Punjab University
Journal of Mathematics (ISSN 1016-2526)
Vol.41 (2009) pp. 23-33

On the Local Convergence of the Gauss–Newton Method

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

Saïd Hilout
Poitiers university
Laboratoire de Mathématiques et Applications
Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179
86962 Futuroscope Chasseneuil Cedex, France
Email: said.hilout@math.univ–poitiers.fr

Abstract. The local convergence of the Gauss–Newton method is studied under a combination of the radius and center–Lipschitz average functions [3], [7], [8]. Using more precise estimates and under the same or less computational cost, we provide an analysis of this method with the following advantages over the corresponding results in [8]: larger convergence ball, and finer error estimates on the distances involved.

AMS (MOS) Subject Classification Codes: 65F20, 65G99, 65H10, 49M15
Key Words: Local Convergence, Fréchet–derivative, Radius and Center–Lipschitz Condition with Average.

1. Introduction

In this study we are concerned with the problem of approximating a solution $x^\star$ of equation

$$F'(x)^T \ F(x) = 0, \tag{1.1}$$

where F is a Fréchet–differentiable operator defined on $\mathcal{X} = \mathbb{R}^n$, with values on $\mathcal{Y} = \mathbb{R}^m$ $(m \geq n)$.

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

We are seeking least–square solutions of (1.1). That is we solve the minimization problem:

$$\min_{x\in\mathcal{X}} \frac{1}{2} F(x)^T F(x). \tag{1.2}$$

We use the famous Gauss–Newton method

$$x_{k+1} = x_k - (F'(x_k)^T F'(x_k))^{-1} F'(x_k)^T F(x_k) \quad (x_0 \in \mathcal{X}), \quad (k \geq 0) \tag{1.3}$$

to generate a sequence approximating a solution $x^\star$ of (1.2).

There is an extensive literature on the local as well as the semilocal convergence analysis of Newton–type methods under various conditions in the more general setting when $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces [1]–[11].

In particular, in the case of Gauss–Newton method (1.3), Li et al. provided a local convergence analysis in [8] using the concept of the generalized Lipschitz condition with L average (inaugurated by Wang in [10]), which unified the Kantorovich–domain–type [1]–[3], [6] approach with the Smale–point–estimate–type approach [3], [9], [10].

Recently, we have successfully used in [1]–[3] a combination of Lipschitz and center–Lipschitz conditions (instead of only Lipschitz conditions as in to provide a finer local and semilocal convergence analysis for Newton–type methods, when F is an isomorphism. The main idea is derived from the observation that more precise upper bounds on the norms $\| F'(x)^{-1} F'(x^\star) \|$ can be obtained if the needed center–Lipschitz condition is used:

$$\begin{aligned} &\| F'(x^\star)^{-1} (F'(x) - F'(x^\star)) \| \leq \ell_0 \| x - x^\star \|, \\ &\text{for all } x \in U(x^\star, r_0) = \{x \in \mathcal{X} : \| x - x^\star \| \leq r_0\} \subseteq \mathcal{X},\ r_0 > 0,\ \ell_0 > 0 \end{aligned} \tag{1.4}$$

instead of the commonly used Lipschitz condition ([4]–[11]):

$$\| F'(x^\star)^{-1} (F'(x) - F'(y)) \| \leq \ell \| x - y \|, \quad \text{for all } x, y \in U(x^\star, r_0),\ \ell > 0. \tag{1.5}$$

If condition (1.5) holds, then, it follows that there exists $\ell_0 \in [0, \ell]$, such that (1.4) is satisfied, and $\dfrac{\ell}{\ell_0}$ can be arbitrarily large [3].

It turn out that these ideas can be used to study the local convergence of the Gauss–Newton method (1.3). In particular, we provide a local convergence analysis with the following advantages over the work by Li et al. [8]:

(1) Larger convergence ball. Enlarging the convergence ball is very important in computational mathematics because it allows for a wider choice of initial guesses in the case of the local convergence of the Gauss–Newton method (1.3).
(2) Finer estimates on the distances involved, which implies that fewer iterations are needed to achieve a desired error tolerance.
(3) An at least as precise information is provided on the uniqueness of the solution.

The above improvements are also obtained under the same computational cost since the computation of the radius Lipschitz condition with L average (see (2.1)) requires that of the center–Lipschitz with L_0 average (see (2.2)).

