A Kantorovich-type analysis for a fast iterative method for solving nonlinear equations

Ioannis K. Argyros

Cameron University, Department of Mathematical Sciences, Lawton, OK 73505, USA

Received 5 January 2006
Available online 7 November 2006
Submitted by V.J. Ervin

Abstract


Keywords: Banach space; Majorant principle; Secant method; Local/semilocal convergence; Radius of convergence; Lipschitz conditions; Divided differences

1. Introduction

In this study we are concerned with the problem of approximating a solution \( x^* \) of the nonlinear equation

\[
F(x) = 0,
\]

(1.1)
where $F$ is a Fréchet-differentiable operator defined on an open subset $D$ of a Banach space $X$ with values in a Banach space $Y$.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is driven by the equation $\dot{x} = B(x)$ (for some suitable operator $B$), where $x$ is the state. Then the equilibrium states are determined by solving Eq. (1.1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative—when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

We are interested in numerical methods that avoid the expensive computation of the Fréchet-derivative $F'(x)$ of operator $F$ at each step. H.T. Kung and J.F. Traub [17] introduced a class of multipoint iterative functions without derivative and D. Chen [11] studied a particular class of these methods which contain the Steffensen method [17] as a special case but only in one dimension. S. Amat, S. Busquier and V. Candela [3] generalized these methods in a Banach space setting. Relevant works can be found in [1,2,4,5,9,10,16,18,21].

In particular they considered the Steffensen-type method

$$
\begin{align*}
y_n &= x_n + \alpha_n F(x_n), \quad \alpha_n \in (0, 1], \\
x_{n+1} &= y_n - [x_n, y_n]^{-1} F(y_n).
\end{align*}
$$

(1.2)

Note that if we set $\alpha_n = 1$ ($n \geq 0$) in (1.2) we obtain the Steffensen method. Under certain Kantorovich-type conditions the quadratic semilocal convergence of method (1.2)–(1.3) was established in [3]. F.A. Potra [19]. Argyros [6–8], Amat et al. [1,2], Hernandez et al. [12–14], Gutíerrez et al. [4] have also introduced methods that avoid the usage of the Fréchet derivative.

In [7] we showed the quadratic convergence of the method:

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

(1.4)

Here, a linear operator from $X$ into $Y$, denoted by $[x, y; F]$ or simply $[x, y]$ which satisfies the condition

$$
[x, y](x - y) = F(x) - F(y),
$$

(1.5)

is called a divided difference of order one [8,14]. Iteration (1.4) has a geometrical interpretation similar to the Secant method in the scalar case.

We provide a local as well as a semilocal convergence analysis for method (1.4). Our approach (conditions) differs from the one in [7]. In particular we use the more concrete Newton–Kantorovich convergence analysis approach based on a cubic scalar majorizing polynomial instead of majorizing sequences.

Here we compare method (1.4) with (1.2)–(1.3) since both methods use two function evaluations per step. An apparent restriction of method (1.2)–(1.3) is that the operator $F$ has to map $X$ into itself. Therefore if this cannot happen our method (1.4) can serve as an alternative. Our
method can also be used instead of (1.2)–(1.3) in cases linear operator \([x_n, y_n]\) is not invertible for all \(n \geq 0\) (see e.g. Example 3.3). According to the hypotheses in [3] method (1.2)–(1.3) cannot be used when \(F'(x_0)\) is smaller than 1 (in norm). Apparently our results do not have such a restriction. Note that a local convergence analysis is given here for method (1.4) but such an analysis was not given in [3,11] for method (1.2)–(1.3).

Finally the radius of convergence is compared favorably to the corresponding ones of our methods.

2. Local convergence analysis of method (1.4)

We can show the following local convergence result for method (1.4).

**Theorem 2.1.** Let \(F\) be a nonlinear operator defined on an open subset \(D\) of a Banach space \(X\) with values in a Banach space \(Y\).

