Local convergence for composite Chebyshev-type methods

Ioannis K. Argyros¹ and Santhosh George²*

Abstract
We replace Chebyshev’s method for solving equations requiring the second derivative by a Chebyshev-type second derivative free method. The local convergence analysis of the new method is provided using hypotheses only on the first derivative in contrast to the Chebyshev method using hypotheses on the second derivative. This way we extend the applicability of the method. Numerical examples are also used to test the convergence criteria and to obtain error bounds and also the radius of convergence.

Keywords: Chebyshev method, Newton method, Local convergence, Fréchet derivative, Divided differences

2010 AMS: Primary 65H10, Secondary 65D15, 65G99, 49M17

¹Department of Mathematical Sciences, Cameron University Lawton, OK 73505, USA
²Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, India-757 025
*Corresponding author:sgeorge@nitk.ac.in

Received: 6 July 2018, Accepted: 17 August 2018, Available online: 30 September 2018

1. Introduction
Let \( \mathbb{B}_1, \mathbb{B}_2 \) be Banach spaces, \( \Omega \subseteq \mathbb{B}_1 \) be nonempty and convex set. Numerous problems can be written in the form

\[
F(x) = 0,
\]

using mathematical modeling, where \( F: \Omega \rightarrow \mathbb{B}_2 \) is a continuously Fréchet differentiable operator. Analytical solutions \( x_\ast \) are not easy or impossible to find in general for equation (1.1). This leads researchers and practitioners to use iterative methods to generate a sequence approximating \( x_\ast \).

Newton’s method defined for \( x_0 \in \Omega \) and for each \( n = 0, 1, 2, \ldots \) by

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

(1.2)
is the most popular method for solving equation (1.1). Newton’s method converges quadratically under certain conditions [1, 2, 3, 11]. Higher convergence order methods have also been suggested such as the cubically convergent Chebyshev’s method defined for each \( n = 0, 1, 2, \ldots \) by

\[
x_{n+1} = x_n - (I + B_n)F'(x_n)^{-1}F(x_n),
\]

(1.3)
where \( B_n = F'(x_n)^{-1}F''(x_n)F'(x_n)^{-1}F(x_n) \). If one considers a system of \( k \) equations in \( k \) unknowns, then \( F'(x) \) is a matrix with \( k^2 \) evaluations whereas \( F''(x) \) requires \( \frac{k^2(k+1)}{2} \) evaluations. That is Chebyshev’s method is expensive to implement. Moreover, the convergence requires conditions of the form [6, 7, 8, 9, 10, 11, 12]

\[
\|F'(x^\ast)^{-1}F''(x)\| \leq a \text{ for each } x \in \Omega
\]
We have that \( h \) does not have bounded third derivative in \( X = [-\frac{3}{2}, \frac{3}{2}] \) by

\[
F(x) = \begin{cases} 
    x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\
    0, & x = 0
\end{cases}
\]

Choose \( x^* = 1 \). We have that

\[
\begin{align*}
F'(x) &= 3x^2 \ln x^2 + 5x^4 - 4x^3 + 2x^2, \\
F''(x) &= 6x \ln x^2 + 20x^3 - 12x^2 + 10x, \\
F'''(x) &= 6 \ln x^2 + 60x^2 - 24x + 22.
\end{align*}
\]

Then, obviously function \( F \) does not have bounded third derivative in \( X \). That is why we suggest the method defined for each \( n = 0, 1, 2, \ldots \) by

\[
\begin{align*}
y_n &= x_n - F'(x_n)^{-1} F(x_n) \\
z_n &= y_n - F'(x_n)^{-1} F(y_n) \\
x_{n+1} &= z_n - C_n F'(x_n)^{-1} F(z_n),
\end{align*}
\]

where \( C_n = 2I - F'(x_n)^{-1} [x_n, y_n; F] \) and \([\ldots; F]: \Omega \times \Omega \rightarrow \mathbb{R}_2\) is a divided difference of order one.

The study of convergence of iterative algorithms is usually centered into two categories: semi-local and local convergence analysis. The semi-local convergence is based on the information around an initial point, to obtain conditions ensuring the convergence of these algorithms, while the local convergence is based on the information around a solution to find estimates of the computed radii of the convergence balls. Local results are important since they provide the degree of difficulty in choosing initial points.

Our local convergence analysis uses only hypotheses on the first Fréchet derivative, whereas the order of convergence is established using (COC) and (ACOC) (see Remark 2.2). Hence, we expand the applicability of method (1.4).

Section 2 contains the local convergence of method (1.4), whereas in the concluding Section 3, we provide numerical examples.

