JOURNAL OF NUMERICAL ANALYSIS AND APPROXIMATION THEORY
J. Numer. Anal. Approx. Theory, vol. 44 (2015) no. 2, pp. 146–153
ictp.acad.ro/jnaat

A SIMPLIFIED PROOF OF THE KANTOROVICH THEOREM FOR SOLVING EQUATIONS USING TELESCOPIC SERIES

IOANNIS K. ARGYROS[1] and HONGMIN REN[2]

Abstract. We extend the applicability of the Kantorovich theorem (KT) for solving nonlinear equations using Newton-Kantorovich method in a Banach space setting. Under the same information but using elementary scalar telescopic majorizing series, we provide a simpler proof for the (KT) [2], [7]. Our results provide at least as precise information on the location of the solution. Numerical examples are also provided in this study.

MSC 2010. 65J15, 65G99, 47H99, 49M15.
Keywords. Newton-Kantorovich method; Banach space; majorizing series, telescopic series, Kantorovich theorem.

1. INTRODUCTION

Newton's method is one of the most fundamental tools in computational analysis, operations research, and optimization [1, 2, 4, 7–11]. One can find applications in management science; industrial and financial research; data mining; linear and nonlinear programming. In particular interior point algorithms in convex optimization are based on Newton's method.

The basic idea of Newton's method is linearization. Suppose $F : \mathbb{R} \to \mathbb{R}$ is a differentiable function, and we would like to solve equation

$$F(x) = 0. \tag{1.1}$$

Starting from an initial guess, we can have the linear approximation of $F(x)$ in the neighborhood of $x_0 : F(x_0 + h) \approx F(x_0) + F'(x_0)h$, and solve the resulting linear equation $F(x_0)+F'(x_0)h = 0$, leading to the recurrent method

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad (n \geq 0). \tag{1.2}$$

This is Newton's method as proposed in 1669 by I. Newton (for polynomial only). It was J. Raphson, who proposed the usage of Newton's method for general functions F. That is why the method is often called the Newton-Raphson method. Later in 1818, Fourier proved that the method converges

[1] Department of Mathematical Sciences, Cameron University, Lawton, Oklahoma 73505-6377, USA, e-mail: ioannisa@cameron.edu.
[2] College of Information and Engineering, Hangzhou Polytechnic, Hangzhou 311402, Zhejiang, PR China, e-mail: rhm65@126.com.

quadratically in a neighborhood of the root, while Cauchy (1829, 1847) provided the multidimensional extension of Newton's method (1.2). In 1948, L. V. Kantorovich published an important paper [7] extending Newton's method for functional spaces (the Newton-Kantorovich method (NKM)). That is $F : D \subseteq X \to Y$, where X, Y are Banach spaces, and D is a open convex set [1, 7]. Ever, since thousands of papers have been written in a Banach space setting for the (NKM) as well as Newton-type methods, and their applications. We refer the reader to the publications [1–11] for recent results (see also, the references there).

It is stated in the (KT) that Newton's method (1.2) converges provided the famous for its simplicity and clarity Kantorovich hypotheses (KH) (see (C_6)) is satisfied. (KH) uses the information (x_0, F, F'). Any successful attempt for weakening (KH) under the same information is extremely important in computational mathematics, since that will imply the extension of the applicability of (NKM). We have already provided conditions weaker than (KH) [1, 2] by introducing the center Lipichitz condition, which is a special case of the Lipschitz condition.

In this study we present a new proof of the (KT) for Newton's method using telescopic majorizing sequences. Our proof is simpler than the corresponding one provided by Kantorovich [7]. Section 2 contains the semilocal convergence of Newton's method (1.2). Numerical examples where our results apply to solve nonlinear equations are provided in Section 3.

2. SEMILOCAL CONVERGENCE OF NEWTON-KANTOROVICH METHOD (NKM)

Let us assume that $F'(x_0)^{-1} \in L(Y, X)$ at some $x_0 \in D$, and the following conditions hold:

(C_1) $\quad 0 < \|F'(x_0)^{-1}\| \leq \beta$,

(C_2) $\quad 0 < \|F'(x_0)^{-1}F(x_0)\| \leq \eta$,

and Lipschitz continuity condition

(C_3) $\quad \|F'(x) - F'(y)\| \leq L\|x - y\| \quad$ for some $L > 0$, and for all $x, y \in D$.

