LOCAL CONVERGENCE
OF A TWO-STEP GAUSS-NEWTON WERNER-TYPE METHOD FOR SOLVING LEAST SQUARES PROBLEMS

IOANNIS K. ARGYROS[*], SANTHOSH GEORGE[‡]

Abstract. The aim of this paper is to extend the applicability of a two-step Gauss-Newton-Werner-type method (TGNWTM) for solving nonlinear least squares problems. The radius of convergence, error bounds and the information on the location of the solution are improved under the same information as in earlier studies. Numerical examples further validate the theoretical results.

MSC. 65G99, 65J15, 65H10, 49M15, 47H17. 65N35, 47H17, 49M15.

Keywords. Gauss-Newton method, Werner's method, local convergence, least squares problem, average Lipschitz condition.

1. INTRODUCTION

Let i, j be natural numbers with $j \geq i$. Let also Ω be an open and convex subset of $\mathbb{R}^j$. We are concerned with the solution p of the least squares problem [4, 5, 6, 7, 8, 9]:

$$(1.1) \qquad \min_{x \in \Omega} f(x) := \tfrac{1}{2} F(x)^T F(x),$$

where $F : \Omega \longrightarrow \mathbb{R}^j$ is a Fréchet-differentiable mapping. Numerous problems can be brought in the form (1.1) using Mathematical Modeling [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. The closed form solutions can only be found in special cases. That explains why most solution methods for problem (1.1) are iterative. Let $x_0, y_0 \in \Omega$ and set $z = \frac{x_0 + y_0}{2}$. In the present study, we provide the local convergence analysis of GNWTM defined for each $n = 0, 1, 2, \ldots$ by

$$(1.2) \qquad \begin{aligned} x_{n+1} &= x_n - A_n F(x_n) \\ y_{n+1} &= x_{n+1} - A_n F(x_{n+1}) \\ z_n &= \tfrac{x_n + y_n}{2}, \end{aligned}$$

where $A_n = [F'(z_n)^T F'(z_n)]^{-1} F'(z_n)^T$. If $i = j$, TGNWTM reduces to a Gauss-Newton-Werner type method [3, 8, 9]. Notice that in each iteration the

---

[*]Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA, e-mail: iargyros@cameron.edu.

[‡]Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, India-575 025, e-mail: sgeorge@nitk.edu.in.

inversion of $[F'(z_n)^T F'(z_n)]^{-1}$ is required only once. Therefore, the computational cost is essentially the same as in the Gauss-Newton method. The LL^T decomposition of $[F'(z_n)^T F'(z_n)]^{-1}$ costs $\mathcal{O}(n^3)$ floating-point operations (Flops) leading to the computation of x_{n+1}. It then follows from the second substep of method (1.2) that $\mathcal{O}(n^2)$ Flops are needed for the computation of y_{n+1}.

The local convergence analysis of method (1.2) was given in the elegant paper by Shakhno *et al.* in [9] (see also related work in [3, 8]). Their convergence analysis uses average Lipschitz continuity condition as well as Lipschitz conditions.

Using the concept of the average Lipschitz continuity [12] and our new idea of restricted convergence domains, we present a local convergence analysis with the following advantages (**A**) over works using the similar information [3, 4, 8, 9, 11, 12, 13]:

(a) Larger radius of convergence;
(b) Tighter error bounds on the distances $\|x_n - p\|$;
(c) An at least as precise information on the location of the solution p.

Achieving (a)–(c) is very important in computational sciences, since: (a)′ We obtain a wider choice of initial guesses; (b)′ Fewer iterates are required to obtain a desired error tolerance; (c)′ Better information about the ball of convergence is obtained.

The rest of the paper is structured as follows: Section 2 contains the local convergence analysis of method (1.2) whereas special cases and the applications are presented in the concluding Section 3.

2. LOCAL CONVERGENCE ANALYSIS

Set $U(w, \rho) = \{v \in \mathbb{R}^j : \|v - w\| < \rho\}$ to be the open ball in $\mathbb{R}^j$ and by $\bar{U}(w, \rho)$ to denote its closure. Let $R > 0$. Define parameter R_1 by $R_1 := \sup\{t \in [0, R] : \bar{U}(p, t) \subset \Omega\}$. The convergence analysis of numerous iterative methods has been given using the following concept due to Wang [12]:

DEFINITION 1. *Mapping $F : \bar{U}(p, R_1) \longrightarrow \mathbb{R}^i$ satisfies the Lipschitz condition with L_1 average on $U(p, R_1)$ if*

$$\|F'(x) - F'(y)\| \leq \int_0^{\|x-y\|} L_1(u) du \text{ for each } x, y \in \bar{U}(p, R_1),$$

where L_1 is a positive non-decreasing function.