Assume:

Equation \(F(x) = 0\) has a solution \(x^* \in D\) at which the Fréchet derivative \(F'(x^*)\) exists, and is invertible;

Operator \(F\) is Fréchet-differentiable with divided difference of order one on \(D_0 \subseteq D\) satisfying the Lipschitz conditions:

\[
\|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq a\|x - x^*\|, \tag{2.1}
\]

\[
\|F'(x^*)^{-1}([x, y] - [x, x^*])\| \leq b\|y - x^*\|, \tag{2.2}
\]

and

\[
\|F'(x^*)^{-1}([y, y] - [2y - x, x])\| \leq c\|y - x\|^2; \tag{2.3}
\]

the ball

\[
U^* = U(x^*, r^*) = \{x \in X \mid \|x - x^*\| < r^*\} \subseteq D_0, \tag{2.4}
\]

where,

\[
r^* = \frac{4}{a + b + \sqrt{(a + b)^2 + 32c}}; \tag{2.5}
\]

for all \(x, y \in D_0\) \(\implies 2y - x \in D_0\). \tag{2.6}

Then, sequence \(\{x_n\} (n \geq 0)\) generated by method (1.4) is well defined, remains in \(U(x^*, r^*)\) for all \(n \geq 0\) and converges to \(x^*\) provided that

\[
x_{-1}, x_0 \text{ belong in } U(x^*, r^*). \tag{2.7}
\]

Its convergence speed can be estimated as:

\[
\|x_{n+1} - x^*\| \leq \frac{b\|x_n - x^*\| + c\|x_{n-1} - x_n\|^2}{1 - a\|x_n - x^*\| - c\|x_{n-1} - x_n\|^2}\|x_n - x^*\| \quad (n \geq 0). \tag{2.8}
\]

**Proof.** Let us denote by \(L = L(x,y)\) the linear operator

\[
L = [2y - x, x]. \tag{2.9}
\]

Assume \(x, y \in U(x^*, r^*)\). We shall show \(L\) is invertible on \(U(x^*, r^*)\), and

\[
\|L^{-1}F'(x^*)\| \leq \left|1 - a\|y - x^*\| - c\|x - y\|^2\right|^{-1} \leq \left|1 - ar^* - 4c(r^*)^2\right|^{-1}. \tag{2.10}
\]
Using (2.1) and (2.3), we obtain in turn:

\[
\| F'(x^*)^{-1}[F'(x^*) - L] \| = \| F'(x^*)^{-1}[[[x^*, x^*] - [y, y]] + ([y, y] - [2y - x, x])] \|
\leq a\|y - x^*\| + c\|y - x\|^2
\leq ar^* + c[\|y - x^*\| + \|x^* - x\|]^2
\leq ar^* + 4c(r^*)^2 < 1,
\]

by the choice of \( r^* \).

It follows from the Banach lemma on invertible operators \([15]\) and (2.11) that \( L^{-1} \) exists on \( U(x^*, r^*) \), so that estimate (2.10) holds. We can also have by (2.2) and (2.3):

\[
\| F'(x^*)^{-1}([y, x^*] - L) \| = \| F'(x^*)^{-1}[[[y, x^*] - [y, y]] + ([y, y] - L)] \|
\leq \| F'(x^*)^{-1}([y, x^*] - [y, y]) \| + \| F'(x^*)^{-1}([y, y] - L) \|
\leq b\|y - x^*\| + c\|y - x\|^2
\leq br^* + 4c(r^*)^2.
\]

Moreover by (1.4) we get

\[
\| x_{n+1} - x^* \| = \| -L_n^{-1}([x_n, x^*] - L_n)(x_n - x^*) \|
\leq \| L_n^{-1}F'(x^*) \| \cdot \| F'(x^*)^{-1}([x_n, x^*] - L_n) \| \cdot \| x_n - x^* \|.
\]

Estimate (2.8) now follows from (2.10), (2.12) and (2.13). Furthermore from (2.8) we get

\[
\| x_{n+1} - x^* \| < \| x_n - x^* \| < r^* \quad (n \geq 0).
\]

Hence, sequence \( \{x_n\} \ (n \geq -1) \) is well defined, remains in \( U(x^*, r^*) \) for all \( n \geq -1 \) and converges to \( x^* \).

