MATHEMATICS OF COMPUTATION
Volume 80, Number 273, January 2011, Pages 327–343
S 0025-5718(2010)02398-1
Article electronically published on July 8, 2010

A SEMILOCAL CONVERGENCE ANALYSIS FOR DIRECTIONAL NEWTON METHODS

IOANNIS K. ARGYROS

Abstract. A semilocal convergence analysis for directional Newton methods in n-variables is provided in this study. Using weaker hypotheses than in the elegant related work by Y. Levin and A. Ben-Israel and introducing the center-Lipschitz condition we provide under the same computational cost as in Levin and Ben-Israel a semilocal convergence analysis with the following advantages: weaker convergence conditions; larger convergence domain; finer error estimates on the distances involved, and an at least as precise information on the location of the zero of the function. A numerical example where our results apply to solve an equation but not the ones in Levin and Ben-Israel is also provided in this study.

1. Introduction

In this study we are concerned with the problem of approximating a zero $x^\star$ of a differentiable function F defined on a convex subset $\mathcal{D}$ of $\mathbb{R}^n$ (n a natural number) with values in $\mathbb{R}$.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations.

More specifically, when it comes to computer graphics, we often need to compute and display the intersection $\mathcal{C} = \mathcal{A} \cap \mathcal{B}$ of two surfaces $\mathcal{A}$ and $\mathcal{B}$ in $\mathbb{R}^3$ (see [5], [6]). If the two surfaces are explicitly given by

$$\mathcal{A} = \{(u, v, w)^T \, : \, w = F_1(u, v)\}$$

and

$$\mathcal{B} = \{(u, v, w)^T \, : \, w = F_2(u, v)\},$$

then the solution $x^\star = (u^\star, v^\star, w^\star)^T \in \mathcal{C}$ must satisfy the nonlinear equation

$$F_1(u^\star, v^\star) = F_2(u^\star, v^\star)$$

and

$$w^\star = F_1(u^\star, v^\star).$$

Hence, we must solve a nonlinear equation in two variables $x = (u, v)^T$ of the form

$$F(x) = F_1(x) - F_2(x) = 0.$$

Received by the editor May 5 2008 and, in revised form, August 10, 2009.
2010 *Mathematics Subject Classification.* Primary 65H05, 65H10; Secondary 49M15.
Key words and phrases. Directional Newton method, systems of equations, Lipschitz/center-Lipschitz condition, Newton–Kantorovich-type hypothesis.

The marching method can be used to compute the intersection $\mathcal{C}$. In this method, we first need to compute a starting point $x_0 = (u_0, v_0, w_0)^T \in \mathcal{C}$, and then compute the succeeding intersection points by succesive updating.

In mathematical programming [9], for an equality-constraint optimization problem, e.g.,

$$\begin{array}{ll} \min & \psi(x) \\ s.t. & F(x) = 0, \end{array}$$

where $\psi,\ F : \mathcal{D} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ are nonlinear functions, we need a feasible point to start a numerical algorithm. That is, we must compute a solution of the equation $F(x) = 0$.

In the case of a system of nonlinear equations $G(x) = 0$, with $G : \mathcal{D} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$, we may solve instead

$$\|G(x)\|_2 = 0$$

if the zero of function G is isolated or locally isolated and if the rounding error is neglected [3], [7], [10], [11], [12].

We use the directional Newton method (DNM) [5] given by

$$x_{k+1} = x_k - \frac{F(x_k)}{\nabla F(x_k)\ \cdot\ d_k}\ d_k \quad (k \geq 0)$$

to generate a sequence $\{x_k\}$ converging to $x^\star$.

Let us explain how (DNM) is conceived. We start with an initial guess $x_0 \in U_0$, where F is differentiable and a direction vector d_0.

Then, we restrict F on the line $\mathcal{A} = \{x_0 + \theta\ d_0,\ \ \theta \in \mathbb{R}\}$, where it is a function of one variable $f(\theta) = F(x_0 + \theta\ d_0)$.

Set $\theta_0 = 0$ to obtain the Newton iteration for f, that is, the next point:

$$v_1 = -\frac{f(0)}{f'(0)}.$$

The corresponding iteration for F is

$$x_1 = x_0 - \frac{F(x_0)}{\nabla F(x_0)\ \cdot\ d_0}\ d_0.$$

Note that $f(0) = F(x_0)$ and $f'(0)$ is the directional derivative

$$f'(0) = F'(x_0, d_0) = \nabla F(x_0)\ \cdot\ d_0.$$

By repeating this process we arrive at (DNM).

If $n = 1$, (DNM) reduces to the classical Newton method [1]–[3], [7].

A semilocal convergence analysis for the (DNM) was provided in the elegant work by Levin and Ben-Israel in [5].

The quadratic convergence of the method was established for directions d_k sufficiently close to the gradients $\nabla F(x_k)$, and under standard Newton–Kantorovich-type hypotheses [1]–[3], [7].

In this study, we are motivated by the paper [5] and optimization considerations. By introducing the center-Lipschitz condition and using it, in combination with the Lipschitz condition (along the lines of our works in [1]–[3]), we provide a semilocal convergence analysis with the following advantages over the work in [5]:

(1) Weaker hypotheses;
(2) Larger convergence domain for (DNM);
(3) Finer error bounds on the distances $\| x_{k+1} - x_k \|$, $\| x_k - x^\star \|$ $(k \geq 0)$;

(4) An at least as precise information on the location of the zero $x^\star$.

A numerical example where our results apply, but the corresponding ones in [5] cannot is also provided in this study.

Throughout the study, we use the Euclidean inner product, the corresponding norm $\| x \|$, and the corresponding matrix norm $\| A \|$, except in Section 3 where the ∞–norm is used for vectors and matrices, denoted by $\| x \|_\infty$, and $\| A \|_\infty$, respectively.

2. Semilocal convergence analysis

We need the following lemma on majorizing sequences for (DNM). The proof can be found in the appendix.

