

ACCESSIBILITY OF SOLUTIONS OF EQUATIONS ON BANACH SPACES BY NEWTON-LIKE METHODS AND APPLICATIONS

BY

IOANNIS K. ARGYROS

Abstract. We provide sufficient conditions for the convergence of Newton-like methods to a locally unique solution of an equation on a Banach space. We use the concept of the degree of logarithmic convexity in connection with the fixed point theorem to extend the region of convergence given so far for these methods. In the case of quadratic equations we find a ring that contains accessibility points for Newton's method lying outside the sphere of convergence given by the Newton-Kantorovich Theorem. Our results extend the region of accessibility of solutions by Newton's method for some quadratic integral equations appearing in radiative transfer [1], [5], [9].

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

$$(1) \quad F(x) = 0$$

where F is an operator defined on a closed convex domain D of a Banach space E_1 with values in a Banach space E_2 . We use Newton-like methods of the form

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

to generate an iteration $\{x_n\}(n \geq 0)$ converging to x^* . Here $A(x) \in L$

Received by the editors March 24, 1997.

AMS 1991 Subject Classification: 65J15, 47H17, 49D15.

Key words and phrases: Fixed point, Newton-like methods, Banach space, logarithmic convexity, region of convergence.

$(E_1, E_2)(x \in D)$ which is the space of bounded linear operators from E_1 into E_2 . For $A(x) = F'(x)(x \in D)$ we obtain Newton's method [3], [4], [8]. Several other choices for A are also possible [4], [8]. We define the operator $P : D \subseteq E_1 \rightarrow E_2$ by

$$(3) \quad P(x) = x - A(x)^{-1}F(x),$$

in which case (2) can also be written as

$$(4) \quad x_{n+1} = P(x_n) \ (n \geq 0)(x_0 \in D).$$

Sufficient conditions for the convergence of iteration $\{x_n\}(n \geq 0)$ to x^* have been given by several authors (see, for example, [4], [8] and the references there).

A solution x^* of equation (1) is said to be accessible from x_0 by Newton-like method (2) if

$$(5) \quad x^* = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} P^n(x_0).$$

The region of accessibility of x^* by method (2) is defined to be the set of all x^0 such that (5) is true.

Let us define operator $L_F \in L(E_1, E_2)$ by

$$(6) \quad L_F(x) = P'(x) \ (x \in D).$$

This operator is the degree of logarithmic convexity of F in x and is a measure of the convexity of the function. It was used in [6], [7] in the special case when $A(x) = F'(x) \ (x \in D)$. These convergence results were used to find starting points x_0 lying outside previously found convergence regions, for which (2) converges to x^* in this case. However this was done only for scalar as well as systems of equations when $E_1 = E_2 = R$. This is because for convergence we need to show $\|L_F(x)\| \leq c < 1$, and this is a very difficult problem in general.

Here we provide sufficient convergent conditions for our method (2) to a locally unique solution x^* of equation (1). Our results reduce to the

corresponding ones in [6]. Moreover we show how to compute c for quadratic equations on E_1 . We also suggest how to compute c for polynomial equations on E_1 of degree $k \in \mathbb{N}$.

Finally we show how to apply our results to solve quadratic integral equations appearing in radiative transfer [1], [3], [4], [5], [8], [9].

2. Convergence analysis. Using contraction mapping techniques we obtain the semilocal convergence results:

Theorem 1. *Let $F : D \subseteq E_1 \rightarrow E_2$ be Fréchet-differentiable on a closed convex domain D , $A(x) \in L(E_1, E_2)$ for all $x \in D$. Assume:*

- (a) *linear operator $\Gamma(x) + A(x)^{-1}$ exists and is differentiable for all $x \in D$;*
- (b) *linear operator $L_F(x)$ exists on D and*

$$(7) \quad \|L_F(x)\| \leq c < 1 \quad \text{for all } x \in D;$$

- (c) *for $x_0 \in D$, $r^* \geq \frac{\|x_0 - P(x_0)\|}{1-c}$, $U(x_0, r^*) = \{x \in E_1 \mid \|x - x_0\| \leq r^*\} \subseteq D$.*