2. Preliminaries

We need to introduce the concept of Lipschitz condition inaugurated in [10]:

Definition 1. The operator $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfies the radius Lipschitz condition with L average on $U(x^\star, r_0)$ if

$$\| F(x) - F(x_\tau) \| \leq \int_{\tau\, s(x)}^{s(x)} L(t)\ dt, \quad \text{for all } x \in U(x^\star, r_0),\ 0 \leq \tau \leq 1, \tag{2.1}$$

where, L is a positive non–decreasing function on $[0, r_0]$, $s(x) = \| x - x^\star \|$, and $x_\tau = x^\star + \tau\ (x - x^\star)$.

Definition 2. The operator $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfies the center–Lipschitz condition with L_0 average on $U(x^\star, r_0)$ if

$$\| F(x) - F(x^\star) \| \leq \int_0^{s(x)} L_0(t)\ dt, \quad \text{for all } x \in U(x^\star, r_0), \tag{2.2}$$

where, L_0 is a positive non–decreasing function on $[0, r_0]$.

Note that in [7], [8], [10], the same function L is used in Definitions 1 and 2. However,

$$L_0(t) \leq L(t) \quad t \in [0, r_0) \tag{2.3}$$

holds in general and $\dfrac{L}{L_0}$ can be arbitrarily large [1]–[3].

We provide an example where strict inequality holds in (2.3).

Example1. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$, be equipped with the Euclidean norm, $x^\star = 0$, and define function F on $U(0, 1)$ by

$$F(z) = (e^x - 1, e^y - 1)^T, \quad z = (x, y)^T. \tag{2.4}$$

Then, using (2.1), and (2.2), we obtain:

$$L(t) = \ell = \sqrt{2}\ e, \quad \text{and} \quad L_0(t) = \ell_0 = \sqrt{2}\ (e - 1) \quad \text{for all } t \in [0, 1]. \tag{2.5}$$

It follows from (2.5) that

$$L_0(t) < L(t) \quad \text{for all } t \in [0, 1]. \tag{2.6}$$

As noted in the introduction using our (2.2) instead of (2.1), which was employed in [8] for the computation of the norms

$$\| (F'(x)^T\ F'(x))^{-1}\ F'(x) \|\ \| F'(x) - F'(x^\star) \|, \quad x \in U(x^\star, r_0),$$

leads to the advantages, as stated at the end of the introduction of this paper.

Let $\mathbb{R}^{m\times n}$ be the set of all $m \times n$ matrices, and A^+ be the generalized inverse of $A \in \mathbb{R}^{m\times n}$. Then, when $m \geq n$, and A is of full rank, we have $A^+ = (A^T\ A)^{-1}\ A^T$.

We need the lemmas:

Lemma 3. [7] *Let $A, E \in \mathbb{R}^{m\times n}$. Assume $B = A + E$, and $\| A^+ \| \, \| E \| < 1$. Then, the following hold:*

$$\text{rank } (B) \geq \text{rank } (A).$$

If rank $(A) = n$, $m \geq n$, then we have rank $(B) = n$.

Lemma 4. [7] *Let $A, E \in \mathbb{R}^{m\times n}$. Assume $B = A + E$, and $\| A^+ \| \, \| E \| < 1$. Then, the following hold:*

$$\| B^+ \| \leq \frac{\| A^+ \|}{1 - \| A^+ \| \, \| E \|},$$

provided that rank $(B) =$ rank (A).

Moreover,

$$\| B^+ - A^+ \| \leq \frac{\sqrt{2} \ \| A^+ \|^2 \, \| E \|}{1 - \| A^+ \| \, \| E \|}$$

provided that rank $(A) =$ rank $(B) = \min \{m, n\}$.

Lemma 5. [7], [8] *Let M be a positive non–decreasing function on $[0, r_0]$. Then, for each $a \geq 0$, the functions*

$$f_a(t) = \frac{1}{t^{1+a}} \ \int_0^t u^a \ M(u) \ du$$

and

$$g(t) = \frac{1}{t^2} \ \int_0^t (2t - u) \ M(u) \ du$$

are non–decreasing on $[0, r_0]$.

