STUDY OF OPTIMAL EIGHTH ORDER WEIGHTED-NEWTON METHODS IN BANACH SPACES

Ioannis K. Argyros, Deepak Kumar, and Janak Raj Sharma

Abstract. In this work, we generalize a family of optimal eighth order weighted-Newton methods to Banach spaces and study its local convergence to approximate a locally-unique solution of a system of nonlinear equations. The convergence in this study is shown under hypotheses only on the first derivative. Our analysis avoids the usual Taylor expansions requiring higher order derivatives but uses generalized Lipschitz-type conditions only on the first derivative. Moreover, our new approach provides computable radius of convergence as well as error bounds on the distances involved and estimates on the uniqueness of the solution based on some functions appearing in these generalized conditions. Such estimates are not provided in the approaches using Taylor expansions of higher order derivatives which may not exist or may be very expensive or impossible to compute. The convergence order is computed using computational order of convergence or approximate computational order of convergence which do not require usage of higher derivatives. This technique can be applied to any iterative method using Taylor expansions involving high order derivatives. The study of the local convergence based on Lipschitz constants is important because it provides the degree of difficulty for choosing initial points. In this sense the applicability of the method is expanded. Finally, numerical examples are provided to verify the theoretical results and to show the convergence behavior.

1. Introduction

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

$$F(x) = 0,$$

where $F$ is a Fréchet-differentiable operator defined on a closed convex subset $D$ of Banach space $X$ with values in a Banach space $Y$. Many problems in computational sciences can be written in the form (1.1) using Mathematical Modelling (see, for example [4, 6, 25, 29]). The solution of these equations can
be found in closed form only in special cases. That explains why most methods for solving these equations are usually iterative. The important part in the development of an iterative method is to study its convergence analysis. This is usually divided into two categories viz. semilocal and local convergence. The semilocal convergence is based on the information around an initial point and gives criteria that ensures the convergence of iteration procedures. Local convergence is based on the information of convergence domain. In general the convergence domain is small. Therefore, it is important to enlarge the convergence domain without additional hypothesis. Another important problem is to find more precise error estimates on $\|x_{n+1} - x_n\|$ or $\|x_n - x^*\|$. There exist many studies which deal with the local and semilocal convergence analysis of iterative methods such as [1,3–9,11,16,18,19,22,24].

The most widely used iterative method for solving (1.1) is the quadratically convergent Newton’s method

$$(1.2) \quad x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \quad n = 0, 1, 2, \ldots,$$

where $F'(x)^{-1}$ is the inverse of first Fréchet derivative $F'(x)$ of the function $F(x)$. In order to accelerate the convergence, researchers have also obtained some modified Newton’s or Newton-like methods (see [2,3,7,8,10–17,20,21,26–28]) and references therein.

It is well-known that a variety of higher order iterative methods are available for solving a scalar equation $f(x) = 0$ (see, for example [23,29]). Contrary to this, higher order methods are rare for multi-dimensional case, that is, for approximating the solution of $F(x) = 0$. One possible reason is that the construction of higher order methods for solving systems is a difficult task. Other reason, which is a fact, is that not every method developed for single equation can be generalized to solve systems of nonlinear equations. Recently, Sharma and Arora [27] have developed a family of optimal eighth order methods for solving a scalar equation $f(x) = 0$, which is given by

$$\begin{align*}
y_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \\
z_n &= \psi_4(x_n, y_n), \\
x_{n+1} &= z_n - \frac{f'(x_n) - f[y_n, x_n] + f[z_n, y_n] f(z_n)}{2 f[z_n, y_n] - f[z_n, x_n]} f'(x_n),
\end{align*}$$

(1.3)

where $\psi_4(x_n, y_n)$ is any optimal fourth order scheme with the base as Newton’s iteration $y_n$ and $f[\cdot, \cdot]$ is Newton’s first order divided difference. In particular, they have considered the following optimal fourth order schemes in the second step of (1.3):

Ostrorowski’s method (see [17]):

$$(1.4) \quad z_n = y_n - \frac{1}{2 f[y_n, x_n] - f'(x_n)} f(y_n).$$
Ostrowski-like method (see [17]):

\[ z_n = y_n - \left( \frac{2}{f[y_n, x_n]} - \frac{1}{f'(x_n)} \right) f(y_n). \]

Sharma-Arora method (see [26]):

\[ z_n = y_n - \left( 3 - 2 \frac{f[y_n, x_n]}{f'(x_n)} \right) f(y_n). \]

It can be observed that the above family of eighth order methods can be easily extendable for solving (1.1). In view of this, here we study the method (1.3) in Banach space. The iterative methods corresponding to the fourth order schemes (1.4), (1.5) and (1.6) in Banach space setting are given by

Method-I (M-I):

\[ y_n = x_n - F'(x_n)^{-1}F(x_n), \]
\[ z_n = y_n - (2F[y_n, x_n] - F'(x_n))^{-1}F(y_n), \]
\[ x_{n+1} = \Psi_8(x_n, y_n, z_n). \]

Method-II (M-II):

\[ y_n = x_n - F'(x_n)^{-1}F(x_n), \]
\[ z_n = y_n - (2F[y_n, x_n] - F'(x_n))^{-1}F(y_n), \]
\[ x_{n+1} = \Psi_8(x_n, y_n, z_n). \]

Method-III (M-III):

\[ y_n = x_n - F'(x_n)^{-1}F(x_n), \]
\[ z_n = y_n - (3I - 2F'(x_n)^{-1}F[y_n, x_n])F'(x_n)^{-1}F(y_n), \]
\[ x_{n+1} = \Psi_8(x_n, y_n, z_n). \]

In the above each case

\[ \Psi_8(x_n, y_n, z_n) = z_n - \frac{1}{F'(x_n)} \left( F'(x_n) - F[y_n, x_n] + F[z_n, y_n] \right) \]
\[ \left( F'(x_n) - F[y_n, x_n] + F[z_n, y_n] \right) F'(x_n)^{-1}F(z_n). \]

In Section 2, the local convergence, including radius of convergence, computable error bounds and uniqueness results of the proposed methods, is presented. In order to verify the theoretical results and to test the performance of the methods, some numerical examples are presented in Section 3.

