

LOCAL CONVERGENCE OF THE GAUSS-NEWTON METHOD FOR INJECTIVE-OVERDETERMINED SYSTEMS

SERGIO AMAT, IOANNIS KONSTANTINOS ARGYROS,
AND ÁNGEL ALBERTO MAGREÑÁN

ABSTRACT. We present, under a weak majorant condition, a local convergence analysis for the Gauss-Newton method for injective-overdetermined systems of equations in a Hilbert space setting. Our results provide under the same information a larger radius of convergence and tighter error estimates on the distances involved than in earlier studies such as [10, 11, 13, 14, 18]. Special cases and numerical examples are also included in this study.

1. Introduction

Let \mathcal{X} and \mathcal{Y} be Hilbert spaces. Let $\mathcal{D} \subseteq \mathcal{X}$ be an open set and $F : \mathcal{D} \rightarrow \mathcal{Y}$ be a continuously Fréchet-differentiable operator. In this study we are concerned with the problem of approximating a locally unique solution x^* of the penalized nonlinear least squares problem

$$(1.1) \quad \min_{x \in \mathcal{D}} \|F(x)\|^2.$$

A solution $x^* \in \mathcal{D}$ of (1.1) is also called a least squares solution of the equation $F(x) = 0$.

Many problems from computational sciences and other disciplines can be brought in a form similar to equation (1.1) using Mathematical Modeling [3, 9, 17]. For example in data fitting, we have $\mathcal{X} = \mathbb{R}^i$, $\mathcal{Y} = \mathbb{R}^j$, i is the number of parameters and j is the number of observations.

The solution of (1.1) can rarely be found in closed form. That is why the solution methods for these equations are usually iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to Newton-type methods [3, 9, 17, 19]. The study about convergence of iterative method is usually divided on two types: semilocal and local convergence analysis. The local convergence matter is based on the information around a

Received October 2, 2013.

2010 *Mathematics Subject Classification.* 65G99, 65K10, 47H17, 49M15, 90C30.

Key words and phrases. the Gauss-Newton method, Hilbert spaces, majorant condition, local convergence, radius of convergence, injective-overdetermined systems.

solution to find estimates of the radii of convergence balls, while the semilocal one is based on the information around the initial point, to give criteria ensuring the convergence of iterative procedures.

A plethora of sufficient conditions for the local as well as the semilocal convergence of Newton-type methods as well as an error analysis for such methods can be found in [1]–[22].

We study the well known Gauss-Newton method defined by

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

where $x_0 \in \mathcal{D}$ is an initial point and $F'(x_n)^+$ is the Moore-Penrose inverse of the linear operator $F'(x_n)$ [7, 9, 12, 14, 16, 18]. In the present paper we use the proximal Gauss-Newton method (to be precised in Section 2, see (2.6)) for solving penalized nonlinear least squares problem (1.1). Notice that if x^* is a solution of (1.1), $F(x^*) = 0$ and $F'(x^*)$ is invertible, then the theories of Gauss-Newton methods merge into those of Newton method.

A survey of convergence results under various Lipschitz-type conditions for Gauss-Newton-type methods can be found in [3, 9] (see also [7, 12, 14, 18]). The convergence of these methods requires among other hypotheses that F' satisfies a Lipschitz condition or F'' is bounded in \mathcal{D} . Several authors have relaxed these hypotheses. In particular, Ferreira et al. [10, 11, 12, 13] used the majorant condition in the local as well as semilocal convergence of Newton-type method. Argyros and Hilout [3, 4, 5, 6, 7, 9] have also used the majorant condition to provide a tighter convergence analysis and weaker convergence criteria for Newton-type method. The local convergence of the Gauss-Newton method was examined by Ferreira et al. [12] using the majorant condition. It was shown that this condition is better than Wang's condition [18], [22] in some sense. A certain relationship between the majorant function and operator F was established that unifies two previously unrelated results pertaining to inexact Gauss-Newton methods, which are the result for analytical functions and the one for operators with Lipschitz derivative.

In this study we are motivated by a recent paper of Gonçalves [14], which weakened earlier convergence conditions [10, 11, 12, 13, 18, 22] for the local convergence analysis of the Gauss-Newton method (1.2) under general majorant condition (see (2.2)). The information used is $I(x^*, F, f)$, where f is a majorant function (to be precised later in (2.2)). Using $I(x^*, F, f)$, Gonçalves [14] provided error estimates on the distances $\|x_n - x^*\|$ ($n \geq 1$) as well as what he claimed to be the best possible convergence radius.

In our analysis we are also motivated by optimization considerations and the work in [14]. Using the same information $I(x^*, F, f)$, we show that in general the radius of convergence given in [14] is not as the best possible but it can be enlarged. We also show that the upper bounds on the distances $\|x_n - x^*\|$ ($n \geq 1$) can be tighter. These observations are very important in computational mathematics, since they allow a wider choice of initial guesses x_0 and fewer iterations to obtain a desired error tolerance $\epsilon > 0$. Note that

similar improvements in both the local and semilocal case of the works in [10, 11, 13, 12, 22], have already been obtained by us in [3, 4, 5, 6, 7, 9] under stronger than (2.2) majorant-type conditions.

The paper is organized as follows: Section 2 contains the local convergence analysis of the Gauss-Newton method (1.2) under weak majorant conditions, whereas in Section 3 we provide special cases and numerical examples further validating the theoretical results.

2. Local convergence analysis

Let \mathcal{X} and \mathcal{Y} be Hilbert spaces. We denote by $U(z, \alpha)$ the open ball centered at $z \in \mathcal{X}$ and of radius $\alpha > 0$, whereas $\overline{U}(z, \alpha)$ denotes its closure.

We state the main local convergence result for the Gauss-Newton method (1.2) under the majorant condition.