3. Local convergence analysis of method (1.3)

We shall show the main local convergence result for the Gauss–Newton method (1.3) using a combination of the radius Lipschitz condition with L average, and the center–Lipschitz condition with L_0 average on $U(x^\star, r_0)$:

Theorem 6. *Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be continuously Fréchet–differentiable on $U(x^\star, r_0)$, where $x^\star$ is a solution of (1.2) and $r_0 > 0$. Set*

$$b = \| (F'(x^\star)^T \ F'(x^\star))^{-1} \ F'(x^\star)^T \|, \quad \text{and} \quad c = \| F(x^\star) \| .$$

Moreover, assume operator $F'(x^\star)$ is of full rank and F' satisfies the radius Lipschitz condition with L average and the center–Lipschitz condition with L_0 average on $U(x^\star, r_0)$.

Furthemore, assume function h_0 has a minimal zero r on $[0, r_0]$, which also satisfies:

$$b \ \int_0^r L_0(t) \ dt < 1, \tag{3.1}$$

where,

$$h_0(p) = \left(\int_0^p L(t) \ t \ dt + (\sqrt{2} \ b \ c + p) \ \int_0^p L_0(t) \ dt \right) b - p. \tag{3.2}$$

Then, sequence $\{x_k\}$ $(k \geq 0)$ generated by the Gauss–Newton method (1.3) is well defined, remains in $U(x^\star, r)$ for all $k \geq 0$, and converges to $x^\star$, provided that $x_0 \in U(x^\star, r)$ with $x_0 \neq x^\star$.

Moreover, the following estimates hold for all $k \geq 1$*:*

$$\begin{aligned} \| x_k - x^\star \| &\leq \alpha \ \| x_{k-1} - x^\star \|^2 + \beta \ \| x_{k-1} - x^\star \| \\ &\leq q^k \ \| x_0 - x^\star \|, \end{aligned} \tag{3.3}$$

where,

$$\alpha = \frac{b \displaystyle\int_0^{s(x_0)} L(t) \ t \ dt}{\left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t) \ dt\right) s(x_0)^2}, \tag{3.4}$$

$$\beta = \frac{\sqrt{2} \ b^2 \ c \displaystyle\int_0^{s(x_0)} L(t) \ dt}{\left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t) \ dt\right) s(x_0)}, \tag{3.5}$$

and

$$0 < q_0 = \alpha \ s(x_0) + \beta < 1. \tag{3.6}$$

Proof. Using (2.2), (3.1), and the choice of r, we obtain in turn:

$$\begin{aligned} \| (F'(x^\star)^T \ F'(x^\star))^{-1} \ F'(x^\star)^T \| \ \| F'(x) - F'(x^\star) \| &\leq b \int_0^{s(x_0)} L_0(t) \ dt \\ &< b \int_0^r L_0(t) \ dt < 1. \end{aligned} \tag{3.7}$$

In view of Lemmas 3, 4 respectively, and (3.7), $F'(x)$ is full rank and satisfies

$$\| (F'(x)^T \ F'(x))^{-1} \ F'(x)^T \| \leq \frac{b}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t) \ dt}, \tag{3.8}$$

and

$$\begin{aligned} &\| (F'(x)^T \ F'(x))^{-1} \ F'(x)^T - (F'(x^\star)^T \ F'(x^\star))^{-1} \ F'(x^\star)^T \| \\ &\leq \frac{\sqrt{2} \ b \displaystyle\int_0^{s(x_0)} L_0(t) \ dt}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t) \ dt} \end{aligned} \tag{3.9}$$

for all $x \in U(x^\star, r)$.

Moreover, in view of Lemma 5 for $a = 0$ and $a = 1$, we get in turn

$$\begin{aligned} \frac{\displaystyle\int_0^{s(x_0)} M(t) \ t \ dt}{s(x_0)} &= s(x_0) \ \frac{\displaystyle\int_0^{s(x_0)} M(t) \ t \ dt}{s(x_0)^2} \\ &\leq r \ \frac{\displaystyle\int_0^r M(t) \ t \ dt}{r^2} = \frac{\displaystyle\int_0^r M(t) \ t \ dt}{r}, \end{aligned} \tag{3.10}$$

and

$$\frac{\displaystyle\int_0^{s(x_0)} M(t)\ dt}{s(x_0)} \leq \frac{\displaystyle\int_0^{r} M(t)\ dt}{r}. \tag{3.11}$$