It is convenient for us to define for $a_0 = b_0 = 1$,

$$\gamma = \beta L \eta, \tag{2.1}$$

and scalar sequences

$$a_{n+1} = \frac{a_n}{1 - \gamma a_n b_n}, \tag{2.2}$$

$$c_n = \tfrac{1}{2} L b_n^2, \tag{2.3}$$

$$b_{n+1} = \beta a_{n+1} c_n \eta. \tag{2.4}$$

We shall find a connection between (NKM) $\{x_n\}$, and scalar sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$. Sequences $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ have been used by Kantorovich [7]. However, we present a new proof for the (KT).

Lemma 2.1. *Under the* $(C_1)-(C_3)$ *conditions further suppose:*
(C_4) $\quad x_n \in D$,
and
(C_5) $\quad \gamma a_n b_n < 1$.
Then, the following estimates hold:
(I_n) $\quad \|F'(x_n)^{-1}\| \leq a_n\beta$,
(II_n) $\quad \|x_{n+1}-x_n\| = \|F'(x_n)^{-1}F(x_n)\| \leq b_n\eta$,
and
(III_n) $\quad \|F(x_{n+1})\| \leq c_n\eta^2$.

Proof. We shall use induction to show items $(I_n)-(III_n)$. (I_0) and (II_0) follow immediately from the initial conditions. To show (III_0), we use (1.2) for $n=0$, (II_0), and (C_3) to obtain

$$\begin{aligned} F(x_1) &= F(x_1)-F(x_0)-F'(x_0)(x_1-x_0) \\ &= \int_0^1 [F'(x_0+t(x_1-x_0))-F'(x_0)](x_1-x_0)dt \end{aligned} \tag{2.5}$$

$\Rightarrow$

$$\begin{aligned} \|F(x_1)\| &= \Big\| \int_0^1 [F'(x_0+t(x_1-x_0))-F'(x_0)](x_1-x_0)dt \Big\| \\ &\leq \int_0^1 \|[F'(x_0+t(x_1-x_0))-F'(x_0)]\|dt\|x_1-x_0\| \\ &\leq L\|x_1-x_0\|^2 \int_0^1 tdt = \tfrac{L}{2}\|x_1-x_0\|^2 \\ &\leq \tfrac{L}{2}b_0^2\eta^2 = c_0\eta^2. \end{aligned} \tag{2.6}$$

If $x_{k+1} \in D$ $(k \leq n)$, then it follows from $(C_3)-(C_5)$ and the induction hypotheses:

$$\begin{aligned} \|F'(x_k)^{-1}\|\|F'(x_{k+1})-F'(x_k)\| &\leq a_k\beta L\|x_{k+1}-x_k\| \\ &\leq a_k\beta L b_k\eta = \gamma a_k b_k < 1. \end{aligned} \tag{2.7}$$

In view of (2.7), and the Banach lemma on invertible operators [1, 2, 7] $F'(x_{k+1})^{-1} \in L(Y,X)$, so that

$$\begin{aligned} \|F'(x_{k+1})^{-1}\| &\leq \frac{\|F'(x_k)^{-1}\|}{1-\|F'(x_k)^{-1}\|\|F'(x_{k+1})-F'(x_k)\|} \\ &\leq \frac{a_k\beta}{1-\gamma a_k b_k} = a_{k+1}\beta, \end{aligned} \tag{2.8}$$

which shows (I_n) for all $n \geq 0$.

As in (2.5), we have

$$\begin{aligned} F(x_{k+1}) &= F(x_{k+1})-F(x_k)-F'(x_k)(x_{k+1}-x_k) \\ &= \int_0^1 [F'(x_k+t(x_{k+1}-x_k))-F'(x_k)](x_{k+1}-x_k)dt \end{aligned} \tag{2.9}$$

and

$$\|F(x_{k+1})\| \leq \tfrac{L}{2}\|x_{k+1}-x_k\|^2 \leq \tfrac{L}{2}b_k^2\eta^2 = c_k\eta^2, \tag{2.10}$$

which shows (III_n) for all $n \geq 0$. Moreover, we get by (1.2), (2.8) and (2.10):
(2.11)

$$\begin{aligned}\|x_{k+2} - x_{k+1}\| &= \|F'(x_{k+1})^{-1}F(x_{k+1})\| \leq \|F'(x_{k+1})^{-1}\|\|F(x_{k+1})\| \\ &\leq \beta a_{k+1} c_k \eta^2 = b_{k+1}\eta.\end{aligned}$$