### 2. Local convergence

Let \( \varphi_0 : I_0 \rightarrow I_0 \) be a continuous and increasing function with \( \varphi_0(0) = 0 \), where \( I_0 = \mathbb{R}_+ \cup \{0\} \). Suppose that equation

\[
\varphi_0(t) = 1,
\]

has at least one positive solution. Denote by \( \rho_0 \) the smallest such solution. Let \( \varphi : [0, \rho_0) \rightarrow I_0 \) be a continuous and increasing function with \( \varphi(0) = 0 \). Define functions \( g_1 \) and \( h_1 \) on \( [0, \rho_0) \) by

\[
g_1(t) = \int_0^1 \frac{\varphi((1 - \theta)t) d\theta}{1 - \varphi_0(t)}
\]

and

\[
h_1(t) = g_1(t) - 1.
\]

We have that \( h_1(0) = -1 \) and \( h_1(t) \rightarrow +\infty \) as \( t \rightarrow \rho_0^- \). It then follows from the intermediate value theorem that equation \( h_1(t) = 0 \) has at least one solution in the interval \((0, \rho_0)\). Denote by \( \rho_1 \) the smallest such solution.

Suppose that

\[
\varphi_0(g_1(t)t) = 1.
\]
has at least one positive solution. Denote by $\rho_1$ the smallest such solution. Let $\varphi_1 : [0, \rho_0) \rightarrow I_0$ be continuous and increasing function. Moreover, define functions $g_2$ and $h_2$ on $[0, \rho)$ by

$$
g_2(t) = \left[ \int_0^1 \varphi_1((1-\theta)g_1(t))d\theta \right] \frac{\varphi_0(g_1(t))}{1-\varphi_0(g_1(t))} + \left( \varphi(t) + \varphi_0(g_1(t)) \right) \int_0^1 \varphi_1(\theta g_1(t))d\theta \right] \frac{\varphi_0(g_1(t))}{(1-\varphi_0(g_1(t)))(1-\varphi_0(t))} \right] g_1(t)
$$

and

$$
h_2(t) = g_2(t) - 1,
$$

where $\rho = \min\{\rho_0, \rho_1\}$. We get that $h_2(0) = -1$ and $h_2(t) \rightarrow +\infty$ as $t \rightarrow \rho^-$. Denote by $r_2$ the smallest solution of equation $h_2(t) = 0$ in the interval $(0, \rho)$. Let $\varphi_2 : [0, \rho) \times [0, \rho) \rightarrow I_0$ be a continuous and increasing function. Furthermore, define functions $g_3$ and $h_3$ on the interval $[0, \rho)$ by

$$
g_3(t) = \left[ 1 + \left( 1 + \frac{\varphi_0(t) + \varphi_2(g_2(t), g_1(t))}{1-\varphi_0(t)} \right) \times \int_0^1 \varphi_1(\theta g_2(t))d\theta \right] \frac{\varphi_0(t)}{1-\varphi_0(t)} \right] g_2(t)
$$

and

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

We obtain $h_3(t) = -1$ and $h_3(t) \rightarrow +\infty$ as $t \rightarrow \rho^-$. Denote by $r_3$ the smallest solution of equation $h_3(t) = 0$ in $(0, \rho)$. Define the radius of convergence $r$ by

$$
r = \min\{r_i\}, i = 1, 2, 3.
$$

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

$$
0 \leq \varphi_0(t) < 1
$$

and

$$
0 \leq \varphi_1(g_1(t)) < 1
$$

and

$$
0 \leq g_i(t) < 1.
$$

Let $B(u,v), \bar{B}(u,v)$ stand for the open and closed balls in $B_1$, respectively with center $u \in B_1$ and of radius $\tau > 0$.

The local convergence of method (1.2) is based on the conditions (A):

(a1) $F : \Omega \subset B_1 \rightarrow B_2$ is a continuously Fréchet-differentiable operator and $[.,.,F] : \Omega \times \Omega \rightarrow \mathcal{L}(B_1, B_2)$ a divided difference of order one for $F$.

(a2) There exists $x_s \in \Omega$ such that $F(x_s) = 0$ and $F(x_s)^{-1} \in \mathcal{L}(B_2, B_1)$.

(a3) There exist a continuous and increasing function $\varphi_0 : I_0 \rightarrow I_0$ such that for each $x \in \Omega$,

$$
\|F'(x)^{-1}(F'(x) - F'(x_s))\| \leq \varphi_0(\|x - x_s\|).
$$

Set $\Omega_0 = \Omega \cap \bar{B}(x_s, \rho_0)$ where $\rho_0$ is given by (2.1).

