A two-step Steffensen’s method under modified convergence conditions

S. Amat*, S. Busquier

Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Spain

Received 29 September 2005
Available online 3 February 2006
Submitted by V.J. Ervin

Abstract

A simplification of a third order iterative method is proposed. The main advantage of this method is that it does not need to evaluate neither any Fréchet derivative nor any bilinear operator. A semilocal convergence theorem in Banach spaces, under modified Kantorovich conditions, is analyzed. A local convergence analysis is also performed. Finally, some numerical results are presented.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Third-order iterative methods; Convergence; Banach spaces

1. Introduction

One of the most studied problems in numerical analysis is the solution of nonlinear equations

\[ F(x) = 0, \tag{1} \]

where \( F \) is a nonlinear operator between Banach spaces.

A powerful tool to solve these equations is by means of iterative methods. Roughly speaking, an iterative method starts from some initial guess \( x_0, x_1, \ldots \), which are improved by means of an iteration

\[ x_{n+1} = \Phi(x_n, \ldots, x_{n-m}). \]
An iterative method is of \( p \)th order if the solution \( x^* \) of (1) satisfies
\[
 x^* = \Phi(x^*, \ldots, x^*),
\]
where \( \Phi'(x^*, \ldots, x^*) = \cdots = \Phi^{(p-1)}(x^*, \ldots, x^*) = 0 \) and \( \Phi^{(p)}(x^*, \ldots, x^*) \neq 0 \). For such a method, the error \( \|x^* - x_{n+1}\| \) is proportional to \( \|x^* - x_n\|^p \) as \( n \to \infty \).

Newton’s method and similar second or less order methods are the most used. Third order methods require more computational cost than other simpler methods, which makes them disadvantageous to be used in general. The main practical difficulty related to the classical third order methods \[7\] is the evaluation of the second order Fréchet derivative. For a nonlinear system of \( N \) equations and \( N \) unknowns, the first Fréchet derivative is a matrix with \( N^2 \) values, while the second Fréchet derivative has \( N^3 \) values. This implies a huge amount of operations in order to evaluate every iteration. Nevertheless, some methods overcome these difficulties by evaluating several times the function and its first derivative. For example, in \[10\], this (two-step) third order recurrence is proposed:
\[
 y_{n+1} = x_n - F'(x_n)^{-1} F(x_n),
\]
\[
 x_{n+1} = y_{n+1} - F'(x_n)^{-1} F(y_{n+1}).
\]

This method is, in general, cheaper than any third order method requiring the evaluation of the second derivative and emerges as a good alternative to Newton’s type methods.

The goal of this paper is to modify the two-step method in order not to evaluate any Fréchet derivative. A bounded linear operator \([x, y; F]\), from \( X \) into \( X \), is called a divided difference of first order for the operator \( F \) on the points \( x \) and \( y \) if
\[
 [x, y; F](x - y) = F(x) - F(y).
\]
If \( F \) is Fréchet differentiable then \( F'(x) = [x, x; F] \) for all \( x \in D \). Using this definition, the following generalization of Steffensen’s method was considered in \[3\]:
\[
 x_{n+1} = x_n - \left[ x_n, x_n + \alpha_n F(x_n); F \right]^{-1} F(x_n), \quad (2)
\]
where \( \alpha_n \in [0, 1] \). In practice, the \( \alpha_n \) are computed such that
\[
 \text{tol}_c \ll \|\alpha_n F(x_n)\| \leq \text{tol}_{\text{user}},
\]
where \( \text{tol}_c \) is related with the computer precision and \( \text{tol}_{\text{user}} \) is a free parameter for the user. The new iterative method is in general a good alternative to Newton’s method, since \([x_n, x_n + \alpha_n F(x_n); F]\) is in all the steps a good approximation to \( F'(x_n) \).

We now consider
\[
 y_{n+1} = x_n - \left[ x_n - \alpha_n F(x_n), x_n + \alpha_n F(x_n); F \right]^{-1} F(x_n),
\]
\[
 x_{n+1} = y_{n+1} - \left[ x_n - \alpha_n F(x_n), x_n + \alpha_n F(x_n); F \right]^{-1} F(y_{n+1}), \quad (3)
\]
where \( \alpha_n \in [0, 1] \) is computed in practice as before.