Using (3.1), (3.6), (3.10), (3.11), and the choice of r, for $M = L_0$ or L, we obtain:

$$\begin{aligned} 0 < q &= \frac{1}{s(x_0)} \left(\frac{b \displaystyle\int_0^{s(x_0)} L(t)\ t\ dt}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} + \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{s(x_0)} L(t)\ dt}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} \right) \\ &\leq \frac{1}{r} \left(\frac{b \displaystyle\int_0^{r} L(t)\ t\ dt}{1 - b \displaystyle\int_0^{r} L_0(t)\ dt} + \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{r} L(t)\ dt}{1 - b \displaystyle\int_0^{r} L_0(t)\ dt} \right) < 1. \end{aligned} \tag{3.12}$$

Using Gauss–Newton method (1.3), we obtain the identity

$$\begin{aligned} x_k - x^\star &= x_{k-1} - x^\star - (F'(x_{k-1})^T\ F'(x_{k-1}))^{-1}\ F'(x_{k-1})^T\ F(x_{k-1}) \\ &= (F'(x_{k-1})^T\ F'(x_{k-1}))^{-1}\ F'(x_{k-1})^T\ (F'(x_{k-1})\ (x_{k-1} - x^\star) - \\ &\quad F(x_{k-1}) + F(x^\star)) + (F'(x^\star)^T\ F'(x^\star))^{-1}\ F'(x^\star)^T\ F(x^\star) - \\ &\quad (F'(x_{k-1})^T\ F'(x_{k-1}))^{-1}\ F'(x_{k-1})^T\ F(x^\star). \end{aligned} \tag{3.13}$$

In particular for $k = 1$ in (3.12), since $x_0 \in U(x^\star, r)$, we obtain in turn using (2.1), (2.2), (3.8), and (3.9):

$$\begin{aligned} &\| x_1 - x^\star \| \\ &\leq \| (F'(x_0)^T\ F'(x_0))^{-1}\ F'(x_0)^T \| \times \\ &\| \int_0^1 (F'(x_0) - F'(x_0 + \tau\ (x^\star - x_0))\ (x_0 - x^\star)\ d\tau \| + \\ &\| (F'(x^\star)^T\ F'(x^\star))^{-1}\ F'(x^\star)^T - (F'(x_0)^T\ F'(x_0))^{-1}\ F'(x_0)^T \|\ \| F(x^\star) \| \\ &\leq \frac{b}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} \int_0^1 \int_{\tau\ s(x_0)}^{s(x_0)} L(t)\ dt\ s(x_0)\ d\tau + \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{s(x_0)} L_0(t)\ dt}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} \\ &= \frac{b}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} \left(\int_0^{s(x_0)} L(t)\ t\ dt + \sqrt{2}\ b\ c \int_0^{s(x_0)} L_0(t)\ dt \right). \end{aligned} \tag{3.14}$$

It follows from (3.6) and (3.14) that:

$$\| x_1 - x^\star \| \leq q_0\ \ \| x_0 - x^\star \|, \tag{3.15}$$

which implies $x_1 \in U(x^\star, r)$.

Hence, (3.3) holds for $k = 0$.

Assume $x_k \in U(x^\star, r)$, by exchanging x_0, x_1 with x_k, x_{k+1}, we obtain:

$$\begin{aligned}
&\| x_{k+1} - x^\star \| \\
&\leq \frac{b \displaystyle\int_0^{s(x_k)} L(t)\ t\ dt}{1 - b \displaystyle\int_0^{s(x_k)} L_0(t)\ dt} + \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{s(x_k)} L_0(t)\ dt}{1 - b \displaystyle\int_0^{s(x_k)} L_0(t)\ dt} \\
&\leq \frac{b \displaystyle\int_0^{s(x_k)} L(t)\ t\ dt\ s(x_k)^2}{s(x_k)^2 \left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt\right)} + \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{s(x_k)} L_0(t)\ dt\ s(x_k)}{s(x_k) \left(1 - b \displaystyle\int_0^{s(x_k)} L_0(t)\ dt\right)} \\
&\leq q_0\ \| x_k - x^\star \| \leq q_0^{k+1}\ \| x_0 - x^\star \| < r
\end{aligned} \tag{3.16}$$

(by Lemma 5, which implies $x_{k+1} \in U(x^\star, r)$, and $\lim_{k \to \infty} x_k = x^\star$.

That completes the induction and the proof of Theorem 6. □

Remark 7. If estimate (2.3) holds as equality, then our Theorem 7 reduces to Theorem 3.1 in [8]. Otherwise, i.e., if (2.3) holds a strict inequality, then our result improves Theorem 3.1 in [8] under the same computational cost, since in practice the evaluation of function L requires that of L_0. Let h, q, r_1 be as h_0, q_0, r respectively by simply replacing L_0 by L.