(a4) There exist functions $\varphi : [0, \rho_1) \rightarrow I_0, \varphi_1 : [0, \rho_1) \rightarrow I_0, \varphi_2 : [0, \rho_1)^2 \rightarrow I_0$ continuous, increasing with $\varphi(0) = \varphi_2(0,0) = 0$ such that for each $x, y, z \in \Omega_0$

$$
\|F'(x_s)^{-1}(F'(x) - F'(y))\| \leq \varphi(\|x - y\|),
$$

$$
\|F'(x_s)^{-1}F'(x)\| \leq \varphi_1(\|x - x_s\|)
$$

and

$$
\|F'(x_s)^{-1}([y,z;F] - F'(x_s))\| \leq \varphi_2(\|x - x_s\|, \|z - x_s\|).$$
(a5) There exist \( \hat{r} \geq r \) such that

\[
\int_0^1 \varphi_0(\theta \hat{r}) d\theta < 1.
\]

Set \( \Omega_1 = \Omega \cap \hat{B}(x_0, \hat{r}) \). Next, the local convergence analysis of method (1.2) follows:

**Theorem 2.1.** Suppose that the conditions (A) hold. Then, sequence \( \{x_n\} \) generated for \( x_0 \in B(x_0, r) - \{x_*\} \) by method (1.2) is well defined in \( B(x_0, r) \), remains in \( B(x_0, r) \) for each \( n = 0, 1, 2, \ldots \) and converges to \( x_* \), so that

\[
\|y_n - x_*\| \leq g_1(\|x_n - x_*\|) \|x_n - x_*\| \leq \|x_n - x_*\| < \rho
\]

(2.7)

\[
\|z_n - x_*\| \leq g_2(\|x_n - x_*\|) \|x_n - x_*\| \leq \|x_n - x_*\|
\]

(2.8)

and

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

(2.9)

where functions \( g_i, i = 1, 2, 3 \) are defined previously and the radius \( r \) is given in (2.3). Moreover, \( x_* \) is the only solution of equation \( F(x) = 0 \) in \( \Omega_1 \).

**Proof.** Inequalities (2.7)-(2.9) are shown using mathematical induction. First, we shall show that iterates \( \{x_n\} \) are well defined and inequation (2.7)-(2.9) are satisfied for \( n = 0 \). Let \( x \in B(x_0, r) - \{x_*\} \). Using (2.1), (2.3) and (2.4), we have in turn that

\[
\|F'(x)^{-1}(F'(x_0) - F'(x_*))\| \leq \varphi_0(\|x - x_*\|) \leq \varphi_0(r) < 1
\]

(2.10)

which together with the Banach Lemma on invertible operators [1, 4, 11] imply that \( F'(x)^{-1} \in \mathcal{L}(\mathcal{B}_2, \mathcal{B}_1) \) and

\[
\|F'(x)^{-1}F'(x)\| \leq \frac{1}{1 - \varphi_0(\|x - x_*\|)}.
\]

(2.11)

Notice that (2.11) holds for \( x = x_0 \), since \( x_0 \in B(x_0, r) \) and \( y_0, z_0 \) are well defined by the first and second sub-step of method (1.2) for \( n = 0 \). We have by the first substep of method (1.2) for \( n = 0 \)

\[
\begin{align*}
y_0 - x_* & = x_0 - x_* - F'(x_0)^{-1}F'(x_0) \\
& = F'(x_0)^{-1} \int_0^1 (F'(x_0 + \theta(x_0 - x_*)) - F'(x_0))(x_0 - x_*) d\theta.
\end{align*}
\]

(2.12)

By (a1)-(a4), (2.3), (2.6) for \( i = 1 \), (2.11) and (2.12), we get in turn that

\[
\begin{align*}
\|y_0 - x_*\| & \leq \|F'(x_0)^{-1}F'(x_0)\| \int_0^1 \|F'(x_0 + \theta(x_0 - x_*)) - F'(x_0)\| \|x_0 - x_*\| d\theta \\
& \leq \int_0^1 \varphi((1 - \theta)\|x - x_*\|) d\theta \|x_0 - x_*\| \\
& \leq \|x_0 - x_*\| < r
\end{align*}
\]

(2.13)

which shows (2.7) for \( n = 0, y_0 \in B(x_0, r) \) and (2.11) hold for \( y = y_0 \). That is

\[
\|F'(y_0)^{-1}F'(x_0)\| \leq \frac{1}{1 - \varphi_0(\|y_0 - x_*\|)}
\]