In semilocal convergence theorems, conditions are imposed on \( x_0, x_{-1}, \ldots \) (and, eventually, on \( F \) or \( \Phi \)) in order to assure the convergence of \( x_n \) to the solution \( x^* \). This analysis, usually known as Kantorovich type, is based on a relationship between the problem in a Banach space and a single nonlinear scalar equation which leads the behavior of the problem. Usually the Fréchet differentiability of some order is assumed. In this study, we present a semilocal convergence theorem under modified Kantorovich conditions \[1,2,8,9\] without assuming any Fréchet differentiability.

On the other hand, in a recent work \[6\] Argyros and Gutiérrez present a unify approach for enlarging the radius of convergence of Newton’s method. Here we introduce similar conditions and we assert a local convergence theorem for the modified two-step method. See \[4\] for secant type methods.
The structure of this paper is as follows: in the next section, we analyze the order of the modified two-step method. In Section 3, we assert a semilocal convergence theorem. In Section 4, a local convergence theorem is performed. Finally, some numerical experiments are presented in Section 5.

2. Preliminary analysis

Let us define \( x_n^- := x_n - \alpha_n F(x_n) \) and \( x_n^+ := x_n + \alpha_n F(x_n) \).
Assuming \( F \) is a sufficiently smooth operator and using Taylor expansion, we have

\[
F(y_{n+1}) = F(x_n) + F'(x_n)(y_{n+1} - x_n) + \frac{1}{2} F''(x_n)(y_{n+1} - x_n)^2 + \cdots
\]

\[
= F(x_n) + F'(x_n)(-\left[x_n^-, x_n^+; F\right]^{-1} F(x_n))
\]

\[
+ \frac{1}{2} F''(x_n)(-\left[x_n^-, x_n^+; F\right]^{-1} F(x_n))^2 + \cdots
\]

\[
= F(x_n) + F'(x_n)(-F'(x_n)^{-1} F(x_n) + O(F(x_n))^2)
\]

\[
+ \frac{1}{2} F''(x_n)(-F'(x_n)^{-1} F(x_n) + O(F(x_n))^2)^2 + \cdots
\]

thus the modified two-step method becomes

\[
x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) - \frac{1}{2} F'(x_n) F''(x_n) F'(x_n) F(x_n) + O(F(x_n)^3).
\]

That is, it is asymptotically the classical Chebyshev method (third order for simple roots),

\[
x_{n+1}^c = x_n^c - F'(x_n^c)^{-1} F(x_n^c) - \frac{1}{2} F'(x_n^c) F''(x_n^c) F'(x_n^c) F(x_n^c).
\]

and its numerical order will be the same as that of Chebyshev. Notice that the classical two-step method admits a similar asymptotical behavior.

3. A semilocal convergence theorem

The convergence theorems for Steffensen’s type methods establish sufficient conditions on the operator and the first approximation to the solution in order to ensure that the sequence of iterates converges to a solution of the equation. The basic assumption is that the divided difference of \( F \) is Lipschitz or Hölder continuous in some ball around the initial iterate. In particular, the Fréchet derivative of \( F \) exists. In some works [1,8,9], the authors relax this requirement and they only assume that the divided difference satisfies

\[
\| [x, y; F] - [v, w; F] \| \leq \omega(\| x - v \|, \| y - w \|), \quad x, y, v, w \in B, \tag{4}
\]

where \( \omega: \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+ \) is a continuous function nondecreasing in both components. In next theorem we extend this theory to our method. We use recurrence relations and we reduce the initial problem to a simpler one with real sequences.

**Theorem 1.** Let \( X \) be a Banach space. Let \( B \) be a convex open subset of \( X \) and suppose that there exists a first order divide difference of \( F: B \subset X \to X \) satisfying condition (4). Let \( \alpha_n \) be such that \( \| \alpha_n F(x_n) \| \leq \text{tol}_{\text{user}} \).