Theorem 2.1. *Let \mathcal{X} and \mathcal{Y} be Hilbert spaces; let $\mathcal{D} \subset \mathcal{X}$ be an open set; and let $F : \mathcal{D} \subset \mathcal{X} \rightarrow \mathcal{Y}$ be a continuously Fréchet-differentiable operator. Let $x^* \in \mathcal{D}$, $\beta := \|F'(x^*)^+\|$, $R > 0$ and $\kappa := \sup\{t \in [0, R) : U(x^*, t) \subset \mathcal{D}\}$. Suppose that $F(x^*) = 0$, $F'(x^*)$ is injective and there exist $f_0, f : [0, R) \rightarrow (-\infty, +\infty)$ continuously differentiable such that*

$$(2.1) \quad \beta \| F'(x^*)^{-1} (F'(x) - F'(x^*)) \| \leq f'_0(\| x - x^* \|) - f'_0(0),$$

$$(2.2) \quad \beta \| F'(x^*)^{-1} (F'(x) - F'(x_\theta)) \| \leq f'(\| x - x^* \|) - f'(\theta \| x - x^* \|)$$

for all $x \in U(x^*, \kappa)$ and $x_\theta = x^* + \theta(x - x^*)$, $\theta \in [0, 1]$,

$$(\mathcal{H}_1) \quad f_0(0) = f(0) = 0 \text{ and } f'_0(0) = f'(0) = -1;$$

$$(\mathcal{H}_2) \quad f'_0, f' \text{ are strictly increasing,}$$

$$(2.3) \quad f_0(t) \leq f(t) \quad \text{and} \quad f'_0(t) \leq f'(t) \quad t \in [0, R).$$

Define parameter ν_0 , function f_1 on $(0, \nu_0)$, parameters ν, ρ_0, r_0 and scalar iteration $\{s_n\}$ by

$$\nu_0 := \sup\{t \in [0, R) : f'_0(t) < 0\},$$

$$(2.4) \quad f_1(t) := \frac{f'(t)}{f'_0(t)},$$

$$\nu := \sup\{t \in [0, R) : f'(t) < 0\},$$

$$\rho_0 := \sup\{\delta \in [0, \nu) : \left(\frac{f(t)}{f'(t)} - t\right) \frac{f_1(t)}{t} < 1, t \in [0, \delta)\},$$

$$r_0 := \min\{\kappa, \rho_0\},$$

$$(2.5) \quad s_0 = \| x_0 - x^* \|, \quad s_{n+1} = \left| \left(s_n - \frac{f(s_n)}{f'(s_n)} \right) f_1(s_n) \right| \quad (n \geq 0).$$

Then, the following assertions hold:

- (a) $\{s_n\}$ is well defined; strictly decreasing; contained in $(0, r_0)$; converges to zero and

$$(2.6) \quad \lim_{n \rightarrow 0} \frac{s_{n+1}}{s_n} = 0.$$

- (b) $\{x_n\}$ generated by the Gauss-Newton method (1.2), starting from $x_0 \in U(x^*, r_0) \setminus \{x^*\}$ is well defined; remains in $U(x^*, r_0)$ for all $n \geq 0$ and converges to x^* , which is the unique solution of equation (1.1) in $U(x^*, \sigma_0)$, where,

$$\sigma_0 := \sup\{t \in [0, \kappa) : f_0(t) < 0\}$$

and

$$(2.7) \quad \lim_{n \rightarrow \infty} \frac{\|x_{n+1} - x_n\|}{\|x_n - x^*\|} = 0.$$

- (c) If

$$\left(\frac{f(\rho_0)}{\rho_0 f'(\rho_0)} - 1\right) f_1(\rho_0) = 1 \quad \text{and} \quad \rho_0 < \kappa,$$

then $r_0 = \rho_0$ is the possible convergence radius.

- (d) If scalar sequence $\{t_n\}$ is given by

$$(2.8) \quad t_0 = \|x_0 - x^*\|, \quad t_{n+1} = \left|t_n - \frac{f(t_n)}{f'(t_n)}\right| \quad (n \geq 0),$$

then

$$(2.9) \quad s_n \leq t_n \quad (n \geq 0)$$

and strict inequality holds for $n > 1$ in (2.9), if $f'_0(t) < f'(t)$, $t \in [0, R)$.

If additionally, given $0 \leq p \leq 1$

$$(H_3) \quad \text{The function } t \rightarrow \left(\frac{f(t)}{f'(t)} - t\right) \frac{f_1(t)}{t^{p+1}} \text{ is strictly increasing on } (0, \nu_0),$$

then,

- (e) The sequence $\left\{\frac{s_{n+1}}{s_n^{p+1}}\right\}$ is strictly decreasing so that

$$(2.10) \quad \|x_{n+1} - x^*\| \leq \frac{s_{n+1}}{s_n^{p+1}} \|x_n - x^*\|^{p+1} \quad (n \geq 0).$$

Furthermore, for $n \geq 0$,

$$(2.11) \quad \|x_n - x^*\| \leq \begin{cases} s_0 \left[\frac{s_1}{s_0}\right]^n & \text{if } p = 0 \\ s_0 \left(\frac{s_1}{s_0}\right)^{((p+1)^n - 1)/p} & \text{if } p \neq 0. \end{cases}$$

Proof of Theorem 2.1. We shall break down the proof into 10 pieces called lemmas.

First we shall show the statements of the theorem involving sequence $\{s_n\}$.