Lemma 2.1. *Assume: there exist constants $L_0 \geq 0$, $L \geq 0$, with $L_0 \leq L$, and $\eta \geq 0$, such that:*

$$q_0 = \overline{L}\ \eta \begin{cases} \leq \dfrac{1}{2}, & if \quad L_0 \neq 0, \\ \\ < \dfrac{1}{2}, & if \quad L_0 = 0, \end{cases} \tag{2.1}$$

where

$$\overline{L} = \frac{1}{8} \left(L + 4\ L_0 + \sqrt{L^2 + 8\ L_0\ L} \right). \tag{2.2}$$

Then, the sequence $\{t_k\}$ $(k \geq 0)$ given by

$$t_0 = 0, \quad t_1 = \eta, \quad t_{k+1} = t_k + \frac{L\ (t_k - t_{k-1})^2}{2\ (1 - L_0\ t_k)} \qquad (k \geq 1) \tag{2.3}$$

is well defined, nondecreasing, bounded above by $t^{\star\star}$, and converges to its unique least upper bound $t^\star \in [0, t^{\star\star}]$, where

$$t^{\star\star} = \frac{2\ \eta}{2 - \delta}, \tag{2.4}$$

$$1 \leq \delta = \frac{4\ L}{L + \sqrt{L^2 + 8\ L_0\ L}} < 2 \quad \text{for } L_0 \neq 0. \tag{2.5}$$

Moreover, the following estimates hold:

$$L_0\ t^\star \leq 1, \tag{2.6}$$

$$0 \leq t_{k+1} - t_k \leq \frac{\delta}{2}\ (t_k - t_{k-1}) \leq \cdots \leq \left(\frac{\delta}{2}\right)^k \eta \quad (k \geq 1), \tag{2.7}$$

$$t_{k+1} - t_k \leq \left(\frac{\delta}{2}\right)^k (2\ q_0)^{2^k - 1}\ \eta \quad (k \geq 0), \tag{2.8}$$

$$0 \leq t^\star - t_k \leq \left(\frac{\delta}{2}\right)^k \frac{(2\ q_0)^{2^k - 1}\ \eta}{1 - (2\ q_0)^{2^k}} \quad (2\ q_0 < 1), \quad (k \geq 0). \tag{2.9}$$

Here, $\angle$ denotes the angle between two vectors u and v, given by

$$\angle(u, v) = \arccos \frac{u \cdot v}{\| u \| \cdot \| v \|}, \quad u \neq 0, \quad v \neq 0.$$

We provide the main semilocal convergence theorem for (DNM):

Theorem 2.2. *Let $F : \mathcal{D} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable function.*
Assume:

(i) *There exists a point $x_0 \in \mathcal{D}$, such that:*

$$F(x_0) \neq 0, \qquad \nabla F(x_0) \neq 0.$$

Let $d_0 \in \mathbb{R}^n$ be such that $\| d_0 \| = 1$, and set

$$h_0 = -\frac{F(x_0)}{\nabla F(x_0) \cdot d_0} \, d_0,$$

$$x_1 = x_0 + h_0.$$

(ii) *For $F \in \mathcal{C}^2[\mathcal{D}]$, there exist constants M_0 and M such that:*

$$\| \nabla F(x) - \nabla F(x_0) \| \leq M_0 \, \| x - x_0 \|, \qquad x \in \mathcal{D}, \tag{2.10}$$

$$\sup_{x \in \mathcal{D}} \| F''(x) \| = M, \tag{2.11}$$

$$p_0 = |F(x_0)| \, \overline{M} \, |\nabla F(x_0) \cdot d_0|^{-2} \leq \frac{1}{2}, \tag{2.12}$$

and

$$U_0 = \{x \in \mathbb{R}^n : \| x - x_0 \| \leq t^\star\} \subseteq \mathcal{D}, \tag{2.13}$$

where

$$\overline{M} = \frac{1}{8} \left(M + 4 \, M_0 + \sqrt{M^2 + 8 \, M_0 \, M} \right). \tag{2.14}$$

(iii) *The sequence $\{x_k\}$ $(k \geq 0)$ given by*

$$x_{k+1} = x_k + h_k, \tag{2.15}$$

where

$$h_k = -\frac{F(x_k)}{\nabla F(x_k) \cdot d_k} \, d_k, \tag{2.16}$$

satisfies

$$\angle(d_k, \nabla F(x_k)) \leq \angle(d_0, \nabla F(x_0)), \qquad k \geq 0, \tag{2.17}$$

where each $d_k \in \mathbb{R}^n$ is such that $\| d_k \| = 1$.

Then the sequence $\{x_k\}$ remains in U_0 for all $k \geq 0$ and converges to a zero $x^\star \in U_0$ of function F.

Moreover, $\nabla F(x^\star) \neq 0$ unless $\| x^\star - x_0 \| = t^\star$.

Furthermore, the following estimates hold for all $k \geq 0$:

$$\| x_{k+1} - x_k \| \leq t_{k+1} - t_k \leq \left(\frac{\delta}{2}\right)^k (2\, p_0)^{2^k - 1} \, \eta \tag{2.18}$$

and

$$\| x_k - x^\star \| \leq t^\star - t_k \leq \left(\frac{\delta}{2}\right)^k \frac{(2\, p_0)^{2^k-1} \, \eta}{1 - (2\, p_0)^{2^k}} \quad (2p_0 < 1), \tag{2.19}$$

where the iteration $\{t_k\}$ is given by (2.3), for

$$L_0 = |\nabla F(x_0)|^{-1} \, M_0, \quad L = |\nabla F(x_0) \cdot d_0|^{-1} \, M, \quad \eta = |\nabla F(x_0) \cdot d_0|^{-1} \, |F(x_0)|.$$

Note that condition (2.17) *is equivalent to*

$$\frac{|\nabla F(x_k)\ \cdot\ d_k|}{\parallel \nabla F(x_k) \parallel} \geq \frac{|\nabla F(x_0)\ \cdot\ d_0|}{\parallel \nabla F(x_0) \parallel}, \qquad k \geq 0. \tag{2.20}$$

Proof. We shall show the following using mathematical induction on $k \geq 0$:

$$\parallel x_{k+1} - x_k \parallel \leq t_{k+1} - t_k \tag{2.21}$$

and

$$\overline{U}(x_{k+1}, t^\star - t_{k+1}) \subseteq \overline{U}(x_k, t^\star - t_k). \tag{2.22}$$

For every $z \in \overline{U}(x_1, t^\star - t_1)$,

$$\begin{aligned} \parallel z - x_0 \parallel &\leq \parallel z - x_1 \parallel + \parallel x_1 - x_0 \parallel \\ &\leq t^\star - t_1 + t_1 - t_0 = t^\star - t_0 \end{aligned}$$

shows that $z \in \overline{U}(x_0, t^\star - t_0)$.