That completes the proof of Theorem 2.1. \( \square \)

Let \( x, y, z \in D_0 \), and define the divided difference of order two of operator \( F \) at the points \( x, y \) and \( z \) denoted by \([x, y, z]\) by

\[
[x, y, z](y - z) = [x, y] - [x, z].
\]

**Remark 2.2.** In order for us to compare method (1.4) with others using divided differences of order one, consider the condition

\[
\| F'(x^*)^{-1}([x, y] - [u, v]) \| \leq \tilde{a}(\| x - u \| + \| y - v \|)
\]

instead of (2.1) and (2.2). Note that (2.16) implies (2.1) and (2.2). Moreover we have:

\[
a \leq 2\tilde{a}
\]
and

\[
b \leq 2\tilde{a}.
\]

Therefore stronger but more popular condition (2.16) can replace (2.1) and (2.2) in Theorem 2.1.
Assuming $F$ has divided differences of order two, condition (2.3) can be replaced by the stronger
\begin{equation}
\|F'(x^*)^{-1}(\{y, x, y\} - [2y - x, x, y])(y - x)\| \leq \tilde{c}\|y - x\|^2,
\end{equation}
(2.19)
or the even stronger
\begin{equation}
\|F'(x^*)^{-1}(\{u, x, y\} - [v, x, y])(y - x)\| \leq \bar{c}\|u - v\|^2.
\end{equation}
(2.20)
Note also that
\begin{equation}
c \leq \bar{c}
\end{equation}
(2.21)
and we can set
\begin{equation}
c = \bar{c}
\end{equation}
(2.22)
despite the fact that $\bar{c}$ (or $\bar{\bar{c}}$) is more difficult to compute since we use divided differences of order two (instead of one). Conditions (2.16) and (2.20) were used in [19] to show method
\begin{equation}
y_{n+1} = y_n - (\{y_n, y_{n-1}\} + \{y_{n-2}, y_n\} - \{y_{n-2}, y_{n-1}\})^{-1} F(y_n) \quad (n \geq 0)
\end{equation}
(2.23)
converges to $x^*$ with order 1.839... which is the solution of the scalar equation
\begin{equation}
t^3 - t^2 - t - 1 = 0.
\end{equation}
(2.24)
Potra in [19] has also shown how to compute the Lipschitz constants appearing here in some cases.

It follows from (2.8) that there exist a constant $c_0$, and $N$ a sufficiently large integer such that:
\begin{equation}
\|x_{n+1} - x^*\| \leq c_0\|x_n - x^*\|^2 \quad \text{for all } n \geq N.
\end{equation}
(2.25)
Hence the order of convergence for method (1.4) is two. Note also that the radius of convergence $r^*$ given by (2.5) is larger than the corresponding one given in [19, estimate (22)]. This observation is very important since it allows a wider choice of initial guesses $x_{-1}$ and $x_0$.

It turns out that our convergence radius $r^*$ given by (2.5) can even be larger than the one given by Rheinboldt [20] (see, e.g., [19, Remark 4.2]) for Newton’s method. Indeed under condition (2.6) radius $r^*_R$ is given by
\begin{equation}
r^*_R = \frac{1}{3a}.
\end{equation}
(2.26)
We showed in [6] that $\frac{\tilde{a}}{a}$ (or $\frac{\bar{a}}{a}$) can be arbitrarily large. Hence we can have:
\begin{equation}
r^*_R < r^*.
\end{equation}
(2.27)
In [6] we also showed that $r^*_R$ is enlarged under the same hypotheses and computational cost as in [20].

We note that condition (2.6) suffices to hold only for $x, y$ being iterates of method (1.4) (see, e.g., Example 3.3).

