}F(x^0)\right\| \leq \eta - t_{0,1} = t_{0,1} - t_{0,0}$$

and

$$\begin{aligned}
\left\|x^{0,1} - x^*\right\| &= \left\|-A(x^0)^{-1}(F(x^0) - F(x^*) - A(x^0)(x^0 - x^*))\right\| \\
&\leq q\left\|\int_0^1 L^{-1}(f'(x^* + t(x^0 - x^*)) - A_1(x^0))dt(x^0 - x^*)\right\| \\
&+ q\left\|L^{-1}(g(x^0) - g(x^*) - A_2(x^0)(x^0 - x^*))\right\| \\
&\leq q\left(\frac{K}{2}\left\|x^0 - x^*\right\| + c + e\right)\left\|x^0 - x^*\right\| \\
&\leq (a\left\|x^0 - x^*\right\| + b)\left\|x^0 - x^*\right\|
\end{aligned}$$

This implies that if $m = 1$, then (2.6) and (2.7) hold for $k = 0$. If $m \geq 2$, then we have by induction on i

$$\begin{aligned}
\left\|x^{0,i} - x^0\right\| &\leq \min\left\{\sum_{j=1}^{i}(t_{0,j} - t_{0,j-1}), \left\|x^{0,i} - x^*\right\| + \left\|x^0 - x^*\right\|\right\} \\
&\leq \min(t_{0,i}, 2r_0) \leq \min(t_{0,m-1}, 2r_0) \leq t^* \\
\left\|x^{0,i-1} - x^{0,i}\right\| &\leq \left\|L^{-1}(F(x^{0,i}) - A(x^0)(x^{0,i} - x^{0,i-1}) - F(x^{0,i-1}))\right\| \\
&\leq q\left(K\int_0^1 \left\|t(x^{0,i} - x^0) + (1-t)(x^{0,i-1} - x^0)\right\| dt + c + e\right) \\
&\quad \times \left\|x^{0,i} - x^{0,i-1}\right\| \\
&\leq q\left(\tfrac{K}{2}\left(t_{0,i} - t_{0,i-1}\right) + c + e\right)\left(t_{0,i} - t_{0,i-1}\right) = t_{0,i+1} - t_{0,i}
\end{aligned}$$

and

$$\begin{aligned}
\left\|x^{0,i+1} - x^*\right\| &= \left\|-A(x^0)^{-1}(F(x^{0,i}) - F(x^*) - A(x^0)(x^{0,i} - x^*))\right\| \\
&\leq q\left(\tfrac{K}{2}(\left\|x^0 - x^*\right\| + \left\|x^{0,i} - x^0\right\|) + c + e\right)\left\|x^{0,i} - x^*\right\| \\
&\leq q\left(\tfrac{K}{2}(\left\|x^0 - x^*\right\| + t^*) + c + e\right)\left\|x^{0,i} - x^*\right\| \\
&\leq (a\left\|x^0 - x^*\right\| + b)(a\left\|x^0 - x^*\right\| + b)^i\left\|x^0 - x^*\right\| \\
&= (a\left\|x^0 - x^*\right\| + b)^{i-1}\left\|x^0 - x^*\right\|.
\end{aligned}$$

This proves (2.6) and (1.1) for the case $k = 0$.

Assume now that (2.6) and (1.1) hold for some k. Then we have

$$x^{k+1,0} = x^{k-1} = x^{k,m} \in U(x^*, r)$$

and

$$\begin{aligned}
&\left\|x^{k+1,1} - x^{k+1,0}\right\| \\
&= \left\|x^{k+1.1} - x^{k+1}\right\| \\
&\leq \left\|A(x^{k+1})^{-1}L\right\|\left\|L^{-1}(F(x^{k,m}) - A(x^k)(x^{k,m} - x^{k,m-1}) - F(x^{k,m-1}))\right\| \\
&\leq q\left(\tfrac{K}{2}(\left\|x^{k,m} - x^k\right\| + \left\|x^{k,m-1} - x^k\right\|) + c + e\right)\left\|x^{k,m} - x^{k,m-1})\right)\right\| \\
&\leq q\left(\tfrac{K}{2}(t_{k,m} + t_{k,m-1}) + e + c\right)\left(t_{k,m} - t_{k,m-1}\right) = t_{k,m+1} - t_{k,m} = \eta_{k-1}.
\end{aligned}$$

By the same argument as for $k = 0$, we can prove that (2.6) and (2.7) hold for $k + 1$. This completes the induction and the proof of the theorem. □

Setting $L = A(x^*)$ in Theorem 1, we obtain the following:

Corollary 2. *Assume that $A(x^*)$ is nonsingular and for any $x \in D$, the following hold:*

$$\begin{aligned}
\left\|A(x^*)^{-1}(f'(x) - A_1(y))\right\| &\leq K\left\|x - y\right\| + c \\
\left\|A(x^*)^{-1}(A(x) - A(x^*))\right\| &\leq L\left\|x - x^*\right\| + d \\
\left\|A(x^*)^{-1}[g(x) - g(x^*) - A_2(x)(x - x^*)]\right\| &\leq e\left\|x - x^*\right\| \\
p &= c + d + e < 1
\end{aligned}$$

Then

(i) *The ball* $U(x^*, r^*)$ *with* $r^* = 2(1-p)/(3K+2L)$ *is a convergence ball for the iterative method (1.3) with any* m*, provided that* $U(x^*, r^*) \subset D$*. The speed of convergence is estimated as follows:*

$$\left\|x^{n+1} - x^*\right\| = \left\|x^{n,m} - x^*\right\| \leq \left(a\left\|x^n - x^*\right\| + b\right)^m \left\|x^n - x^*\right\| \leq p^m \left\|x^n - x^*\right\|$$

where

$$a = \tfrac{3K}{2(1-Lr-d)}, \quad b = \tfrac{c+e}{1-Lr-d}$$

$$p = a\left\|x^0 - x^*\right\| + b < 1$$

(ii) *The ball* $U(x^*, r^*)$ *with* $r^* = 2(1-p)/(K+2L)$ *is convergence ball for the iteration (1.4) and*

$$\left\|x^{n+1} - x^*\right\| \leq \frac{1}{1-Lr-d}\left(\tfrac{K}{2}\left\|x^n - x^*\right\| + c + e\right)\left\|x^n - x^*\right\|$$

provided that $U(x^*, r^*) \subset D$*.*

Proof. (see Corollary 1 in [5, p.19]). □

Remark 3. If we set

$$A_2 = 0 \quad \text{and } A_1 = A \tag{2.8}$$

our results reduce to the corresponding ones in [4]. Otherwise our results have the following advantages over the ones in [4]: more flexible choices of operator A (i.e A_1 and A_2); finer error bounds on the distances $\left\|x^{n+1} - x^*\right\|$;and a larger radius of r^*. That is we can obtain a desired error tolerance ε with fewer computations, a larger m can be used and there is a wider choice of initial guesses x^0 available. Such an information is important in computational mathematics and scientific computing [1], [2]. In what follows we provide an example. For simplicity we take $m = 1$, and $A(x) = L$.

Example 4. *Let* $X = Y = (\mathbf{R}^2, \|\cdot\|_\infty)$*. Consider the system* [3]*:*

$$\begin{aligned} 3x^2y + y^2 - 1 + |x-1| &= 0 \\ x^4 + xy^3 - 1 + |y| &= 0. \end{aligned} \tag{2.9}$$

It can easily be seen that the solution of (2.9) is given by

$$x^* = (.8946553334687, \ .327826521746298) \tag{2.10}$$

Set for $v = (v_1, v_2)$, $\|v\|_\infty = \|(v_1, v_2)\|_\infty = \max\{|v_1|, |v_2|\}$, $F(v) = f(v) + g(v)$, $f(v) = (f_1, f_2)$, $g(v) = (g_1, g_2)$.

Define

$$\begin{aligned} &f_1(v) = 3v_1^2 v_2 + v_2^2 - 1, \quad f_2(v) = v_1^4 + v_1 v_2^3 - 1, \\ &g_1(v) = |v_1 - 1|, g_2(v) = |v_2|. \end{aligned}$$

We shall take divided differences of order one $[x, y; f], [x, y; g] \in M_{2\times 2}(\mathbf{R})$ *to be for* $w = (w_1, w_2)$*:*

$$\begin{aligned} [v, w; f]_{i,1} &= \frac{f_i(w_1, w_2) - f_i(v_1, w_2)}{w_1 - v_1} \\ [v, w; f]_{i,2} &= \frac{f_i(v_1, w_2) - f_i(v_1, v_2)}{w_2 - v_2} \end{aligned}$$

provided that $w_1 \neq v_1$ *and* $w_2 \neq v_2$. *If* $w_1 = v_1$ *or* $w_2 = v_2$ *replace* $[x, y, f]$ *by* f'. *Similarly we define*

$$[v, w; g]_{i,1} = \frac{g_i(w_1, w_2) - g_i(v_1, w_2)}{w_1 - v_1}$$
$$[v, w; g]_{i,2} = \frac{g_i(v_1, w_2) - g_i(v_1, v_2)}{w_2 - v_2}$$

for $w_1 \neq v_1$ *and* $w_2 \neq v_2$. *If* $w_1 = v_1$ *or* $w_2 = v_2$ *replace* $[x, y; g]$ *by the zero* 2×2 *matrix in* $M_{2\times 2}(\mathbf{R})$. *We consider a possible choice for operator* A *as suggested by the hypotheses in* [5]*:*