Since, also

$$\parallel x_1 - x_0 \parallel = \parallel h_0 \parallel \leq \eta = t_1 - t_0,$$

estimates (2.21) and (2.22) hold for $k = 0$.

Assume that (2.21) and (2.22) hold for all $i \leq k$. Then we have:

$$\begin{aligned} \parallel x_{k+1} - x_0 \parallel &\leq \parallel x_{k+1} - x_k \parallel + \parallel x_k - x_{k-1} \parallel + \cdots + \parallel x_1 - x_0 \parallel \\ &\leq (t_{k+1} - t_k) + (t_k - t_{k-1}) + \cdots + (t_1 - t_0) = t_{k+1} \end{aligned} \tag{2.23}$$

and

$$\parallel x_k + t\,(x_{k+1} - x_k) - x_0 \parallel \leq t_k + t\ (t_{k+1} - t_k) \leq t^\star, \quad t \in [0, 1].$$

Using condition (2.10) for $x = x_k$, we get in turn:

(2.24)

$$\begin{aligned} \parallel \nabla F(x_k) \parallel &\geq \parallel \nabla F(x_0) \parallel - \parallel \nabla F(x_k) - \nabla F(x_0) \parallel \\ &\geq \parallel \nabla F(x_0) \parallel - M_0 \parallel x_k - x_0 \parallel \\ &\geq \parallel \nabla F(x_0) \parallel - M_0 (t_k - t_0) \\ &\geq \parallel \nabla F(x_0) \parallel - M_0 t_k > 0 \qquad \text{(by (2.22) and Lemma 2.1).} \end{aligned}$$

We have the identity

$$\begin{aligned} \int_{x_{k-1}}^{x_k} (x_k - x)\ F''(x)\ dx &= -(x_k - x_{k-1})\ \nabla F(x_{k-1}) + F(x_k) - F(x_{k-1}) \\ &= -h_{k-1}\ \nabla F(x_{k-1}) + F(x_k) - F(x_{k-1}) \\ &= \frac{F(x_{k-1})}{(F(x_{k-1}) \cdot d_{k-1})} (d_{k-1} \cdot \nabla F(x_{k-1})) + F(x_k) - F(x_{k-1}) \\ &= F(x_k). \end{aligned} \tag{2.25}$$

We prefer the integration to be from 0 to 1. That is why we introduce a change of variable given by $x = x_{k-1} + t\ h_{k-1}$, $t \in [0, 1]$. We can write

$$x_k - x = x_k - x_{k-1} - t\ h_{k-1} = h_{k-1} - t\ h_{k-1} = (1 - t)\ h_{k-1}, \qquad dx = h_{k-1}\ dt.$$

Then (2.25) can be written as:

$$F(x_k) = \int_0^1 (1-t)\ h_{k-1}\ F''(x_{k-1} + \theta\ h_{k-1})\ h_{k-1}\ d\theta. \tag{2.26}$$

Using (2.11), (2.15)–(2.20), we get in turn:

$$\begin{aligned}
\|x_{k+1} - x_k\| \ = \ \|h_k\| \quad &= \quad \frac{|F(x_k)|}{|\nabla F(x_k)\ \cdot\ d_k|} \\
&\leq \quad \frac{|\displaystyle\int_0^1 (1-t)\ h_{k-1}\ F''(x_0 + \theta\ h_{k-1})\ h_{k-1}\ d\theta|}{|\nabla F(x_k)\ \cdot\ d_k|} \\
&\leq \quad \frac{M\ \ \|\ h_{k-1}\ \|^2}{2\ |\nabla F(x_k)\ \cdot\ d_k|} \\
&\leq \quad \frac{M\ \ \|\ h_{k-1}\ \|^2}{2\ \ \|\ \nabla F(x_k)\ \|}\ \frac{\|\ \nabla F(x_0)\ \|}{|\nabla F(x_0)\ \cdot\ d_0|} \\
&\leq \quad \frac{M\ \ \|\ h_{k-1}\ \|^2\ \|\ \nabla F(x_0)\ \|}{2\ (\|\ \nabla F(x_0)\ \| - M_0\ t_k)\ |\nabla F(x_0)\ \cdot\ d_0|} \\
&\leq \quad \frac{M\ \ \|\ h_{k-1}\ \|^2}{2\ (1 - \|\ \nabla F(x_0)\ \|^{-1}\ \ M_0\ t_k)\ |\nabla F(x_0)\ \cdot\ d_0|} \\
&\leq \quad \frac{M\ (t_k - t_{k-1})^2}{2\ (1 - |\nabla F(x_0) \cdot d_0|^{-1}\ M_0\ t_k)\ |\nabla F(x_0)\ \cdot\ d_0|} \\
&= \quad t_{k+1} - t_k,
\end{aligned}$$

which shows (2.21) for all $k \geq 0$.

Then, for every $w \in \overline{U}(x_{k+2}, t^\star - t_{k+2})$, we obtain:

$$\begin{aligned}
\|\ w - x_{k+1}\ \| \quad &\leq \quad \|\ w - x_{k+2}\ \| + \|\ x_{k+2} - x_{k+1}\ \| \\
&\leq \quad t^\star - t_{k+2} + t_{k+2} - t_{k+1} = t^\star - t_{k+1},
\end{aligned}$$

showing (2.22) for all $k \geq 0$.