Then Newton-like iteration $\{x_n\}$ ($n \geq 0$) generated by (2) is well defined, remains in $U(x_0, r^)$ for all $n \geq 0$ and converges to a fixed point x^* of P in $U(x_0, r^*)$ which is unique in D . Moreover the following error bounds are true for all $n \geq 0$*

$$(8) \quad \|x_n - x^*\| \leq c^n r^*.$$

Proof. Newton-like iteration $\{x_n\}$ ($n \geq 0$) is well defined on D for all $x_0 \in D$ since linear operator $A(x)$ is invertible on D . Using induction on $n \geq 0$ we can show

$$(9) \quad x_n \in U(x_0, r^*) \quad \text{and} \quad \|x_n - x_0\| \leq (1 - c^n) r^* < r^*.$$

For $n = 1$ and hypothesis (c) we have $\|x_1 - x_0\| \leq (1 - c) r^* < r^*$, which shows (9) in this case. Assume that (9) is true for all positive integers smaller or equal to n . Then we must show

$$x_{n+1} \in U(x_0, r^*) \quad \text{and} \quad \|x_{n+1} - x_0\| \leq (1 - c^{n+1}) r^* < r^*.$$

By hypothesis (b) and (4) we get

$$\begin{aligned}
 (10) \quad & \|x_{n+1} - x_n\| = \|P(x_n) - P(x_{n-1})\| \\
 & \leq \sup_{z \in [x_{n-1}, x_n]} \|P'(z)\| \|x_n - x_{n-1}\| \leq c \|x_n - x_{n-1}\| \\
 & \leq \cdots \leq c^n \|x_1 - x_0\| = c^n (1 - c)r^*,
 \end{aligned}$$

and

$$\begin{aligned}
 \|x_{n+1} - x_0\| & \leq \|x_{n+1} - x_n\| + \|x_n - x_0\| \leq c^n (1 - c)r^* + (1 - c^n)r^* \\
 & = (1 - c^{n+1})r^* < r^*,
 \end{aligned}$$

which completes the induction. Moreover by (10) we obtain, for $n, m \in N$

$$(11) \quad \|x_{n+m} - x_n\| \leq (1 - c^m)c^n r^*.$$

Estimate (11) shows that $\{x_n\}$ ($n \geq 0$) is a Cauchy sequence in a Banach space E_2 and as such it converges to a limit $x^* \in U(x_0, r^*)$ (since $U(x_0, r^*)$ is a closed set). By taking the limit as $n \rightarrow \infty$ in (4) and using the continuity of F and $A(x)$ we deduce $P(x^*) = x^*$. To show uniqueness, let $y^* \in D$ with $P(y^*) = y^*$. Then we can get

$$\begin{aligned}
 \|x^* - y^*\| & = \|P(x^*) - P(y^*)\| \leq \sum_{z \in [x^*, y^*]} \|P'(z)\| \|x^* - y^*\| \\
 & \leq c \|x^* - y^*\|
 \end{aligned}$$

which implies that $x^* = y^*$ (since $c \in [0, 1)$).

Finally, letting $n \rightarrow \infty$ in (11) we obtain (8).

That completes the proof of the Theorem.