Lemma 2.2. *The constants κ, ν, σ_0 are positive and $(t - \frac{f(t)}{f'(t)}) f_1(t) < 0$ for all $t \in (0, \nu)$.*

Proof. The set \mathcal{D} is open and $x^* \in \mathcal{D}$, so we deduce that κ is positive. Since f' is continuous in 0 with $f'(0) = -1$, there exists $\delta > 0$ such that $f'(t) < 0$ for all $t \in (0, \delta)$. Thus $\nu > 0$. Now, because $f(0) = 0$ and $f'(0) = -1$, there exists $\delta > 0$ such that $f(t) < 0$ for all $t \in (0, \delta)$. Hence, we have $\sigma = \sup\{t \in [0, \kappa] : f(t) < 0\} > 0$ and by (\mathcal{H}_2) : $\sigma_0 \geq \sigma > 0, f_0(t) < 0, t \in (0, \sigma_0)$.

It also follows from (\mathcal{H}_1) and (\mathcal{H}_2) that $0 = f(0) > f(t) - t f'(t)$ for $t \in (0, R)$. If $t \in (0, \nu)$, then $f'(t) < 0$, which together with (2.4) complete the proof of Lemma 2.2. □

According to (\mathcal{H}_2) , the definition of ν_0 and ν , we have that $f'_0(t) < 0$ and $f'(t) < 0$ for all $t \in [0, \nu)$, since $\nu \leq \nu_0$. Moreover, the function f_1 is well defined on $(0, \nu_0)$. Therefore, the Newton-Gauss iteration function

$$(2.12) \quad \begin{aligned} \eta_{f,f_0} : [0, \nu) &\longrightarrow (-\infty, 0] \\ t &\longrightarrow \left(t - \frac{f(t)}{f'(t)}\right) f_1(t) \end{aligned}$$

is well defined.

Lemma 2.3. *The following assertions hold:*

$$(2.13) \quad \lim_{t \rightarrow 0} \frac{\eta_{f,f_0}(t)}{t} = 0,$$

$$(2.14) \quad \rho_0 > 0$$

and

$$(2.15) \quad |\eta_{f,f_0}(t)| < t \quad \text{for all } t \in (0, \rho_0).$$

Proof. Using definition (2.12), Lemma 2.2, (\mathcal{H}_1) and the definition of ν , a simple algebraic manipulation gives

$$(2.16) \quad \begin{aligned} \frac{|\eta_{f,f_0}(t)|}{t} &= \left(\frac{f(t)}{f'(t)} - t\right) \frac{f_1(t)}{t} \\ &= \left(\frac{1}{f'(t)} \frac{f(t) - f(0)}{t - 0} - 1\right) f_1(t) \quad \text{for all } t \in (0, \nu), \end{aligned}$$

which leads to (2.13) if we let $t \rightarrow 0$ in (2.16). It then follows from (2.13) and the first equality in (2.16) that there exists $\delta > 0$ such that

$$(2.17) \quad 0 < \left(\frac{f(t)}{f'(t)} - t\right) \frac{f_1(t)}{t} < 1 \quad \text{for all } t \in (0, \delta).$$

Hence, we deduce that $\rho_0 > 0$. Finally, the first equality in (2.16) together with the definition of ρ_0 imply (2.15).

That completes the proof of Lemma 2.3. □

In view of (2.12), sequence $\{s_n\}$ can be defined as:

$$(2.18) \quad s_0 = \|x_0 - x^*\|, \quad s_{n+1} = |\eta_{f,f_0}(s_n)| \quad (n \geq 0).$$

Replace η_f by η_{f,f_0} in the proof of [11] Corollary 5 to obtain:

Lemma 2.4. *Sequence $\{s_n\}$ is well defined, strictly decreasing and contained in $(0, \rho_0)$. Moreover, $\{s_n\}$ converges to zero with superlinear rate, i.e.,*

$$\lim_{n \rightarrow \infty} \frac{s_{n+1}}{s_n} = 0.$$

Furthermore, if (\mathcal{H}_3) holds, then sequence $\left\{ \frac{s_{n+1}}{s_n^{p+1}} \right\}$ is strictly decreasing.

Secondly, we need relationships between the majorant function f and non-linear operator F .

It is convenient for us to state some properties of the Moore-Penrose inverse. More properties can be found in [3, 9, 17]. Let $M : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous, linear and injective with closed image. The Moore-Penrose inverse $M^+ : \mathcal{Y} \rightarrow \mathcal{X}$ of M is defined by

$$A^+ := (A^*A)^{-1}A^*,$$

where A^* denotes the adjoint of the linear operator M . We also need the following Banach-type perturbation Lemma.

Lemma 2.5 ([3, 9, 17]). *Let $M_1, M_2 : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous, linear operators with closed images. Suppose that M_1 is injective and that $\|M_1^+\| \|M_1 - M_2\| < 1$. Then M_2 is injective and*

$$\|M_2^+\| \leq \frac{\|M_1^+\|}{1 - \|M_1^+\| \|M_1 - M_2\|}.$$

We provide in the following lemma a perturbation result.

Lemma 2.6. *If $x \in U(x^*, t)$, $t \in [0, \min\{\kappa, \nu_0\})$, $\|x - x^*\| \leq \min\{\kappa, \nu_0\}$, then the following assertions hold: $F'(x)^*F'(x)$ is invertible and*

$$(2.19) \quad \|F'(x)^+\| \leq -\frac{\beta}{f'_0(\|x - x^*\|)} \leq -\frac{\beta}{f'_0(t)}.$$

In particular, $F'(x)^*F'(x)$ is invertible in $U(x^*, t)$.