Lemma 2.1 implies that $\{t_n\}$ is a Cauchy sequence. It then follows from (2.21) and (2.22) that $\{x_n\}$ is a Cauchy sequence too, and as such it converges to some $x^\star \in U_0$ (since U_0 is a closed set).

The point $x^\star$ is a zero of F, since

$$0 \leq |F(x_k)| \leq \frac{1}{2}\ M(t_k - t_{k-1})^2 \longrightarrow 0 \quad \text{as} \quad k \to \infty.$$

Furthermore, we prove that $\nabla F(x^\star) \neq 0$, except if $\|\ x^\star - x_0\ \| = t^\star$.

Using (2.10) for $x \in U_0$, (2.6), and the definition of the constant L_0, we get

$$\begin{aligned}
\|\ \nabla F(x) - \nabla F(x_0)\ \| \quad &\leq \quad M_0\ \|\ x - x_0\ \| \\
&\leq \quad M\ t^\star \leq |\nabla F(x_0) \cdot d_0| \leq \|\ \nabla F(x_0)\ \|\ .
\end{aligned}$$

If $\|\ x - x_0\ \| < t^\star$, then by (2.10), we obtain:

$$\|\ \nabla F(x) - \nabla F(x_0)\ \| \leq M_0\ \|\ x - x_0\ \| < M_0\, t^\star \leq \|\ \nabla F(x_0)\ \|,$$

or

$$\| \nabla F(x_0) \| > \| \nabla F(x) - \nabla F(x_0) \|,$$

which shows $\nabla F(x) \neq 0$.

The left-hand side inequality in (2.20) follows from (2.19) by using standard majorization techniques [3], [6], [8], [9].

That completes the proof of Theorem 2.2. □

Note that $t^\star$ in (2.13) can be replaced by $t^{\star\star}$ given in closed form by (2.4).

It follows from the proof of Theorem 2.2, and Lemma 2.1, that

$$\| h_{k+1} \| \leq \frac{\delta}{2} \| h_k \| \quad (k \geq 0). \tag{2.27}$$

Therefore, if we define the nested balls

$$\mathcal{S}_k = \{x \in \mathbb{R}^n : \| x - x_{k+1} \| \leq \| h_k \|\},$$

the proof given in [5] applies for Theorem 2.2 by simply replacing $\frac{1}{2}$ by $\frac{\delta}{2}$.

However, we decided to provide a proof for Theorem 2.2 different than the corresponding one in [5].

Definition 2.3. Let $F : \mathcal{D} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable function, where $\mathcal{D}$ is a convex region. The function ∇F is said to be Lipschitz continuous if there exists a constant $M \geq 0$ such that:

$$\| \nabla F(x) - \nabla F(y) \| \leq M \| x - y \| \quad \text{for all} \quad x, y \in \mathcal{D}. \tag{2.28}$$

Note that in view of (2.28), there exists a center-Lipschitz constant $M_0 \geq 0$ such that (2.10) holds.

Clearly,

$$M_0 \leq M \tag{2.29}$$

holds in general, and $\frac{M}{M_0}$ can be arbitrarily large [1]–[3].

Remark 2.4. If F is twice differentiable, the Lipschitz constant M in (2.28) can replace the corresponding constant in hypothesis (2.11). Note also that in view of the proof of Theorem 2.2, the constant L_0 can be defined by the more precise:

$$L_0 = \| \nabla F(x_0) \|^{-1} M_0.$$

Remark 2.5. Our Theorem 2.2 improves Theorem 1 in [5, pages 252–253].

(1) *Case* $M = M_0$: Theorem 1 in [5] uses the stronger, and more difficult to verify than (2.17) condition:

$$\angle(d_{k+1}, \nabla F(x_{k+1}) \leq \angle(d_k, \nabla F(x_k) \qquad (k \geq 0), \tag{2.30}$$

or equivalently

$$\frac{|\nabla F(x_{k+1}) \cdot d_{k+1}|}{\| \nabla F(x_{k+1}) \|} \geq \frac{|\nabla F(x_k) \cdot d_k|}{\| \nabla F(x_k) \|} \qquad (k \geq 0). \tag{2.31}$$

(1) *Case* $M_0 < M$: Theorem 1 in [5] again uses (2.30) instead of the weaker condition (2.17).

Theorem 1 in [5] uses the condition

$$p = |F(x_0)|\ M\ |\nabla F(x_0)\ \cdot\ d_0|^{-2} \leq \frac{1}{2} \tag{2.32}$$

corresponding to our condition (2.12).

But, we have

$$\overline{M} < M. \tag{2.33}$$

That is,

$$p \leq \frac{1}{2} \Longrightarrow p_0 \leq \frac{1}{2}, \tag{2.34}$$

but not necessarily vice versa (unless $M_0 = L$).

Define the sequence $\{s_k\}$ $(k \geq 0)$ by

$$s_0 = 0, \qquad s_{k+1} = s_k + \parallel h_k \parallel . \tag{2.35}$$

It was shown in [5] (under (2.32)) that

$$\parallel h_{k+1} \parallel \leq \frac{1}{2} \parallel h_k \parallel, \tag{2.36}$$

i.e.,

$$s_{k+2} - s_{k+1} \leq \frac{1}{2}\ (s_{k+1} - s_k). \tag{2.37}$$

It turns out from the proof of our Theorem 2.2 that a finer sequence $\{\overline{t}_k\}$ than $\{s_k\}$ could have been used in [5], given by

$$\overline{t}_0 = 0, \quad \overline{t}_1 = \eta, \quad \overline{t}_{k+1} = \overline{t}_k + \frac{L\ (\overline{t}_k - \overline{t}_{k-1})^2}{2\ (1 - L_0\ \overline{t}_k)} \qquad (k \geq 1). \tag{2.38}$$

Note that $L_0 \leq L$, since $M_0 \leq M$. We showed in [2, page 392]:

$$t_k \leq \overline{t}_k, \tag{2.39}$$

$$t_{k+1} - t_k \leq \overline{t}_{k+1} - \overline{t}_k, \tag{2.40}$$

and

$$t^\star \leq \overline{t}^\star = \lim_{k\to\infty} \overline{t}_k. \tag{2.41}$$

Moreover, strict inequality holds in (2.39) and (2.40) if $M_0 < M$ (i.e., if $L_0 < L$).