2. Local convergence analysis

We present the local convergence analysis of the methods M-I, M-II and M-III in this section. We shall find the radius of convergence, computable error bounds on the distances \( \|x_n - x^*\| \) and then establish the uniqueness of the solution \( x^* \) inside a certain ball based on some Lipschitz constants.
2.1. Convergence of M-I

Let $L_0 > 0$, $L > 0$, $M \geq 0$ be given parameters. It is convenient for the local convergence analysis that follows to produce some functions and parameters. Define the functions $g_1$ and $h_p$ on interval $[0, \frac{1}{L_0})$ by

$$g_1(t) = \frac{Lt}{2(1 - L_0 t)},$$
$$p(t) = (2L_1(1 + g_1(t)) + L_0)t,$$
$$h_p(t) = p(t) - 1$$

and parameter $r_1$ by

$$(2.1) r_1 = \frac{2}{2L_0 + L} < \frac{1}{L_0}.$$  

Then, we have that $h_p(0) = -1 < 0$ and $h_p(t) \to +\infty$ as $t \in [0, \frac{1}{L_0})$. The intermediate value theorem guarantees that $h_p(t)$ has zeros in interval $[0, \frac{1}{L_0})$. Let $r_p$ be the smallest such zero. Moreover define functions $g_2$, $q$, $h_2$ and $h_q$ on interval $[0, r_p)$ by

$$g_2(t) = \left(1 + \frac{M}{1 - p(t)}\right)g_1(t),$$
$$q(t) = L_1(1 + 2g_1(t) + 3g_2(t))t,$$
$$h_2(t) = g_2(t) - 1 \quad \text{and} \quad h_q(t) = q(t) - 1.$$

Then, we have that $h_2(0) = h_q(0) = -1 < 0$, $h_q(t) \to +\infty$ and $h_2(t) \to +\infty$ as $t \to r_p^-$. It follows from the intermediate theorem that functions $h_2$, $h_q$ have zeros in the interval $(0, r_p)$. Denote by $r_2$ and $r_q$ the smallest such zeros. Finally define functions $g_3$ and $h_3$ on the interval $[0, r_q)$ by

$$g_3(t) = \left(1 + \frac{M}{(1 - L_0 t)(1 - q(t))}\right)(L_0 + L_1)t + L_1 t(2g_1(t) + g_2(t)) + 1\right)g_2(t),$$

and

$$h_3(t) = g_3(t) - 1.$$  

Now, we have that $h_3(0) = -1 < 0$ and $h_3(t) \to +\infty$ as $t \to r_q^-$. It follows from the intermediate theorem that function $h_3$ has zeros in the interval $(0, r_q)$. Denote by $r_3$ the smallest such zero of function $h_3$ on interval $[0, r_q)$. Set:

$$(2.2) r = \min\{r_i\}, \quad i = 1, 2, 3.$$  

Then we have that

$$(2.3) 0 < r \leq r_q.$$  

Then, for each $t \in [0, r)$,

$$(2.4) 0 \leq g_1(t) \leq 1,$$

$$(2.5) 0 \leq p(t) \leq 1,$$

$$(2.6) 0 \leq g_2(t) \leq 1.$$
and

\[(2.7) \quad 0 \leq g_3(t) \leq 1.\]

Let $U(v, \rho)$ and $\bar{U}(v, \rho)$ denote the open and closed ball in $X$, respectively with center $v \in X$ and of radius $\rho > 0$. Let also $\mathcal{L}(X, Y)$ be the set of bounded linear operators between $X$ and $Y$.

Next, we present the local convergence analysis of M-I using the preceding notations.

**Theorem 2.1.** Let $F : D \subseteq X \to Y$ be a Fréchet-differentiable operator and $F[\cdot, \cdot] : D \times D \to \mathcal{L}(X, Y)$ be a divided difference operator of $F$. Suppose that there exist $x^* \in D$, $L_0 > 0$, $L > 0$, $L_1 > 0$ and $M \geq 1$ such that for each $x, y \in D$

\[(2.8) \quad F(x^*) = 0, \quad F'(x^*)^{-1} \neq 0,\]

\[(2.9) \quad \|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq L_0\|x - x^*\|,\]

\[(2.10) \quad \|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq L\|x - y\|,\]

\[(2.11) \quad \|F'(x^*)^{-1}F'(x)\| \leq M,\]

\[(2.12) \quad \|F'(x^*)^{-1}(F[x, y] - F'(x^*))\| \leq L_1(\|x - x^*\| + \|y - x^*\|),\]

and

\[(2.13) \quad \bar{U}(x^*, r) \subseteq D,\]

where the radius $r$ is defined in (2.2). Then, the sequence $\{x_n\}$ generated by M-I for $x_0 \in U(x^*, r) - \{x^*\}$ is well defined, remains in $U(x^*, r)$ for each $n = 0, 1, \ldots$ and converges to $x^*$. Moreover, the following estimates hold

\[(2.14) \quad \|y_n - x^*\| \leq g_1(\|x_n - x^*\|)\|x_n - x^*\| \leq \|x_n - x^*\| < r,\]

\[(2.15) \quad \|z_n - x^*\| \leq g_2(\|x_n - x^*\|)\|x_n - x^*\| \leq \|x_n - x^*\|\]

and

\[(2.16) \quad \|x_{n+1} - x^*\| \leq g_3(\|x_n - x^*\|)\|x_n - x^*\|,\]

where the “$g$” functions are defined previously. Furthermore, for $T \in [r, \frac{2r}{L_0})$ the limit point $x^*$ is the only solution of equation $F(x) = 0$ in $D_0 = \bar{U}(x^*, T) \cap D$.