Proof. Let $x \in U(x^*, t)$, $t \in [0, \min\{\kappa, \nu_0\})$. Using $f'_0(0) = -1$, (2.1) and the fact that f'_0 is strictly increasing, we obtain in turn

$$(2.20) \quad \begin{aligned} \beta \|F'(x^*)^{-1}(F'(x) - F'(x^*))\| &\leq f'_0(\|x - x^*\|) - f'_0(0) \\ &= f'_0(\|x - x^*\|) + 1 \leq f'_0(t) + 1 < 1. \end{aligned}$$

The last inequality in (2.20) holds by the definitions of κ, ν_0 and the choice of t . It then follows from (2.20) and Lemma 2.5, that $F'(x)^*F'(x)$ is invertible so that (2.19) holds. That completes the proof of Lemma 2.6. \square

The Newton-Gauss iteration at a point is a zero of the linearization of F at such a point. Hence, we shall study the linearization error at a point in \mathcal{D} :

$$(2.21) \quad E_F(x, y) := F(y) - (F(x) + F'(x)(x - y)) \quad \text{for all } x, y \in \mathcal{D}.$$

We shall bound this error by the error in linearization of the majorant function f :

$$(2.22) \quad e_f(t, u) := f(u) - (f(t) + f'(t)(u - t)) \quad \text{for all } t, u \in [0, R].$$

Lemma 2.7. *If $\|x^* - x\| < \kappa$, then the following assertion holds*

$$\beta \|E_F(x, x^*)\| \leq e_f(\|x - x^*\|, 0).$$

Proof. The proof of Lemma 2.7 is given in Lemma 7 [10]. □

Lemma 2.6 guarantees the invertibility of F' and consequently

$$(2.23) \quad \begin{array}{ccc} N_F : U(x^*, r_0) & \longrightarrow & \mathcal{Y} \\ x & \longrightarrow & x - F'(x)^{-1} F(x) \end{array}$$

is a well defined operator.

Lemma 2.8. *If $\|x - x^*\| < r_0$, then the following assertions hold*

$$(2.24) \quad \|N_F(x) - x^*\| \leq |\eta_{f, f_0}(\|x - x^*\|)|$$

and

$$(2.25) \quad N_F(U(x^*, r_0)) \subset U(x^*, r_0).$$

Proof. It follows from $F'(x^*)^+ F(x^*) = 0$ that the first inequality is trivial for $x = x^*$. If $0 < \|x - x^*\| < r_0$, Lemma 2.6 implies that $F'(x^*)^+ F'(x)$ is invertible. Using $F(x^*) = 0$ and (2.23), we obtain the approximation

$$(2.26) \quad \begin{aligned} x^* - N_F(x) &= -F'(x)^+ (F(x^*) - F(x) - F'(x)(x^* - x)) \\ &= -F'(x)^+ E_F(x, x^*). \end{aligned}$$

Using Lemma 2.6, Lemma 2.7 and (2.26), we get in turn

$$(2.27) \quad \begin{aligned} \|x^* - N_F(x)\| &\leq \|F'(x)^+\| \|E_F(x, x^*)\| \\ &\leq \beta \frac{E_F(x, x^*)}{|f'_0(\|x - x^*\|)|} \\ &\leq \frac{e_f(\|x - x^*\|, 0)}{|f'_0(\|x - x^*\|)|}. \end{aligned}$$

By the definition of e_f , η_{f, f_0} and hypothesis $f(0) = 0$, we have

$$(2.28) \quad \frac{e_f(\|x - x^*\|, 0)}{|f'_0(\|x - x^*\|)|} = |\eta_{f, f_0}(\|x - x^*\|)|,$$

which together with (2.27) show (2.24). Let $x \in U(x^*, r_0)$. It follows from $\|x - x^*\| < r_0 \leq \rho_0$, (2.24) and Lemma 2.3 that

$$\|N_F(x) - x^*\| \leq |\eta_{f, f_0}(\|x - x^*\|)| < \|x - x^*\|,$$

which shows (2.25). That completes the proof of Lemma 2.8. □

Lemma 2.9. *If (\mathcal{H}_3) holds and*

$$(2.29) \quad \|x - x^*\| \leq t < r_0,$$

then the following assertion holds

$$(2.30) \quad \|N_F(x) - x^*\| \leq \frac{|\eta_{f,f_0}(t)|}{t^{p+1}} \|x - x^*\|^{p+1}.$$

Proof. Estimate (2.30) is trivial, if $x = x^*$. Assume $\|x - x^*\| \leq t$, then (\mathcal{H}_3) and (2.12) imply

$$(2.31) \quad \frac{|\eta_{f,f_0}(\|x - x^*\|)|}{\|x - x^*\|^{p+1}} \leq \frac{|\eta_{f,f_0}(t)|}{t^{p+1}}.$$

The result follows from Lemma 2.8 and (2.31). That completes the proof of Lemma 2.9. □

Next, we shall establish the uniqueness and optimal convergence radius. The proof of the next two results can be found in the analogous Lemma 10 and Lemma 11, respectively [10, 14].

Lemma 2.10. *The point x^* is the unique zero of operator F in $U(x^*, \sigma_0)$.*

Lemma 2.11. *If*

$$\left(\frac{f(\rho_0)}{\rho_0 f'(\rho_0)} - 1 \right) f_1(\rho_0) = 1$$

and $\rho_0 < \kappa$, then $r_0 = \rho_0$ is the optimal convergence radius.

Finally, we shall show the statements of Theorem 2.1 involving Newton’s method sequence $\{x_n\}$. It follows from (1.2) and (2.23) that Newton’s method can be written as:

$$(2.32) \quad x_{n+1} = N_F(x_n) \quad (n \geq 0).$$

Lemma 2.12. *Sequence $\{x_n\}$ is well defined, is contained in $U(x^*, r_0)$ and converges to the point x^* , which is the unique zero of F in $U(x^*, \sigma_0)$. Moreover,*

$$(2.33) \quad \lim_{n \rightarrow \infty} \frac{\|x_{n+1} - x^*\|}{\|x_n - x^*\|} = 0.$$

Furthermore, if (\mathcal{H}_3) holds, then so do (2.10) and (2.11).