$$A(v) = A_1(v) = F'(v), \text{ and } A_2 = 0.$$

Then, using Newton's method (1.2) in this case for $x^0 = (1, 0)$, *we obtain Table 1. Moreover, if we choose:* $A(v, w) = A_1(v, w) = [v, w; g]$, *and* $A_2 = 0$, *i.e. the method of Chord or Secant method (1.2), we obtain Table 2, for* $x^{-1} = (5, 5)$, *and* $x^0 = (1, 0)$. *Furthermore if we choose:* $A = A_1 + A_2$, *where* $A_1(v, v) = F'(v) = [v, v; f]$, *and* $A_2(v, w) = [v, w; g]$ *for* $x^{-1} = (5, 5)$, *and* $x^0 = (1, 0)$ *our method (1.2) provides Table 3. Tables 2 and 3 show the superiority of the results obtained here, over the results in* [5] *using Table 1. Finally, although the superiority of our results over the ones in* [5] *has already been established, we note that if e.g., we let* $x^{-1} = x_7$, $x^0 = x_8$ *(chosen from Table 3), then hypotheses of Theorem 1 hold for* $K = q = 1, e = .25, c = 0, \eta = r_0 = 1.077E - 14, r^* = r_0$, *and* $t^* = 2r_0$.

Table 1.

n	$x_n^{(1)}$	$x_n^{(2)}$	$\|x_n - x_{n-1}\|$
0	1	0	
1	1	0.333333333333333	3.333E-1
2	0.906550218340611	0.354002911208151	9.344E-2
3	0.885328400663412	0.338027276361322	2.122E-2
4	0.891329556832800	0.326613976593566	1.141E-2
5	0.895238815463844	0.326406852843625	3.909E-3
6	0.8951546711372635	0.327730334045043	1.323E-3
7	0.894673743471137	0.327979154372032	4.809E-4
8	0.894598908977448	0.327865059348755	1.140E-4
9	0.894643228355865	0.327815039208286	5.002E-5
10	0.894659993615645	0.327819889264891	1.676E-5
11	0.894657640195329	0.327826728208560	6.838E-6
12	0.894655219565091	0.327827351826856	2.420E-6
13	0.894655074977661	0.327826643198819	7.086E-7
...			
39	0.89455373334687	0.327826521746298	5.149E-19

Table 2.

n	$x_n^{(1)}$	$x_n^{(2)}$	$\|x_n - x_{n-1}\|$
-1	5	5	
0	1	0	5.000E+00
1	0.989800874210782	0.021627489072365	1.262E-02
2	0.921814765493287	0.307939916152262	2.953E-01
3	0.900073765669214	0.325927010697792	2.174E-02
4	0.894939851625105	0.327725437396226	5.133E-03
5	0.894658420586013	0.327825363500783	2.814E-04
6	0.894655375077418	0.327826521051833	3.045E-04
7	0.894655373334698	0.327826521746293	1.742E-09
8	0.894655373334687	0.327826521746298	1.076E-14
9	0.894655373334687	0.327826521746298	5.421E-20

Table 3.

n	$x_n^{(1)}$	$x_n^{(2)}$	$\|x_n - x_{n-1}\|$
-1	5	5	
0	1	0	5
1	0.909090909090909	0.363636363636364	3.636E-01
2	0.894886945874111	0.329098638203090	3.453E-02
3	0.894655531991499	0.327827544745569	1.271E-03
4	0.894655373334793	0.327826521746906	1.022E-06
5	0.894655373334687	0.327826521746298	6.089E-13
6	0.894655373334687	0.327826521746298	2.710E-E20

References

[1] I. K. Argyros, *A unifying local-semilocal convergence analysis and applications for two point Newton-like methods in Banach space*, J. Math. Anal. and Applic. 298, 2, (2004), 374-397.

[2] I. K. Argyros, *Newton Methods*, Nova Science Publ. Corp., New York, 2005.

[3] E. Cătinaş, *On some iterative methods for solving nonlinear equations*, Rev. Anal. Numer. Theor. Approx. **23**, 1, (1994), 47-53.

[4] X. Chen and T. Yamamoto, *Convergence domains of certain iterative methods for solving nonlinear equations*, Numer. Funct. Anal. Optimiz. **10**(1989), 37-48.

[5] X. Chen and T. Yamamoto, *A convergence ball for multistep Newton-like methods*, Numer. Funct. Anal. Optim. **14**, (1 and 2), (1993), 15-24.

[6] J. E. Dennis Jr., *Towards a unified convergence theory for Newton-like methods*, In: Nonlinear Functional Analysis and Applications (Ed. L.B. Rall), Academic Press, New York 1971, 425-472.

[7] P.P. Zabrejko and D. F. Nguen, *The majorant method in the theory of Newton-Kantorovich method for an equation with a non-differentiable term,* (Russian), Ukr. Math. Zhurn. **34**, (1982), 365-369.