**Proof.** We shall show the estimates (2.14)-(2.16) using mathematical induction. Using (2.1), (2.9) and the hypotheses $x_0 \in U(x^*, r) - \{x^*\}$, we get that

\[(2.17) \quad \|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \leq L_0\|x_0 - x^*\| < L_0r < 1.\]

It follows from (2.17) and the Banach Lemma on invertible operators [6] that $F'(x_0)^{-1} \neq 0$ and

\[(2.18) \quad \|F'(x_0)^{-1}F'(x^*)\| \leq \frac{1}{1 - L_0\|x_0 - x^*\|} < \frac{1}{1 - L_0r}.\]
Hence, $y_0$ is well defined by the first step of method M-I for $n = 0$. Then, we have by equations (2.1), (2.4), (2.10) and (2.18) that
\[
\|y_0 - x^*\| \leq \|x_0 - x^* - F'(x_0)^{-1}F(x_0)\|
\leq \|F'(x_0)^{-1}F'(x^*)\|
\leq \int_0^1 F'(x^*)^{-1}[F'(x^* + \theta(x_0 - x^*)) - F'(x_0)](x_0 - x^*)\|d\theta
\leq \frac{L\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)}
\]
which shows (2.14) for $n = 0$ and $y_0 \in U(x^*, r)$. We can write from (2.8) that
\[
F(x_0) = F(x_0) - F(x^*) = \int_0^1 F'(x^* + \theta(x_0 - x^*)) (x_0 - x^*) d\theta.
\]
Notice that for each $\theta \in [0, 1]$ and $\|x^* + \theta(x_0 - x^*) - x^*\| = \theta\|x_0 - x^*\| < r$. That is $x^* + \theta(x_0 - x^*) \in U(x^*, r)$. Then using (2.11) and (2.19), we get that
\[
\|F'(x^*)^{-1}F(x_0)\| = \left\| \int_0^1 F'(x^*)^{-1}F'(x^* + \theta(x_0 - x^*)) (x_0 - x^*) d\theta \right\|
\leq M\|x_0 - x^*\|.
\]
Similarly, we obtain that
\[
\|F'(x^*)^{-1}F(y_0)\| \leq M\|y_0 - x^*\|.
\]
Next we shall show that $2F[y_0, x_0] - F'(x_0)$ is invertible. By using (2.8), we obtain
\[
A_0 = 2F[y_0, x_0] - F'(x_0)
= 2F[y_0, x_0] - 2F'(x^*) + 2F'(x^*) - F'(x_0)
= 2(F[y_0, x_0] - F'(x^*)) + F'(x^*) - F'(x_0) + F'(x^*).
\]
Using the equations (2.2), (2.9), (2.12) (2.19) and (2.20), we obtain that
\[
\|F'(x^*)^{-1}(A_0 - F'(x^*))\|
\leq \|F'(x^*)^{-1}(2F[y_0, x_0] - F'(x^*)) + F'(x^*) - F'(x_0)\|
\leq 2\|F'(x^*)^{-1}(2F[y_0, x_0] - F'(x^*))\| + \|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\|
\leq 2L_1(\|x_0 - x^*\| + \|y_0 - x^*\|) + L_0\|x_0 - x^*\|
\leq 2L_1(\|x_0 - x^*\| + g_1(\|x_0 - x^*\|)\|x_0 - x^*\|) + L_0\|x_0 - x^*\|
\leq (2L_1(1 + g_1(\|x_0 - x^*\|) + L_0)\|x_0 - x^*\|
\leq p(\|x_0 - x^*\|) \leq p(r) < 1.
\]
Hence, we get that
\begin{equation}
\|A_0^{-1}F'(x^*)\| = \frac{1}{1 - p(\|x_0 - x^*\|)}.
\end{equation}

Therefore, \(z_0\) is well defined by method M-I for \(n = 0\). Then using the equation (2.2), (2.19), (2.24) and (2.25), we obtain that
\begin{align*}
\|z_0 - x^*\| & \leq \|y_0 - x^*\| + \|A_0^{-1}F'(x^*)\| \|F'(x^*)^{-1}F(y_0)\| \\
& \leq \|y_0 - x^*\| + \frac{M\|y_0 - x^*\|}{1 - p(\|x_0 - x^*\|)} \\
& \leq \left(1 + \frac{M}{1 - p(\|x_0 - x^*\|)}\right)\|y_0 - x^*\| \\
& \leq \left(1 + \frac{M}{1 - p(\|x_0 - x^*\|)}\right)g_1(\|x_0 - x^*\|)\|x_0 - x^*\| \\
& = g_2(\|x_0 - x^*\|)\|x_0 - x^*\| \\
& \leq \|x_0 - x^*\||< r,
\end{align*}
which proves the equation (2.15) for \(n = 0\) and \(z_0 \in U(x^*, r)\). We also have as in (2.22) for \(z_0 = x_0\)
\begin{equation}
\|F'(x^*)^{-1}F(z_0)\| \leq M\|z_0 - x^*\|.
\end{equation}