Proof. Let $x_0 \in U(x^*, r_0)$ and $r_0 \leq \nu_0$. Using Lemmas 2.6, 2.8 and (2.32), we deduce that sequence $\{x_n\}$ is well defined and remains in $U(x^*, r_0)$ for all $n \geq 0$. Using Lemmas 2.3, 2.8 and (2.32), we obtain in turn

$$(2.34) \quad \begin{aligned} \|x_{n+1} - x^*\| &\leq \|N_F(x_n) - x^*\| \\ &\leq |\eta_{f,f_0}(\|x_n - x^*\|)| \|x_n - x^*\| \quad (n \geq 0). \end{aligned}$$

Hence $\{\|x_n - x^*\|\}$ is strictly decreasing and converges to some α . Since $\|x_n - x^*\|$ is inside $(0, \rho_0)$ and strictly decreasing, we obtain $0 \leq \alpha < \rho_0$. It then follows from (2.34) and the continuity of η_{f,f_0} in $[0, \rho_0)$ that $0 \leq \alpha =$

$|\eta_{f,f_0}(\alpha)|$ and from Lemma 2.3, we get $\alpha = 0$. The uniqueness part was shown in Lemma 2.10. Next we shall show (2.33) that

$$(2.35) \quad \frac{\|x_{n+1} - x^*\|}{\|x_n - x^*\|} \leq \frac{|\eta_{f,f_0}(\|x_n - x^*\|)|}{\|x_n - x^*\|} \quad (n \geq 0)$$

since, $\lim_{n \rightarrow \infty} \|x_n - x^*\| = 0$, (2.33) follows from Lemma 2.3. We shall show by induction

$$(2.36) \quad \|x_n - x^*\| \leq s_n \quad (n \geq 0).$$

Since $s_0 = \|x_0 - x^*\|$, (2.36) holds as equality for $n = 0$. Assume $\|x_k - x^*\| \leq s_k$. In view of (2.32), Lemma 2.9, the induction hypothesis and (2.18), we obtain in turn

$$(2.37) \quad \begin{aligned} \|x_{k+1} - x^*\| &= \|N_F(x_k) - x^*\| \\ &\leq \frac{|\eta_{f,f_0}(s_k)|}{s_k^{p+1}} \|x_k - x^*\|^{p+1} \leq |\eta_{f,f_0}(s_k)| = s_{k+1}. \end{aligned}$$

which completes the induction for (2.36). Hence (2.10) follows from (2.32), (2.36), Lemma 2.9 and (2.18). Finally, to show (2.11) notice that since the sequence $\left\{ \frac{s_{k+1}}{s_k^{p+1}} \right\}$ is strictly decreasing, we have

$$(2.38) \quad \frac{s_{k+1}}{s_k^{p+1}} \leq \frac{s_1}{s_0^{p+1}} \quad (k \geq 0).$$

It then follows from (2.10) that

$$(2.39) \quad \|x_{k+1} - x^*\| \leq \frac{s_1}{s_0^{p+1}} \|x_k - x^*\|^{p+1} \quad (k \geq 0).$$

The first inequality in (2.11) follows from (2.39) if $p = 0$, whereas the second inequality is also derived from (2.39) if $0 < p \leq 1$. That completes the proof of Lemma 2.12. □

The proof of Theorem 2.1 now follows the above lemmas. □

Remark 2.13. (a) If $f_0(t) = f(t)$ ($t \in [0, R)$), then our Theorem 2.1 reduces to [14, Theorem 2, p. 491]. Moreover, in this case, we have

$$s_n = t_n \quad (n \geq 0), \quad \rho_0 = \rho, \quad \sigma_0 = \sigma$$

and

$$r_0 = r,$$

where ρ, r are defined as ρ_0, r_0 , respectively by replacing (f_0, f'_0) by (f, f') . Otherwise, it constitutes an improvement with advantages as already stated in the introduction of this study (see also (2.9)).

- (b) Theorem 2.1 uses the same information (x^*, F, f) as [14, Theorem 2, p. 491], since f_0 is a special case of f . In practice, the computation of f requires that of f_0 . Note also that the existence of function f_0 is

implied by (2.2). Hence, (2.1) is not an additional hypothesis. We also have:

$$f'_0(t) \leq f'(t) \quad t \in [0, R)$$

hold in general and $\frac{f'(t)}{f'_0(t)}$ can be arbitrarily large [3, 9]. The proof of [14, Theorem 2, p. 493] uses (2.2) to obtain the estimate

$$\| F'(x)^+ \| \leq -\frac{\beta}{f'(\| x - x^* \|)} \leq -\frac{\beta}{f'(t)}$$

corresponding to (2.19). However, we note that (2.19) is a tighter estimate than the above.

- (c) If $F'(x)^+ = F'(x)^{-1}$, then the results further specialize to the ones in [10, 11]. Clearly the results in [10, 11] are also improved.

In the next Section we provide special cases and numerical examples.