It turns out that the convergence analysis of iterative methods based on the preceding notion can be improved as follows:

DEFINITION 2. *The mapping $F : \bar{U}(p, R_1) \longrightarrow \mathbb{R}^i$ satisfies the center-Lipschitz condition with L_0 average on $U(p, R_1)$ if*

$$\|F'(x) - F'(p)\| \leq \int_0^{\|x-p\|} L_0(u) du \text{ for each } x \in \bar{U}(p, R_1),$$

where L_0 is a positive non-decreasing function.

Clearly, we have that

$$(2.1) \qquad L_0(u) \le L_1(u) \text{ for each } u \in [0, R_1],$$

and $\frac{L_0}{L_1}$ can be arbitrary small [2, 3, 4]. Let $\beta > 0$ be a parameter. Suppose that equation

$$(2.2) \qquad \beta \int_0^t L_0(u) du = 1$$

has positive solutions. Denote by R_0 the smallest such solution. Notice for example that R_0 exists, if

$$(2.3) \qquad \beta L_0(R_1) R_1 \ge 1.$$

Indeed, function $g(t) := \beta \int_0^t L(u) du - 1$ is such that $g(0) = -1 < 0$ and $g(R_1) = \beta L(R_1) R_1 - 1 > 0$. The existence of R_0 follows from the intermediate value theorem.

DEFINITION 3. *The mapping $F : \bar{U}(p, R_1) \longrightarrow \mathbb{R}^i$ satisfies the restricted Lipschitz condition with L average on $U(p, R_0)$ if*

$$\|F'(x) - F'(y)\| \le \int_0^{\|x-y\|} L(u) du \text{ for each } x, y \in \bar{U}(p, R_0),$$

where L is a positive non-decreasing function.

We have that

$$(2.4) \qquad L(u) \le L_1(u) \text{ for each } u \in [0, R_0],$$

since $R_0 \le R_1$. Throughout this paper, we suppose that

$$(2.5) \qquad L_0(u) \le L(u) \text{ for each } u \in [0, R_0],$$

unless, otherwise stated. Otherwise, *i.e.*, if

$$(2.6) \qquad L(u) \le L_0(u) \text{ for each } u \in [0, R_0],$$

then the results that follow hold with L_0 replacing L. Moreover, we need the definitions:

DEFINITION 4 ([12]). *Let $F : \bar{U}(p, R_1) \longrightarrow \mathbb{R}^i$ be a twice Fréchet-differentiable mapping. We say that mapping F satisfies the Lipschitz condition with M_1 average on $U(p, R_1)$ if*

$$\|F''(x) - F''(y)\| \le \int_0^{\|x-y\|} M_1(u) du \text{ for each } x, y \in \bar{U}(p, R_1),$$

where M_1 is a positive nondecreasing function.

DEFINITION 5. *Let $F : \bar{U}(p, R_0) \longrightarrow \mathbb{R}^i$ be twice Fréchet-differentiable mapping. We say that mapping F satisfies the restricted Lipschitz condition with M average on $U(p, R_0)$ if*

$$\|F''(x) - F''(y)\| \leq \int_0^{\|x-y\|} M(u)du \text{ for each } x, y \in \bar{U}(p, R_0),$$

where M is a positive nondecreasing function.

We have that

(2.7) $$M(u) \leq M_1(u) \text{ for each } u \in [0, R_0].$$

It is worth noticing that the definition of functions L and M (based on L_0 and R_0) was not possible in earlier studies using L_1 and M_1. That is, $L = L(L_0, R_0, R_1), M = M(L_0, R_0, R_1)$, whereas $L_1 = L_1(R_1)$ and $M_1 = M_1(R_1)$. It turns out that L_0 can replace the less precise L in the computation of the upper bounds on the inverses of the operators involved and $\bar{U}(p, R_0), L, M$ can replace $\bar{U}(p, R_1), L_1, M_1$, respectively in the proofs of such results. Moreover, notice that the iterates x_n lie in $\bar{U}(p, R_0)$ which is a more precise location than $\bar{U}(p, R_1)$ used in earlier studies [2, 3, 4, 8, 11, 12, 13]. We shall make the paper as self contained as possible by stating some standard auxiliary concepts and results.

