J. Korean Soc. Math. Educ. Ser. B: Pure Appl. Math. ISSN 1226-0657
Volume 17, Number 4 (November 2010), Pages 307–319

ON THE SEMILOCAL CONVERGENCE OF THE GAUSS-NEWTON METHOD USING RECURRENT FUNCTIONS

IOANNIS K. ARGYROS [a] AND SAÏD HILOUT [b]

ABSTRACT. We provide a new semilocal convergence analysis of the Gauss–Newton method (GNM) for solving nonlinear equation in the Euclidean space. Using our new idea of recurrent functions, and a combination of center–Lipschitz, Lipschitz conditions, we provide under the same or weaker hypotheses than before [7]–[13], a tighter convergence analysis. The results can be extented in case outer or generalized inverses are used. Numerical examples are also provided to show that our results apply, where others fail [7]–[13].

1. INTRODUCTION

In this study, we are concerned with the problem of finding $x^\star \in \mathbb{R}^i$, minimizing the objective function:

$$G(x) := \frac{1}{2} \parallel F(x) \parallel^2 = \frac{1}{2} F(x)^T F(x), \tag{1.1}$$

where $\| \, . \, \|$ denotes the Euclidean norm, and F is a Fréchet–differentiable function, defined on a convex subset $\mathcal{D}$ of $\mathbb{R}^i$, with value in $\mathbb{R}^j$ $(i \leq j)$. Many problems in applied mathematics, and also in engineering are solved by finding such solutions $x^\star$ [1]–[14].

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 the solution of the equation. Iteration methods are also used for solving optimization problems like (1.1).

Iteration sequences converge to an optimal solution of the problem at hand. In particular, here for $x^\star$ to be a local minimum it is necessary to be a zero of the

Received by the editors May 8, 2010. Accepted November 21, 2010.
2000 *Mathematics Subject Classification.* 65F20, 65G99, 65H10, 49M15.
Key words and phrases. Gauss–Newton method, semilocal convergence, Fréchet–derivative, More–Penrose pseudo–inverse.

gradien ∇G of G, too:

$$\nabla G(x^\star) = \mathcal{J}^T(x^\star)\ F(x^\star) = 0, \tag{1.2}$$

with

$$\mathcal{J}(x) = F'(x) \qquad (x \in \mathcal{D}). \tag{1.3}$$

The iterative method for computing such zero is so–called Gauss–Newton method (GNM), as introduced by Ben–Israel [7]:

$$x_{n+1} = x_n - \mathcal{J}^+(x_n)\ F(x_n) \quad (n \geq 0), \quad (x_0 \in \mathcal{D}), \tag{1.4}$$

where, $\mathcal{J}^+$ denotes the well known Moore–Penrose–pseudoinverse of $\mathcal{J}$ [5] (see also Definition 2.1). There is an extensive literature on convergence results for the (GNM). We refer the reader to [1]–[14], and the reference there. In particular, Häuβler [11] provided a Kantorovich–type semilocal convergence analysis for (GNM).

Using the center–Lipschitz conditions (instead of Lipschitz conditions used in [11]) to find more precise upper bounds on the inverses of the mappings involved, and our new idea of recurrent functions, we provide a analysis for (NGM) with the following advantages (under the same or weaker computational cost and hypotheses):

(a) finer estimates on the distances $\| x_{n+1} - x_n \|$, $\| x_n - x^\star \|$, $(n \geq 0)$;
(b) an at least as precise information on the distances involved.

Numerical examples are provided to show that our results apply, where the corresponding ones in [7]–[13] do not.

