REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Tome 29, N° 2, 2000, pp. 119–127

A NEW CONVERGENCE THEOREM FOR THE STEFFENSEN METHOD IN BANACH SPACE AND APPLICATIONS

IOANNIS K. ARGYROS

Abstract. We approximate a locally unique solution of a nonlinear operator equation in a Banach space using the Steffensen method. A new semilocal convergence theorem is provided using Lipschitz conditions on the second Fréchet-derivative of the operator involved. Earlier results have used Lipschitz conditions only on the first divided difference. This way our conditions are different from earlier ones. Hence, they have theoretical and practical value. A numerical example is also provided to show that our results apply to solve a nonlinear equation, where earlier ones fail.

Key words and phrases: Banach space, Steffensen method, divided difference, Fréchet-derivative, Newton-Kantorovich hypothesis.

1. INTRODUCTION

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

(1) $$F(x) = 0,$$

where F is a twice Fréchet-differentiable operator defined on a convex subset D of a Banach space E with values on itself.

We use the Steffensen method

(2) $$x_{n+1} = x_n - [x_n, B(x_n); F]^{-1} F(x_n) \quad (x_0 \in D)\ (n \geq 0)$$

to generate a sequence $\{x_n\}$ $(n \geq 0)$ converging to x^*. Here, $[x, y; F]$ denotes a divided difference of order one at the points $x, y \in D$, which is an element of $L(E, E)$, the space of bounded linear operators from E into itself. B is a continuous operator defined on D with values in E, usually related to F by $B(x) = x - F(x)$ $(x \in D)$.

Sufficient convergence conditions for the Steffensen method were given in [1], [2], [4], [5], [6], [9]–[11] using mainly Lipschitz-type conditions on the

AMS subject classification: 65B05, 47H17, 49D15.

first divided difference of F. In our study, we use Lipschitz conditions on the second Fréchet-derivative to obtain a new semilocal convergence theorem for the Steffensen method. This way, our convergence conditions differ from earlier ones. Therefore our results have theoretical as well as practical value.

Finally, we complete this study by providing a numerical example to show that the Steffensen method starting from an initial guess x_0 converges to x^*, whereas the same is not guaranteed by existing conditions [1], [2], [4], [5], [6], [9]–[11] .

2. CONVERGENCE ANALYSIS

From now on we will set $A(x) = [x, B(x); F]$ $x \in D$, and further denote $A(x_n)$ by A_n $(n \geq 0)$ for simplicity.

Let a, b, c, d, R be given nonnegative constants with $c \in [0,1)$, $x_0 \in D$ such that $A_0^{-1} = A(x_0)^{-1} \in L(E,E)$, f be an increasing real function, which is continuous and nonvanishing on $[0,R]$. Define the polynomial p by

$$p(t) = \frac{1}{6}at^3 + \frac{1}{2}bt^2 - (1-c)t + d, \tag{3}$$

the constants α, β by

$$\alpha = \frac{2(1-c)}{b + \sqrt{b^2 + 2a(1-c)}}, \tag{4}$$

$$\beta = (1-c)\alpha - \frac{1}{6}a\alpha^3 - \frac{1}{2}b\alpha^2, \tag{5}$$

and the iteration $\{t_n\}$ $(n \geq 0)$, by

$$t_{n+1} = t_n - \frac{p(t_n)}{f(t_n)}, \; t_0 = 0 \;\; (n \geq 0). \tag{6}$$

We need the lemmas:

LEMMA 1. *The real polynomial p has two positive zeros r_1, r_2 with $r_1 \leq r_2$ and a negative zero $-r_3$ $(r_3 > 0)$ if and only if*

$$d \leq \beta. \tag{7}$$

Proof. Polynomial p has a negative zero $-r_3$, since $p(0) = d > 0$, and $p(t) < 0$ as $t \to -\infty$. Moreover, $p'(0) = -(1-q) < 0$, and $p'(t) > 0$ as $t \to +\infty$. Hence, there exists a zero of p' in $(0,\infty)$, which by the form of p is given by (4). Thus, p has two positive zeros if and only if

$$p(\alpha) \leq 0, \tag{8}$$

which is equivalent to condition (7).