3. Special cases and application

I. Convergence under Hölder-type condition

Proposition 3.1. *Let \mathcal{X} and \mathcal{Y} be Hilbert spaces and let $\mathcal{D} \subseteq \mathcal{X}$ be an open set; let $F : \mathcal{D} \subset \mathcal{X} \rightarrow \mathcal{Y}$ be a continuously Fréchet-differentiable operator such that F' has a closed image in \mathcal{D} . Let $x^* \in \mathcal{D}$, $\beta := \| F(x^*)^+ \|$, $R > 0$ and $\kappa := \sup\{t \in [0, R) : U(x^*, t) \subset \mathcal{D}\}$. Suppose $F(x^*) = 0$, $F'(x^*)$ is injective and there exist $L_0 > 0$, $L > 0$ and $0 < p \leq 1$ such that*

$$\beta \| F'(x) - F'(x^*) \| \leq L_0 \| x - x^* \|^p,$$

$$\beta \| F'(x) - F'(x_\theta) \| \leq L(1 - \theta^p) \| x - x^* \|^p,$$

for all $x \in U(x^*, \kappa)$ and $x_\theta = x^* + \theta(x - x^*)$, $\theta \in [0, 1]$. Let

$$r_0 := \min\left\{\kappa, \left(\frac{p+1}{L+L_0(p+1)}\right)^{1/p}\right\},$$

$$x_0 \in U(x^*, r_0) \setminus \{x^*\}, \quad s_0 = \| x_0 - x^* \|, \quad s_{n+1} = \frac{L p s_n^{p+1}}{(p+1)(1 - L_0 s_n^p)}.$$

Then, the following assertions hold:

- (a) *The sequence $\{s_n\}$ is well defined; strictly decreasing; contained in $(0, r_0)$; converges to zero and*

$$\lim_{n \rightarrow \infty} \frac{s_{n+1}}{s_n} = 0.$$

- (b) *The sequence $\{x_n\}$ given by the Gauss-Newton method (1.2), starting from $x_0 \in U(x^*, r_0) \setminus \{x^*\}$ is well defined; remains in $U(x^*, r_0)$ for all $n \geq 0$ and converges to x^* , which is the unique solution of (1.1) in $U\left(x^*, \left(\frac{p+1}{L_0}\right)^{1/p}\right)$, so that for $n \geq 0$:*

$$\| x_{n+1} - x^* \| \leq \frac{L p}{(p+1)(1 - L_0 s_n^p)} \| x_n - x^* \|^{p+1}$$

and

$$\|x_n - x^*\| \leq \left(\frac{Lp \|x_0 - x^*\|^p}{(p+1)(1 - L_0 \|x_0 - x^*\|^p)} \right)^{((p+1)^n - 1)/p} \|x_0 - x^*\|.$$

Furthermore, if

$$\varrho_0 = \left(\frac{p+1}{Lp + L_0(p+1)} \right)^{1/p} < \kappa,$$

then $r = \varrho_0$ is the best possible convergence radius.

Proof. Use Theorem 2.1 for functions $f_0, f : [0, \kappa] \rightarrow \mathbb{R}$ defined by

$$f_0(t) = \frac{L_0 t^{p+1}}{p+1} - t \quad \text{and} \quad f(t) = \frac{L t^{p+1}}{p+1} - t.$$

That completes the proof of Proposition 3.1. □

Remark 3.2. If $L = L_0$, our results reduce to the ones in [14, Theorem 13] (see also [4, 10, 11, 12, 13, 16]). Moreover, if $L_0 < L$, we have

$$\varrho = \left(\frac{p+1}{(2p+1)L} \right)^{1/p} < \varrho_0$$

and

$$\|x_{n+1} - x^*\| \leq \frac{Lp}{(p+1)(1 - L_0 t_n^p)} \|x_n - x^*\|^{p+1} \quad (n \geq 0).$$

That is our results provide a larger convergence radius and tighter error bounds than the ones in [10, 11, 12, 13, 14, 16, 18]. Note also that we have $\frac{\varrho}{\varrho_0} \rightarrow \left(\frac{p}{2p+1}\right)^{1/p}$ as $\frac{L_0}{L} \rightarrow 0$. So, our approach provides a radius of convergence at most $\left(\frac{p}{2p+1}\right)^{-1/p}$ times larger than the one in [10, 11, 12, 13, 14, 16, 18]. If the Lipschitz condition

$$\beta \|F'(x) - F'(y)\| \leq L \|x - y\|,$$

holds for $x, y \in \mathcal{D}$, then $p = 1$ and we have

$$\varrho = \frac{2}{3L} \leq \varrho_0 = \frac{2}{2L_0 + L}.$$

The radius of convergence ϱ was obtained by Rheinboldt [21], when $F'(x)^+ = F'(x)^{-1}$.

Example 3.3 ([3, 9]). Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$. Define function F on $\mathcal{D} = (-1, 1)$, given by

$$(3.1) \quad F(x) = e^x - 1.$$

Then, for $x^* = 0$, using (3.1), we have $F(x^*) = 0$ and $F'(x^*) = e^0 = 1$. Moreover, hypotheses of Proposition 3.1 hold for $\beta = p = 1, L = e > L_0 = e - 1$. Note that

$$\frac{L}{L_0} = \frac{e}{e - 1} = 1.581976707$$

and

$$\varrho = \frac{2}{3L} = .2452529608 < \varrho_0 = \frac{2}{2L_0 + L} = .3249472314.$$

We also can provide the comparison table using the software Maple 13. Using (2.5) and (2.8) for $x_0 = .7158$.

Comparison table.

k	(1.2) $\ x_{k+1} - x_k\ $	(2.5) s_k	(2.8) t_k
0	.2473838936	.2842	.2842
1	.03614663422	.2145495033	.4826134043
2	.0006692478074	.09909547154	1.015025071
3	2.239999498e-7	.01608560415	.7960154923
4	0	.0003616695761	.7399991923
5	~	1.778927982e-7	.7357830417
6	~	4.299999235e-14	.7357588833
7	~	0	.7357588824
8	~	~	.7357588825

The table shows that our error bounds (2.5) are tighter than (2.8). Note that hypothesis (\mathcal{H}_3) of Theorem 2.1 does not hold, since Γ is not increasing on $(0, \nu_0)$ for all $\nu_0 > 0$ (see Figure 1), where,

$$f_0(t) = \frac{(e - 1)t^2}{2} - t, \quad f(t) = \frac{et^2}{2} - t$$

and

$$\Gamma(t) = \left(\frac{f(t)}{f'(t)} - t \right) \frac{f_1(t)}{t^2} = \left(\frac{.5et - 1}{.5(e - 1)t - 1} - t \right) \frac{et - 1}{(e - 1)t^3 - t^2}$$