2. Semilocal Convergence Analysis of (GNM)

We need the following definition:

Definition 2.1. *$\mathcal{M}^+$ is the Moore–Penrose–pseudoinverse of matrix $\mathcal{M}$ if the following four axioms hold:*

$$(\mathcal{M}^+\ \mathcal{M})^T = \mathcal{M}^+\ \mathcal{M},$$

$$(\mathcal{M}\ \mathcal{M}^+)^T = \mathcal{M}\ \mathcal{M}^+,$$

$$\mathcal{M}^+\ \mathcal{M}\ \mathcal{M}^+ = \mathcal{M}^+,$$

and

$$\mathcal{M}\ \mathcal{M}^+\ \mathcal{M} = \mathcal{M}.$$

In the case of a full rank (m, n) matrix $\mathcal{M}$, with rank rank $\mathcal{M} = n$, the pseudo-inverse is given by:

$$\mathcal{M}^+ = (\mathcal{M}^T\,\mathcal{M})^{-1}\,\mathcal{M}^T.$$

We need also the following result on majorizing sequences for (GNM).

Lemma 2.2. *Let $\beta > 0$, $\gamma_0 > 0$, $\gamma > 0$, with $\gamma_0 \le \gamma$, and $\eta \in [0, 1)$ be given.*

Let

$$\delta_0 = \frac{\gamma\,\beta + 2\,\eta}{1 - \gamma_0\,\beta}, \tag{2.1}$$

$$\alpha = \frac{-\gamma + \sqrt{\gamma^2 + 8\,\gamma_0\,\gamma}}{4\,\gamma_0}, \tag{2.2}$$

$$\beta^\star = \min\left\{\frac{\alpha\,(\eta + 1) - \eta - \alpha^2}{\gamma_0}, \frac{2\,(\alpha - \eta)}{\gamma + 2\,\alpha\,\gamma_0}\right\}. \tag{2.3}$$

Assume that the following hold:

$$\eta < \alpha \quad \text{and} \quad \beta \le \beta^\star. \tag{2.4}$$

Then, scalar sequence $\{t_n\}$ $(n \ge 0)$ generated by

$$t_0 = 0,\ t_1 = \beta,\ t_{n+2} = t_{n+1} + \frac{\gamma\,(t_{n+1} - t_n) + 2\,\eta}{2\,(1 - \gamma_0\,t_{n+1})}\,(t_{n+1} - t_n) \tag{2.5}$$

is increasing, bounded from above by

$$t^{\star\star} = \frac{\eta}{1 - \alpha}, \tag{2.6}$$

and

converges to its unique least upper bound $t^\star$ such that

$$t^\star \in [0, t^{\star\star}]. \tag{2.7}$$

Moreover, the following estimates hold for all $n \ge 0$:

$$0 \le t_{n+2} - t_{n+1} \le \alpha\,(t_{n+1} - t_n) \le \alpha^{n+1}\,\beta, \tag{2.8}$$

and

$$t^\star - t_n \le \frac{\beta\,\alpha^n}{1 - \alpha}. \tag{2.9}$$

Proof. We note that $\beta^\star > 0$ by the choice of η, and (2.4). Moreover, we have also $\delta_0 \leq 2\,\alpha$ by (2.4). We shall show using induction on the integer m:

$$0 < t_{m+2} - t_{m+1} = \frac{\gamma\ (t_{m+1} - t_m) + 2\ \eta}{2\ (1 - \gamma_0\ t_{m+1})}\ (t_{m+1} - t_m) \leq \alpha\ (t_{m+1} - t_m), \tag{2.10}$$

and

$$\gamma_0\ t_{m+1} < 1. \tag{2.11}$$

If (2.10), and (2.11) hold, then we have (2.8) holds, and

$$\begin{aligned} t_{m+2} &\leq t_{m+1} + \alpha\ (t_{m+1} - t_m) \\ &\leq t_m + \alpha\ (t_m - t_{m-1}) + \alpha\ (t_{m+1} - t_m) \\ &\leq \eta + \alpha\ \beta + \cdots + \alpha^{m+1}\ \beta \\ &= \frac{1 - \alpha^{m+2}}{1 - \alpha}\ \beta < \frac{\eta}{1 - \alpha} = t^{\star\star} \qquad \text{(by (2.6))}. \end{aligned} \tag{2.12}$$

Estimates (2.10) and (2.11) hold for $m = 0$, by the initial conditions, (2.4), and the choices of α, and δ_0:

$$\frac{\gamma\ (t_1 - t_0) + 2\ \eta}{1 - \gamma_0\ t_1} = \frac{\gamma\ \beta + 2\ \eta}{1 - \gamma_0\ \beta} = \delta_0 \leq 2\,\alpha,$$

$$\gamma_0\ t_1 = \gamma_0\ \beta < 1.$$

Let us assume (2.8), (2.10), and (2.11) hold for all $m \leq n + 1$.