Denote by $\mathbb{R}^{i \times j}$ the set of all $i \times j$ matrices. The Moore-Penrose pseudo-inverse is defined by $A^\dagger = (A^T A)^{-1} A^T$ for each full rank $A \in \mathbb{R}^{i \times j}$ [6].

LEMMA 2.1 ([2, 6]). *Let $A, A_1 \in \mathbb{R}^{m \times n}$. Assume that $A_2 = A + A_1, \|A^\dagger\|\|A_1\| < 1$, and $\mathrm{rank}(A) = \mathrm{rank}(A_2)$. Then,*

$$\|A_2^\dagger\| \leq \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|A_1\|}.$$

If $\mathrm{rank}(A) = \mathrm{rank}(A_2) = \min\{m, n\}$, the following holds

$$\|A_2 - A^\dagger\| \leq \frac{\sqrt{2}\|A^\dagger\|^2\|A_1\|}{1 - \|A^\dagger\|\|A_1\|}.$$

LEMMA 2.2 ([5]). *Let $A, A_1 \in \mathbb{R}^{m \times n}$. Assume that $A_2 = A + A_1, \|A_1 A^\dagger\| < 1$, and $\mathrm{rank}(A) = n$, then $\mathrm{rank}(A_2) = n$.*

LEMMA 2.3 ([12]). *Let $\varphi(t) = \frac{1}{t} \int_0^t P(u)du, 0 \leq t \leq \rho$, where $P(u)$ is a positive integrable function and monotonically non-decreasing on $[0, \rho]$. Then, $\varphi(t)$ is monotonically non-decreasing with respect to t.*

LEMMA 2.4 ([12]). *Let $\psi(t) = \frac{1}{t^3} \int_0^t Q(u)du, 0 \leq t \leq \rho$, where $Q(u)$ is a positive integrable function and monotonically non-decreasing on $[0, \rho]$. Then, $\psi(t)$ is monotonically non-decreasing with respect to t.*

As in [9], it is convenient for the local convergence analysis that follows to introduce some functions and parameters:

$$\alpha = \|F(p)\|, \ \beta = \|(F'^T F'^{-1} F'^T\|,$$

$$
\begin{aligned}
d(x) &= \|x - p\|, \; s_0 = \max\{d(x_0), d(y_0)\}, \\
\mu(t) &= \mu(L_0, L, M)(t) \\
&= \frac{\beta}{8} \int_0^t M(u)(t-u)^2 du \\
&\quad + \beta t \left(\int_0^{\frac{3}{2}t} L(u) du + \int_0^t L_0(u) du \right) + \sqrt{2}\alpha\beta^2 \int_0^t L_0(u) du - t, \\
\gamma &= \gamma(L_0, L, M) \\
&= \frac{\beta \int_0^{d(x_0)} M(u)(d(x_0)-u)^2 du}{8 d(x_0) \left(1 - \beta \int_0^{d(x_0)} L_0(u) du \right)} + \frac{\beta d(x_0) \int_0^{\frac{d(x_0)+d(y_0)}{2}} L(u) du}{\frac{2d(x_0)+d(y_0)}{3} \left(1 - \beta \int_0^{d(z_0)} L_0(u) du \right)} \\
&\quad + \frac{\sqrt{2}\alpha\beta^2 \int_0^{d(z_0)} L_0(u) du}{d(z_0)(1 - \beta \int_0^{d(z_0)} L_0(u) du)} < 1, \\
\delta &= \frac{\beta \int_0^{d(x_0)} M(u)(d(x_0)-u)^2 du}{8 d(x_0)^3 \left(1 - \beta \int_0^{d(z_0)} L_0(u) du \right)}, \\
\tau &= \frac{\sqrt{2}\alpha\beta^2 \int_0^{d(z_0)} L_0(u) du}{d(z_0) \left(1 - \beta \int_0^{d(z_0)} L_0(u) du \right)}, \\
\lambda &= \frac{\beta \int_0^{\frac{d(x_0)+d(y_0)}{2}} L(u) du}{\frac{2d(x_0)+d(y_0)}{3} \left(1 - \beta \int_0^{d(z_0)} L_0(u) du \right)}, \\
e_{n+1}^1 &= \delta d(x_n)^3 + \lambda d(x_n) d(y_n) + \tau d(z_n), \\
e_{n+1}^2 &= \delta d(x_{n+1})^3 + \frac{\lambda}{3}(d(x_n) + d(y_n) + d(x_{n+1})) d(x_{n+1}) + \tau d(z_n),
\end{aligned}
$$

and

$$
s_{n+1} = \max\{d(x_{n+1}), d(y_{n+1})\}.
$$

Notice that if $L_0 = L = L_1$ and $M = M_1$, then the preceding definitions reduce to the corresponding ones in [9].