That completes the induction for (II_n) and the proof of the lemma. □

We need to show the convergence of sequence $\{x_n\}$, which is equivalent to proving that $\{b_n\}$ is a Cauchy sequence. To this effect we need the following auxiliary results:

LEMMA 2.2. *Suppose:*
there exists $x_0 \in D$ such that
(C_6) $2\gamma \leq 1$, *where, γ is given by* (2.1)
and Condition (C_4) holds. Then, the following assertions hold:
(a) Scalar sequence $\{a_n\}$ increases. In particular, (C_5) holds.
(b) $\lim_{n\to\infty} b_n = 0$.
(c)

$$r := \sum_{k=0}^{\infty} b_k = \frac{2}{1 + \sqrt{1 - 2\gamma}}. \tag{2.12}$$

(d) (C_4) holds if $\overline{U}(x_0, r\eta) \subseteq D$.

Proof. (a) We shall show using induction that $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ are positive sequences.

In view of the initial conditions, a_0, b_0, c_0 are positive, and $1 - \gamma a_0 b_0 > 0$. Assume a_k, b_k, c_k and $1 - \gamma a_k b_k$ are positive for $k \leq n$. By (2.2), $a_{k+1} > 0$, consequently $b_{k+1} > 0$. Moreover, $1 - \gamma a_{k+1} b_{k+1} > 0$ by (C_6) (see also [2], [7]). The induction is completed.

Solving (2.2) for b_n, we obtain

$$b_n = \frac{1}{\gamma}\left(\frac{1}{a_n} - \frac{1}{a_{n+1}}\right). \tag{2.13}$$

By telescopic sum, we have:

$$\sum_{k=0}^{n-1} b_k = \frac{1}{\gamma}\left(\frac{1}{a_0} - \frac{1}{a_n}\right) = \frac{1}{\gamma}\left(1 - \frac{1}{a_n}\right), \quad \textit{since } a_0 = 1, \tag{2.14}$$

or

$$a_n = \frac{1}{1 - \gamma \sum_{k=0}^{n-1} b_k}. \tag{2.15}$$

But $1 - \gamma \sum_{k=0}^{n-1} b_k$ decreases, so $\{a_n\}$ given by (2.15) increases. Note also that $a_n \geq a_0 = 1$.

(b) By (a) $a_n \geq 1$ is increasing, so that

$$0 < \frac{1}{a_n} \leq 1. \tag{2.16}$$

Therefore, $\{\frac{1}{a_n}\}$ is monotonic on the compact set $[0,1]$ and as such it converges to some limit $\overline{a}$. By letting $n \to \infty$ in (2.13), we get

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{1}{\gamma}(\frac{1}{a_n} - \frac{1}{a_{n+1}}) = \frac{1}{\gamma}(\overline{a} - \overline{a}) = 0. \tag{2.17}$$

(c) It follows from (b) and (2.14) that there exists $r = \sum_{k=0}^{\infty} b_k$. The value of r is well known to be given by (2.12) [2], [7].

(d) We have $\|x_1 - x_0\| \leq b_0\eta = \eta \Rightarrow x_1 \in U(x_0, r\eta)$. Assume $x_k \in U(x_0, r\eta) \subseteq D$ for all $k \leq n$. Then, we have by Lemma 2.1 in turn that

$$\|x_{k+1} - x_0\| \leq \|x_{k+1} - x_k\| + \cdots + \|x_1 - x_0\| \leq (b_k + \cdots + b_0)\eta < r\eta,$$

$\Rightarrow$

$$x_{k+1} \in U(x_0, r\eta) \subseteq D.$$

That completes the proof of the lemma. □

Remark 2.3. In view of (C_3) there exists $L_0 > 0$ such that

$$\|F'(x) - F'(x_0)\| \leq L_0\|x - x_0\| \quad \text{for all } x \in D.$$

Note that

$$L_0 \leq L$$

holds in general and $\frac{L}{L_0}$ can be arbitrarily large [2].