Then, we have:

$$r_1 < r \tag{3.17}$$

and

$$q_0 < q. \tag{3.18}$$

Since, r_1 and q were used in [8], it follows from (3.17) and (3.18) that in this case a larger convergence ball is obtained and smaller ratio than in [8].

We state the local convergence result for the Gauss–Newton method (1.3) using only the weaker center–Lipschitz condition with L_0 average on $U(x^\star, r_0)$.

Theorem 8. *Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be continuously Fréchet–differentiable on $U(x^\star, r_0)$, where $x^\star$ is a solution of (1.2) and $r_0 > 0$.*

Moreover, assume operator $F'(x^\star)$ is of full rank and F' satisfies the center–Lipschitz condition with L_0 average on $U(x^\star, r_0)$.

Furthemore, assume function H_0 has a minimal zero r on $[0, r_0]$, which also satisfies:

$$b \int_0^r L_0(t)\ dt < 1,$$

where, b is defined in Theorem 6, and H_0 has the following form:

$$H_0(p) = \left(\int_0^p (2\ p - t)\ L_0(t)\ dt + (\sqrt{2}\ b\ c + p) \int_0^p L_0(t)\ dt \right) b - p. \tag{3.19}$$

Then, sequence $\{x_k\}$ $(k \geq 0)$ generated by the Gauss–Newton method (1.3) is well defined, remains in $U(x^\star, r)$ for all $k \geq 0$, and converges to $x^\star$, provided that $x_0 \in U(x^\star, r)$

with $x_0 \neq x^\star$.

Moreover, the following estimates hold for all $k \geq 1$*:*

$$\begin{array}{rcl} \| x_k - x^\star \| & \leq & \alpha_0 \ \| x_{k-1} - x^\star \|^2 + \beta_0 \ \| x_{k-1} - x^\star \| \\ & \leq & \overline{q}^k \ \| x_0 - x^\star \|, \end{array} \tag{3.20}$$

where,

$$\alpha_0 = \frac{b \displaystyle\int_0^{s(x_0)} (2\ s(x_0) - t)\ L_0(t)\ dt}{\left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt\right)\ s(x_0)^2},$$

$$\beta_0 = \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{s(x_0)} L_0(t)\ dt}{\left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt\right)\ s(x_0)},$$

and

$$0 < \overline{q} = \alpha_0\ s(x_0) + \beta_0 < 1.$$

Proof. We shall show (3.20) for all $k \geq 1$.

As in the proof of Theorem 6, using (1.3) for $k = 1$, (2.1), and (2.2), we get:

$$\begin{aligned}
&\| x_1 - x^\star \| \\
&\leq \| (F'(x_0)^T\ F'(x_0))^{-1}\ F'(x_0)^T \| \times \\
&\| \int_0^1 (F'(x_0) - F'(x_0 + \tau\ (x^\star - x_0))\ (x_0 - x^\star)\ d\tau \| + \\
&\| (F'(x^\star)^T\ F'(x^\star))^{-1}\ F'(x^\star)^T - (F'(x_0)^T\ F'(x_0))^{-1}\ F'(x_0)^T \| \ \| F(x^\star) \| \\
&\leq b \int_0^1 \| (F'(x_0) - F'(x^\star)) + (F'(x^\star) - F'(x_0 + \tau\ (x^\star - x_0)) \| \ \| x_0 - x^\star \| \ d\tau + \\
&\| (F'(x^\star)^T\ F'(x^\star))^{-1}\ F'(x^\star)^T - (F'(x_0)^T\ F'(x_0))^{-1}\ F'(x_0)^T \| \ \| F(x^\star) \| \\
&\leq \frac{b}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} \int_0^1 \left(\int_0^{s(x_0)} L_0(t)\ dt + \int_0^{\tau\ s(x_0)} L_0(t)\ dt \right) s(x_0)\ d\tau \\
&+ \frac{\sqrt{2}\ b^2\ c \displaystyle\int_0^{s(x_0)} L_0(t)\ dt}{1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt} \\
&= \frac{b}{s(x_0)^2 \left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt\right)} \left(\int_0^{s(x_0)} (2\ s(x_0) - t)\, L_0(t)\, dt\ \| x_0 - x^\star \|^2 + \right. \\
&\left. \sqrt{2}\, b\ c\, s(x_0) \int_0^{s(x_0)} L_0(t)\, dt\ \| x_0 - x^\star \| \right) \\
&= \overline{q}\ \| x_0 - x^\star \| .
\end{aligned}$$