That completes the proof of Lemma 1. □

LEMMA 2. *Assume condition (7) holds, and*

$$f(t) \neq 0, \; f(t) \leq p'(t) \; \textit{for all} \; t \in [0, r_1]. \tag{9}$$

Then iteration $\{t_n\}$ $(n \geq 0)$ given by (10) is monotonically increasing and converges to r_1.

Proof. Define function g by

$$g(t) = t - \frac{p(t)}{f(t)}. \tag{10}$$

Then by differentiating function g we get

$$g'(t) = \frac{f(t)(f(t) - p'(t)) + f'(t)p(t)}{f(t)^2}. \tag{11}$$

It follows from the proof of Lemma 1, (3) and (9) that $p'(t) < 0$, $p(t) > 0$, $f(t) < 0$, and $f'(t) > 0$ for all $t \in [0, r_1]$. Hence, by (11) function g increases on $[0, r_1]$. So, if $t_k \in [0, r_1]$ for some k, then

$$t_k \leq t_k - \frac{p(t_k)}{f(t_k)} = t_{k+1}, \text{ and } t_{k+1} \leq t_k - \frac{p(t_k)}{f(t_k)} \leq r_1 - \frac{p(r_1)}{f(r_1)} = r_1$$

That completes the proof of Lemma 2. □

Remark 1. It can easily be seen by (11) that condition (9) can be replaced by the weaker

$$f(t) \neq 0, \; f(t)(f(t) - p'(t)) + f'(t)p(t) \geq 0 \; \textit{for all} \; t \in [0, r_1].$$

We can now prove the semilocal convergence theorem for Steffensen method (2).

THEOREM 1. *Let $F : D \subseteq E \to E$ be a twice Fréchet-differentiable operator. Assume:*

(a) there exists $x_0 \in D$ such that $A_0^{-1} = A(x_0)^{-1} \in L(E,E)$;

(b) for all $x \in \overline{U}(x_0, R) = \{x \in E \mid \|x - x_0\| \leqslant R\}$ there exist constants a, b, c such that

$$\left\| A_0^{-1}\left(F''(x) - F''(x_0)\right)\right\| \leqslant a\,\|x - x_0\|, \tag{12}$$

$$\left\| A_0^{-1} F''(x_0)\right\| \leqslant b, \tag{13}$$

$$\left\| A_0^{-1}\left(F'(x) - A(x)\right)\right\| \leqslant c; \tag{14}$$

(c) conditions (7) and (9) are satisfied for some continuous monotonically increasing and nonvanishing function f on $[0, r_1]$ such that

$$\left\| A_0^{-1}\left(A(x) - A_0\right)\right\| \leqslant f\left(\|x - x_0\| + 1\right) < 1, \;\; \textit{for all} \;\; x \in \overline{U}(x_0, r_1) \tag{15}$$

and

(16) $$d \leqslant \left\| A_0^{-1} F(x_0) \right\| ;$$

(d) the following hold:

(17) $$c \in [0, 1),$$

(18) $$r_2 \leqslant R,$$

(19) $$\overline{U}(x_0, R) \subseteq D,$$

where r_1 and r_2 are the positive zeros of equation $p(t) = 0$, and polynomial p is given by (3).

Then, the Steffensen method $\{x_n\}$ $(n \geqslant 0)$ generated by (2) is well defined, remains in $\overline{U}(x_0, r_1)$ for all $n \geqslant 0$, and converges to a solution $x^* \in \overline{U}(x_0, r_1)$ of equation $F(x) = 0$. If $r_1 < r_2$ the solution x^* is unique in $U(x_0, r_2)$, whereas if $r_1 = r_2$, x^* is unique in $\overline{U}(x_0, r_1)$.