We can show the semilocal convergence result for (NKM) (1.2).

Theorem 2.4. *Under conditions* $(C_1) - (C_3)$, (C_6) *further assume*

(C_7) $\overline{U}(x_0, r\eta) = \{x \in X | \|x - x_0\| \leq r\eta\} \subseteq D$ *where,* r *is given by* (2.12).

Then, sequence $\{x_n\}$ *generated by (NKM)* (1.2) *is well defined, remains in* $\overline{U}(x_0, r\eta)$ *for all* $n \geq 0$ *and converges to a solution* $x^\star \in \overline{U}(x_0, r\eta)$ *of equation* $F(x) = 0$. *Moreover, the following estimate holds:*

$$\|x_n - x^\star\| \leq \sum_{k=n}^{\infty} b_k\eta < r\eta. \tag{2.18}$$

Furthermore, $x^\star$ *is the only solution of equation* $F(x) = 0$ *in* $D_0 \bigcap D = D_1$ *where,* $D_0 = U(x_0, \frac{1}{\beta L_0})$.,

Proof. It follows from Lemmas 2.1 and 2.2 (see also (II_n)) that $\{x_n\}$ is a Cauchy sequence in a Banach space X and as such it converges to some $x^\star$. We have $\lim_{n\to\infty} b_n = 0$, which implies by (2.3) that $\lim_{n\to\infty} c_n = 0$. By letting $n \to \infty$ in (2.10) and using the continuity of operator F, we obtain $F(x^\star) = 0$.

By (C_7), we obtain

$$\|x_{n+1} - x_0\| \leq \sum_{k=0}^{n} \|x_{k+1} - x_k\| \leq \sum_{k=0}^{n} b_k\eta < r\eta. \tag{2.19}$$

$\Rightarrow x_{n+1} \in \overline{U}(x_0, r\eta)$

$\Rightarrow x^\star = \lim_{n\to\infty} x_n \in \overline{U}(x_0, r\eta)$ (since $\overline{U}(x_0, r\eta)$ is a closed set).

Let $m > n$. Then, we have

$$\|x_n - x_m\| \le \sum_{k=n}^{m-1} \|x_k - x_{k+1}\| \le \sum_{k=n}^{m-1} b_k \eta \le \sum_{k=n}^{\infty} b_k \eta < r\eta. \tag{2.20}$$

By letting $m \to \infty$ in (2.20), we obtain (2.18). Finally, to show uniqueness, let $y^\star \in D_0$ be a solution of equation $F(x) = 0$. Define linear operator

$$M = \int_0^1 F'(x^\star + t(y^\star - x^\star))dt. \tag{2.21}$$

We have using (C_3):

$$\begin{aligned} \|F'(x_0)^{-1}\|\|M - F'(x_0)\| &\le \beta L_0 \int_0^1 \|x^\star + t(y^\star - x^\star) - x_0\|dt \\ &\le \beta L_0 \int_0^1 [(1-t)\|x^\star - x_0\| + t\|y^\star - x_0\|]dt \\ &< \frac{\beta L_0}{2}(r\eta + \frac{1}{\beta L_0}) \le \frac{1}{2} + \frac{1}{2} = 1. \end{aligned} \tag{2.22}$$

It follows from (2.22), and the Banach lemma on invertible operators that $M^{-1} \in L(Y, X)$. In view of (2.21), we get

$$0 = F(y^\star) - F(x^\star) = M(y^\star - x^\star),$$

which implies $x^\star = y^\star$. That completes the proof of the theorem. □

Remark 2.5. If $L_0 = L$, Theorem 2.4 reduces to the (KT). Otherwise, (i.e if $L_0 < L$), then our theorem provides a more precise information on the location of the solution.

3. APPLICATIONS

In the first example we show that the Kantorovich hypothesis (see (3.2)) is satisfied with the bigger uniqueness ball of solution than before [2], [7].