Note that the convergence in Theorem 2.2 is quadratic, whereas it was only shown to be linear in the corresponding Theorem 1 in [5] (see also (2.8), (2.9), and (2.36)).

Finally, it follows from Lemma 2.1 and the proof of Theorem 2.2 that the sequence

$$\gamma_0 = 0, \quad \gamma_1 = \eta, \quad \gamma_{k+1} = \gamma_k + \frac{L_1\ (\gamma_k - \gamma_{k-1})^2}{2\ (1 - L\ \gamma_k)} \qquad (k \geq 1) \tag{2.42}$$

is a finer majorizing sequence for $\{x_k\}$ than $\{t_k\}$, where

$$L_1 = \begin{cases} L_0, & \text{if} \quad k = 1, \\ L, & \text{if} \quad k > 1. \end{cases} \tag{2.43}$$

We can also show the following convergence result for (DNM):

Theorem 2.6. *Under the hypotheses of Theorem 2.2 for the operator $F \in \mathcal{C}^1[\mathcal{D}]$, and (2.11) replaced by (2.28), further assume that there exist constants $M_0 > 0$, $M > 0$, $\beta_0 > 0$, $\beta > 0$, such that:*

$$|\nabla F(x_0) \cdot d_0| \geq \frac{1}{\beta_0}, \tag{2.44}$$

$$|\nabla F(x) \cdot d| \geq \frac{1}{\beta}, \quad \textit{for all} \quad x \in U \subset \mathcal{D},\ d \in \mathbb{R}^n,\ \| d \| = 1, \tag{2.45}$$

and

$$\alpha = \frac{M\ \beta\ t_1}{2} \leq 1. \tag{2.46}$$

Set

$$a = \frac{M_0\ \beta_0}{M\ \beta}. \tag{2.47}$$

Then, the following estimates hold:

$$\| x_{k+1} - x_k \| \leq a^{2^{k-1}} \alpha^{2^k - 1} t_1 \tag{2.48}$$

and

$$\| x^\star - x_k \| \leq \frac{a^{2^{k-1}} \alpha^{2^k - 1} t_1}{1 - a^{2^{k-1}} \alpha^{2^k - 1}} \quad \text{for} \quad \text{all} \quad k \geq 1 \quad (\alpha < 1). \tag{2.49}$$

Proof. According to Theorem 2.2: $x_k \in U_0$ $(k \geq 0)$. Using (2.15), (2.16), (2.44), and (2.45), we get in turn:

$$\begin{aligned} \| x_{k+1} - x_k \| &= \frac{|F(x_k)|}{|\nabla F(x_k) \cdot d_k|} \\ &\leq \beta_1\ |F(x_1)| \\ &= \beta_1\ |F(x_k) - F(x_{k-1}) - \nabla F(x_{k-1})\ (x_k - x_{k-1})|, \end{aligned} \tag{2.50}$$

where

$$\beta_1 = \begin{cases} \beta_0, & \text{if} \quad k = 1, \\ \beta, & \text{if} \quad k > 1. \end{cases} \tag{2.51}$$

Multiplying both sides of

$$x_k - x_{k-1} = - \frac{F(x_{k-1})}{\nabla F(x_{k-1}) \cdot d_{k-1}}\ d_{k-1} \tag{2.52}$$

by $\nabla F(x_{k-1})$, we obtain

$$\begin{gathered} \| x_2 - x_1 \| \leq \frac{\beta_1\ M_0}{2} \| x_1 - x_0 \|^2 = \frac{a\ \beta_1\ M}{2} \| x_1 - x_0 \|^2 = a^{2^0} \alpha^{2^1 - 1} t_1, \\ \| x_3 - x_2 \| \leq \frac{\beta_1\ M}{2} \| x_2 - x_1 \|^2 \leq \frac{\beta_1\ M}{2}\ a\ t_1\ \alpha^2 = a^{2^1} \alpha^{2^2 - 1} t_1, \\ \| x_{k+1} - x_k \| \leq a^{2^{k-1}} \alpha^{2^k - 1} t_1. \end{gathered} \tag{2.53}$$

That is, we showed

$$\| x_{k+1} - x_k \| \quad \leq \quad a^{2^{k-1}} \alpha^{2^k - 1} t_1 \quad \text{for} \quad k = 1, 2. \tag{2.54}$$

Assume it is correct up to $k-1$. Then, we have:

$$\begin{aligned} \| x_{k+1} - x_k \| &\leq \frac{\beta_1\ M}{2}\ \| x_k - x_{k-1} \|^2 \\ &\leq \frac{\beta_1\ M}{2} \left(a^{2^{k-2}} \alpha^{2^{k+1}-1} t_1 \right)^2 = a^{2^{k-1}} \alpha^{2^k-1} t_1, \end{aligned}$$

which completes the induction for (2.54) and shows (2.48).

Let $m > k$. Then using (2.48), we can obtain in turn for $\alpha < 1$:

$$\begin{aligned} \| x_m - x_k \| &\leq \| x_m - x_{m-1} \| + \| x_{m-1} - x_{m-2} \| + \cdots + \| x_{k+1} - x_k \| \\ &\leq \left(a^{2^{m-2}} \alpha^{2^{m-1}-1} + a^{2^{m-3}} \alpha^{2^{m-2}-1} + \cdots + a^{2^{k-1}} \alpha^{2^k-1} \right) t_1 \\ &\leq a^{2^{k-1}} \alpha^{2^k-1} \left(1 + a^{2^{k-1}} \alpha^{2^k} + \left(a^{2^{k-1}} \alpha^{2^k} \right)^2 + \cdots \right. \\ &\qquad \left. + \left(a^{2^{k-1}} \alpha^{2^k} \right)^{m-k-1} \right) t_1 \qquad (2.55) \\ &= a^{2^{k-1}} \alpha^{2^k-1}\ \frac{1 - \left(a^{2^{k-1}} \alpha^{2^k} \right)^{m-k}}{1 - a^{2^{k-1}} \alpha^{2^k}}\ \eta \\ &\leq \frac{a^{2^{k-1}} \alpha^{2^k-1}\ \eta}{1 - a^{2^{k-1}} \alpha^{2^k}}. \end{aligned}$$

By letting $m \longrightarrow \infty$ in (2.55), we obtain (2.49).