Therefore, $x_1 \in U(x^\star, r)$. Clearly, (3.20) holds for $k=0$. Assume that $x_k \in U(x^\star, r)$, (3.20) holds for $k>0$, and $s(x_k)$ decreases monotonically. Then, using Lemmas 3, and 5, we have in turn:

$$\begin{aligned}
&\| x_{k+1} - x^\star \| \\
&\leq \frac{b}{s(x_k)^2 \left(1 - b \displaystyle\int_0^{s(x_k)} L_0(t)\ dt\right)} \left(\int_0^{s(x_k)} (2\ s(x_k) - t)\, L_0(t)\, dt\ \| x_k - x^\star \|^2 + \right. \\
&\left. \sqrt{2}\, b\ c\, s(x_k) \int_0^{s(x_k)} L_0(t)\, dt\ \| x_k - x^\star \| \right) \\
&\leq \frac{b}{s(x_0)^2 \left(1 - b \displaystyle\int_0^{s(x_0)} L_0(t)\ dt\right)} \left(\int_0^{s(x_0)} (2\ s(x_0) - t)\, L_0(t)\, dt\ \| x_k - x^\star \|^2 + \right. \\
&\left. \sqrt{2}\, b\ c\, s(x_0) \int_0^{s(x_0)} L_0(t)\, dt\ \| x_k - x^\star \| \right) \\
&\leq \overline{q}\ \| x_k - x^\star \| \leq \overline{q}^{k+1}\ \| x_0 - x^\star \| .
\end{aligned}$$

That completes the induction, and the proof of Theorem 8. □

Remark 9. If estimate (2.3) holds as equality, then our Theorem 8 reduces to Theorem 3.2 in [8]. Otherwise, i.e., if (2.3) holds a strict inequality, then our result improves Theorem 3.2 in [8] under less computational cost, since the evaluation of L is more expensive than the evaluation of L_0. Let H, $\overline{\overline{q}}$, r_2 be as H_0, $\overline{q}$, r respectively by simply replacing L_0 by L.

Then, we have:

$$r_2 < r$$

and

$$\overline{q} < \overline{\overline{q}}.$$

4. Applications

In this section we apply the results of Section 3 in a concrete case. Assume: $L_0(t) = L_0$, and $L(t) = L$ on $[0, \infty)$. That is consider the Kantorovich–type case. Then in the case of Theorem 6, and also using the notation introduced in Remark 7, we can easily obtain:

$$r = \frac{2\ (1 - \sqrt{2}\ b^2\ c\ L_0)}{(2\ L_0 + L)\ b},$$

$$q_0 = \frac{(L\ s(x_0) + 2\ \sqrt{2}\ b\ c\ L_0)\ b}{2\ (1 - b\ L_0\ s(x_0))},$$

$$r_1 = \frac{2\ (1 - \sqrt{2}\ b^2\ c\ L)}{3\ b\ L} \leq r \tag{4.1}$$

and

$$q = \frac{(s(x_0) + 2\ \sqrt{2}\ b\ c)\ b\ L}{2\ (1 - b\ L\ s(x_0))} \geq q_0. \tag{4.2}$$

Note that strict inequality holds in (4.1), and (4.2), if $L_0 < L$.

We provide a numerical example using the above values.

Example2. Return back to Example 1. Then, we have $c = 0$, $b = \sqrt{2}$, and for $x_0 = .2$,

$$r_1 = .122626481, \qquad q = .595665552$$

$$r = .162473616 \quad \text{and} \quad q_0 = .414155282.$$

That is, we conclude that estimates (4.1), and (4.2) hold as strict inequalities.

Below, we provide a comparison between error bounds obtained in Theorem 3.1 [8], and Theorem 6 in this study.

Comparison table

k	$q^{k+1} \parallel x_0 - x^\star \parallel$	$q_0^{k+1} \parallel x_0 - x^\star \parallel$
0	.119130759	.082831056
1	.070960689	.03430492
2	.042268004	.014207564
3	.025177097	.005884138
$\vdots$	$\vdots$	$\vdots$
20	.000006323	.000000004
22	.000002243	0

Finally, we provide an example where (2.2) holds, but (2.1) is violated.