Let \( x_0 \in B \) and assume that:

\[
\]
(1) $\|\Gamma_0^{-1} := [x_0^-, x_0^+; F]^{-1}\| \leq \beta$.

(2) $\|\Gamma_0^{-1} F(x_0)\| \leq \eta$.

(3) Let $m = \beta w(\eta + \text{tol}_\text{user}, \text{tol}_\text{user})$. Assume that the equation
\[ t \left( 1 - \frac{m}{1 - \beta w(t + 2 \text{tol}_\text{user}, t + 2 \text{tol}_\text{user})} \right) - \eta = 0 \] has a minimum positive root $R$.

If
\[ \beta w(R + 2 \text{tol}_\text{user}, R + 2 \text{tol}_\text{user}) < 1, \]
\[ M := \frac{m}{1 - \beta w(R + 2 \text{tol}_\text{user}, R + 2 \text{tol}_\text{user})} \quad \text{and} \quad \overline{B(x_0, R)} \subset B \]
then the modified two-step method (3) is well defined, remains in $B(x_0, R)$ and converges to the unique solution of $F(x) = 0$ in $\overline{B(x_0, R)}$.

**Proof.** From the initial hypothesis $y_1$ and $x_1$ are well defined and $\|y_1 - x_0\| \leq \eta < R$, then $y_1 \in B(x_0, R)$.

From the definition of divided difference and iteration (3),
\[ \| F(y_1) \| \leq \|[y_1, x_0; F] - \Gamma_0\| \|y_1 - x_0\| \]
\[ \leq w(\eta + \text{tol}_\text{user}, \text{tol}_\text{user}) \eta, \]
then
\[ \|x_1 - x_0\| \leq \|x_1 - y_1\| + \|y_1 - x_0\| \]
\[ \leq \beta w(\eta + \text{tol}_\text{user}, \text{tol}_\text{user}) \eta + \eta \]
\[ = (\beta w(\eta + \text{tol}_\text{user}, \text{tol}_\text{user}) + 1) \eta \]
\[ < R, \]
that is, $x_1 \in B(x_0, R)$.

Since $w$ is a nondecreasing function
\[ \| I - \Gamma_0^{-1} \Gamma_1 \| \leq \|\Gamma_0^{-1}\| \|\Gamma_0 - \Gamma_1\| \]
\[ \leq \beta w(\eta + 2 \text{tol}_\text{user}, \eta + 2 \text{tol}_\text{user}) \]
\[ \leq \beta w(R + 2 \text{tol}_\text{user}, R + 2 \text{tol}_\text{user}) \]
\[ < 1, \]
then using Banach lemma $\Gamma_1^{-1}$ exists and
\[ \|\Gamma_1^{-1}\| \leq \frac{\beta}{1 - \beta w(R + 2 \text{tol}_\text{user}, R + 2 \text{tol}_\text{user})}. \]

We have that $y_2$ and $x_2$ are well defined.

Moreover,
\[ \| F(x_1) \| \leq \|[x_0^-, x_0^+; F] - [x_1, y_1; F]\| \|x_1 - y_1\| \]
\[ \leq w(\eta + \text{tol}_\text{user}, \text{tol}_\text{user}) \|x_1 - y_1\|, \]
and
\[ \| F(y_2) \| \leq \| [x_0^-, x_0^+; F] - [y_2, x_1; F] \| \| y_2 - x_1 \| \leq w(\eta + \text{tol}_{\text{user}}) \| y_2 - x_1 \|. \]

Then
\[ \| y_2 - x_1 \| = \| -[x_1^-, x_1^+; F]^{-1} F(x_1) \| \leq \frac{m}{1 - \beta w(R + 2 \text{tol}_{\text{user}}, R + 2 \text{tol}_{\text{user}})} \| x_1 - y_1 \| \leq M \| x_1 - y_1 \|. \]

analogously
\[ \| x_2 - y_2 \| = \| -[x_1^-, x_1^+; F]^{-1} F(y_2) \| \leq \frac{m}{1 - \beta w(R + 2 \text{tol}_{\text{user}}, R + 2 \text{tol}_{\text{user}})} \| y_2 - x_1 \| \leq M \| y_2 - x_1 \| \leq M^2 \| x_1 - y_1 \|. \]

Using the above bounds,
\[ \| x_2 - x_0 \| \leq \| x_2 - y_2 \| + \| y_2 - x_1 \| + \| x_1 - y_1 \| + \| y_1 - x_0 \| \leq (M^3 + M^2 + M + 1) \eta < R, \]

that is, \( x_2 \in B(x_0, R) \).