(2.14)

We can write

\[
F(x_0) = F(x_0) - F(x_*) = \int_0^1 F'(x_0 + \theta(x_0 - x_*))(x_0 - x_*) d\theta.
\]

(2.15)
In view of (a4) (second condition) and (2.15), we obtain
\[
\|F'(x_i)^{-1}F(x_0)\| = \|\int_0^1 F'(x_i)^{-1}F'(x_i + \theta(x_0 - x_i))d\theta(x_0 - x_i)\|
\leq \int_0^1 \varphi_1(\|x_0 - x_i\|)d\theta\|x_0 - x_i\|.
\]
(2.16)

Then, using the second substep of method (1.2), (2.3), (2.6) (for \(i = 2\)), (2.13) (for \(x_0 = y_0\)), (2.14) and (2.16) (for \(y_0 = x_0\), we have in turn from
\[
0 = F(y_0) - F(x_0) = F(x_0)^{-1}(F(y_0) - F(x_0))F(x_0)^{-1}F(y_0),
\]
(2.17)
so
\[
\|z_0 - x_*\| \leq \|y_0 - x_*\| + \|F'(y_0)^{-1}F'(x_0)\|
\times \left[\|F'(x_0)^{-1}(F'(x_0) - F'(x_*))\| + \|F'(x_0)^{-1}(F'(y_0) - F'(x_*))\|\right]
\times \left[\int_0^1 \varphi_1((1 - \theta))(y_0 - x_*)d\theta\right]
\times \left[\int_0^1 \varphi_1(\|x_0 - x_*\|)d\theta\right]
\leq g_2(\|y_0 - x_*\|)\|x_0 - x_*\| \leq \|x_0 - x_*\| < r
\]
(2.18)
which shows (2.8) for \(n = 0\) and \(z_0 \in B(x_*, r)\). The third substep of method (1.4) together with (2.3), (2.6) (for \(i = 3\), (2.15) (for \(x_0 = z_0\)), the third hypothesis in (a4) and (2.18), we get
\[
\|x_1 - x_*\| \leq \|z_0 - x_*\|
\times \left[1 + \|F'(x_0)^{-1}(F'(x_0) - [z_0, y_0; F])F'(x_0)^{-1}F(y_0)\|
\times \left[\int_0^1 \varphi_1(\|x_0 - x_*\|)d\theta\right]
\leq g_3(\|z_0 - x_*\|)\|x_0 - x_*\| \leq \|x_0 - x_*\| < r
\]
(2.19)
which shows (2.9) and \(z_0 \in B(x_*, r)\). The induction for inequation (2.7)-(2.9) is completed replacing \(x_0, y_0, z_0, x_1\) by \(x_m, y_m, z_m, x_{m+1}\) in the preceding estimates. We then also have that
\[
\|x_{m+1} - x_*\| \leq g\|x_m - x_*\| < r
\]
(2.20)
where \(g = g_3(\|x_0 - x_*\|) \in (0, 1)\), leading to \(\lim_{m \to +\infty} x_m = x_*\) and \(x_{m+1} \in B(x_*, r)\). The, uniqueness part is shown as follows: Let \(Q = \int_0^1 F'(x_i + \theta(y_i - x_i))d\theta\) for some \(y_1 \in \Omega_1\) with \(F(y_1) = 0\). The condition (a5) gives
\[
\|F'(x_i)^{-1}(Q - F'(x_*))\| \leq \int_0^1 \varphi_0(\|y_1 - x_*\|)d\theta \leq \int_0^1 \varphi_0(\|r\|)d\theta < 1,
\]
(2.21)
so \(Q^{-1} \in \mathcal{L}(\mathcal{B}_2, \mathcal{B}_1)\) and from the identity
\[
0 = F(y_1) - F(x_1) = Q(y_1 - x_*),
\]
(2.22)
we deduce that \(x_* = y_*\). □
Remark 2.2. 1. The second condition in (a4) can be dropped, since this condition follows from (a3), if we set
\[ \varphi_1(t) = 1 + \varphi_0(t). \]

2. The results obtained here can be used for operators \( F \) satisfying autonomous differential equations [11] of the form
\[ F'(x) = P(F(x)) \]
where \( P \) is a continuous operator. Then, since \( F'(x^*) = P(F(x^*)) = P(0) \), we can apply the results without actually knowing \( x^* \). For example, let \( F(x) = e^x - 1 \). Then, we can choose: \( P(x) = x + 1 \).