Example3. Let $\mathcal{X} = \mathbb{R}$, and $\mathcal{Y} = \mathbb{R}^2$. Define

$$g_1(x) = \int_0^x \left(1 + \frac{x}{4} \sin\frac{\pi}{x}\right) dx, \quad g_2(x) = \frac{1}{8}\, x^2 \quad \text{for all} \quad x \in \mathcal{X}$$

and

$$g = (g_1, g_2)^T.$$

Then, we get

$$g_1'(x) = \begin{cases} 1 + \dfrac{x}{4} \sin\dfrac{\pi}{x}, & if \quad x \neq 0 \\ 1, & if \quad x = 0 \end{cases}$$

and

$$g_2'(x) = \frac{x}{4} \quad \text{for all} \quad x \in \mathcal{X}.$$

Clearly, $x^\star = 0$ is a solution of (1.2) (with F is replaced by g).

We also have

$$\parallel g'(x) - g'(x^\star) \parallel = \frac{1}{2} \parallel x - x^\star \parallel \quad \text{for all} \quad x \in \mathcal{X}.$$

That is, we can set

$$L_0(t) = \frac{1}{2}.$$

Then, since $b = \sqrt{2}$, and $c = 0$, Theorem 8 for any $x_0 \in U(x^\star, \dfrac{2\sqrt{5}}{5})$, guarantees that Gauss–Newton method converges to $x^\star$.

However, there is no positive integrable function L such that (2.1) holds. Indeed, we have

$$\parallel g_1'(x) - g_1'(x_\tau) \parallel = \frac{1}{4} \; |x \; \sin\frac{\pi}{x} - \tau \; x \; \sin\frac{\pi}{x}| = \frac{1}{4 \; (2 \; k+1)}$$

for $x = \dfrac{1}{k}, \tau = \dfrac{2}{2 \; k+1}, k = 1, 2, \cdots$

That is, if there was a positive integrable function L, such that (2.1) holds on $U(x^\star, r)$ for some $r > 0$, then, there exists $k_0 > 1$, such that

$$\begin{aligned} \int_0^r L(t) \; dt &\geq \sum_{k=k_0}^{\infty} \int_{2/(2k+1)}^{1/k} L(t) \; dt \\ &\geq \parallel g'(x) - g'(x_\tau) \parallel \\ &\geq \parallel g_1'(x) - g_1'(x_\tau) \parallel \geq \sum_{k=k_0}^{\infty} \frac{1}{4 \; (2 \; k+1)} = \infty, \end{aligned}$$

which is a contradiction.

References

[1] I.K. Argyros, *A convergence analysis of Newton–like methods for singular equations using outer or generalized inverses*, Applicationes Mathematicae, **32** (2005), 37–49.

[2] I.K. Argyros, *A unifying local–semilocal convergence analysis and applications for two–point Newton–like methods in Banach space*, J. Math. Anal. and Appl., **298** (2004), 374–397.

[3] I.K. Argyros, *Convergence and applications of Newton–type iterations*, Springer Verlag Publ., New York, 2008.

[4] J.M. Gutierrez, *A new semilocal convergence theorem for Newton's method*, **79** (1997), 131–145.

[5] Z. Huang, *The convergence ball of Newton's method and the uniqueness ball of equations under Hölder continuous derivatives*, Comput. Appl. Math., **47** (2004), 247–251.

[6] L.V. Kantorovich, G.P. Akilov, *Functional analysis in normed spaces*, Pergamon Press, New York, 1982.

[7] C. Li, X. Wang, *Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces II*, Acta Math. Sinica, New Series, **18** (2003), 405–412.

[8] C. Li, W. Zhang, X. Jin, *Convergence and uniqueness properties of Gauss–Newton's method*, Comput. Math. Appl., **47** (2004), 1057–1067.

[9] S. Smale, *Newton's method estimates from data at a point*, The merging of disciplines: New directions in Pure, Applied and Computational Mathematics, R. Ewing, K. Gross and C. Martin eds, New–York, Springer Verlag Publ., (1986), 185–196.

[10] X. Wang, *Convergence of Newton's method and uniqueness of the solution of equations in Banach space*, IMA J. Numer. Anal., **20** (2000), 123–134.

[11] Y. Yuan, W. Sun, *Optimization theory and methods. Nonlinear Programming*, Springer Optimization and Its Applications, Springer, New York, 2006.