The local convergence analysis is based on the conditions ($\mathcal{C}$):

($\mathcal{C}_1$) Mapping $F : \bar{U}(p, R_1) \longrightarrow \mathbb{R}^i$ is twice Fréchet-differentiable, $F'(p)$ has full rank and p solves problem (1.1).

($\mathcal{C}_2$) $F'(x)$ satisfies: the center-Lipschitz condition with L_0 average on $\bar{U}(p, R_1)$ and the restricted Lipschitz condition with L average on $\bar{U}(p, R_0)$; $F''(x)$ satisfies the restricted Lipschitz condition with M average on $\bar{U}(p, R_0)$, where L_0, L and M are positive non-decreasing functions on $[0, \frac{3R_0}{2}]$.

($\mathcal{C}_3$) Function μ has a minimal zero R^* in $[0, R_0]$, which also satisfies

$$
\beta \int_0^{R^*} L_0(u) du < 1.
$$

Then, we can show the following local convergence result for TGNWTM under the conditions ($\mathcal{C}$) and the preceding notation.

THEOREM 2.5. *Suppose that conditions ($\mathcal{C}$) hold. Then, sequences $\{x_n\}$, $\{y_n\}, \{z_n\}$ generated for $x_0, y_0 \in \bar{U}(p, R^*) - \{p\}$ by TGNWTM are well defined in $\bar{U}(p, R^*)$, remain in $\bar{U}(p, R^*)$ for each $n = 0, 1, 2, \ldots$ and converge to p. Moreover, the following estimates hold*

$$(2.8) \qquad\qquad d(x_{n+1}) \le e_{n+1}^1,$$

$$(2.9) \qquad\qquad d(y_{n+1}) \le e_{n+1}^2,$$

and

$$(2.10) \qquad\qquad s_{n+1} \le \gamma s_k \le \ldots \le \gamma^{n+1} s_0.$$

Proof. The proof follows the corresponding one in [9] but there are differences where we use (L_0, L), M instead of L_1, M_1, respectively used in [9]. We shall use mathematical induction to show that iterates $\{x_k\}$, $\{y_k\}, \{z_k\}$ are well defined converge to p and the error estimates (2.8)–(2.10) are satisfied. Using TGNWTM for $n = 0$, we can write

$$
\begin{aligned}
x_1 - p &= x_0 - p - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0) F(x_0) \\
&= [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T [F'(z_0)(x_0 - p) - F(x_0) + F(p)] \\
&\quad + [F'^T F'^{-1} F'^T F(p) - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T F(p) \\
&= [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T J(x_0) \\
&\quad + [F'^T F'^{-1} F'^T F(p) - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T F(p),
\end{aligned}
$$

and

$$
\begin{aligned}
y_1 - p &= x_1 - p - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T F(x_1) \\
&= [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T [F'(z_0)(x_1 - p) - F(x_1) + F(p)] \\
&\quad + [F'^T F'^{-1} F'^T F(p) - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T F(p) \\
&= [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T J(x_1) \\
&\quad + [F'^T F'^{-1} F'^T F(p) - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0) F(p),
\end{aligned}
$$

where

$$
\begin{aligned}
J(x_i) &= F'\left(\tfrac{x_i + p}{2}\right)(x_i - p) - F(x_i) + F(p) \\
&\quad + \left(F'(z_0) - F'\left(\tfrac{x_i + p}{2}\right)\right)(x_i - p).
\end{aligned}
$$

In view of the estimate

$$
\begin{aligned}
F(x) - F(y) - F'\left(\tfrac{x+y}{2}\right)(x - y) = \int_0^1 \tfrac{1-t}{4} &\left(F''\left(\tfrac{x+y}{2} + \tfrac{t}{2}(x - y)\right) \right. \\
&\left. - F''\left(\tfrac{x+y}{2} + \tfrac{t}{2}(y - x)\right) \right)(x - y)^2 dt,
\end{aligned}
$$