That completes the proof of Theorem 2.6. □

Remark 2.7. If $\beta_0 = \beta$ and $M_0 = M$, then $a = 1$, and our Theorem 2.6 reduces to a weaker version of Theorem 2 in [5].

Note that even in this case it is weaker, since (2.20) is used instead of (2.31). Otherwise (i.e., if $\beta_0 < \beta$, $M_0 = M$, or if $\beta_0 = \beta$, $M_0 < M$, or if $\beta_0 < \beta$, $M_0 < M$) the ratio of the quadratic convergence is improved since $a \in (0,1)$.

3. Special cases and applications

The gradient method is now stated as a special case of Theorem 2.2:

Corollary 3.1. Define direction d_k and step h_k for all $k \geq 0$ by:

$$d_k = \frac{\nabla F(x_k)}{\| \nabla F(x_k) \|}$$

and

$$h_k = -\frac{F(x_k)}{\| \nabla F(x_k) \|^2}\ \nabla F(x_k).$$

Let F, x_0, U_0, M_0, M, β_0, β, p_0 be as in Theorems 2.2 and 2.6. Then, the following gradient method:

$$x_{k+1} = x_k - \frac{F(x_k)}{\| \nabla F(x_k) \|^2}\ \nabla F(x_k) \quad (k \geq 0)$$

satisfies the conclusions of Theorems 2.2 and 2.6.

Following [5], as a second application, the direction d_k in each iteration is chosen as the unit vector along the maximal value $|\frac{\partial F}{\partial x_k}|$. The proof of Theorems 2.2 and 2.6 applies here, by replacing each occurrence of $|\nabla F(x_k) \cdot d_k|$ by $\| \nabla F(x_k) \|_\infty$, and by using the ∞–norm instead of the Euclidean norm. The following is the analog of Theorem 2.2.

Theorem 3.2. *Let $F : \mathcal{D} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differentiable function. If $x_0 \in \mathcal{D}$ is such that*

$$F(x_0) \neq 0, \qquad \nabla F(x_0) \neq 0,$$

let $m(0)$ be an index such that

$$|F'_\star(x_0)| = |\frac{\partial F}{\partial x_{m(0)}}(x_0)| = \max_{j=1,\cdots,n} |\frac{\partial F}{\partial x_j}(x_0)|.$$

Define:

$$h_0[k] = \begin{cases} -\dfrac{F(x_0)}{|F'_\star(x_0)|} & \text{if} \quad k = m(0), \\ \\ 0 & \text{if} \quad k \neq m(0), \end{cases}$$

$$x_1 = x_0 + h_0,$$

and

$$U_0 = \{x : \| x - x_0 \|_\infty \leq t^\star\}.$$

Assume $F \in \mathcal{C}^2[\mathcal{D}]$,

$$\sup_{x \in \mathcal{D}} \| F''(x) \|_\infty = M,$$

$$\| \nabla F(x) - \nabla F(x_0) \|_\infty \leq M_0 \| x - x_0 \|_\infty \quad \text{for} \quad \text{all} \quad x \in \mathcal{D},$$

$$\overline{p_0} = |F(x_0)| \ \overline{M} \ |\frac{\partial F}{\partial x_{m(0)}}(x_0)|^{-2} \leq \frac{1}{2},$$

and

$$U_0 \subseteq \mathcal{D},$$

where $\overline{M}$ is given by (2.14).

Define sequences $\{x_k\}$, $\{h_k\}$ as follows:

Let $m(k)$ be an index of the maximal modulus of $\frac{\partial F}{\partial x_{m(k)}}(x_k)$,

$$|F'_\star(x_k)| = |\frac{\partial F}{\partial x_{m(k)}}(x_k)| = \max_{j=1,\ldots,n} |\frac{\partial F}{\partial x_j}(x_k)|,$$

$$h_k[j] = \begin{cases} -\dfrac{F(x_k)}{|F'_\star(x_k)|}, & \text{if} \quad j = m(k), \\ \\ 0, & \text{if} \quad j \neq m(k), \end{cases}$$

and

$$x_{k+1} = x_k + h_k.$$

Then, the sequence $\{x_k\}$ remains in U_0 for all $k \geq 0$ and converges to a zero $x^\star \in U_0$ of function F.

Moreover, $\nabla F(x^\star) \neq 0$ unless $\| x^\star - x_0 \|_\infty = t^\star$. Furthermore, the following estimates hold for all $k \geq 0$:

$$\begin{aligned} \| x_{k+1} - x_k \|_\infty &\leq \frac{M}{2 \ \| \nabla F(x_k) \|_\infty} \ \| x_k - x_{k-1} \|_\infty^2 \\ &\leq \ t_{k+1} - t_k \leq \left(\frac{\delta}{2} \right)^k \ (2\,\overline{p_0})^{2^k - 1} \ \eta, \end{aligned}$$

and

$$\begin{aligned} \| x_k - x^\star \|_\infty &\leq \frac{M}{2 \ \| \nabla F(x_k) \|_\infty} \ \| x_k - x_{k-1} \|_\infty^2 \\ &\leq \ t^\star - t_k \leq \left(\frac{\delta}{2} \right)^k \ \frac{(2\,\overline{p_0})^{2^k - 1} \ \eta}{1 - (2\,\overline{p_0})^{2^k}} \quad (2 \ \overline{p_0} < 1), \end{aligned}$$

where the iteration $\{t_k\}$ is given by (2.3)*, for*

$$L_0 = \| \frac{\partial F}{\partial x_{m(0)}}(x_0) \|_\infty^{-1} \ M_0, \quad L = \| \frac{\partial F}{\partial x_{m(0)}}(x_0) \|_\infty^{-1} \ M,$$

$$\eta = \| \frac{\partial F}{\partial x_{m(0)}}(x_0) \|_\infty^{-1} \ |F(x_0)|.$$

A result similar to Theorem 2.6 can be stated under analogous conditions and proof.