Estimate (2.10) can be re–written as:

$$\gamma\ (t_{m+1} - t_m) + 2\ \eta + \gamma_0\ \delta\ t_{m+1} - 2\,\alpha \leq 0,$$

or

$$\gamma\ \alpha^m\ \beta + 2\ \gamma_0\ \alpha\ \frac{1 - \alpha^{m+1}}{1 - \alpha}\ \beta + 2\ \eta - 2\,\alpha \leq 0. \tag{2.13}$$

Estimate (2.13) motivates us to introduce functions f_m on $[0, +\infty)$ $(m \geq 1)$ for $s = \alpha$ by:

$$f_m(s) = \gamma\ s^m\ \beta + 2\ \gamma_0\ s\ (1 + s + s^2 + \cdots + s^m)\ \beta - 2\ s + 2\ \eta. \tag{2.14}$$

Estimate (2.13) certainly holds, if:

$$f_m(\alpha) \leq 0 \quad \text{for all} \quad m \geq 1. \tag{2.15}$$

We need to find a relationship between two consecutive polynomials f_m:

$$\begin{aligned} f_{m+1}(s) &= \gamma\ s^{m+1}\ \beta + 2\ \gamma_0\ s\ (1 + s + s^2 + \cdots + s^m + s^{m+1})\ \beta - 2\ s + 2\ \eta \\ &= \gamma\ s^m\ \beta - \gamma\ s^m\ \beta + \gamma\ s^{m+1}\ \beta + \\ &\quad\ 2\ \gamma_0\ s\ (1 + s + s^2 + \cdots + s^m)\ \beta + 2\ \gamma_0\ s^{m+2}\ \beta - 2\ s + 2\ \eta \\ &= f_m(s) + g(s)\ \beta\ s^m, \end{aligned} \tag{2.16}$$

where,

$$g(s) = 2\,\gamma_0\,s^2 + \gamma\,s - \gamma. \tag{2.17}$$

Note that function g has a unique positive root α given by (2.2).

In view of (2.16), and (2.17), we have

$$f_m(\alpha) = f_1(\alpha) \qquad (m \geq 1). \tag{2.18}$$

Moreover, define

$$f_\infty(\alpha) = \lim_{m\to\infty} f_m(\alpha), \quad s \in [0,1), \quad (m \geq 1). \tag{2.19}$$

Then, we have by (2.18) that

$$f_\infty(\alpha) = f_m(\alpha) \qquad (m \geq 1). \tag{2.20}$$

In view of (2.20), we can show, instead of (2.15), since,

$$f_\infty(\alpha) = 2\left(\frac{\gamma_0\,\beta\,\alpha}{1-\alpha} + \eta - \alpha\right)$$

that

$$f_\infty(\alpha) \leq 0,$$

which is true by (2.4). That completes the induction.

Estimate (2.9) follows from (2.8) by using standard majorization techniques [5], [13]. Finally, sequence $\{t_n\}$ is non–decreasing, bounded from above by $t^{\star\star}$, and as such it converges to its unique least upper bound $t^\star$.

That completes the proof of Lemma 2.2. □

We need the following standard perturbation lemma [5], [11], [14].