We can write as \(A_1 = F'(x_0) - F[y_0, x_0] + F[z_0, x_0]\) from M-I and by using (2.9), (2.12) and (2.21), we get that
\begin{align*}
\|F'(x^*)^{-1}(A_1)\| & = \|F'(x^*)^{-1}(F'(x_0) - F[y_0, x_0] + F[z_0, x_0])\| \\
& \leq L_0\|x_0 - x^*\| + L_1(\|x_0 - x^*\| + \|y_0 - x^*\|) \\
& \quad + L_1(\|z_0 - x^*\| + \|y_0 - x^*\|) + 1 \\
& \leq (L_0 + L_1)\|x_0 - x^*\| + 2L_1\|y_0 - x^*\| + L_1\|z_0 - x^*\| + 1 \\
& \leq (L_0 + L_1)\|x_0 - x^*\| + 2L_1g_1(\|x_0 - x^*\|)\|x_0 - x^*\| \\
& \quad + L_1g_2(\|x_0 - x^*\|)\|x_0 - x^*\| + 1. \tag{2.28}
\end{align*}

We must show that \(2F[z_0, y_0] - F[z_0, x_0]\) is invertible. As in (2.23), we have in turn that
\begin{align*}
\|F'(x^*)^{-1}(2F[z_0, y_0] - F[z_0, x_0] - F'(x^*))\| & \leq 2\|F'(x^*)^{-1}(F[z_0, y_0] - F'(x^*))\| + \|F'(x^*)^{-1}(F[z_0, x_0] - F'(x^*))\| \\
& \leq 2L_1(\|z_0 - x^*\| + \|y_0 - x^*\|) + L_1(\|z_0 - x^*\| + \|x_0 - x^*\|) \\
& \leq L_1(3g_2(\|x_0 - x^*\|) + 2g_1(\|x_0 - x^*\|) + 1)\|x_0 - x^*\| \\
& \leq q(\|x_0 - x^*\|) < q(r) < 1, \tag{2.29}
\end{align*}
so
\begin{equation}
\|(2F[z_0, y_0] - F[z_0, x_0])^{-1}F'(x^*)\| = \frac{1}{1 - q(\|x_0 - x^*\|)}.
\end{equation}
Hence, $x_1$ is well defined by last substep of M-I for $n = 0$. Then by using (2.18), (2.27), (2.28), (2.29) and (2.30), we get that

$$
\|x_1 - x^*\| \leq \|x_0 - x^*\| + \|A_0^{-1}F'(x^*)\|\|F'(x^*)^{-1}A_1\|
\|F'(x_0)^{-1}F'(x^*)\|\|F'(x^*)^{-1}F(z_0)\|
= \|x_0 - x^*\| + \frac{1}{1-q(\|x_0 - x^*\|)}(L_0 + L_1)\|x_0 - x^*\|
+ 2L_1g_1(\|x_0 - x^*\|)\|x_0 - x^*\|
+ L_1g_2(\|x_0 - x^*\|)\|x_0 - x^*\| + 1) \times \frac{M\|z_0 - x^*\|}{1-L_0\|x_0 - x^*\|}
= \left(1 + \frac{1}{1-q(\|x_0 - x^*\|)}(L_0 + L_1)\|x_0 - x^*\|
+ 2L_1g_1(\|x_0 - x^*\|)\|x_0 - x^*\|
+ L_1g_2(\|x_0 - x^*\|)\|x_0 - x^*\| + 1) \times \frac{M\|x_0 - x^*\|}{1-L_0\|x_0 - x^*\|}
\right)
\|x_0 - x^*\||x_0 - x^*|
(2.31)
= g_3(\|x_0 - x^*\|)\|x_0 - x^*\| \leq \|x_0 - x^*\| \leq r,
$$
which proves the estimate (2.16) for $n = 0$ and $x_1 \in U(x^*, r)$. By simply replacing $x_0$, $y_0$, $z_0$, $x_1$ by $x_n$, $y_n$, $z_n$, $x_{n+1}$ in the preceding estimates we arrive at (2.14)-(2.16). Then, from the estimates $\|x_{n+1} - x^*\| \leq c\|x_n - x^*\| < r$, where $c = g_3(\|x_0 - x^*\|) \in [0, 1)$ we deduce that $\lim_{n \to \infty} x_n = x^*$ and $x_{n+1} \in U(x^*, r)$.

Finally, we show the uniqueness part, let $Q = \int_0^1 F'(y^* + t(x^* - y^*))dt$ for some $y^* \in D_0$ with $F(y^*) = 0$. Using (2.13), we get that

$$
\|F'(x^*)^{-1}(Q - F'(x^*))\| \leq \int_0^1 L_0\|y^* + t(x^* - y^*) - x^*\|dt
\leq \int_0^1 (1-t)\|x^* - y^*\|dt
(2.32)
\leq \frac{L_0}{2} T < 1.
$$
It follows from (2.32) that $Q$ is invertible. Then, from the identity $0 = F(x^*) - F(y^*) = Q(x^* - y^*)$, we deduce that $x^* = y^*$. \qed

2.2. Convergence of M-II

We present the local convergence analysis of M-II along the same lines of M-I. Here we use the functions $g_1$, $g_3$ and $r_1$, $h_3$ as defined in Subsection 2.1. Define functions $g_2(t)$ and $h_2(t)$ on interval $[0, \frac{1}{M})$ by

$$
g_2(t) = \left(1 + \frac{(2L_0 + L_1 + L_1g_1(t))t + 1)M}{(1 - L_0t)(1 - p(t))}\right)g_1(t), \quad p(t) = L_1(1 + g_1(t))t
$$
and
\[ h_2(t) = g_2(t) - 1. \]
Then, we have that \( h_2(0) = -1 < 0 \) and \( h_2(t) \to +\infty \) as \( t \to r_p^- \). It follows from the intermediate theorem that function \( h_2 \) has zeros in the interval \((0, r_p)\).
Denote by \( r_2 \) the smallest such zero.