Moreover, the following error bounds hold for all $n \geqslant 0$

(20) $$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n$$

and

(21) $$\|x_n - x^*\| \leqslant r_1 - t_n.$$

Proof. We first show linear operator $A(x)$ is invertible for all $x \in \overline{U}(x_0, \alpha)$, where α is given by (4). It follows from (15) the Banach lemma on invertible operators [4], [8], the estimate

$$\left\| A_0^{-1} (A(x) - A_0) \right\| \leqslant f(\|x - x_0\|) + 1 < 1,$$

that $A(x)^{-1} \in L(E_2, E_1)$ and

(22) $$\left\| A(x)^{-1} A_0 \right\| \leqslant -f(\|x - x_0\|)^{-1} \leqslant -f(\alpha)^{-1}.$$

We must show that estimate (20) holds for all $n > 0$. First, note that x_1 is defined, and by using (2), (6), and (16) we get $\|x_1 - x_0\| = \left\| -A_0^{-1} F(x_0) \right\| \leqslant d = t_1 - t_0$, which shows (20) for n = 0. It follows from (22) that linear operator $A(x_1)^{-1} \in L(E, E)$, and hence x_2 can then be defined by (2). Let $x \in [x_0, x_1] = \{x : x = \lambda x_1 + (1 - \lambda) x_0,\ 0 \leqslant \lambda \leqslant 1\}$. By Taylor's formula [3], [4], [7] for a twice Fréchet-differentiable operator G on D, we can write

(23) $$G(x) = G(x_0) + G'(x_0)(x - x_0) + \frac{1}{2} G''(x_0)(x - x_0)^2 + \int_{x_0}^{x} \left[G''(y) - G''(x_0) \right] (x - y)\, dy.$$

Using approximation (23) for $G(x) = A_0^{-1} F(x)$ $(x \in D)$, we can get

(24) $$\begin{aligned} A_0^{-1} F(x) &= A_0^{-1} F(x_0) + (x - x_0) + A_0^{-1} \left(F'(x_0) - A_0 \right) + \\ &+ \frac{1}{2} A_0^{-1} F''(x_0)(x - x_0)^2 + \\ &+ \int_{x_0}^{x} A_0^{-1} \left(F''(y) - F''(x_0) \right)(x - y)\, dy. \end{aligned}$$

Let $\lambda d = s$, then by using (3), (12)–(14), (16), approximation (24) gives

(25) $$\left\| A_0^{-1} F(x) \right\| \leqslant (1 - \lambda) d + c\lambda d + \frac{1}{2} b \lambda^2 d^2 + \frac{1}{6} a \lambda^3 d^3 = p(s)$$

(since $x - x_0 = \lambda (x_1 - x_0) = -\lambda A_0^{-1} F(x_0)$).

Moreover, by (2), (22) and (25), we get

$$\|x_2 - x_1\| \leqslant \left\| A(x_1)^{-1} A_0 \right\| \cdot \left\| A_0^{-1} F(x_1) \right\| \leqslant -\frac{p(t_2)}{f(t_1)} = t_2 - t_1.$$

Similarly, we can show (20) for all $n \geqslant 0$. Estimate (20) and Lemma 2 imply that Steffensen method $\{x_n\}$ $(n \geqslant 0)$ is Cauchy in a Banach space E, and as such it converges to some $x^* \in \overline{U}(x_0, s)$ (since $\overline{U}(x_0, s)$ is a closed set). From (25) and the continuity of F we get $F(x^*) = 0$. Furthermore, estimate (21) follows immediately from (20) by using standard majorization techniques [3], [4], [8]. To show uniqueness, let $z \in U(x_0, r_2)$ with $F(z) = 0$. Using (24) for $x = 0$, we obtain,