Lemma 2.3. *Let A and B be $(m \times n)$ matrices. Assume:*

$$rank\,(A) \leq rank\,(B) = r \leq i \quad (r \geq 1), \tag{2.21}$$

and

$$\parallel A - B \parallel\, \parallel B^+ \parallel < 1. \tag{2.22}$$

Then, the following hold:

$$rank\,(A) = r, \tag{2.23}$$

and

$$\parallel A^+ \parallel \leq \frac{\parallel B^+ \parallel}{1 - \parallel B^+ \parallel\, \parallel A - B \parallel}. \tag{2.24}$$

We can show the semilocal convergence result for (GNM):

Theorem 2.4. *Let $F \in \mathcal{C}^1(\mathcal{D}_0)$, $\mathcal{D}_0 \subseteq \mathcal{D} \subseteq \mathbb{R}^i$, and $\mathcal{D}_0$ be a convex set.*

Assume:

there exist $x_0 \in \mathcal{D}_0$, and constants $\beta > 0$, $\beta_0 > 0$ $K > 0$, $K_0 > 0$, and $\eta : \mathcal{D}_0 \longrightarrow \mathbb{R}^+$, such that for all $x, y \in \mathcal{D}_0$:

$$rank\,(\mathcal{J}(x_0)) = r \leq i \quad r \geq 1, \tag{2.25}$$

$$rank\,(\mathcal{J}(x)) \leq r, \tag{2.26}$$

$$\| \mathcal{J}^+(x_0)\ F(x_0) \| \leq \beta, \tag{2.27}$$

$$\| \mathcal{J}(x) - \mathcal{J}(y) \| \leq K \ \| x - y \|, \tag{2.28}$$

$$\| \mathcal{J}(x) - \mathcal{J}(x_0) \| \leq K_0 \ \| x - x_0 \|, \tag{2.29}$$

$$\| \mathcal{J}^+(x_0) \| \leq \beta_0, \tag{2.30}$$

$$\| \mathcal{J}^+(y)\ r(x) \| \leq \eta(x) \ \| x - y \|, \tag{2.31}$$

with

$$r(x) = (\mathcal{I} - \mathcal{J}(x)\ \mathcal{J}^+(x))\ F(x), \tag{2.32}$$

$$\eta(x) \leq \eta < 1, \tag{2.33}$$

$$\overline{U}(x_0, t^\star) \subseteq \mathcal{D}_0, \tag{2.34}$$

where, $t^\star$ is given in (2.7),

and

hypotheses of Lemma 2.2 hold, for

$$\gamma_0 = \beta_0\ K_0, \quad \text{and} \quad \gamma = \beta_0\ K. \tag{2.35}$$

Then, the following hold:

$$rank\,(\mathcal{J}(x)) = r \qquad x \in U(x_0, t^\star); \tag{2.36}$$

Sequence $\{x_n\}$ $(n \geq 0)$ generated by (GNM) is well defined, remains in $\overline{U}(x_0, t^\star)$ for all $n \geq 0$, and converges to a zero $x^\star$ of $\mathcal{J}^+(x)\ F(x)$ in $\overline{U}(x_0, t^\star)$;

$$\| x_{n+1} - x_n \| \leq t_{n+1} - t_n, \tag{2.37}$$

and

$$\| x_n - x^\star \| \leq t^\star - t_n, \tag{2.38}$$

where, sequence $\{t_n\}$ is given in Lemma 2.2.

Moreover, the following equality holds

$$\text{(2.39)} \qquad rank\,(\mathcal{J}(x^\star)) = r,$$

and, if rank $(\mathcal{J}(x_0)) = i$, and $F(x^\star) = 0$, then, $x^\star$ is unique in $U(x_0, t^{\star\star})$, and also $x^\star$ is the unique zero of $\mathcal{J}^+(x)\ F(x)$ in $U(x_0, t^\star)$ too.

Proof. By hypothesis $x_1 \in \overline{U}(x_0, t^\star)$, since $\| x_1 - x_0 \| \le \beta \le t^\star$. Then, (2.37) holds for $n = 0$.

Assume $x_m \in \overline{U}(x_0, t^\star)$, and (2.37) holds for $m \le n$.

Using (2.29), and (2.11), we get:

$$\text{(2.40)} \qquad \begin{aligned} \| \mathcal{J}(x_m) - \mathcal{J}(x_0) \| &\le K_0\ \| x_m - x_0 \| \\ &\le K_0\ (t_m - t_0) = K_0\ t_m < \frac{1}{\beta_0}. \end{aligned}$$

It follows from (2.40), Lemma 2.3, that (2.36), (2.39), and

$$\text{(2.41)} \qquad \| \mathcal{J}^+(x_m) \| \le \frac{\beta_0}{1 - \beta_0\ K_0\ \| x_m - x_0 \|} \le \frac{\beta_0}{1 - \gamma_0\ t_m}$$

hold.