Using the same arguments as before and an induction strategy, we can prove:

- \( x_n \) is well defined and
  \[ \| x_n - x_0 \| \leq \sum_{k=0}^{n+1} M^k \eta < R \]
  that is, \( x_n \in B(x_0, R) \).

- \( \| x_n - x_{n-1} \| \leq (M^2 + M) \| x_n - y_n \| \leq (M^2 + M)(M^2)^n \| x_1 - y_1 \|. \)

Consequently, \( \{ x_n \} \) is a Cauchy sequence and converges to \( x^* \in B(x_0, R) \).

- Since
  \[ \| F(x_n) \| \leq w(\eta + \text{tol}_{\text{user}}, \text{tol}_{\text{user}}) \| x_n - y_n \| \]
  and \( \| x_n - y_n \| \to 0 \) when \( n \to \infty \), we obtain \( F(x^*) = 0 \). Moreover, if \( y^* \) is other root of \( F(x) \) in \( B(x_0, R) \), we have
  \[ \| I - I_0^{-1} [x^*, y^*; F] \| \leq \| I_0^{-1} \| \| I_0 - [x^*, y^*; F] \| \leq \beta w(R + 2 \text{tol}_{\text{user}}, R + 2 \text{tol}_{\text{user}}) \leq 1, \]
  in consequence the operator \([x^*, y^*; F]\) is invertible and \( x^* = y^* \). \( \square \)
4. A local convergence theorem

In this section, following [6], we obtain a local convergence theorem for the modified two-step method. The goal is to enlarge the radius of convergence, without increasing the necessary hypothesis.

**Theorem 2.** Let \( B \) be a convex open subset of \( X \). Let \( F : B \subset X \to X \) be a Fréchet-differentiable operator and let \( x^* \) be a simple zero of \( F \). Assume that:

(a) There exist functions \( f : [0, \infty) \to [0, \infty) \) and \( g : [0, \infty) \to [0, \infty) \) nondecreasing such that

\[
\|F'(x^*)^{-1}([x, x^*; F] - [z, y; F])\| \leq f\left(\max(\|x - z\|, \|x^* - y\|)\right),
\]

\[
\|F'(x^*)^{-1}([x^*, x^*; F] - [x, y; F])\| \leq g\left(\max(\|x^* - x\|, \|x^* - y\|)\right),
\]

for all \( x, y \in B \).

(b) Equation

\[
g(r) + f(r) = 1,
\]

has a minimum positive zero \( R \).

(c) \( \overline{B(x^*, r^*)} = \{x \in X : \|x^* - x\| \leq r^*\} \subset B \) for \( r^* \in [0, R) \).

Then, the modified two-step method \( \{x_n\} \) (3) is well defined, remains in \( \overline{B(x^*, r^*)} \) for all \( n > 0 \) and converges to \( x^* \) provided that \( x_0 \in \overline{B(x^*, r^*)} \). Moreover, the following error bounds hold for all \( n > 0 \):

\[
\|y_n - x^*\| \leq \frac{f(\max(\|x_n^- - x_n\|, \|x_n^+ - x^*\|))}{1 - g(\max(\|x_n^- - x^*\|, \|x_n^+ - x^*\|))} \|x_n - x^*\|,
\]

\[
\|x_{n+1} - x^*\| \leq \frac{f(\max(\|x_n^- - y_n\|, \|x_n^+ - x^*\|))}{1 - g(\max(\|x_n^- - x^*\|, \|x_n^+ - x^*\|))} \|x_n - x^*\|.
\]