(26) $$\begin{aligned} x_1 - z &= A_0^{-1} \left(F'(x_0) - A_0 \right)(z - x_0) + \frac{1}{2} A_0^{-1} F''(x_0)(z - x_0)^2 + \\ &+ \int_{x_0}^{x} A_0^{-1} \left[F''(y) - F''(x_0) \right](z - y)\, dy. \end{aligned}$$

We get $\|z - x_0\| \leqslant r_1 - t_0$ if $z \in \overline{U}(x_0, r_1)$, and $\|z - x_0\| = \mu (r_2 - t_0)$, $0 < \mu < 1$, if $z \in U(x_0, r_2)$. Hence, by (21) and (26), we get for all $n \geqslant 0$: $\|z - x_n\| \leqslant r_1 - t_n$, if $z \in \overline{U}(x_0, r_1)$, and $\|z - x_n\| \leqslant \mu^n (r_2 - t_n)$, if $z \in U(x_0, r_2)$. In either case, we get $\lim_{n \to \infty} x_n = z$, which yields $x^* = z$.

That completes the proof of Theorem 1. □

We now state a theorem by Pavaloiu [9] for comparison:

Theorem 2. *Let $B, F : D \subseteq E \to E$ be continuous operators with divided differences of order one $[x, y; B]$ and $[x, y; F]$ respectively. Assume:*

(a)

(27) $$B(x) = x - F(x)\ (x \in D);$$

(b) *There exists* $x_0 \in D$ *such that* $\Gamma_0 = [x_0, B(x_0); F]^{-1} \in L(E,E)$ *and*

(28) $$\|\Gamma_0\| \leqslant h_1;$$

(c)

(29) $$\max\{\|x_1 - x_0\|, \|x_1 - B(x_0)\|\} \leqslant h_0;$$

(d)

(30) $$\|[x,y;B] - [y,v;B]\| \leqslant h_2 \|x - v\| \text{ for all } x, y, v \in D;$$

(e)

(31) $$h_3 = h_0 h_1 h_2 < \frac{1}{4};$$

and

(f) $$\overline{U}(x_0, 2h_0) \subseteq D.$$

Then, Steffensen method generated by (2) is well defined, remains in $\overline{U}(x_0, 2h_0)$ *for all* $n \geqslant 0$, *and converges to a solution* x^* *of equation* $F(x) = 0$. *Moreover, the following error bounds hold for all* $n \geqslant 0$:

$$\|x^* - x_n\| \leqslant \frac{h_0}{2^{n-1}},$$

and

$$\|x^* - x_n\| \leqslant h_1^n h_2 \|x^* - x_{n-1}\| \|x^* - B(x_{n-1})\|,$$

where

$$h_1^n \geqslant \left\|[x_n, B(x_n); F]^{-1}\right\| \ (n \geqslant 0).$$

Furthermore,

(g) *if* $\|[x,y;B]\| \leqslant h_4 < 1$ *for all* $x, y \in\in \overline{U}(x_0, 2h_0)$, *then* x^* *is the unique solution of equation (1) in* $\overline{U}(x_0, 2h_0)$.

We provide an example to show that under the conditions of Theorem 1, Steffensen method converges to a solution x^* of equation (1), whereas the same is not guaranteed under the conditions of Theorem 2.

Example. Let $E = \mathbb{R}$, $D = [-1, 1]$, $x_0 = 0$. Define function F on D by

(32) $$F(x) = \frac{1}{6}x^3 + \frac{1}{6}x^2 - \frac{5}{6}x + \frac{1}{4},$$

and the divided difference $[x, y; F]$ by

$$[x, y; F] = \frac{F(y) - F(x)}{y - x}, \ x, y \in D, \ x \neq y.$$

Set also $B(x) = x - F(x)$ $(x \in D)$. Then, using (4), (5), (12) – (16), (27) – (30) and (33), we obtain

$$h_1 = 1.1566265,\ h_2 = \frac{2}{3},\ h_0 = .5391566,\ h_3 = .4157352 > .25,$$
$$a = 1.1566265,\ b = .3855422,\ c = .0119947,\ d = .2891566$$
$$\alpha = 1.0155688,\ \beta = .5392822.$$

Condition (31) of Theorem 2 does not hold. That is, Theorem 2 cannot guarantee that Steffensen method starting from $x_0 = 0$ converges to a solution of equation $F(x) = 0$, where function F is given by (32). However all conditions of Theorem 1 are satisfied. Indeed from the above we have that condition (7) is satisfied.