Using (1.4), (2.5), (2.28), (2.31)–(2.35), (2.41), and the induction hypotheses, we obtain in turn:

(2.42)

$$\begin{aligned} \| x_{m+1} - x_m \| &= \| \mathcal{J}^+(x_m) \int_0^1 (\mathcal{J}(x_{m-1} + \theta\ (x_m - x_{m-1})) - \mathcal{J}(x_{m-1})) \\ &\qquad \cdot (x_m - x_{m-1}) d\theta + \mathcal{J}^+(x_m)(\mathcal{I} - \mathcal{J}(x_{m-1})\mathcal{J}^+(x_{m-1}))F(x_{m-1}) \| \\ &\le \frac{1}{1 - \gamma_0\ t_m} \left(\frac{1}{2}\ \gamma\ \| x_m - x_{m-1} \| + \eta \right) \| x_m - x_{m-1} \| \\ &\le \frac{1}{2\ (1 - \gamma_0\ t_m)}\ (\gamma\ (t_m - t_{m-1}) + \eta)\ (t_m - t_{m-1}) = t_{m+1} - t_m, \end{aligned}$$

which completes the induction for (2.37).

Note also that (2.37), implies:

$$\| x_{k+1} - x_0 \| \le t_{k+1} \quad \text{for} \quad k = 1, \cdots, m+1.$$

That is $x_{m+1} \in \overline{U}(x_0, t^\star)$.

In view of Lemma 2.2, sequence $\{x_n\}$ is Cauchy in $\mathbb{R}^i$, and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star)$ (since $\overline{U}(x_0, t^\star)$ is a closed set).

We claim: $x^\star$ is a zero of $\mathcal{J}^+(x)\ F(x)$. Indeed, we get:

(2.43)
$$\begin{aligned}\| \mathcal{J}^+(x^\star)\ F(x_m) \| &\leq \| \mathcal{J}^+(x^\star)\ (\mathcal{I} - \mathcal{J}(x_m)\ \mathcal{J}^+(x_m))\ F(x_m) \| + \\ &\quad \| \mathcal{J}^+(x^\star) \|\ \| \mathcal{J}(x_m)\ \mathcal{J}^+(x_m)\ F(x_m) \| \\ &\leq \eta\ \| x_m - x^\star \| + \| \mathcal{J}^+(x^\star) \|\ \| \mathcal{J}(x_m) \|\ \| x_{m+1} - x_m \| .\end{aligned}$$

By using (2.43), and the continuilty of mapping $\mathcal{J}(x)$, $F(x)$, we justify the claim.

Finally, estimate (2.38) follows from (2.37) by using standard majorization techniques [5], [13].

The uniqueness part as identical to Lemma 2.9 in [11, p. 122] is omitted.

That completes the proof of Theorem 2.4. □

We can now state Häuβler's result for comparison purposes:

Theorem 2.5 ([11]). *Under hypotheses (2.25)-(2.33) (excluding (2.29)), further assume:*

(2.44)
$$h_H = \beta\ \gamma \leq \frac{1}{2}\ (1 - \eta)^2,$$

and

(2.45)
$$\overline{U}(x_0, v^\star) \subseteq \mathcal{D}_0,$$

where,

(2.46)
$$v^\star = \lim_{n \longrightarrow \infty} v_n,$$

(2.47)
$$v_0 = 0,\ v_1 = \beta,\ v_{n+2} = v_{n+1} + \frac{\gamma\ (v_{n+1} - v_n) + 2\ \eta}{2\ (1 - \gamma\ v_{n+1})}\ (v_{n+1} - v_n).$$

Then, the conclusions of Theorem 2.4 hold, with $v^\star$, $\{v_n\}$ replacing $t^\star$, $\{t_n\}$ ($n \geq 0$), respectively.