Set:
\[ r = \min\{r_i\}, \quad i = 1, 2, 3. \]
Then we have that
\[ 0 < r \leq r_q. \]
Then, for each \( t \in [0, r) \).
\[ 0 \leq g_1(t) \leq 1, \]
\[ 0 \leq g_2(t) \leq 1 \]
and
\[ 0 \leq g_3(t) \leq 1. \]

Next, we present the local convergence analysis of M-II.

**Theorem 2.2.** Suppose that the hypotheses of Theorem 2.1 are satisfied but \( r \) is defined by \((2.33)\). Then, the conclusions of Theorem 2.1 hold with M-II replacing M-I.

**Proof.** According to the proof of Theorem 2.1 we only need to show using mathematical induction that
\[ \|z_n - x^*\| \leq g_2(\|x_0 - x^*\|)\|x_0 - x^*\| \leq \|x_0 - x^*\|. \]
Hence, \( z_0 \) is well defined by the second substep of method M-II for \( n = 0 \).

We must show that \( F[y_0, x_0] \) is invertible. Indeed, we have that
\[ \|F'(x^*)^{-1}(F[y_0, x_0] - F'(x^*))\| \leq L_1(\|y_0 - x^*\| + \|x_0 - x^*\|) \]
\[ \leq L_1(1 + g_1(\|x_0 - x^*\|))\|x_0 - x^*\| \]
\[ = p(\|x_0 - x^*\|) < p(r) < 1, \]
so
\[ \|F[y_0, x_0]^{-1}F'(x^*)\| \leq \frac{1}{1 - p(\|x_0 - x^*\|)}. \]
We also need the estimate
\[ \|F'(x^*)^{-1}(2F'(x_0) - F[y_0, x_0])\| \]
\[ = \|F'(x^*)^{-1}(2F'(x_0) - 2F'(x^*) + 2F'(x^*) - F[y_0, x_0])\| \]
\[ \leq 2L_0\|x_0 - x^*\| + L_1(\|y_0 - x^*\| + \|x_0 - x^*\|) + 1 \]
\[ \leq (2L_0 + L_1g_1(\|x_0 - x^*\|))\|x_0 - x^*\| + 1, \]
so we have the estimate
\[
\| (2F[y_0, x_0] - F(x_0)^{-1}) F'(x^*) F'(x^*)^{-1} F(y_0) \|
\leq \| F[y_0, x_0] - F'(x^*) \| \| F'(x^*)^{-1} (2F'(x_0) - F[y_0, x_0]) \|
\leq \| F'(x_0) - F'(x^*) \| \| F'(x^*)^{-1} F(y_0) \|
\]
(2.41)
leading to
\[
\| z_0 - x^* \|
\leq \| y_0 - x^* \| + \frac{(2L_0 + L_1 + L_1g_1(\|x_0 - x^*\|))\|x_0 - x^*\| + 1} M \|y_0 - x^*\|
\leq \left( 1 + \frac{(2L_0 + L_1 + L_1g_1(\|x_0 - x^*\|))\|x_0 - x^*\| + 1} M \right) g_1(\|x_0 - x^*\|)\|x_0 - x^*\|
= g_2(\|x_0 - x^*\|)\|x_0 - x^*\|\]
(2.42)
which shows (2.38) for \( n = 0 \) and \( z_0 \in U(x^*, r) \). The rest of proof follows as the proof of Theorem 2.1.

\[2.3. \text{Convergence of M-III}\]

We use the definition of functions \( g_1, g_3 \) and \( r_1, h_3 \) given in Subsection 2.1. Then, we define functions \( g_2(t) \) and \( h_2(t) \) on interval \( (0, \frac{1}{L_0}) \) by
\[
g_2(t) = \left( 1 + \frac{(3L_0 + 2L_1 + 2L_1g_1(t))t + 1} M \right)
\]
and
\[
h_2(t) = g_2(t) - 1.
\]
Then, we have that \( h_2(0) = -1 < 0 \) and \( h_2(t) \to \infty \) as \( t \to \frac{1}{L_0} \). It follows from the intermediate theorem that function \( h_2 \) has zeros in the interval \( (0, \frac{1}{L_0}) \). Denote by \( r_2 \) the smallest such zero.

Set:
\[
r = \min\{r_i\}, \quad i = 1, 2, 3.
\]
Then we have that
\[
0 < r \leq r_q.
\]
Then, for each \( t \in [0, r) \),
\[
0 \leq g_1(t) \leq 1,
\]
\[
0 \leq p(t) \leq 1
\]
(2.47) \(0 \leq g_2(t) \leq 1\)

and

(2.48) \(0 \leq g_3(t) \leq 1\).

The local convergence analysis of M-III is presented in an analogous way to M-I using the preceding notations.

**Theorem 2.3.** Suppose that the hypotheses of Theorem 2.1 are satisfied but \(r\) is defined by (2.43). Then, the conclusions of Theorem 2.1 hold with M-III replacing M-I.