Following [4], [6], [7], the region of accessibility to x^* is extended to a closed ball around x_0 as the following result indicates:

Theorem 2. *Consider the iteration $y_{n+1} = P(y_n)$ for $y_0 \in U(x_0, r^*)$ under the hypotheses of Theorem 1. Then iteration $\{y_n\}$ ($n \geq 0$) is well defined, remains in $U(x_0, r^*)$ and converges to a unique fixed point x^* of P in $U(x_0, r^*)$. Moreover the following error bounds are true for all $n \geq 0$:*

$$\|y_n - x^*\| \leq \frac{c^n}{1-c} \|y_1 - y_0\| \quad \text{and} \quad \|y_n - x^*\| \leq c^n \|x^* - y_0\|.$$

Proof. The result follows immediately by the contraction mapping principle [4], [8] provided we show that operator P maps $U(x_0, r^*)$ into itself. Indeed let $x \in U(x_0, r^*)$, then we obtain

$$\|x_1 - P(x)\| = \|P(x_0) - P(x)\| \leq c\|x_0 - x\| \leq cr^*.$$

That completes the proof of the Theorem.

For $D = U(v_0, r_0)$ we can obtain immediately from Theorem 2.

Corollary. *If $\|v_0 - x_0\| \leq r_0 - r^*$ under the hypotheses of Theorem 2 Newton-like iteration $\{x_n\}$ ($n \geq 0$) converges to x^* for any starting point in $U(x_0, r^*)$.*

In terms of the degree of logarithmic convexity we have the following result concerning the convergence on Newton-like method (2).

Theorem 3. *Assume that hypothesis (b) of Theorem 1 holds on $D = U(v_0, r_0)$. If $\|A(v_0)^{-1}F(v_0)\| \leq (1-c)r_0$, Newton-like method $\{x_n\}$ ($n \geq 0$) generated by (2) converges to the unique solution x^* of equation $F(x) = 0$ in D for any $x_0 \in D$.*

Proof. We note that a fixed point P is a solution of equation $F(x) = 0$. The result now follows immediately from the proofs of the previous Theorems.

Remark 1. For $A(x) = F'(x)$ ($x \in D$) Theorem 1, 2, Corollary, Theorem 3 reduce respectively to Theorems 2.1, 2.2, Corollary, Theorem 2.4 in [6].

3. Applications to quadratic equations. The verification of condition (b) of Theorem 1 is a very hard problem in general. In [6], [7] and the references there the authors verified this condition for scalar as well as systems of real or complex equations. Here we suggest a possible extension of our results in the case of quadratic equations of the form

$$(12) \quad F(x) = y + B(x, x) - x$$

where B is a bounded symmetric bilinear operator on $D \subset E_1$ and $y \in E_1$ is fixed. Hence in the case of F given by (12) we obtain from (6) for $A(x) = F'(x)(x \in D)$

$$(13) \quad L_F(x)(z) = 2(2B(x) - I)^{-1}B(2B(x) - I)^{-1}(z)(-x + y + B(x, x)).$$

Let $x_0 \in D$ be such that $(F'(x_0))^{-1} = (2B(x_0) - I)^{-1}$ exists and set $b \geq \|(2B(x_0) - I)^{-1}B\| \neq 0$. Let $r \in [0, \frac{1}{2b})$, and assume $U(x_0, r) \subseteq D$. Then for $x \in U(x_0, r)$ we have

$$(14) \quad \begin{aligned} 2B(x) - I &= (2B(x) - I) - (2B(x_0) - I) + (2B(x_0) - I) \\ &= (2B(x_0) - I)[I + 2(2B(x_0) - I)^{-1}B(x - x_0)], \end{aligned}$$

and

$$(15) \quad \|2(2B(x_0) - I)^{-1}B(x - x_0)\| \leq 2\|(2B(x_0) - I)^{-1}B\|\|x - x_0\| \leq 2br < 1.$$