for $x = p$ and $y = x_0$, we obtain in turn

$$\left\| F(p) - F(x_0) - F'\left(\tfrac{x_0+p}{2}\right)(p - x_0) \right\| =$$

$$= \tfrac{1}{4}\left\| \int_0^1 (1-t)\left[F''\left(\tfrac{x_0+p}{2} + \tfrac{t}{2}(p-x_0)\right) - F''\left(\tfrac{x_0+p}{2} + \tfrac{t}{2}(x_0-p)\right) \right](p-x_0)^2\, dt \right\|$$

$$\leq \int_0^1 (1-t) \int_0^{t\|x_0-p\|} M(u)\, du\, \|x_0 - p\|^2\, dt$$

$$= \tfrac{1}{8}\int_0^{d(x_0)} M(u)\left(1 - \tfrac{u}{d(x_0)}\right)^2 du\, d(x_0)^2$$

$$= \tfrac{1}{8}\int_0^{d(x_0)} M(u)(d(x_0) - u)^2\, du,$$

and

$$\left\| F'\left(\tfrac{x_0+y_0}{2}\right) - F'\left(\tfrac{x_0+p}{2}\right) \right\| \leq \int_0^{d(y_0)/2} L(u)du.$$

By the central Lipschitz condition, we have that

$$\left\| \left(F'(p)^T F'(p)\right)^{-1} F'(p)^T \right\| \|F'(x) - F'(p)\| \leq \beta \int_0^{d(x)} L_0(u)du.$$

Moreover, by Lemma 2.1 and Lemma 2.2 and $(\mathcal{C}_1)$, for all $x \in U(p, R^*)$, we get

$$\left\| \left(F'(x)^T F'(x)\right)^{-1} F'(x)^T \right\| \leq \frac{\beta}{1 - \beta \int_0^{d(x)} L_0(u)du}$$

and

$$\left\| \left(F'(x)^T F'(x)\right)^{-1} F'(x)^T - \left(F'(p)^T F'(p)\right)^{-1} F'(p)^T \right\| \leq \frac{\sqrt{2}\beta^2 \int_0^{d(x)} L_0(u)du}{1 - \beta \int_0^{d(x)} L_0(u)du}.$$

By the monotonicity of $L(u)$ and $M(u)$ with Lemma 2.3 and Lemma 2.4, functions $\tfrac{1}{t}\int_0^t L(u)du$ and $\tfrac{1}{t^3}\int_0^t M(u)(t-u)^2 du$ are non-decreasing in t. That is, by $(\mathcal{C}_3)$

$$\gamma = \tfrac{1}{R_0}\left[\frac{\beta \int_0^{R_0} M(u)(R_0 - u)^2 du}{8(1 - \beta \int_0^{R_0} L(u)du)} + \frac{\beta R_0 \int_0^{\frac{3}{2}R_0} L(u)du}{1 - \beta \int_0^{R_0} L_0(u)du} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{R_0} L_0(u)du}{1 - \beta \int_0^{R_0} L_0(u)du} \right]$$

$$< \tfrac{1}{R}\left[\frac{\beta \int_0^{R} M(u)(R - u)^2 du}{8(1 - \beta \int_0^{R} L_0(u)du)} + \frac{\beta R \int_0^{\frac{3}{2}R} L(u)du}{1 - \beta \int_0^{R} L_0(u)du} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{R} L_0(u)du}{1 - \beta \int_0^{R} L_0(u)du} \right]$$

$$\leq 1.$$

Thus, by Lemma 2.1–2.9 and condition $(\mathcal{C}_2)$, we have in turn

$$\begin{aligned} \|x_1 - p\| &\leq \|[F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T \| \\ &\quad \times \left\| \int_0^1 (1-t)\left[F''\left(\tfrac{x_0+p}{2} + \tfrac{t}{2}(p-x_0)\right) \right. \right. \end{aligned}$$

$$-F''\left(\tfrac{x_0+p}{2} + \tfrac{t}{2}(x_0 - p)\right)\Big](p - x_0)^2\,dt\Big\|$$

$$+\|[F'^T F'^{-1} F'^T F(p) - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T F(p)\|$$

$$\leq \quad \frac{\beta d(x_0)^3 \int_0^{d(x_0)} M(u)(d(x_0) - u)^2 du}{8d(x_0)^3(1 - \beta \int_0^{d(x_0)} L_0(u)du)}$$

$$+\frac{\beta d(x_0)d(y_0) \int_0^{d(y_0)/2} L(u)du}{d(y_0)(1 - \beta \int_0^{d(x_0)} L_0(u)du)} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{d(x_0)} L_0(u)du}{d(z_0)(1 - \beta \int_0^{d(x_0)} L_0(u)du)}$$

$$< \quad \delta d(x_0)^3 + \lambda d(x_0)d(y_0) + \tau d(z_0) < \delta R_0 < R.$$