Define function f so that $1 + f(t) = \left|A_0^{-1}(A(t) - A_0)\right|$ $(t \in D)$. It can easily be seen that the left-hand side inequality in (15) is satisfied as equality, whereas the right-hand side is smaller than 1. Furthermore it is simple algebra to show that condition (9) holds also. Conditions (18) and (19) are needed to show uniqueness of the solution x^* in $U(x_o, r_2)$. They can certainly be replaced by $r_1 \leqslant R$ and $\overline{U}(x_o, r_1) \subseteq D$. Uniqueness is then guaranteed only in $\overline{U}(x_o, r_1)$. Since $p(0) = d > 0$ and $p(1) = -.3133065$, it follows that for $r_1 = R$, $\overline{U}(x_o, r_1) \subseteq D$. That is the hypotheses of Theorem 1 are satisfied. Hence, Steffensen method (2), indeed starting from $x_0 = 0$, converges to a solution $x^* \in \overline{U}(x_o, r_1) \subseteq D$ of equation $F(x) = 0$, where F is given by (32). Similar favorable comparisons can be made with the results in [1], [2], [5], [6], [10], [11].

Remark 2. *Condition (7) can be replaced by a stronger, but easier to check Newton-Kantorovich type hypothesis* [4], [8] *as follows: Define polynomials* p_1, f_1 *by* $p_1(t) = \frac{b_1}{2}t^2 - (1 - c)t + d$, $f_1(t) = b_1 t - b_2$, *where* $b_1 = \frac{1}{3}ad + b$ *and* $b_2 = 1 - c$. *It can easily be seen that the conclusions of Theorem 1 hold in the balls* $\overline{U}(x_0, r_4)$, $\overline{U}(x_0, r_5)$, *provided that the Newton-Kantorovich hypothesis*

$$2b_1 d \leqslant (1 - c)^2$$

holds, and r_4, r_5 $(r_4 \leqslant r_5)$ *are the nonnegative zeros of the equation* $p_1(t) = 0$. *Since* $p(t) \leqslant p_1(t)$ *for all* $t \in [0, r_2]$, *we have* $r_1 \leqslant r_4 \leqslant r_5 \leqslant r_2$.

Remark 3. *The results obtained in Theorem 2 can be extended for Steffensen-Aitken method*

(33) $$y_{n+1} = y_n - [B_1(y_n), B_2(y_2); F]^{-1} F(y_n) \ (n \geqslant 0),$$

where $B_1, B_2 : D \subseteq E \to E$ *are continuous operators related to* F [4], [10], [11]. *Simply replace* $A(x) = [x, B(x); F]$ $(x \in D)$ *to obtain a Theorem* 1′ *holding for method (33).*

Remark 4. *Condition (14) can be replaced by*

$$\left\|A_0^{-1}\left(F'(x)-A(x)\right)\right\| \leqslant c_0+c_1\left\|x-x_0\right\| \text{ for some } c_0 \geqslant 0,\ c_1 \geqslant 0$$

and all $\overline{U}(x_0,R)$. *We can also set* $c=c_0+c_1R$. *Condition (18) can be replaced by* $\alpha \leqslant R$, *but uniqueness is then guaranteed only in* $\overline{U}(x_0,\alpha)$.