Remark 2.6. Note that in general

(2.48)
$$\gamma_0 \leq \gamma$$

holds in general, and $\dfrac{\gamma}{\gamma_0}$ can be arbitrarily large [3]–[5].

Using induction on integer, we can easily show:

Proposition 2.7. *Under only hypotheses of Theorem 2.5, or Theorems 2.4 and 2.5, the following hold for all $n \geq 0$:*

(2.49)
$$\| x_{n+1} - x_n \| \leq t_{n+1} - t_n \leq v_{n+1} - v_n,$$

$$t_n \leq v_n \quad (n \geq 2), \tag{2.50}$$

and

$$\| x_n - x^\star \| \leq t^\star - t_n \leq v^\star - v_n. \tag{2.51}$$

Note also that if $\gamma_0 < \gamma$, *then, strict inequality holds in (2.49), and (2.50) for* $n \geq 2$.

Remark 2.8. By Proposition 2.7, the error estimates of Theorem 2.4 can certainly be improved under the same computational cost, since in practice, the computation of γ requires that of γ_0.

In the next section, we shall show:

(a) conditions of Lemma 2.2 are always weaker than (2.44), when $\gamma_0 < \gamma$, and $i = j$ (i.e., when $\mathcal{J}(x) = F'(x)^{-1}$ $(x \in \mathcal{D}_0)$, in the case of Newton's method), where as they coincide, when $\gamma_0 = \gamma$;

(b) conditions of Lemma 2.2 can be weaker than (2.44), when $\gamma_0 < \gamma$.

3. Special Cases and Applications

Application 3.1. *(Newton's method). That is* $\eta = 0$.

Hypothesis (see [10]*)*

$$h_G = \beta\, \gamma \leq \frac{(1-\eta)^2}{2} \tag{3.1}$$

reduces to the famous for its simplicity and clarity Newton–Kantorovich hypothesis [4], [13] *for solving nonlinear equations:*

$$h_K = \gamma\, \beta \leq \frac{1}{2}. \tag{3.2}$$

Note that in this case, polynomials f_m $(m \geq 1)$ *should be:*

$$f_m(s) = \left(\gamma\, s^{m-1} + 2\, \gamma_0\, (1 + s + s^2 + \cdots + s^m)\right)\, \beta - 2, \tag{3.3}$$

and

$$f_{m+1}(s) = f_m(s) + g(s)\, s^{m-1}\, \beta. \tag{3.4}$$

It is then simple algebra to show that condition of Lemma 2.2 reduces to:

$$h_A = \alpha\, \beta \leq \frac{1}{2}, \tag{3.5}$$

where,

$$\alpha = \frac{1}{8}\left(\gamma + 4\, \gamma_0 + \sqrt{\gamma^2 + 8\, \gamma_0\, \gamma}\right). \tag{3.6}$$

In view of (3.2), (3.5), and (3.6), we get:

$$h_K \leq \frac{1}{2} \Longrightarrow h_A \leq \frac{1}{2}, \tag{3.7}$$

but not necessarily vice verca unless if $\gamma = \gamma_0$.

Moreover, if $\gamma_0 < \gamma$, *Condition (3.5) is also weaker than*

$$h_{HSL} = \frac{\gamma_0 + \gamma}{2}\ \beta \leq \frac{1}{2}, \tag{3.8}$$

provided in [12] *for nonsingular operators. Note that condition (3.8) was first given by us in* [2], [4] *for the case when linear operator* $F'(x_0)$ *is invertible.*

We provide examples, where $\gamma_0 < \gamma$, or (3.5) holds but (3.2) is violated.

Example 3.2. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$, equipped with the max–norm, and

$$x_0 = (1,1)^T, \quad U_0 = \{x : \| x - x_0 \| \leq 1 - p\}, \quad p \in \left[0, \frac{1}{2}\right).$$

Define function F on U_0 by

$$F(x) = (\xi_1^3 - p, \xi_2^3 - p), \qquad x = (\xi_1, \xi_2)^T. \tag{3.9}$$

The Fréchet–derivative of operator F is given by

$$F'(x) = \begin{bmatrix} 3\ \xi_1^2 & 0 \\ 0 & 3\ \xi_2^2 \end{bmatrix}.$$

Case 1: $\eta = 0$.