Condition (2.6) can be removed if $D_0 = X$. In this case (2.4) is also satisfied.

Finally delicate condition (2.6) can also be replaced by a stronger but more practical one which we decided not to introduce originally in Theorem 2.1, so we can leave the result as uncluttered-general as possible.

Indeed, define ball $U_1$ by
\begin{equation}
U_1 = U(x^*, R^*) \quad \text{with } R^* = 3r^*.
\end{equation}
(2.28)
If \( x_{n-1}, x_n \in U^* (n \geq 0) \) then we conclude \( 2x_n - x_{n-1} \in U_1 (n \geq 0) \). This is true since it follows from the estimates
\[
\|2x_n - x_{n-1} - x^*\| \leq \|x_n - x^*\| + \|x_n - x_{n-1}\| \\
\leq 2\|x_n - x^*\| + \|x_{n-1} - x^*\| < 3r^* = R^* \quad (n \geq 0).
\]
Hence the proof of Theorem 2.1 goes through if both conditions (2.4), (2.6) are replaced by
\[
U_1 \subseteq D_0. \tag{2.29}
\]
We complete this section with a numerical example to justify estimate (2.27).

**Example 2.3.** Let \( X = Y = \mathbb{R} \), \( x^* = 0 \), \( D = U(0, 1) \) and define function \( F \) on \( D \) by
\[
F(x) = e^x - 1. \tag{2.30}
\]
Using (2.1)–(2.3), (2.16) and (2.30), we obtain
\[
a = b = e - 1, \quad c = e \quad \text{and} \quad \bar{a} = \frac{e}{2}. \tag{2.31}
\]
In view of (2.5) and (2.26), we have
\[
r_R = 0.24525296 < 0.299040145 = r^*. \tag{2.32}
\]
We can also set \( R^* = 3r^* = 0.897120435 \).

### 3. Semilocal convergence of method (1.4)

We can show the following result for the semilocal convergence of method (1.4).

**Theorem 3.1.** Let \( F \) be a nonlinear operator defined on an open set \( D \) of a Banach space \( X \) with values in a Banach space \( Y \).

Assume:

- Operator \( F \) has divided differences of order one and two on \( D_0 \subseteq D \);
- There exist points \( x_{-1}, x_0 \) in \( D_0 \) such that \( 2x_0 - x_{-1} \in D_0 \) and \( A_0 = [2x_0 - x_{-1}, x_{-1}] \) is invertible on \( D_0 \);
- Set \( A_n = [2x_n - x_{n-1}, x_{n-1}] \) \( (n \geq 0) \).
- There exist constants \( \alpha, \beta \) such that:
  \[
  \|A_0^{-1}([x, y] - [u, v])\| \leq \alpha(\|x - u\| + \|y - v\|), \tag{3.1}
  \]
  \[
  \|A_0^{-1}([y, x, y] - [2y - x, x, y])\| \leq \beta\|x - y\| \tag{3.2}
  \]
  for all \( x, y, u, v \in D_0 \), and condition (2.6) holds;
- Define constants \( \gamma, \delta \) by
  \[
  \|x_0 - x_{-1}\| \leq \gamma, \tag{3.3}
  \]
  \[
  \|A_0^{-1}F(x_0)\| \leq \delta, \tag{3.4}
  \]
  \[
  2\beta\gamma^2 \leq 1; \tag{3.5}
  \]

Moreover define \( \theta, r, h \) by
\[ \theta = \left\{ (\alpha + \beta \gamma)^2 + 3\beta(1 - \beta \gamma^2) \right\}^{1/2}, \tag{3.6} \]
\[ r = \frac{1 - \beta \gamma^2}{\alpha + \beta \gamma + \theta}, \tag{3.7} \]

and
\[ h(t) = -\beta t^3 - (\alpha + \beta \gamma)t^2 + (1 - \beta \gamma^2)t, \tag{3.8} \]
\[ \delta \leq h(r) = \frac{1}{3} \left( 1 - 2\beta \gamma^2 \right) r; \tag{3.9} \]
\[ U_0 = U(x_0, r_0) \subseteq D_0, \tag{3.10} \]

where \( r_0 \in (0, r) \) is the unique solution of equation
\[ h(t) = (1 - 2\beta \gamma^2)\delta \tag{3.11} \]
on interval \((0, r)\).