Example 3.1. Let $X = Y = \mathbb{R}$ be equipped with the max-norm. Let $x_0 = 1$, $D = U(x_0, 1-q)$, $q \in [0, 1)$ and define function F on D by

$$F(x) = x^3 - q. \tag{3.1}$$

Then, we obtain that $\beta = \frac{1}{3}$, $L = 6(2-q)$, $L_0 = 3(3-q)$ and $\eta = \frac{1}{3}(1-q)$. Then, the famous for its simplicity and clarity Kantorovich hypothesis for solving equations using (NKM) [1, 2, 7] is satisfied, say for $q = .6$, since

$$h = 2L\beta\eta = \frac{4}{3}(2-q)(1-q) = 0.746666\ldots < 1. \tag{3.2}$$

Hence, (NKM) converges starting at $x_0 = 1$. We also have that $r = 1.330386707$, $\eta = .133\ldots$, $r\eta = .177384894$, $L_0 = 7.2 < L = 8.4$ and $\frac{1}{\beta L_0} = .41666\ldots$. That is our Theorem 2.4 guarantees the convergence of

(NKM) to $x^\star = \sqrt[3]{0.49} = .788373516$ and the uniqueness ball is better than the one given in (KT).

In the second example we apply Theorem 2.4 to a nonlinear integral equation of Chandrasekhar-type.

EXAMPLE 3.2. Let us consider the equation

$$x(s) = 1 + \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}dt, \quad s \in [0,1]. \tag{3.3}$$

Note that solving (3.3) is equivalent to solving $F(x) = 0$, where $F : C[0,1] \to C[0,1]$ defined by

$$[F(x)](s) = x(s) - 1 - \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}dt, \quad s \in [0,1]. \tag{3.4}$$

Using (3.4), we obtain that the Fréchet-derivative of F is given by

$$[F'(x)y](s) = y(s) - \frac{s}{4}y(s)\int_0^1 \frac{x(t)}{s+t}dt - \frac{s}{4}x(s)\int_0^1 \frac{y(t)}{s+t}dt, \quad s \in [0,1]. \tag{3.5}$$

Let us choose the initial point $x_0(s) = 1$ for each $s \in [0,1]$. Then, we have that $\beta = 1.534463572$, $\eta = .2659022747$, $L_0 = L = \ln 2 = .693147181$, $h = 2\gamma = .392066334$ and $r = 1.23784269$ (see also [1,2,3,7]). Then, hypotheses of Theorem 2.4 are satisfied. In consequence, equation $F(x) = 0$ has a solution $x^\star$ in $U(1,\rho)$, where $\rho = r\eta = .298816793$.

ACKNOWLEDGEMENTS. This work was supported by National Natural Science Foundation of China (Grant No. 10871178).

REFERENCES

[1] I.K. ARGYROS, *Polynomial equations in abstract spaces and applications*, St. Lucie/CRC/Lewis Publ. Mathematics Series, 1998, Boch Raton Florida.

[2] I.K. ARGYROS, *Computational theory of iterative methods*, Series: Studies in Computational Mathematics 15, Editors, C.K. Chui and L. Wuytack, Elservier Publ. Co. New York, USA, 2007.

[3] S. CHANDRASEKHAR, *Radiative Transfer*, Dover. Publ, 1960, New York.

[4] P. DEUFLHARD, *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, Springer-Verlag, Berlin, Heidelberg, 2004.

[5] J.A. EZQUERRO, M.A. HERNÁNDEZ, *New Kantorovich-type conditions for Halley's method*, Appl. Num. Anal. Comp. Math., 2 (2005), pp. 70–77.

[6] J.A. EZQUERRO, M.A. HERNÁNDEZ, *An optimization of Chebyshev-method*, J. Complexity, 25 (2009), pp. 343–361.

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

[8] J.H. MATHEWS, *Bibligraphy for Newton's method*, Available from: `http://mathfaculty.fullerton.edu/mathews/n2003/newtonsmethod/Newton'sMethodBib/Links/Newton'sMethodBib_lnk_3.html`.

[9] Yu. Nesterov, *Introductory Lectures on Convex Programming*, Kluwer, Boston, 2004.
[10] S.J. Wright, *Primal-Dual Interior Point Methods*, SIAM, Philadelphia, 1997.
[11] T.J. Ypma, *Historical developments of the Newton-Raphson method*, SIAM Review, 37 (1995), pp. 531–551.

Received by the editors: September 12, 2012.