Using hypotheses of Theorem 2.4, we get:

$$\beta = \frac{1}{3}\ (1-p), \quad \gamma_0 = 3 - p, \quad \text{and} \quad \gamma = 2\ (2-p).$$

The Kantorovich condition (3.2) is violated, since

$$\frac{4}{3}\ (1-p)\ (2-p) > 1 \quad \text{for all} \quad p \in \left[0, \frac{1}{2}\right).$$

Hence, there is no guarantee that Newton's method converges to $x^\star = (\sqrt[3]{p}, \sqrt[3]{p})^T$, starting at x_0.

However, our condition (3.5) is true for all $p \in I = \left[.450339002, \frac{1}{2}\right)$. Hence, the conclusions of our Theorem 2.4 can apply to solve equation (3.9) for all $p \in I$.

Case 2: $0 \neq \eta = 0.01$.

Choose $p = .49$, then we get

$$\gamma_0 = 2.51 < \gamma = 3.02, \qquad \beta = .17,$$

$$\frac{\delta}{2} = .033058514 < \alpha = .53112045,$$

and

$$\delta_0 = .3347085 < 2\ \alpha.$$

Note also that condition (3.1) is violated no matter how η is chosen in $(0,1)$.

Finally, by comparing (3.5) with (2.44), we see that our condition is weaker provided that

$$a < \frac{\gamma}{(1-\eta)^2}, \tag{3.10}$$

which can certainly happen.

For example, if $\gamma_0 \approx 0$, then $\alpha \approx 0$, in which case (3.10) holds.

Application 3.3. In the case $\mathcal{X} = \mathcal{Y} = \mathbb{R}^j$ (j fixed in $\mathbb{N}$), we can split matrix $F'(x_n)$ into $F'(x_n) = B_n - C_n$, to obtain the inner–outer iteration:

$$x_{n+1} = x_n - (H_n^{m_n-1} + \cdots + H_n + \mathcal{I})\ B_n^{-1}\ F(x_n), \quad (n \geq 0), \tag{3.11}$$

$$H_n = B_n^{-1}\ C_n, \tag{3.12}$$

where, m_n is the number of inner iterations. Let us assume $m_n = m$ in iteration (3.11). We can obtain result concerning the estimation of the number of inner iterations under the conditions of Theorem 2.4:

Theorem 3.4. *Under the hypotheses of Theorem 2.4, further assume:*

$$\| B_0^{-1}\ F'(x_0) \| \leq q,$$

$$a_0\ h^m + m\ b\ h^{m-1} \leq \eta_n, \qquad \sup_n \| H_n \| \leq h < 1,$$

where,

$$\begin{aligned} a_0 &= \frac{3 - 2\ \eta + 2\ \beta\ \gamma^n}{\eta^2}, \\ b &= \frac{2-\eta}{\eta}\ \frac{q\ (q+1)\ \gamma_0}{[1-(1-\eta)\ \gamma_0\ q]^2}\ \left[\frac{(1-\eta)^2}{2\gamma} + \frac{1-\eta}{\gamma} + \beta\right]; \end{aligned} \tag{3.13}$$

the matrix norm has the property:

$$\| F'(x_0)^{-1}\ R \| \leq \| F'(x_0)^{-1}\ S \|$$

with R any submatrix of S;

$$\overline{U}(x_0, t^\star) \subseteq \mathcal{D};$$

and

hypotheses of Lemma 2.2 hold.

Then the conclusions of Theorem 2.4 hold true for inexact iteration (1.4).