Then, sequence \( \{x_n\} (n \geq -1) \) generated by method (1.4) is well defined, remains in \( U(x_0, r_0) \) for all \( n \geq -1 \) and converges to a solution \( x^* \) of equation \( F(x) = 0 \).

Moreover its speed of convergence can be estimated for all \( n \geq -1 \) as:
\[ \|x_{n+1} - x_n\| \leq t_n - t_{n+1}, \tag{3.12} \]
\[ \|x_n - x^*\| \leq t_n, \tag{3.13} \]

where,
\[ t_{-1} = r_0 + \gamma, \quad t_0 = r_0, \tag{3.14} \]
\[ \gamma_0 = \alpha + 3\beta r_0 + \beta \gamma, \quad \gamma_1 = 3\beta r_0^2 - 2\gamma_0 r_0 - \beta \gamma^2 + 1, \tag{3.15} \]

for \( n \geq 0 \)
\[ t_{n+1} = \frac{\gamma_0 t_n - (t_n - t_{n-1})^2 \beta - 2\beta t_n^2}{\gamma_1 + 2\gamma_0 t_n - (t_n - t_{n-1})^2 - 3\beta t_n^2} \cdot t_n. \tag{3.16} \]

Furthermore if
\[ r_0 \leq r_1, \tag{3.17} \]

and
\[ \alpha(2\gamma + r_0 + r_1) < 1, \tag{3.18} \]

\( x^* \) is the unique solution of Eq. (1.1) in \( \overline{U}(x_0, r_1) \).

**Proof.** Sequence \( \{t_n\} (n \geq -1) \) generated by (3.14) and (3.16) can be obtained if method (1.4) is applied to the scalar polynomial \( f(t) = -\beta t^3 + \gamma_0 t^2 + \gamma_1 t \), where,
\[ [2y - x, x] = \frac{f(2y - x) - f(x)}{2(y - x)}. \tag{3.19} \]

It is simple calculus to show sequence \( \{t_n\} (n \geq -1) \) converges monotonically to zero (decreasingly).

We can have:
\[ t_{k+1} - t_{k+2} = \frac{2(t_{k+1} - t_k)}{f(2t_{k+1} - t_k) - f(t_k)} f(t_{k+1}) \] (3.20)
\[ = \frac{[(\eta_0 - (2t_k + t_{k+1})\beta)(t_k - t_{k+1}) + (t_k - t_{k-1})^2\beta](t_k - t_{k+1})}{1 - \beta^2 - 2(t_0 - t_{k+1})\alpha - [3(t_0 - t_{k+1})(3t_0 + t_{k+1}) - (t_k - t_{k+1})^2]\beta} \]
\[ \geq \frac{(t_k - t_{k+1})\alpha + (t_{k-1} - t_k)^2\beta}{1 - 2(t_0 - t_{k+1})\alpha - \beta^2}(t_k - t_{k+1}). \] (3.21)