3. The radius \( r \) was shown by us to be the convergence radius of Newton’s method [1, 2, 3, 4, 5]

\[ x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad \text{for each } n = 0, 1, 2, \ldots \]  \hspace{1cm} (2.23)

under the conditions (a1)–(a4) for \( \varphi_0(t) = L_0 t \) and \( \varphi(t) = L t \). It follows from the definition of \( r \) that the convergence radius \( r_1 \) of the method (1.4) cannot be larger than the convergence radius \( r_1 \) of the second order Newton’s method (2.23). As already noted in [11] \( r_1 \) is at least as large as the convergence ball given by Rheinboldt [11]

\[ r_R = \frac{2}{3L}. \] \hspace{1cm} (2.24)

In particular, for \( L_0 < L \) we have that

\[ r_R < r_1 \]

and

\[ \frac{r_R}{r_1} \rightarrow \frac{1}{3} \quad \text{as } \frac{L_0}{L} \rightarrow 0. \]

That is our convergence ball \( r_1 \) is at most three times larger than Rheinboldt’s. The same value for \( r_R \) was given by Traub [13].

4. It is worth noticing that method (1.2) is not changing when we use the conditions of Theorem 2.1 instead of the stronger conditions used in [14]. Moreover, we can compute the computational order of convergence (COC) defined by

\[ \xi = \ln \left( \frac{|x_{n+1} - x_s|}{|x_n - x_s|} \right) / \ln \left( \frac{|x_n - x_s|}{|x_{n-1} - x_s|} \right) \]

or the approximate computational order of convergence

\[ \xi_1 = \ln \left( \frac{|x_{n+1} - x_n|}{|x_n - x_{n-1}|} \right) / \ln \left( \frac{|x_n - x_{n-1}|}{|x_{n-1} - x_{n-2}|} \right). \]

This way we obtain in practice the order of convergence in a way that avoids the bounds involving estimates using estimates higher than the first Fréchet derivative of operator \( F \).

3. Numerical examples

In this Section the divided difference is given by [x, y; F] = \( \int_0^1 F'(y + \theta(x - y))d\theta \).

Example 3.1. Returning back to the example in the introduction, we have for \( \varphi_0(t) = \varphi(t) = 147t, \varphi_1(t) = 1 + \varphi_0(t), \varphi_2(s, t) = \frac{1}{2}(\varphi_0(s) + \varphi_0(t)) \). Using the definition of \( r \) we obtain

\[ r_1 = 0.0045, r_2 = 0.0029 = r, r_3 = 0.0039. \]

Example 3.2. Let \( \mathcal{X} = \mathcal{Y} = \mathbb{R}^3, \Omega = \bar{U}(0, 1), x^* = (0, 0, 0)^T \). Define function \( F \) on \( \Omega \) for \( w = (x, y, z)^T \) by

\[ F(w) = (e^x - 1, \frac{e - 1}{2}y^2 + y, z)^T. \]
Then, the Fréchet-derivative is defined by

\[
F'(v) = \begin{bmatrix}
  e^x & 0 & 0 \\
  0 & (e - 1)y + 1 & 0 \\
  0 & 0 & 1
\end{bmatrix}.
\]

Then, we have \( \varphi_0(t) = (e - 1)t, \varphi(t) = e^{\frac{1}{3}t}, \varphi_1(t) = 1 + \varphi_0(t), \varphi_2(s,t) = \frac{1}{2}(\varphi_0(s) + \varphi_0(t)). \) Using the definition of \( r \) we obtain

\[
r_1 = 0.4977, \quad r_2 = 0.3731 = r, \quad r_3 = 0.4951.
\]

**Example 3.3.** Let \( X = Y = C[0,1] \), be the space of continuous functions on \( [0,1] \) equipped with the max-norm. Let \( \Omega = \bar{U}(0,1) \). Define \( F \) on \( \Omega \) by

\[
F(\varphi)(x) = \varphi(x) - 10 \int_0^1 x\theta\varphi(\theta)^3 d\theta.
\]

We have that

\[
[F'(\varphi(\xi))](x) = \xi(x) - 30 \int_0^1 x\theta\varphi(\theta)^2 d\theta, \text{ for each } \xi \in D.
\]

Then, we get that \( x^* = 0, \varphi_0(t) = 15t, \varphi(t) = 30t, \varphi_1(t) = 1 + \varphi_0(t), \varphi_2(s,t) = \frac{1}{2}(\varphi_0(s) + \varphi_0(t)). \) We obtain

\[
r_1 = 0.0333, \quad r_2 = 0.0197 = r, \quad r_3 = 0.0350.
\]

**References**