It follows from the Banch Lemma on invertible operators [4], [8] that $F'(x) = 2B(x) - I$ is invertible on $U(x_0, r)$ and

$$(16) \quad \|(2B(x) - I)^{-1}(2B(x_0) - I)\| \leq (1 - 2br)^{-1}.$$

Moreover we have by (12)

$$\begin{aligned} (2B(x_0) - I)^{-1}F(x) &= (2B(x_0) - I)^{-1}[(x_0 - x) + (y - x_0) \\ &\quad + B((x - x_0) + x_0, (x - x_0) + x_0)] \\ &= (2B(x_0) - I)^{-1}(2B(x_0) - I)(x - x_0) \\ &\quad + [(2B(x_0) - I)^{-1}B](x - x_0, x - x_0) \\ &\quad + (2B(x_0) - I)^{-1}(y - x_0 + B(x_0, x_0)). \end{aligned}$$

By taking norms, using the triangle inequality and (4) we get

$$(17) \quad \begin{aligned} &\|(2B(x_0) - I)^{-1}F(x)\| \\ &\leq \|x - x_0\| + \|(2B(x_0) - I)^{-1}B\| \cdot \|x - x_0\|^2 + \|x_0 - P(x_0)\| \\ &\leq r + br^2 + \|x_0 - P(x_0)\|. \end{aligned}$$

It follows from (13), (16) and (17) that

$$(18) \quad \|L_F(x)\| \leq c(r), \quad r \in \left[0, \frac{1}{2b}\right),$$

where

$$(19) \quad c(r) = \frac{2b}{(1-2br)^2}(r + \delta + br^2), \quad \delta \geq \|x_0 - P(x_0)\|.$$

Define the scalar function h on $[0, +\infty)$ by

$$(20) \quad h(r) = c_1 r^2 + c_2 r + c_3$$

where

$$c_1 = -2b^2, \quad c_2 = 6br \quad \text{and} \quad c_3 = 2b\delta - 1.$$

It is simple algebra to show that $c(r) \in [0, 1)$ if

$$(21) \quad h(r) < 0 \quad \text{and} \quad r \in \left[0, \frac{1}{2b}\right).$$

It can easily be seen that (21) is true if

$$(22) \quad 2b\delta \leq q < 1$$

and

$$r \in [0, a)$$

where

$$(23) \quad a = \frac{3 - \sqrt{7 + 2q}}{2b}.$$

Note that estimate (22) is the Newton-Kantorovich hypothesis for equation (12) and a is the smallest zero of the scalar equation $h(r) = 0$ where h is given by (20) [4], [8].

Hence we arrive at:

Theorem 4. *Let $F : U^0(x_0, a) \subseteq E_1 - E_2$ be given by (12), and $A(x) = F'(x)(x \in D)$ in (2). Assume that the Newton-Kantorovich hypothesis (22)*

is true for some $x_0 \in U^0(x_0, a)$ at which $F'(x_0)$ is inverible. Then (7) is true for all $r \in [0, a)$. Moreover if there exists a minimum nonnegative number $r^* \in [0, a)$ satisfying the inequality

$$(24) \quad r \geq \frac{\delta}{1 - c(r)},$$

then thhe conclusions of Theorem 1 for equation (12) and iteration (2) are true.

Proof. It follows immediately from the above discussion, the proof of Theorem 1 and the observation that (24) is true if $g(r^*) \geq 0$ where g is a function defined on $[0, +\infty)$ given by

$$(25) \quad g(r) = d_1 r^3 + d_2 r^2 + d_3 r + d_4,$$

$$d_1 = 2b^2, d_2 = -2b(3 + q), d_3 = 1 + q \text{ and } d_4 = -\delta.$$

Remark 2. By Descarte's rule of signs the equation $g(r) = 0$ has three positive zeros or one. Let s denote the smallest such zero in either case. We note that equation (25) can have zeros in $[0, a)$ even if $g(a) < 0$. However it is simple algebra to check that $g(a) < 0$. Hence we can set $r^* = s$ in this case.