**Proof.** According to the proof of Theorem 2.1 we only need to show using mathematical induction that

(2.49) \[\|z_n - x^*\| \leq g_2(\|x_0 - x^*\|)\|x_0 - x^*\| \leq \|x_0 - x^*\|,\]

We have the estimate

\[
\begin{align*}
\|z_0 - x^*\| &\leq \|y_0 - x^*\| + \|F'(x_0)^{-1}F'(x^*)\|\left(3\|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| + 1\right) \\
&\quad + 2\|F'(x^*)^{-1}(F[y_0, x_0] - F'(x^*))\| \\
&\quad \leq \left(1 + \frac{((3L_0 + 2L_1 + 2L_1 g_1(\|x_0 - x^*\|))\|x_0 - x^*\| + 1)M}{(1 - L_0\|x_0 - x^*\|)^2} \right) \\
&\quad \quad \times g_1(\|x_0 - x^*\|)\|x_0 - x^*\| \\
(2.50) &\quad = g_2(\|x_0 - x^*\|)\|x_0 - x^*\| \leq \|x_0 - x^*\| < r,
\end{align*}
\]

which shows (2.49) for \(n = 0\) and \(z_0 \in U(x^*, r)\). The rest of proof follows as the proof of Theorem 2.1. \(\square\)

**Remark 2.4.** It is worth noticing that the methods M-I, M-II and M-III are not changing when we use the conditions of the Theorems 2.1, 2.2 and 2.3 instead of stronger conditions used in ([27], Theorem 1). Moreover, we can compute the computational order of convergence (COC) [30] defined by

(2.51) \[\text{COC} = \ln\left(\frac{\|x_{n+1} - x^*\|}{\|x_n - x^*\|}\right)\left/\ln\left(\frac{\|x_n - x^*\|}{\|x_{n-1} - x^*\|}\right)\right.,\]

or the approximate computational order of convergence (ACOC) [12], given by

(2.52) \[\text{ACOC} = \ln\left(\frac{\|x_{n+1} - x_n\|}{\|x_n - x_{n-1}\|}\right)\left/\ln\left(\frac{\|x_n - x_{n-1}\|}{\|x_{n-1} - x_{n-2}\|}\right)\right.,\]

In this way we obtain in practice the order of convergence.
3. Numerical examples

Here, we shall demonstrate the theoretical results which we have shown in Section 2. We use the divided difference given by
\[ F(x, y) = \frac{1}{2} (F'(x) + F'(y)) \]
or
\[ F(x, y) = \int_0^1 (F'(y + \tau(x - y))d\tau. \]

**Example 3.1.** Suppose that the motion of an object in three dimensions is governed by system of differential equations
\[
\begin{align*}
\frac{df_1}{dx}(x) - f_1(x) - 1 &= 0, \\
\frac{df_2}{dy}(y) - (e - 1)y - 1 &= 0, \\
\frac{df_3}{dz}(z) - 1 &= 0
\end{align*}
\]
with \(x, y, z \in D\) for \(f_1(0) = f_2(0) = f_3(0) = 0\). Then, the solution of the system is given for \(v = (x, y, z)\) by function \(F := (f_1, f_2, f_3) : D \rightarrow \mathbb{R}^3\) defined by
\[
F(v) = \left( e^x - 1, -\frac{e - 1}{2} y^2 + y, z \right)^t.
\]
The Fréchet-derivative is given by
\[
F'(v) = \begin{bmatrix}
e^x & 0 & 0 \\
0 & (e - 1)y + 1 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]
Then for \(x^* = (0, 0, 0)^t\) we have that \(L_0 = e - 1, L = e, L_1 = \frac{L}{2}\) and \(M = 2\). The parameters \(r_1, r_2\) and \(r_3\) using methods M-I, M-II and M-III are given in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>M-I</th>
<th>M-II</th>
<th>M-III</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r_1)</td>
<td>0.324948</td>
<td>0.324948</td>
<td>0.324948</td>
</tr>
<tr>
<td>(r_2)</td>
<td>0.122514</td>
<td>0.113310</td>
<td>0.100785</td>
</tr>
<tr>
<td>(r_3)</td>
<td>0.051586</td>
<td>0.048325</td>
<td>0.045283</td>
</tr>
<tr>
<td>(r_q)</td>
<td>0.159265</td>
<td>0.156776</td>
<td>0.144734</td>
</tr>
<tr>
<td>(r)</td>
<td>0.051586</td>
<td>0.048325</td>
<td>0.045283</td>
</tr>
</tbody>
</table>

Theorems 2.1, 2.2 and 2.3 guarantee the convergence of M-I, M-II and M-III to \(x^* = 0\) provided that \(x_0 \in U(x^*, r)\). This condition yields very close initial approximation.

**Example 3.2.** Let \(X = C[0, 1]\) be the space of continuous functions defined on the interval \([0, 1]\) and be equipped with max norm. Let \(D = U(0, 1)\). Define function \(F\) on \(D\) by
\[
F(\phi)(x) = \phi(x) - 10 \int_0^1 x \phi'(\theta)^3 d\theta.
\]
We have that
\[ F'(\varphi(\xi))(x) = \xi(x) - 30 \int_0^1 x\theta\varphi(\theta)^2\xi(\theta)d\theta \]
for each \( \xi \in D \).

Then for \( x^* = 0 \) we have that \( L_0 = 15, L = 30, L_1 = \frac{L_0}{2} \) and \( M = 1.85 \). The parameters \( r_1, r_2 \) and \( r_3 \) using M-I, M-II and M-III are given in Table 2.