We show (3.12) holds for all \( k \geq -1 \). Using (3.3)–(3.8) and
\[ t_0 - t_1 = \left[ 1 - \frac{\gamma_0 t_0 - (t_0 - t_{-1})^2\beta - 2\beta t_0^2}{\gamma_1 + 2\gamma_0 t_0 - (t_0 - t_{-1})^2\beta - 3\beta t_0^2} \right] t_0 = \frac{h(r_0)}{1 - 2\beta^2} = c \] (3.22)
we conclude that (3.12) holds for \( n = -1, 0 \). Assume (3.12) holds for all \( n \leq k \) and \( x_k \in U(x_0, r_0) \). By (2.6) and (3.12) \( x_{k+1} \in U(x_0, r_0) \). By (3.1), (3.2) and (3.12)
\[ \| A_0^{-1}(A_0 - A_{k+1}) \| \]
\[ = \| A_0^{-1}([(2x_0 - x_{-1}, x_{-1}] - [x_0, x_{-1}] + [x_0, x_{-1}] - [x_0, x_0]) + [x_k, x_0] - [x_{k+1}, x_0] + [x_{k+1}, x_0] - [x_{k+1}, x_k]) \|
\]
\[ = \| A_0^{-1}(([2x_0 - x_{-1}, x_{-1}] - [x_0, x_{-1}] + [x_0, x_{-1}] - [x_0, x_0]) + ([x_{k+1}, x_0] - [x_{k+1}, x_k]) + ([x_{k+1}, x_k] - [2x_{k+1} - x_k, x_k]) \| \]
\[ \leq \beta^2 + (\| x_0 - x_{k+1} \| + \| x_0 - x_k \| + \| x_k - x_{k+1} \|)\alpha \]
\[ \leq \beta^2 + 2(t_0 - t_{k+1})\alpha < 2\beta^2 + 2\alpha r \leq 1. \] (3.23)

It follows by the Banach lemma on invertible operators and (3.23) that \( A_{k+1}^{-1} \) exists, so that
\[ \| A_{k+1}^{-1} A_0 \| \leq [1 - \beta^2 - (\| x_0 - x_{k+1} \| + \| x_0 - x_k \| + \| x_k - x_{k+1} \|)\alpha]^{-1}. \] (3.24)

We can also obtain
\[ \| A_{k+1}^{-1}([x_{k+1}, x_k] - A_k) \|
\]
\[ = \| A_{k+1}^{-1}([x_{k+1}, x_k] - [x_k, x_k] + \| x_k, x_k \| + \| x_k, x_k \| - [x_k, x_k] - [2x_k - x_k, x_k]) \|
\]
\[ = \| A_{k+1}^{-1}([(x_{k+1}, x_k] - [x_k, x_k]) + ([x_k, x_k] - [2x_k - x_k, x_k]) \| (x_k - x_{k-1})) \| \]
\[ \leq \alpha \| x_k - x_{k+1} \| + \beta \| x_{k-1} - x_k \|^2. \] (3.25)

Using (1.5), (3.24) and (3.25) we get
\[ \| x_{k+2} - x_{k+1} \| = \| A_{k+1}^{-1} F(x_{k+1}) \|
\]
\[ = \| A_{k+1}^{-1} (F(x_{k+1}) - F(x_k) - A_k(x_{k+1} - x_k)) \|
\]
\[ \leq \| A_{k+1}^{-1} A_0 \| \| A_0^{-1}([x_{k+1}, x_k] - A_k) \| \cdot \| x_k - x_{k+1} \|
\]
\[ \leq \frac{\alpha \| x_k - x_{k+1} \| + \beta \| x_{k-1} - x_k \|^2}{1 - \beta^2 - \alpha (\| x_0 - x_{k+1} \| + \| x_0 - x_k \| + \| x_k - x_{k+1} \|)} \| x_k - x_{k+1} \| \]
\[
\leq \frac{[\alpha (t_k - t_{k+1}) + \beta (t_{k-1} - t_k)^2](t_k - t_{k+1})}{1 - \beta \gamma^2 - 2(t_0 - t_{k+1})\alpha} \leq t_{k+1} - t_{k+2}, \quad (3.26)
\]

which together with (3.11) completes the induction.

It follows from (3.12) that sequence \{\(x_n\)\} \((n \geq -1)\) is Cauchy in a Banach space \(X\) and as such it converges to some \(x^* \in \overline{U}(x_0, r_0)\) (since \(\overline{U}(x_0, r_0)\) is a closed set). By letting \(k \to \infty\) in (3.26) we obtain \(F(x^*) = 0\).