Example 3.4 ([9, 11]). Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$. Define function F on $\mathcal{D} = (1, 3)$, given by

$$(3.2) \quad F(x) = \frac{2}{3}x^{3/2} - x.$$

Then, the zero of F is $x^* = \frac{9}{4} = 2.25$. Using (3.2) and hypotheses of Proposition 3.1, $F'(x^*) = .5, L = 2 > L_0 = 1$ and $p = .5$. Moreover, we have

$$\varrho = .1406250000 < \varrho_0 = .1836734694.$$

Example 3.5. Let $\mathcal{X} = \mathbb{R}^3, \mathcal{Y} = \mathbb{R}^2, D = U(0, 1), x^* = (0, 0, 0)$ and define function F on D by

$$(3.3) \quad F(x, y, z) = (e^x - 1, \frac{e - 1}{2}y^2 + y).$$

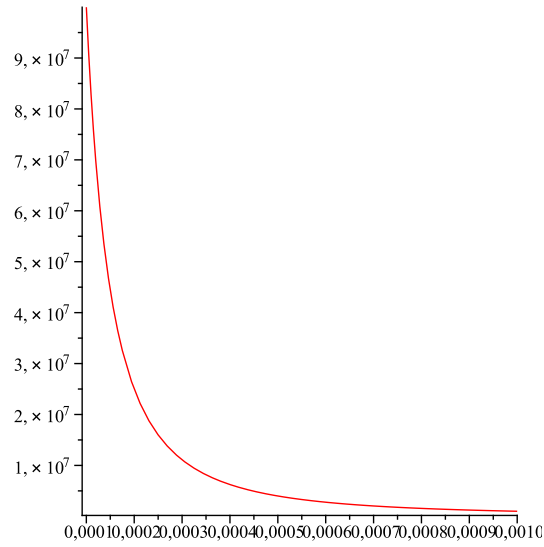


FIGURE 1. Function Γ on interval $(.0001, .001)$

We have that for $u = (x, y, z)$

$$(3.4) \quad F'(u) = \begin{pmatrix} e^x & 0 \\ 0 & (e - 1)y + 1 \\ 0 & 0 \end{pmatrix},$$

Using the norm of the maximum of the rows and (3.3)–(3.4) we can define parameters L_0 and L by

$$L_0 = e - 1 < L = e.$$

II. Convergence under generalized Lipschitz condition

Proposition 3.6. *Let \mathcal{X} and \mathcal{Y} be Hilbert spaces; let $\mathcal{D} \subseteq \mathcal{X}$ be an open and convex set; and let $F : \mathcal{D} \subset \mathcal{X} \rightarrow \mathcal{Y}$ be a continuously Fréchet-differentiable operator such that F' has a closed image in \mathcal{D} . Let $x^* \in \mathcal{D}$, $\beta := \|F(x^*)^+\|$, $R > 0$ and $\kappa := \sup\{t \in [0, R) : U(x^*, t) \subset \mathcal{D}\}$. Suppose $F(x^*) = 0$, $F'(x)$ is injective and there exist positive integrable functions $L_0, L : [0, R) \rightarrow \mathbb{R}$ such that*

$$\beta \|F'(x) - F'(x^*)\| \leq \int_0^{\|x-x^*\|} L_0(u) du,$$

$$\beta \|F'(x^*)^{-1} (F'(x) - F'(x_\theta))\| \leq \int_{\theta \|x-x^*\|}^{\|x-x^*\|} L(u) du,$$

for all $x \in U(x^*, \kappa)$ and $x_\theta = x^* + \theta(x - x^*)$, $\theta \in [0, 1]$. Let $\nu_0 > 0$, $\rho_0 > 0$ and $r_0 > 0$ be the constants defined by

$$\begin{aligned} \nu_0 &= \sup\{t \in [0, R) : \int_0^t L_0(u) \, du - 1 < 0\}, \\ \rho_0 &= \sup\{t \in [0, \nu_0) : \frac{\int_0^t L(u) \, du}{t \left(1 - \int_0^t L_0(u) \, du\right)} < 1\}, \\ r_0 &= \min\{\kappa, \rho_0\}. \end{aligned}$$

Let

$$x_0 \in U(x^*, r_0) \setminus \{x^*\}, \quad s_0 = \|x_0 - x^*\|, \quad s_{n+1} = \frac{\int_0^{s_n} L(u) \, u \, du}{1 - \int_0^{s_n} L_0(u) \, du}.$$

Then, the following assertions hold:

- (a) The sequence $\{s_n\}$ is well defined; strictly decreasing; contained in $(0, r_0)$; converges to zero and

$$\lim_{n \rightarrow \infty} \frac{s_{n+1}}{s_n} = 0.$$

- (b) The sequence $\{x_n\}$ given by the Gauss-Newton method (1.2), starting from $x_0 \in U(x^*, r_0) \setminus \{x^*\}$ is well defined; remains in $U(x^*, r_0)$ for all $n \geq 0$ and converges to x^* , which is the unique solution of (1.1) in $U(x^*, \sigma_0)$, so that

$$\lim_{n \rightarrow \infty} \frac{\|x_{n+1} - x^*\|}{\|x_n - x^*\|} = 0,$$

where,

$$\sigma_0 = \sup\{t \in [0, \kappa) : \int_0^t L_0(u) (t - u) \, du - t < 0\}.$$

Moreover, if

$$\varrho_0 = \frac{\int_0^{\rho_0} L(u) \, u \, du}{\rho_0 \left(1 - \int_0^{\rho_0} L_0(u) \, du\right)} = 1$$

and $\rho_0 < \kappa$, then $r_0 = \rho_0$ is the best possible convergence radius. Furthermore, if (\mathcal{H}_3) of Theorem 2.1 holds for

$$f(t) = \int_0^t L(u) (t - u) \, du - t,$$

$$f_0(t) = \int_0^t L_0(u) (t - u) \, du - t,$$

then estimate (2.10) and (2.11) also hold.