<table>
<thead>
<tr>
<th>M-I</th>
<th>M-II</th>
<th>M-III</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_1 = 0.033333 )</td>
<td>( r_1 = 0.033333 )</td>
<td>( r_1 = 0.033333 )</td>
</tr>
<tr>
<td>( r_2 = 0.012743 )</td>
<td>( r_2 = 0.011754 )</td>
<td>( r_2 = 0.010525 )</td>
</tr>
<tr>
<td>( r_3 = 0.005415 )</td>
<td>( r_3 = 0.005093 )</td>
<td>( r_3 = 0.004794 )</td>
</tr>
<tr>
<td>( r_q = 0.017279 )</td>
<td>( r_q = 0.016969 )</td>
<td>( r_q = 0.015737 )</td>
</tr>
<tr>
<td>( r = 0.005415 )</td>
<td>( r = 0.005093 )</td>
<td>( r = 0.004794 )</td>
</tr>
</tbody>
</table>

It is clear that the convergence of M-I, M-II and M-III is guaranteed to \( x^* = 0 \) provided that \( x_0 \in U(x^*, r) \).

**Example 3.3.** Let us consider the function \( F := (f_1, f_2, f_3): D \to \mathbb{R}^3 \) defined by
\begin{equation}
F(x) = (10x_1 + \sin(x_1 + x_2) - 1, 8x_2 - \cos^2(x_3 - x_2) - 1, 12x_3 + \sin(x_3) - 1)^t,
\end{equation}
where \( x = (x_1, x_2, x_3)^t \).

The Fréchet-derivative is given by
\[
F'(x) = \begin{bmatrix}
10 + \cos(x_1 + x_2) & \cos(x_1 + x_2) & 0 \\
0 & 8 + \sin(2(x_2 - x_3)) & -2\sin(x_2 - x_3) \\
0 & 0 & 12 + \cos(x_3)
\end{bmatrix}.
\]

With the initial approximation \( x_0 = \{0, 0.5, 0.1\}^t \), we obtain the root \( x^* \) of the function (3.4)
\[
x^* = \{0.06897 \ldots, 0.24644 \ldots, 0.07692 \ldots\}^t.
\]

Then we get that \( L_0 = L = 0.269812, L_1 = 1.08139 \) and \( M = 13.0377 \). The parameters \( r_1, r_2 \) and \( r_3 \) using methods M-I, M-II and M-III are given in Table 3.

<table>
<thead>
<tr>
<th>M-I</th>
<th>M-II</th>
<th>M-III</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_1 = 2.470856 )</td>
<td>( r_1 = 2.470856 )</td>
<td>( r_1 = 2.470856 )</td>
</tr>
<tr>
<td>( r_2 = 0.225329 )</td>
<td>( r_2 = 0.245951 )</td>
<td>( r_2 = 0.251249 )</td>
</tr>
<tr>
<td>( r_3 = 0.031669 )</td>
<td>( r_3 = 0.031338 )</td>
<td>( r_3 = 0.031008 )</td>
</tr>
<tr>
<td>( r_q = 0.226114 )</td>
<td>( r_q = 0.237689 )</td>
<td>( r_q = 0.239995 )</td>
</tr>
<tr>
<td>( r = 0.031669 )</td>
<td>( r = 0.031338 )</td>
<td>( r = 0.031008 )</td>
</tr>
</tbody>
</table>
Example 3.4. Lastly, we apply the methods M-I, M-II and M-III to solve systems of nonlinear equations in $\mathbb{R}^2$. The performance is also compared with some existing methods. For example, we choose Newton method (NM), sixth-order methods proposed by Grau et al. [17] and Sharma and Arora [26], and eighth-order method by Noor and Noor [21]. These methods are given as follows:

**Grau-Grau-Noguera method (GGNM-I):**

\[
y_n = x_n - F'(x_n)^{-1}F(x_n), \\
z_n = y_n - \left(2[y_n, x_n; F] - F'(x_n)\right)^{-1}F(y_n), \\
x_{n+1} = z_n - \left(2[y_n, x_n; F] - F'(x_n)\right)^{-1}F(z_n).
\]

**Grau-Grau-Noguera method (GGNM-II):**

\[
y_n = x_n - F'(x_n)^{-1}F(x_n), \\
z_n = y_n - \left(2[y_n, x_n; F] - F'(x_n)^{-1}\right)F(y_n), \\
x_{n+1} = z_n - \left(2[y_n, x_n; F] - F'(x_n)^{-1}\right)F(z_n).
\]

**Sharma-Arora Method (SAM):**

\[
y_n = x_n - F'(x_n)F(x_n), \\
z_n = y_n - \left(3I - 2F'(x_n)^{-1}[y_n, x_n; F]\right)F'(x_n)^{-1}F(y_n), \\
x_{n+1} = z_n - \left(3I - 2F'(x_n)^{-1}[y_n, x_n; F]\right)F'(x_n)^{-1}F(z_n).
\]

**Noor-Noor Method (NNM):**

\[
y_n = x_n - F'(x_n)^{-1}F(x_n), \\
z_n = y_n - F'(y_n)^{-1}F(y_n), \\
x_{n+1} = z_n - F'(z_n)^{-1}F(z_n).
\]

Let us consider the system of nonlinear equations:

\[\sum_{j=1, j \neq i}^{m} x_j - e^{-x_i} = 0, \quad 1 \leq i \leq m,\]

with initial value $x_0 = \{-1, -1, m+\text{times}, -1\}^t$ towards the required solution of the systems for $m = 8, 25, 50, 100$. The corresponding solutions are:

\[x^* = (0.125951, \ldots, 8, 0.125951 \ldots)^t, (0.040031, \ldots, 25, 0.040031 \ldots)^t, (0.020003, \ldots, 50, 0.020003 \ldots)^t\] and $(0.010000, \ldots, 100, 0.010000 \ldots)^t$.