Finally to show uniqueness, define operator
\[
M = \begin{bmatrix} y^*, x^* \end{bmatrix} \quad (3.27)
\]
where \(y^*\) is a solution of Eq. (1.1) in \(\overline{U}(x_0, r_1)\). We can have
\[
\|A_0^{-1}(A_0 - M)\| \leq \alpha \left[\|y^* - (2x_0 - x_{-1})\| + \|x^* - x_{-1}\|\right]
\leq \alpha \left[\|(y^* - x_0) - (x_0 - x_{-1})\| + \|(x^* - x_0) + (x_0 - x_{-1})\|\right]
\leq \alpha \left[\|y^* - x_0\| + 2\|x_0 - x_{-1}\| + \|x^* - x_0\|\right]
\leq \alpha (2\gamma + r_0 + r_1) < 1. \quad (3.28)
\]

It follows from the Banach lemma on invertible operators and (3.28) that linear operator \(M\) is invertible.

We deduce from (3.27) and the identity
\[
F(x^*) - F(y^*) = M(x^* - y^*) \quad (3.29)
\]
that
\[
x^* = y^*. \quad (3.30)
\]

The proof of Theorem 3.1 is now complete. \(\square\)

**Remark 3.2.** (a) It follows from (3.12), (3.13), (3.21), and (3.26) that the order of convergence of scalar sequence \{\(t_n\)\} and iteration \{\(x_n\)\} is two.

(b) The conclusions of Theorem 3.1 hold in a weaker setting. Indeed assume:
\[
\|A_0^{-1}([x_0, x_0] - [x, x_0])\| \leq \alpha_0\|x - x_0\|, \quad (3.31)
\]
\[
\|A_0^{-1}([x, x_0] - [x, y])\| \leq \alpha_1\|y - x_0\|, \quad (3.32)
\]
\[
\|A_0^{-1}([y, x] - [2y - x, x])\| \leq \alpha_2\|y - x\|, \quad (3.33)
\]
\[
\|A_0^{-1}([y, x] - [x, x])\| \leq \alpha_3\|y - x\|, \quad (3.34)
\]
\[
\|A_0^{-1}([2x_0 - x_{-1}, x_0] - [x, y])\| \leq \alpha_4(\|2x_0 - x_{-1} - x\| + \|x_0 - y\|) \quad (3.35)
\]
and
\[
\|A_0^{-1}([2x_0 - x_{-1}, x_{-1}, x_0] - [x_0, x_{-1}, x_0])\| \leq \beta_0\|x_0 - x_{-1}\| \quad (3.36)
\]
for all \(x, y \in D_0\).

It follows from (3.1), (3.2) and (3.31)–(3.36) that
\[
\alpha_i \leq 2\alpha, \quad i = 1, 2, 3, 4, \quad (3.37)
\]
and
\[
\beta_0 \leq \beta. \quad (3.38)
\]
For the derivation of: (3.24), we can use (3.31)–(3.33) and (3.36) instead of (3.1) and (3.2), respectively; (3.25), we can use (3.34) instead of (3.1); (3.28), we can use (3.36) instead of (3.1). The resulting majorizing sequence call it \( \{ \tilde{t}_n \} \) is also converging to zero and is finer than \( \{ t_n \} \) because of (3.37) and (3.38).

Therefore if (2.6), (3.31)–(3.36) are used in Theorem 3.1 instead of (3.1) we draw the same conclusions but with weaker conditions, and corresponding error bounds are such that:

\[
\| x_{n+1} - x_n \| \leq \tilde{t}_n - \tilde{t}_{n+1} \leq t_{n+1} - t_n
\]

(3.39)

and

\[
\| x_n - x^* \| \leq \tilde{t}_n \leq t_n
\]

(3.40)

for all \( n \geq 0 \).