In an analogous way, we get in turn

$$\|y_1 - p\| \leq$$

$$\leq \|[F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T\|$$

$$\times \left\| F'\left(\tfrac{x_1+p}{2}\right)(x_1 - p) - F(x_1) + F(p) + \left(F'(z_0) - F'\left(\tfrac{x_1+p}{2}\right)\right)(x_1 - p)\right\|$$

$$+ \left\| [F'^T F'^{-1} F'^T F(p) - [F'(z_0)^T F'(z_0)]^{-1} F'(z_0)^T F(p)\right\|$$

$$\leq \frac{\beta d(x_1)^3 \int_0^{d(x_1)} M(u)(d(x_1) - u)^2 du}{8d(x_1)^3 \left(1 - \beta \int_0^{d(x_1)} L_0(u)du\right)}$$

$$+ \frac{\beta d(x_1)d(z_0') \int_0^{d(z_0)/2} L(u)du}{d(z_0') \left(1 - \beta \int_0^{d(z_0)} L_0(u)du\right)} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{d(z_0)} L_0(u)du}{d(z_0) \left(1 - \beta \int_0^{d(z_0)} L_0(u)du\right)}$$

$$\leq \delta d(x_1)^3 + \tfrac{\lambda}{3} d(x_1)(d(x_0) + d(y_0) + d(x_1)) + \tau d(z_0)$$

$$< \delta d(x_0)^3 + \tfrac{\lambda}{3} d(x_0)(d(x_0) + d(y_0)) + \tau d(z_0) < \delta R_0 < R,$$

hold, where $d(z_0') = (d(x_0) + d(y_0) + d(x_1))/2$, so $x_1, y_1 \in U(p, R)$ and we also have that

$$R_1 = \max\{\|x_1 - p\|, \|y_1 - p\|\} \leq \delta R_0,$$

so (2.10) is satisfied. Suppose that $x_k, y_k \in U(p, R)$ and (2.10) hold for $k > 0$. By TGNWTM for $k + 1$ we get in turn that

$$\|x_{k+1} - p\| \quad \leq \quad \frac{\beta d(x_k)^3 \int_0^{d(x_k)} M(u)(d(x_k) - u)^2 du}{8d(x_k)^3 \left(1 - \beta \int_0^{d(x_k)} L_0(u)du\right)}$$

$$+ \frac{\beta d(x_k)d(y_k) \int_0^{d(y_k)/2} L(u)du}{d(y_k) \left(1 - \beta \int_0^{d(y_k)} L_0(u)du\right)} + \frac{\sqrt{2}\alpha\beta^2 d(z_k) \int_0^{d(z_k)} L_0(u)du}{d(z_k) \left(1 - \beta \int_0^{d(z_k)} L_0(u)du\right)}$$

$$\leq \quad \frac{\beta d(x_k)^3 \int_0^{d(x_0)} M(u)(d(x_0) - u)^2 du}{8d(x_0)^3 \left(1 - \beta \int_0^{d(x_0)} L_0(u)du\right)}$$

$$+\frac{\beta d(x_k)d(y_k)\int_0^{d(y_0)/2} L_0(u)du}{d(y_0)\left(1-\beta\int_0^{d(y_0)} L_0(u)du\right)} + \frac{\sqrt{2}\alpha\beta^2 d(z_k)\int_0^{d(z_0)} L_0(u)du}{d(z_0)\left(1-\beta\int_0^{d(z_0)} L_0(u)du\right)}$$

$$\leq \quad \delta d(x_k)^3 + \lambda d(x_k)d(y_k) + \tau d(z_k) < \delta R_k < R$$

and

$$\|y_{k+1}-p\| \quad \leq \quad \frac{\beta d(x_{k+1})^3\int_0^{d(x_{k+1})} M(u)(d(x_{k+1})-u)^2 du}{8d(x_{k+1})^3\left(1-\beta\int_0^{d(x_k)} L_0(u)du\right)}$$

$$+\frac{\beta d(x_{k+1})d(z_k')\int_0^{d(z_k')/2} L(u)du}{d(z_k')\left(1-\beta\int_0^{d(z_k)} L_0(u)du\right)} + \frac{\sqrt{2}\alpha\beta^2 d(z_k)\int_0^{d(z_k)} L(u)du}{d(z_k)\left(1-\beta\int_0^{d(z_k)} L_0(u)du\right)}$$