Proof. Hypotheses (\mathcal{H}_1) – (\mathcal{H}_3) can easily be verified with the above choices of functions f_0 and f . □

Remark 3.7. If $L = L_0$, the results of Proposition 3.6 reduce the ones [10, 14]. Otherwise they constitute an improvement with advantages as already stated in Remark 3.2.

Acknowledgements. Research of the first author supported in part by FEDER MTM 2010-17508 and 08662/PI/08. Research of the third author supported in part by MTM2011-28636-C02 -01.

References

- [1] S. Amat, S. Busquier, and J. M. Gutiérrez, *Geometric constructions of iterative functions to solve nonlinear equations*, J. Comput. Appl. Math. **157** (2003), no. 1, 197–205.
- [2] J. Appel, E. De Pascale, J. V. Lysenko, and P. P. Zabrejko, *New results on Newton–Kantorovich approximations with applications to nonlinear integral equations*, Numer. Funct. Anal. Optimiz. **18** (1997), no. 1-2, 1–17.
- [3] I. K. Argyros, *Computational Theory of Iterative Methods*, Studies in Computational Mathematics, 15, Editors: K. Chui and L. Wuytach. Elsevier, 2007.
- [4] ———, *Concerning the semilocal convergence of Newton’s method and convex majorants*, Rend. Circ. Mat. Palermo (2) **57** (2008), no. 3, 331–341.
- [5] ———, *Concerning the convergence of Newton’s method and quadratic majorants*, J. Appl. Math. Comput. **29** (2009), no. 1-2, 391–400.
- [6] ———, *A semilocal convergence analysis for directional Newton methods*, Math. Comp. **80** (2011), no. 273, 327–343.
- [7] I. K. Argyros and S. Hilout, *Extending the applicability of the Gauss–Newton method under average Lipschitz-conditions*, Numer. Algorithms **58** (2011), no. 1, 23–52.
- [8] ———, *Improved local convergence of Newton’s method under weak majorant condition*, J. Comput. Appl. Math. **236** (2012), no. 7, 1892–1902.
- [9] ———, *Computational Methods in Nonlinear Analysis*, World Scientific Publ. Comp., New Jersey, 2013.
- [10] O. P. Ferreira, *Local convergence of Newton’s method in Banach space from the viewpoint of the majorant principle*, IMA J. Numer. Anal. **29** (2009), no. 3, 746–759.
- [11] ———, *Local convergence of Newton’s method under majorant condition*, J. Comput. Appl. Math. **235** (2011), no. 5, 1515–1522.
- [12] O. P. Ferreira, M. L. N. Gonçalves, and P. R. Oliveira, *Local convergence analysis of the Gauss–Newton method under a majorant condition*, J. Complexity **27** (2011), no. 1, 111–125.
- [13] O. P. Ferreira and B. F. Svaiter, *Kantorovich’s majorants principle for Newton’s method*, Comput. Optim. Appl. **42** (2009), no. 2, 213–229.
- [14] M. L. N. Gonçalves, *Local convergence of the Gauss–Newton method for injective-overdetermined systems of equations under a majorant condition*, Comput. Math. Appl. **66** (2013), no. 4, 490–499.
- [15] J. M. Gutiérrez, M. A. Hernández, and M. A. Salanova, *Accessibility of solutions by Newton’s method*, Inter. J. Comput. Math. **57** (1995), no. 3-4, 239–247.
- [16] W. M. A. Häussler, *Kantorovich-type convergence analysis for the Gauss-Newton method*, Numer. Math. **48** (1986), no. 1, 119–125.
- [17] L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1964.
- [18] C. Li and K. F. Ng, *Majorizing functions and convergence of the Gauss-Newton method for convex composite optimization*, SIAM J. Optim. **18** (2007), no. 2, 613–692.
- [19] Á. A. Magreñán, *Estudio de la dinámica del método de Newton amortiguado*, PhD Thesis, Servicio de Publicaciones, Universidad de La Rioja, 2013.

- [20] P. D. Proinov, *General local convergence theory for a class of iterative processes and its applications to Newton's method*, J. Complexity **25** (2009), no. 1, 38–62.
- [21] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*, Polish Academy of Science, Banach Ctr. Publ. **3** (1977), no. 1, 129–142.
- [22] X. Wang, *Convergence on Newton's method and inverse function theorem in Banach space*, Math. Comput. **68** (1999), no. 225, 169–186.

SERGIO AMAT
DEPARTAMENTO DE MATEMÁTICA APLICADA Y ESTADÍSTICA
UNIVERSIDAD POLITÉCNICA DE CARTAGENA
30203 CARTAGENA, MURCIA, ESPAÑA
E-mail address: sergio.amat@upct.es

IOANNIS KONSTANTINOS ARGYROS
DEPARTMENT OF MATHEMATICS SCIENCES
CAMERON UNIVERSITY
LAWTON, OK 73505, USA
E-mail address: iargyros@cameron.edu

ÁNGEL ALBERTO MAGREÑÁN
DEPARTAMENTO DE TFG/TFM
UNIVERSIDAD INTERNACIONAL DE LA RIOJA
26002 LOGROÑO, LA RIOJA, SPAIN
E-mail address: angel.magrenan@unir.net