The advantages of our results over the corresponding ones in [5, see Section 3, Theorem 3] have already been stated in Remarks 2.5 and 2.7.

We now refer the motivated reader to Section 4 in [5] for further applications (see also [7]–[12]). Clearly, the applicability of the results listed in [5] has now been expanded in view of our results. More applications and other relevant work can be found in [7]–[12]. Maple programs for the methods mentioned here can be downloaded from [4].

We provide an example where our Theorem 2.2 can apply to solve an equation, but not the corresponding Theorem 1 in [5].

Example 3.3. Let $n = 2$. Here, we use the Euclidean inner product and the corresponding norm for both the vector and matrix. Choose:

$$x_0 = (1,1)^T, \quad \mathcal{D} = \{x \ : \| \ x - x_0 \ \| \leq 1 - b\} \quad \text{for} \quad b \in [0,1),$$

and define function F on $\mathcal{D}$ by

$$F(x) = \frac{\theta_1^3 + \theta_2^3}{2} - b, \quad x = (\theta_1, \theta_2)^T. \tag{3.1}$$

Then, the gradient ∇F of the operator F is given by

$$\nabla F(x) = \frac{3}{2} \ (\theta_1^2, \theta_2^2)^T. \tag{3.2}$$

Using (2.10), (2.11), (3.1), and (3.2), we obtain the parameters:

$$M = 3 \ (2 - b) \ \sqrt{2}, \quad M_0 = \frac{3 \ (3 - b) \ \sqrt{2}}{2}, \quad \| \ \nabla F(x_0) \ \| = \frac{3 \ \sqrt{2}}{2},$$

and

$$F(x_0) = 1 - b \quad \text{and} \quad \eta = \frac{\sqrt{2}}{3} \ F(x_0).$$

We can choose the directions d_k by

$$d_k = \frac{\nabla F(x_k)}{\| \nabla F(x_k) \|},$$

so that condition (2.20) is satisfied as an equality.

Then, condition (2.32) used in [5] is violated for say $b = .6166$, since

$$L = 2.7668, \quad \eta = .180736493,$$

and

$$p = L\eta = \frac{2\sqrt{2}}{3}\,(2-b)\,(1-b) = .500061729 > .5.$$

Hence, there is no guarantee that (DNM) starting at x_0 converges to a zero $x^\star$ of the function F.

However, our conditions hold.

We have:

$$L_0 = 2.3834, \quad \overline{L} = 2.50910088, \quad \delta = 1.050097978.$$

That is,

$$q_0 = .453486093 < .5 \quad \textit{and} \quad t^{\star\star} = .38053712 < 1 - b = .3834.$$

Hence, the hypotheses of Theorem 2.2 are satisfied. We found $x^\star = (.851140338, .851140338)$.

That is, our Theorem 2.2 guarantees the existence of a zero $x^\star$ in U_0 of function F, obtained as the limit of (DNM) starting at x_0.

Remark 3.4. The results can be extended in a Hilbert space setting. Indeed, let F be a differentiable operator defined on a convex subset $\mathcal{D}$ of a Hilbert space $\mathcal{H}$ with values in $\mathbb{R}$. Here $x \cdot y$ denotes the inner product of elements x and y in $\mathcal{H}$, and $\| x \| = (x \cdot x)^{1/2}$.

Moreover, instead of condition (2.17) or (2.20), assume:

$$\| d_k \| = 1,$$

and, there exists $\xi \in [0,1]$ such that:

$$|\nabla F(x_k) \cdot d_k| \geq \xi \, \| \nabla F(x_k) \| \, .$$

Note that in the case of (2.17), we can always set:

$$\xi = \frac{|\nabla F(x_k) \cdot d_k|}{\| \nabla F(x_0) \|} \leq 1.$$

In the case of Theorems 2.2 and 2.6, set

$$L_0 = \frac{M_0}{\xi \ \| \nabla F(x_0) \|}, \quad L = \frac{M}{\xi \ \| \nabla F(x_0) \|}, \quad \text{and} \quad \eta = \frac{|F(x_0)|}{\xi \ \| \nabla F(x_0) \|}.$$

Then, due to the proofs of Theorems 2.2 and 2.6, the results in Section 2 hold in this more general setting. A similar extension can follow for the results of this section.

Appendix

Proof of Lemma 2.1. If $L_0 = 0$, then for $L > 0$, an induction argument shows that

$$t_{k+1} - t_k = \frac{2}{L}\ (2\ q_0)^{2^k} \qquad (k \geq 0).$$

Therefore, we get

$$t_{k+1} = t_1 + (t_2 - t_1) + \cdots + (t_{k+1} - t_k) = \frac{2}{L}\ \sum_{m=0}^{k} (2\ q_0)^{2^m}$$

and

$$t^\star = \lim_{k \to \infty} t_k = \frac{2}{L}\ \sum_{k=0}^{\infty} (2\ q_0)^{2^k}.$$

Clearly, this series converges, since $k \leq 2^k$, $2q_0 < 1$, and is bounded above by the number

$$\frac{2}{L}\ \sum_{k=0}^{\infty} (2\ q_0)^k = \frac{4}{L\ (2 - L\ \eta)}.$$

If $L = 0$, since $0 \leq L_0 \leq L$, we deduce: $L_0 = 0$, and $t^\star = t_k = \eta$ $(k \geq 1)$.

In the rest of the proof, we assume that $L_0 > 0$.

The result until estimate (2.7) follows from Lemma 1 in [2] (see also [1], [3]).

Note that, in particular, Newton–Kantorovich-type convergence condition (2.1) is given in [2, page 387, Case 3 for δ given by (2.5). The factor η is missing from the left-hand side of the inequality three lines before the end of page 387].