**Proof.** Given \( x, y \in \overline{B(x^*, r^*)} \), using (7) and (8), we obtain

\[
\|I - F'(x^*)^{-1}[x, y; F]\| = \|F'(x^*)^{-1}([x^*, x^*; F] - [x, y; F])\|
\leq g\left(\max(\|x^* - x\|, \|x^* - y\|)\right)
\leq g(r^*)
< 1.
\]

Using the Banach lemma on invertible operators and (11) it follows that \( [x, y; F]^{-1} \) exists and

\[
\|[x, y; F]^{-1}F'(x^*)\| \leq \frac{1}{1 - g(\max(\|x^* - x\|, \|x^* - y\|))}, \quad x, y \in \overline{B(x^*, r^*)}.
\]

Choose \( x_0 \in \overline{B(x^*, r^*)} \). Let us assume \( x_k \in \overline{B(x^*, r^*)} \) for all \( k = 0, 1, \ldots, n \). Indeed, denoting \( \Gamma_n^{-1} : [x_n^-, x_n^+; F]^{-1} \) and using relations (2)–(6)–(7) and (12), we obtain

\[
y_{n+1} - x^* = x_n - \Gamma_n^{-1}F(x_n) - x^*
=- \Gamma_n^{-1}\left(-F(x^*) + F(x_n) - [x_n^-, x_n^+; F](x_n - x^*)\right)
=- \Gamma_n^{-1}\left([x_n, x^*; F](x_n - x^*) - [x_n^-, x_n^+; F](x_n - x^*)\right)
\]
and
\[ \| y_{n+1} - x^* \| \leq \frac{f(\max(\| x_n^- - x_n \|, \| x_n^+ - x^* \|))}{1 - g(\max(\| x_n^- - x^* \|, \| x_n^+ - x^* \|))} \| x_n - x^* \|, \]

which shows (9). Moreover, by the definition of \( f \) and \( g \), and relation (8),
\[ \frac{f(\max(\| x_n^- - x_n \|, \| x_n^+ - x^* \|))}{1 - g(\max(\| x_n^- - x^* \|, \| x_n^+ - x^* \|))} \leq \frac{f(\max(\| x_0^- - x_0 \|, \| x_0^+ - x^* \|))}{1 - g(\max(\| x_0^- - x^* \|, \| x_0^+ - x^* \|))} < 1. \]

Similarly,
\[ x_{n+1} - x^* = x_n - \Gamma_n^{-1} F(y_n) - x^* \]
\[ = -\Gamma_n^{-1}(\{ \{ F(x^*) + F(y_n) - [x_n^- , x_n^+ ; F](x_n - x^*) \}) \]
\[ = -\Gamma_n^{-1}(\{ \{ y_n , x^* ; F \}(x_n - x^*) \} - [x_n^- , x_n^+ ; F](x_n - x^*)) \]

and
\[ \| y_{n+1} - x^* \| \leq \frac{f(\max(\| x_n^- - y_n \|, \| x_n^+ - x^* \|))}{1 - g(\max(\| x_n^- - x^* \|, \| x_n^+ - x^* \|))} \| x_n - x^* \|, \]

which shows (10). As before,
\[ \frac{f(\max(\| x_n^- - y_n \|, \| x_n^+ - x^* \|))}{1 - g(\max(\| x_n^- - x^* \|, \| x_n^+ - x^* \|))} \leq \frac{f(\max(\| x_0^- - x_0 \|, \| x_0^+ - x^* \|))}{1 - g(\max(\| x_0^- - x^* \|, \| x_0^+ - x^* \|))} < 1. \]

In particular, there exists a positive constant \( c \in [0, 1) \) such that
\[ \| x_{n+1} - x^* \| \leq c \| x_n - x^* \| \leq \cdots \leq c^{n+1} \| x_0 - x^* \| \leq c^{n+1} r^*, \]

that is, \( x_{n+1} \in B(x^*, r^*) \) and \( \lim_{n \to \infty} x_n = x^* \). \( \square \)

For different choices of relations (6)–(7) we refer [6]. The simplest one, for Steffensen’s type
methods, could be
\[ \| F'(x^*)^{-1}(x, x^* ; F) \| \leq l_0 \max(\| x - z \|, \| x^* - y \|), \]
\[ \| F'(x^*)^{-1}(x, x^* ; F) \| \leq l_1 \max(\| x^* - x \|, \| x^* - y \|). \]