Proof. It follows exactly as in Corollary 3.3 in [10], and our Theorem 3.7 in [6]. Here are the changes (with γ_0 replacing γ in the proof):

$$\begin{aligned}
\| F'(x_0)^{-1} F'(x_n) \| &\leq 1+\gamma_0 \| x_n - x_0 \|, \\
\| F'(x_n)^{-1} F'(x_0) \| &\leq \frac{1}{1-\gamma_0 \| x_n - x_0 \|}, \\
\| F'(x_0)^{-1} F(x_n) \| &\leq \frac{\gamma}{2} \| x_n - x_0 \|^2 + \| x_n - x_0 \| + \beta, \\
\| F'(x_0)^{-1} (B_n - B_{n-1}) \| &\leq \gamma \| x_n - x_{n-1} \|,
\end{aligned}$$

and

$$\| B_n^{-1} F'(x_0)^{-1} \| \leq \frac{q}{1-\gamma_0 \| x_n - x_0 \| q}.$$

□

The constant $\bar{b}$ defined in [10] (for $\gamma_0 = \gamma$) is larger than b, which is an advantage of our approach for the selection of a smaller η, when $\gamma < \gamma_0$.

Note that the hypotheses of Theorem 3.4 are simpler than the hypotheses of our Theorem 3.7 in [6], and weaker than Corollary 3.3 in [10].

Hence, all the above justify the claims made.

Note that in the case $i = j$, the results can be provided in affine–invariant form by simply replacing $F(x)$ by $F'(x_0)^{-1} F(x)$ for $x \in \mathcal{D}_0$, and setting $\beta_0 = 1$. The advantages of this approach have been explained in [5], [9].

Finally, our results immediately extend to the more general case of outer or generalized inverses, by simply replacing perturbation Lemma 2.3 by its analog in [8, Lemma 2.2, p. 238], (see also [1]–[5]), and using the same approach as in this paper. Note that the crucial majorizing sequence (2.5) remains the same in this new setting. We leave the details in the motivated reader.

REFERENCES

1. Argyros, I.K.: Convergence domains for some iterative processes in Banach spaces using outer or generalized inverses. *J. Comput. Anal. Appl.* **1** (1999), 87–104.
2. ———: On the Newton–Kantorovich hypothesis for solving equations. *J. Comput. Appl. Math.* **169** (2004), 315–332.
3. ———: A convergence analysis for Newton–like methods for singular equations using outer or generalized inverses. *Applicationes Mathematicae* **32** (2005), 37–49.

4. ______: On the semilocal convergence of the Gauss–Newton method. *Adv. Nonlinear Var. Inequal.* **8** (2005), 93–99.
5. ______: *Convergence and applications of Newton-type iterations.* Springer–Verlag, 2008, New York.
6. ______: On the semilocal convergence of inexact methods in Banach spaces. *J. Comput. Appl. Math.* **228** (2009), 434–443.
7. Ben–Israel, A.: A Newton–Raphson method for the solution of systems of equations. *J. Math. Anal. Appl.* **15** (1966), 243–252.
8. Chen, X. & Nashed, M.Z.: Convergence of Newton–like methods for singular operator equations using outer inverses. *Numer. Math.* **66** (1993), 235–257.
9. Deuflhard, P. & Heindl, G., Affine invariant convergence theorems for Newton's method and extensions to related methods. *SIAM J. Numer. Anal.* **16** (1979), 1–10.
10. Guo, X.: On semilocal convergence of inexact Newton methods. *J. Comput. Math.* **25** (2007), 231–242.
11. Häuβler, W.M.: A Kantorovich–type convergence analysis for the Gauss–Newton–method. *Numer. Math.* **48** (1986), 119–125.
12. Hu, N., Shen, W. & Li, C.: Kantorovich's type theorems for systems of equations with constant rank derivatives. *J. Comput. Appl. Math.* **219** (2008), 110–122.
13. Kantorovich, L.V. & Akilov, G.P.: Functional analysis in normed spaces. Pergamon Press, Oxford, 1982.
14. Penrose, R.: A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* **51** (1955), 406–413.

[a]Cameron University, Department of Mathematics Sciences, Lawton, OK 73505, USA
Email address: iargyros@cameron.edu

[b]Poitiers University, Laboratoire de Mathématiques et Applications, Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France
Email address: said.hilout@math.univ-poitiers.fr