Remark 3. Another approach will be to define the function g_1 on $[0, +\infty)$ by $g_1(r) = d_2 r^2 + d_3 r + d_4$. We note that by (25) $g_1(r) \geq 0$ implies $g(r) \geq 0$ for all $r \in [0, +\infty)$. The discriminant of this quadratic polynomial is nonnegative if $q \in [0, \frac{2\sqrt{7}-5}{3}]$. Let $t_1 \leq t_3$ be the real zeros of the equation $g_1(r) = 0$. Then we easily deduce $t_1 < a$. Hence in this case we can set $r^* = t_1$ in Theorem 4.

The Newton-Kantorovich Theorem [4], [8] for equation (12) asserts that if hypothesis (22) is satisfied then $x^* \in U(x_0, r_k)$ where $r_k = \frac{1-\sqrt{1-q}}{b}$. We easily show:

- (i) if $q \in [0, \frac{2\sqrt{13}-5}{9})$ then $r_k < a$;
- (ii) if $q \in [\frac{2\sqrt{13}-5}{9}, 1]$ then $r_k \geq a$;
- (iii) $g(r_k) \leq 0$ and $g(r_k) = 0$ if $q = 0$ ($r_k = 0$ in this case).

At the end of this study we provide an example where $x^* \in U(x_0, r_k) \subseteq U(x_0, t_1)$.

Remark 4. Theorem 4 is a crude application of Theorem 1. In practice one hopes that (7) will be satisfied in cases that do not imply the Newton-Kantorovich hypothesis (22). Examples where Newton's method converges but (22) is violated where given in [6], [7] for scalar or systems of real equations and in [1], [3], [4] for quadratic integral equations on various Banach spaces. See also the example that follows.

Remark 5. Concluding we note that both Newton-Kantorovich and Theorem 4 apply if condition (22) is satisfied. However the balls centered at the same point x_0 that contain the solution x^* are not of the same radius.

Let the ring $U = U(x_0, r^*) - U(x_0, r_k) \neq \emptyset$ then there exists a starting point $w_0 \notin U(x_0, r_k)$ such that Newton's method (2) converges. However the Newton-Kantorovich Theorem [4], [8] does not guarantee convergence in this case. Hence there exists a region of accessibility for the convergence of Newton's method that is missed by the Newton-Kantorovich Theorem [1], [3], [4], [5], [8], [9], [10]. We confront such a case in the example that follows.

Example. Special cases of (12) are quadratic integral equations of the form

$$(26) \quad F(x)(s) = y(s) + \lambda x(s) \int_0^1 k(s, t)x(t)dt - x(s) = 0$$

in the space $E_2 = E_1 = C[0, 1]$ of all functions continuous on the interval $[0, 1]$ with norm

$$\|x\| = \max_{s \in [0, 1]} |x(s)|.$$

Here we assume that λ is a number called the "abledo" for scattering and the kernel $k(s, t)$ is a continuous function of two variables $s, t \in [0, 1]$ satisfying

$$0 \leq k(s, t) \leq 1, \quad k(0, 0) = 1 \quad \text{and} \quad k(s, t) + k(t, s) = 1, \quad s, t \in (0, 1).$$

Theu function $y(s)$ is a given continuos function defined on $[0, 1]$, and $x(s)$ is the unknown function sought in $[0, 1]$.

Equations of this type are closely related with the works of S. Chandrasekhar [5], (Nobel prize of Physics 1983), and arise in the theories of radiative transfer, neutron transport and in the kinetic theory of gases in connection with the problem of determination of the angular distribution of the radiant flux emerging from a plane radiation field [1], [3], [4], [5], [9].

To apply theorem 4 we need to compute b and δ initially. For example choose $K(s, t) = \frac{s}{s+t}$, $K(0, 0) = 1$, $s, t \in (0, 1]$, and define

$$B(x, y)(s) = \frac{1}{2}\lambda \left[x(s) \int_0^1 \frac{s}{s+t} y(t) dt + y(s) \int_0^1 \frac{s}{s+t} x(t) dt \right].$$

Then B is a symmetric, bounded, bilinear operrator on E_1 with

$$(27) \quad B(x, x)(s) = \lambda x(s) \int_0^1 \frac{s}{s+t} x(t) dt$$

and

$$(28) \quad \|B\| = |\lambda| \max_{s, t \in [0, 1]} \left| \int_0^1 \frac{s}{s+t} dt \right| = |\lambda| \ln 2.$$

Note that the choice of B given by (27) shows that (26) is indeed a special case of equation (12). The values of b , δ for various choices of λ have been given in [1], [3], [4], [5], [9] and the references there.