$$\leq \quad \frac{\beta d(x_{k+1})^3\int_0^{d(x_0)} M(u)(d(x_0)-u)^2 du}{8d(x_0)^3\left(1-\beta\int_0^{d(x_0)} L_0(u)du\right)}$$

$$+\frac{\beta d(x_{k+1})d(z_k')\int_0^{d(z0')} L(u)du}{d(z_0')\left(1-\beta\int_0^{d(z_0)} L_0(u)du\right)} + \frac{\sqrt{2}\alpha\beta^2 d(z_k)\int_0^{d(z_0)} L_0(u)du}{d(z_0)\left(1-\beta\int_0^{d(z_0)} L_0(u)du\right)}$$

$$\leq \quad \delta d(x_{k+1})^3 + \tfrac{\lambda}{3}(d(x_k)+d(y_k)+d(x_{k+1}))d(x_{k+1}) + \tau d(z_k)$$

$$< \quad \delta R_k < R,$$

where $d(z_k') = (d(x_k)+d(y_k)+d(x_{k+1}))/2$. Furthermore, we obtain

$$R_{k+1} = \max\{\|x_{k+1}-p\|,\|y_{k+1}-p\|\} \leq \delta R_k \leq \delta^2 R_{k-1} \leq \ldots \delta^{k+1} R_0,$$

so

$$x_{k+1}, y_{k+1} \in U(p,R), (2.8), (2.10) \text{ hold},$$

$\lim_{k\longrightarrow\infty} x_k = p$ and $\lim_{k\longrightarrow\infty} y_k = p$. $\qquad\qquad\square$

Concerning the uniqueness of the solution p we have:

PROPOSITION 2.6. *Under the conditions* (ℭ) *further suppose that*

$$(2.11) \qquad \frac{\beta}{R^*}\int_0^{R^*} L_0(u)(R^*-u)du + \frac{\alpha\bar{\beta}}{R^*}\int_0^{R^*} L_0(u)du < 1,$$

where $\bar{\beta} = \|[F'^T F'^{-1}]\|$. *Then, limit point* p *is the only solution of problem* (1.1) *in* $\bar{U}(p,R^*)$.

The proof follows from the corresponding one in [5] but we only use the center-Lipschitz condition.

3. SPECIAL CASES AND APPLICATIONS

REMARK 3.1. (a) *Set* $\alpha = \|F(p)\| = 0$ *in* Theorem 2.5 *and* Proposition 2.6 *to obtain the results in the case of zero residual.*
(b) *If* L_0, L, M *are constants, then we can obtain results of special cases.*

(c) *In the literature functions L_1 and M_1 are used instead of L and M, respectively* [3, 5, 8, 9, 12]. *Let us compare ratios and ball of convergence. Notice that in view of* (2.1), (2.4), (2.5) *and* (2.7), *we have*

$$(3.1) \qquad \mu(L_0, L, M)(t) \leq \mu(L_1, L_1, M_1)(t),$$

and

$$(3.2) \qquad \gamma(L_0, L, M)(t) \leq \gamma(L_1, L_1, M_1)(t),$$

so

$$(3.3) \qquad R^*(L_1, L_1, M_1)(t) \leq R^*(L_0, L, M)(t).$$

Therefore, our radius of convergence is larger and our ratio of convergence is smaller. Moreover the information on the location of the solution p is more precise, since only L is used in (2.11) [9]. *Notice that these advantages are obtained under the same computational cost, since in practice the computation of L_1 and M_1 require the computation of the rest of the functions L_0, L and M as special cases.*

REMARK 3.2. *In particular, using the error estimates, it follows that for $\alpha = 0$ we have $\tau = 0$ and*

$$
\begin{aligned}
d(x_{k+1}) &\leq d(x_k)(\delta d(x_k)^2 + \lambda d(y_k)) \\
d(y_{k+1}) &\leq d(x_{k+1})\left[\delta d(x_{k+1})^2 + \tfrac{\lambda}{3}\left(d(x_k) + d(x_{k+1}) + d(y_k)\right)\right] \\
&\leq d(x_{k+1})\left[\left(\delta d(x_k) + \tfrac{2\lambda}{3}\right)d(x_k) + \tfrac{\lambda d(y_k)}{3}\right] \\
&\leq d(x_{k+1})d(x_k)(\delta R^* + \lambda) \\
&= d(x_{k+1})d(x_k)\ell_1.
\end{aligned}
$$