All computations are performed in the programming package Mathematica [31] using multiple-precision arithmetic. For every method, we record the number of iterations ($n$) needed to converge to the solution such that the stopping criterion

\[\|x_{n+1} - x_n\| + \|F(x_n)\| < 10^{-400}\]
is satisfied. In order to verify the theoretical order of convergence, we calculate the approximate computational order of convergence (ACOC) using the formula (2.52). In the comparison of performance of methods, we also include CPU time utilized in the execution of program which is computed by the *Mathematica* command “TimeUsed[]”. For the computation of divided difference we use the formula (see [17])

$$F[x, y]_{ij} = \frac{f(x_1, \ldots, x_{i-1}, y_{i+1}, \ldots, y_m) - f(x_1, \ldots, x_{i-1}, y_i, \ldots, y_m)}{x_j - y_j}, \quad 1 \leq i, j \leq m.$$  

Table 4. Comparison of performance of methods.

<table>
<thead>
<tr>
<th>Methods</th>
<th>NM</th>
<th>GGNM-I</th>
<th>GGNM-II</th>
<th>SAM</th>
<th>NNM</th>
<th>M-I</th>
<th>M-II</th>
<th>M-III</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 8$</td>
<td>20</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>x_{n+1} - x_n</td>
<td></td>
<td>$</td>
<td>5.09(−385)</td>
<td>2.21(−183)</td>
<td>5.49(−247)</td>
<td>8.14(−212)</td>
</tr>
<tr>
<td>ACOC</td>
<td>2.000</td>
<td>6.000</td>
<td>6.000</td>
<td>6.000</td>
<td>6.000</td>
<td>8.000</td>
<td>8.000</td>
<td>8.000</td>
</tr>
<tr>
<td>CPU-time</td>
<td>5.2214</td>
<td>5.2625</td>
<td>5.5476</td>
<td>3.5224</td>
<td>4.4756</td>
<td>3.9017</td>
<td>3.4055</td>
<td>5.2092</td>
</tr>
<tr>
<td>$n = 25$</td>
<td>38</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>x_{n+1} - x_n</td>
<td></td>
<td>$</td>
<td>1.36(−329)</td>
<td>1.52(−237)</td>
<td>1.79(−248)</td>
<td>3.61(−256)</td>
</tr>
<tr>
<td>ACOC</td>
<td>2.000</td>
<td>6.000</td>
<td>6.000</td>
<td>6.000</td>
<td>8.000</td>
<td>8.000</td>
<td>8.000</td>
<td></td>
</tr>
<tr>
<td>CPU-time</td>
<td>35.7182</td>
<td>16.5626</td>
<td>16.3124</td>
<td>15.6871</td>
<td>30.6092</td>
<td>15.3262</td>
<td>15.7533</td>
<td>15.7190</td>
</tr>
<tr>
<td>$n = 50$</td>
<td>63</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>21</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>x_{n+1} - x_n</td>
<td></td>
<td>$</td>
<td>8.06(−239)</td>
<td>1.24(−273)</td>
<td>7.70(−279)</td>
<td>1.03(−284)</td>
</tr>
<tr>
<td>ACOC</td>
<td>2.000</td>
<td>6.000</td>
<td>6.000</td>
<td>6.000</td>
<td>8.000</td>
<td>8.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU-time</td>
<td>253.994</td>
<td>48.9901</td>
<td>51.5638</td>
<td>56.5483</td>
<td>247.235</td>
<td>46.1577</td>
<td>47.2342</td>
<td>50.9234</td>
</tr>
<tr>
<td>$n = 100$</td>
<td>failure</td>
<td>failure</td>
<td>failure</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$</td>
<td></td>
<td>x_{n+1} - x_n</td>
<td></td>
<td>$</td>
<td>–</td>
<td>2.54(−311)</td>
<td>8.00(−314)</td>
<td>1.87(−316)</td>
</tr>
<tr>
<td>ACOC</td>
<td>–</td>
<td>6.000</td>
<td>6.000</td>
<td>6.000</td>
<td>–</td>
<td>8.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU-time</td>
<td>–</td>
<td>232.828</td>
<td>268.125</td>
<td>255.641</td>
<td>–</td>
<td>221.281</td>
<td>238.641</td>
<td>245.187</td>
</tr>
</tbody>
</table>

Numerical results are displayed in Table 4, which include:

- The dimension ($m$) of the system of equations.
- The required number of iterations ($n$).
- The error ($||x_{n+1} - x_n||$) of approximation to the corresponding solution of considered problems, where $A(-h)$ denotes $A \times 10^{-h}$.
- The approximate computational order of convergence (ACOC).
- The elapsed CPU time (CPU-time) in seconds.

It is clear from the numerical results shown in Table 4 that the methods show stable convergence behavior. From the calculation of computational order of convergence, it is also verified that order of convergence is preserved. In the case of $m = 100$, NM and NNM are not converging to the solution. In other cases also they are very slow. Elapsed CPU time shows the efficient nature of present methods. Similar numerical experimentations, carried out for a number of problems of different type, confirmed the above conclusions to a large extent.
References


Ioannis K. Argyros
Department of Mathematics Sciences
Cameron University
Lawton, OK 73505, USA
Email address: iargyros@cameron.edu

Deepak Kumar
Department of Mathematics
Sant Longowal Institute of Engineering and Technology
Longowal, Sangrur, 148106, India
Email address: deepak.babbi@gmail.com

Janak Raj Sharma
Department of Mathematics
Sant Longowal Institute of Engineering and Technology
Longowal, Sangrur, 148106, India
Email address: jrshira@yahoo.co.in