(c) Condition (3.2) can be replaced by the stronger (not really needed in the proof) but more popular [17],

\[
\| A_0^{-1}([v, x, y] - [u, x, y]) \| \leq \beta_1 \| u - v \|
\]

(3.41)

for all \( v, u, x, y \in D_0 \).

(d) As already noted at the end of Remark 2.2, conditions (2.6) and (3.10) can be replaced by

\[
U_2 = U(x_0, R_0) \subseteq D_0 \quad \text{with} \quad R_0 = 3r_0
\]

(3.42)

provided that \( x_{-1} \in U_2 \).

Indeed if \( x_{n-1}, x_n \in U_0 \) \( (n \geq 0) \) then

\[
\| 2x_n - x_{n-1} \| \leq 2\| x_n - x_0 \| + \| x_{n-1} - x_0 \| < 3r_0.
\]

That is \( 2x_n - x_{n-1} \in U_2 \) \( (n \geq 0) \).

A simple numerical example follows to show:

(a) method (1.2)–(1.3) says in the case of Steffensen’s method, i.e. when \( \alpha_n = 1 \) \( (n \geq 0) \) cannot be used to solve a simple scalar equation,

(b) method (1.4) can coincide with Newton’s method

\[
x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad (n \geq 0).
\]

(3.43)

Note that the analytical representation of \( F'(x_n) \) may be complicated which makes the use of method (1.4) very attractive.

**Example 3.3.** Let \( X = Y = \mathbb{R} \), and define function \( F \) on \( D_0 = D = (0.4, 1.5) \) by

\[
F(x) = x^2 - 6x + 5.
\]

(3.44)

Moreover define divided difference of order one appearing in method (1.4) by (3.19). In this case method (1.4) becomes

\[
x_{n+1} = \frac{x_n^2 - 5}{2(x_n - 3)},
\]

(3.45)

and coincides with Newton’s method (3.43) applied to \( F \). Choose \( x_{-1} = 0.9 \) and \( x_0 = 0.1 \). Then using (1.4), we obtain \( x_1 = x^* = 1 \).
Note however that Steffensen’s method cannot be used since
\[ F(x_n + F(x_n)) - F(x_n) = x_n^4 - 10x_n^3 + 28x_n^2 - 14x_n - 5 = 0, \]
when \( n = 0 \).

In order for us to further compare our method (1.4) with (1.2)–(1.3) we use a numerical example already considered in [3].

**Example 3.4.** Let \( X = Y = \mathbb{R}^2 \), \( x_{-1} = (3.9, 0.9) \), \( x_0 = (4, 1) \) and define operator \( F \) on \( X \) by
\[
F(x, y) = (y^2 - 4, x^2 - 2y - 21).
\]
(3.46)
Then for \( \alpha_n = \alpha^* \) fixed and small, say e.g. \( \alpha = 10^{-8} \) we cannot compute the iterates of method (1.2)–(1.3), whereas our method (1.4) generates the solution \( x^* = (5, 2) \) after 5 iterations.

We conclude this section with an example involving a nonlinear integral equation:

**Example 3.5.** Let \( H(x, t, x(t)) \) be a continuous function of its arguments which is sufficiently many times differentiable with respect to \( x \). It can easily be seen that if operator \( F \) in (1.1) is given by
\[
F(x(s)) = x(s) - \frac{1}{0} H(s, t, x(t)) dt,
\]
(3.47)
then divided difference of order one appearing in (1.5)–(1.6) can be defined as
\[
h_n(s, t) = \frac{H(s, t, 2x_n(t) - x_{n-1}(t)) - H(s, t, x_{n-1}(t))}{2(x_n(t) - x_{n-1}(t))},
\]
(3.48)
provided that if for \( t = t_m \) we get \( x_n(t) = x_{n-1}(t) \), then the above function equals \( H'_x(s, t_m, x_n(t_m)) \). Note that this way \( h_n(s, t) \) is continuous for all \( t \in [0, 1] \).

**References**