Also, for sufficiently large k,

$$
\begin{aligned}
d(x_{k+1}) &\leq d(x_k)(\delta d(x_k)^2 + \lambda d(y_k)) \\
&\leq d(x_k)(\delta d(x_k)^2 + \lambda \ell_1 d(x_k) d(x_{k-1})) \\
&\leq d(x_k)^2 d(x_{k-1})(\delta + \lambda \ell_1) \\
&= d(x_k)^2 d(x_{k-1})\ell_2,
\end{aligned}
$$

leading to the equation

$$x^2 - 2x - 1 = 0,$$

so the order of iterative method (1.2) *is the positive root of the preceding equation which is $1 + \sqrt{2}$.*

Next, we present an example to show that (3.1)–(3.3) hold as strict inequalities justifying the advantages as claimed at the introduction of this study.

EXAMPLE 3.3. *Let $X = \mathbb{R}^3, D = \bar{U}(0,1), p = (0,0,0)^T$. Define function F on D for $w = (x, y, z)^T$ by*

$$F(w) = \left(e^x - 1, \tfrac{e-1}{2}y^2 + y, z\right)^T.$$

Then, the Fréchet-derivative is given by

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

Notice that using the (2.9) *conditions, we get* $L_0 = e-1, L = M = e^{1/L_0}, L_1 = M_1 = e, \beta = 1, i = j = 3, \alpha = 0.$ *Then*

$$R^*(L_1, L_1, M_1)(t) = 0.1468 < R^*(L_0, L, M)(t) = 0.2263,$$

which justify the improvements as stated in the introduction of this paper.

REFERENCES

[1] S. Amat, M.A. Hernández, N. Romero, *Semilocal convergence of a sixth order iterative method for quadratic equations*, Appl. Numer. Math., **62** (2012), pp. 833–841. `http://doi.org/10.1016/j.apnum.2012.03.001`

[2] I.K. Argyros, S. Hilout, *Extending the applicability of the Gauss-Newton method under average Lipschitz-type conditions*, Numer. Algor., **58** (2011), pp. 23–52. `https://doi.org/10.1007/s11075-011-9446-9`

[3] I.K. Argyros, S. Hilout, *On the semilocal convergence of Werner's method for solving equations using recurrent functions*, Punj. Univ. J. Math., **43** (2011), pp. 19–28.

[4] I.K. Argyros, A. Magreñan, *Iterative methods and their dynamics with applications*, CRC Press, New York, 2017.

[5] J. Chen, W. Li, *Convergence of Gauss-Newton methods and uniqueness of the solution*, Appl. Math. Comput., **170** (2005), pp. 686–705. `http://doi.org/10.1016/j.amc.2004.12.055`

[6] J.M. Dennis, R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, N.Y., 1983.

[7] W.M. Häubler, *A Kantorovich-type convergence analysis for the Gauss-Newton method*, Numer. Math., **48** (1986), pp. 119–125. `http://doi.org/10.1007/BF01389446`

[8] R. Iakymchuk, S.M. Shakhno, *On the convergence analysis of a two-step modification of the Gauss-Newton method*, PAMM, **14** (2014), pp. 813–814. `http://doi.org/10.1002/pamm.201410387`

[9] R.P. Iakymchuk, S.M. Shakhno, H.P. Yarmola, *Convergence analysis of a two-step modification of the Gauss-Newton method and its applications*, J. Numer. Appl. Math., **3(126)** (2017), pp. 61–74.

[10] A.A. Magreñan, *A new tool to study real dynamics: The convergence plane*, Appl. Math. Comput., **248** (2014), pp. 29–38. `http://doi.org/10.1016/j.amc.2014.09.061`

[11] G.N. Silva, *Local convergence of Newton's method for solving generalized equations with monotone operator*, Appl. Anal., **97** (2018), pp. 1094–1105. `http://doi.org/10.1080/00036811.2017.1299860`

[12] X. Wang, *Convergence of Newton's method and uniqueness of the solution of equations in Banach space*, IMA J. Numer. Anal., **20** (2000), pp. 123–134. `http://doi.org/10.1093/imanum/20.1.123`

[13] W. Werner, *Über ein Verfahren der Ordnung* $1 = \sqrt{2}$ *zur Nullstellenbestimmung*, Numer. Math., **32** (1979), pp. 333–342. `http://doi.org/10.1007/BF01397005`