Remark 5. *The results obtained in Theorem 1 can be extended so as to hold a more general setting as follows:*

(a) Let c_0, c_1 *be nonnegative constants;* v_1, v_2 *be positive monotonically increasing functions of one variable on* $[0,R]$ *with* $\lim_{t\to 0} v_1(t)=\lim_{t\to 0} v_2(t)=0$ *such that*

$$\left\|A_0^{-1}\left(F''(x)-F''(x_0)\right)\right\| \leqslant v_1\left(\left\|x-x_0\right\|\right),$$

and

$$\left\|A_0^{-1}\left(F'(x)-A(x)\right)\right\| \leqslant c_0+c_1 v_2\left(\left\|x-x_0\right\|\right)$$

for all $x \in U(x_0,R)$.

(b) Function $\overline{p}$ *given by*

$$\overline{p}(t)=\int_0^t (t-r)\, v_1(r)\,\mathrm{d}r+\frac{1}{2}bt^2-\left(1-c_0-c_1v_2(t)\right)t+d \text{ on } [0,R],$$

has a unique zero $\varepsilon_0 \in [0,R]$, *and* $\overline{p}(R) \leqslant 0$.

Moreover, set $\varepsilon_0=r_1$, and $R=r_2$. Furthermore, replace conditions (12), (14) by (a), (7) by (b), and polynomial $\overline{p}$ by function p above. Then, under the rest of the hypotheses, as it can easily be seen from the proof, the conclusions of Theorem 1 hold in this more general setting. Call such a result Theorem 1″.

Finally, note that for $v_1(t)=at$, $c_0=c$, and $v_1(t)=0$ (or $c_1=0$), $t\in[0,R]$ function p reduces to polynomial p and Theorem 1″ to Theorem 1.

REFERENCES

[1] ARGYROS, I. K., *On the solution of nonlinear equations with a nondifferentiable term*, Revue d'analyse numerique et de theorie de l'approximation, **22**, no. 2, pp. 125–135, 1993.

[2] ARGYROS, I. K., *An error analysis for the Steffensen method under generalized Zabrejko-Nguen-type assumptions*, Revue d'analyse numerique et de theorie de l'approximation, **25**, nos. 1-2, pp. 11-22, 1996.

[3] ARGYROS, I. K., *Polynomial Operator Equations in Abstract Spaces and Applications*, CRC Press LLC, Boca Raton, Florida, 1998.

[4] ARGYROS, I. K. and SZIDAROVSZKY, F., *The theory and Applications of Iteration Methods*, C.R.C. Press, Boca Raton, Florida, 1993.

[5] BALĂZS, M. and GOLDNER, G., *On the approximate solution of equations in Hilbert space by a Steffensen-type method*, Revue d'analyse numerique et de theorie de l'approximation, **17**, no. 1, pp. 19–23, 1988.

[6] CĂTINAŞ, E., *On Some Steffensen-type iterative methods for a class of nonlinear equations*, Revue d'analyse numerique et de theorie de l'approximation, **24**, nos. 1–2, pp. 37–43, 1995.

[7] GRAVES, L. M., *Riemann integration and Taylor's theorem in general analysis*, Trans. Amer. Math. Soc., **29**, pp. 163–177, 1927.

[8] KANTOROVICH, L. V. and AKILOV, G. P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[9] PĂVĂLOIU, I., *Sur la méthode de Steffensen pour la résolution des équations_ opérationnelles non linéaires*, Rev. Roum. Math. Pure et Appl., **XIII**, no. 6, pp. 857–861, 1968.

[10] PĂVĂLOIU, I., *Sur une généralisation de la methode de Steffensen*, Revue d'analyse numerique et de theorie de l'approximation, **21**, no. 1, pp. 59–65, 1992.

[11] PĂVĂLOIU, I., *Bilateral approximations for the solutions of scalar equations*, Revue d'analyse numerique et de theorie de l'approximation, **23**, no. 1, pp. 95–100, 1994.

Received January 27, 2000

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