In order to compare our results with earlier ones, notice that the classical assumptions are
\[ \| F'(x^*)^{-1}(x, y ; F) \| \leq l_2 \max(\| u - x \|, \| v - y \|). \]

and in general \( l_2 \geq l_0, l_1 \), that is, our radius of convergence is at least as large as the classical
one. For instance, for equation \( e^x - 1 = 0 \) in \([-1, 1]\), it is easy to check that \( l_2 = 2 \cdot e, l_1 = e \) and
\( l_0 = e/2. \)

5. Numerical experiments

In order to see the performance of the introduced iterative method, we have tested it on some
nonlinear equations. We present a comparison with three classical iterative methods. In our ex-
periments, we consider \( \text{tol}_{\text{user}} = 10^{-4}, \) this number is small enough but without numerical instability
(we have obtained similar results with smaller values of \( \text{tol}_{\text{user}} \) such that \( \text{tol}_c \ll \text{tol}_{\text{user}} \)).
Let us consider the Hammerstein equation

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

studied in [11].

Using the trapezoidal rule of integration with step \( h = 1/m \), we obtain the following system of nonlinear equations:

\[ 0 = x^i - 1 + \frac{1}{4m} \left( \frac{1}{2} t_i + t_i x^0 + \sum_{k=0}^{n} \frac{t_i}{t_i + t_k} \frac{1}{x^k} + \frac{1}{2} \frac{t_i}{t_i + t_m} x^m \right), \quad i = 0, 1, \ldots, m, \]  

where \( t_j = j/m \).

In this case, the second Fréchet derivative is diagonal by blocks.

We consider \( m = 20 \) in the quadrature trapezoidal formula. The exact solution is computed numerically by Newton’s method until convergence.

The three iterative methods have similar convergence properties, see Table 1.

Next, we consider quadratic equations of the type

\[ F(x) = x^T A x + B x + C = 0, \]  

where \( \dim(A) = (m \times m) \times m \), \( \dim(B) = m \times m \) and \( \dim(C) = \dim(x) = m \).

The above kind of equations may come from the discretization of equilibrium problems, where interacting forces between particles determine the output. However, the actual case we are going to analyze is prepared to get an exact solution in order to make it easy the evaluation of the errors.

We randomly generate \( A \) and \( B \), and then we determine \( C \) such that \( x^*(i) = 2, \ i = 1, 2, \ldots, m, \) is a solution of (18).

Notice that the second Fréchet derivative is constant \( F''(x) = A + A^T \).

For quadratic examples Chebychev’s method [5] is nothing else than the two-step method.

In Table 2, the dimension we consider is \( m = 100 \).

Finally, we study the system of nonlinear equations

\[ 3x^2 + y^2 - 1 + |x - 1| = 0, \]
Table 3

\[ \|\text{Error}\|_{\infty}, \ (x_{-1}, y_{-1}) = (5, 5), \ (x_0, y_0) = (1, 0) \]

<table>
<thead>
<tr>
<th>Iterations</th>
<th>Secant</th>
<th>Mod. two-step</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$3.15e-01$</td>
<td>$2.56e-03$</td>
</tr>
<tr>
<td>2</td>
<td>$2.71e-02$</td>
<td>$1.34e-08$</td>
</tr>
<tr>
<td>3</td>
<td>$5.41e-03$</td>
<td>$0.00e+00$</td>
</tr>
<tr>
<td>4</td>
<td>$2.84e-04$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>$3.05e-06$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

\[ x^4 + xy^3 - 1 + |y| = 0 \]

analyzed in [9].

The approximate solution

\[ (x^*, y^*) = (0.894655373346867, 0.3278265117462974) \]

is considered.

In this case, the operator is not Fréchet differentiable and we cannot apply classical third order methods.

The convergence of the modified two-step method is better than secant’s method (see Table 3).

Summing up, in this paper we have studied a third order method, without evaluating neither any Fréchet derivative nor any bilinear operator. We have established two convergence theorems and we have analyzed its numerical behavior.

References