Choose as an example $x_0(s) = y(s) = 1$ for all $s \in [0, 1]$, $\lambda = .2$, then from (22), (23), Remark 3, (30) we obtain,

$$\|F'(1)^{-1}\| \leq 1.3836213, \quad a = .8513131,$$

$$b = \delta \leq .1918106, \quad q = .0735826, \quad r_k = .1954752,$$

$$t_1 = .242025 = r^* \quad \text{and} \quad c(r^*) = .2074761.$$

All hypotheses of Theorem 4 are satisfied for $f^* = t_1$ and $U(x_0, r_k) \subseteq U(x_0, t_1)$ (also see Remar 5).

Finally the results of Theorem 4 can be extended to include polynomial operator equations of degree $k \in N$ given by

$$F(x) = M_0 + M_1(x) + M_2(x, x) + \cdots + M_k(x, x, \dots, x) - x = 0,$$

where M_i is a bounded, symmetric i -linear operator $i = 1, 2, \dots, k$ and $M_0 \in E_1$ is fixed [1], [2], [4], [8], [9]. For the computational details in deriving the crucial functions $c(r)$ and $h(r)$ in this case we refer the reader especially to [2].

Conclusion. We provide sufficient conditions for the convergence of Newton-like methods to a locally unique solution of an equation on a Banach space. We use the concept of the degree of logarithmic convexity in connection with the fixed point theorem to extend the region of convergence given so far for these methods. In the case of quadratic equations we find a ring that contains accessibility points for Newton's method lying outside the sphere of convergence given by the Newton-Kantorovich Theorem. Some applications of our results to the solution of quadratic integral equations appearing in radiative transfer in connection with the problem of determination of the angular distribution of the radiant flux emerging from a plane radiation field are given. Relevant work can be found especially in [1], [4], [5], [9], [10].

References

1. I. K. Argyros, *Quadratic equations and applications to Chandrasekhar's and related equations*. Bull. Austral. Math. Soc., **32** (1985), 275-292.
2. I. K. Argyros, *On the approximation of some nonlinear equations*, Aequationes Mathematicae, **32** (1987), 87-95.
3. I. K. Argyros, *On a class of nonlinear equations arising in neutron transport*, Aequationes Mathematicae, **36** (1988), 99-111.
4. I. K. Argyros and F. Szidarovszky, *The Theory and Applications of Iteration Methods*, C. R. C. Press, Inc., Boca Raton, Florida, 1993.
5. S. Chandrasekhar, *Radiative Transfer*, Dover Publ., New York, 1960.
6. J. M. Gutierrez, M. A. Hernandez and M. A. Salanova, *Accessibility of solutions by Newton's method*, Intern. J. Computer. Math., **57** (1995), 239-247.
7. M. A. Hernandez, *A note on Halley's method*, Numer. Math., **54** (1991), 273-276.

8. L. V. Kantorovich and G. P. Akilov, *Funnnctional Analysis*, Pergamon Press, New York, 1982.
9. D. K. Ruch, *On uniformly contractive systems and quadratic equations in Banach spaces*, Bull. Austral. Math. Soc., **32** (1985), 275-292.
10. F. Szidarovszky and O. Palusinski, *A special matrix equation and its applications in microelectronics*, Appl. Math. Comp., **64** (1994), 115-119.

Department of Mathematics, Cameron University, Lawton, OK 73505, U.S.A.