In order for us to show (2.8), we first need the estimate:

$$\frac{1 - \left(\dfrac{\delta}{2}\right)^{k+1}}{1 - \dfrac{\delta}{2}}\ \eta \leq \frac{1}{L_0}\ \left(1 - \left(\frac{\delta}{2}\right)^{k-1}\ \frac{L}{4\ \overline{L}}\right) \qquad (k \geq 1). \tag{A.1}$$

For $k = 1$, estimate (A.1) becomes

$$\left(1 + \frac{\delta}{2}\right)\ \eta \leq \frac{4\ \overline{L} - L}{4\ \overline{L}\ L_0}$$

or

$$\left(1 + \frac{2\ L}{L + \sqrt{L^2 + 8\ L_0\ L}}\right)\ \eta \leq \frac{4\ L_0 - L + \sqrt{L^2 + 8\ L_0\ L}}{L_0\ (4\ L_0 + L + \sqrt{L^2 + 8\ L_0\ L})}.$$

In view of hypothesis (2.1), it suffices to show:

$$\frac{L_0\ (4\ L_0 + L + \sqrt{L^2 + 8\ L_0\ L})\ (3\ L + \sqrt{L^2 + 8\ L_0\ L})}{(L + \sqrt{L^2 + 8\ L_0\ L})\ (4\ L_0 - L + \sqrt{L^2 + 8\ L_0\ L})} \leq 2\ \overline{L},$$

which is true as an equality.

Let us now assume that estimate (A.1) is true for all integers smaller than or equal to k. We must show that (A.1) holds for k replaced by $k + 1$:

$$\frac{1 - \left(\dfrac{\delta}{2}\right)^{k+2}}{1 - \dfrac{\delta}{2}}\ \eta \leq \frac{1}{L_0}\ \left(1 - \left(\frac{\delta}{2}\right)^{k}\ \frac{L}{4\ \overline{L}}\right) \qquad (k \geq 1)$$

or

$$\left(1+\frac{\delta}{2}+\left(\frac{\delta}{2}\right)^2+\cdots+\left(\frac{\delta}{2}\right)^{k+1}\right)\eta \le \frac{1}{L_0}\left(1-\left(\frac{\delta}{2}\right)^k \frac{L}{4\,\overline{L}}\right). \tag{A.2}$$

By the induction hypothesis to show estimate (A.2), it suffices to have:

$$\frac{1}{L_0}\left(1-\left(\frac{\delta}{2}\right)^{k-1}\frac{L}{4\,\overline{L}}\right)+\left(\frac{\delta}{2}\right)^{k+1}\eta \le \frac{1}{L_0}\left(1-\left(\frac{\delta}{2}\right)^k \frac{L}{4\,\overline{L}}\right)$$

or

$$\left(\frac{\delta}{2}\right)^{k+1}\eta \le \frac{1}{L_0}\left(\left(\frac{\delta}{2}\right)^{k-1}-\left(\frac{\delta}{2}\right)^k\right)\frac{L}{4\,\overline{L}}$$

or

$$\delta^2\,\eta \le \frac{L\,(2-\delta)}{2\,\overline{L}\,L_0}.$$

In view of hypothesis (2.1), we can show instead:

$$\frac{2\,\overline{L}\,L_0\,\delta^2}{L\,(2-\delta)} \le 2\,\overline{L},$$

which holds as an equality by the choice of δ given in (2.5).

That completes the induction for estimate (A.1).

We shall show (2.8) using induction on $k \ge 0$: estimate (2.8) is true for $k=0$ by (2.1), (2.3), and (2.5). In order for us to show estimate (2.8) for $k=1$, since $t_2-t_1=\dfrac{L\,(t_1-t_0)^2}{2\,(1-L_0\,t_1)}$, it suffices that

$$\frac{L\,\eta^2}{2\,(1-L_0\,\eta)} \le \delta\,\overline{L}\,\eta^2$$

or

$$\frac{L}{1-L_0\,\eta} \le \frac{8\,\overline{L}\,L}{L+\sqrt{L^2+8\,L_0\,L}} \quad (\eta \ne 0)$$

or

$$\eta \le \frac{1}{L_0}\left(1-\frac{L+\sqrt{L^2+8\,L_0\,L}}{8\,\overline{L}}\right) \quad (L_0 \ne 0,\ L \ne 0).$$

But by (2.1) we have:

$$\eta \le \frac{4}{L+4\,L_0+\sqrt{L^2+8\,L_0\,L}}.$$

It then suffices to show that

$$\frac{4}{L+4\,L_0+\sqrt{L^2+8\,L_0\,L}} \le \frac{1}{L_0}\left(1-\frac{L+\sqrt{L^2+8\,L_0\,L}}{8\,\overline{L}}\right)$$

or

$$\frac{L+\sqrt{L^2+8\,L_0\,L}}{8\,\overline{L}} \le 1-\frac{4\,L_0}{L+4\,L_0+\sqrt{L^2+8\,L_0\,L}}$$

or

$$\frac{L+\sqrt{L^2+8\,L_0\,L}}{8\,\overline{L}} \le \frac{L+\sqrt{L^2+8\,L_0\,L}}{L+4\,L_0+\sqrt{L^2+8\,L_0\,L}},$$

which is true as an equality by (2.2).

Let us assume that (2.8) holds for all integers smaller than or equal to k. We shall show that (2.8) holds for k replaced by $k+1$.

Using (2.3), and the induction hypothesis, we have in turn

$$\begin{aligned} t_{k+2} - t_{k+1} &= \frac{L}{2\,(1 - L_0\ t_{k+1})}\ (t_{k+1} - t_k)^2 \\ &\le \frac{L}{2\,(1 - L_0\ t_{k+1})} \left(\left(\frac{\delta}{2}\right)^k (2\ q_0)^{2^k - 1}\, \eta \right)^2 \\ &\le \frac{L}{2\,(1 - L_0\ t_{k+1})} \left(\left(\frac{\delta}{2}\right)^{k-1} (2\ q_0)^{-1}\, \eta \right) \left(\left(\frac{\delta}{2}\right)^{k+1} (2\ q_0)^{2^{k+1} - 1}\, \eta \right) \\ &\le \left(\frac{\delta}{2}\right)^{k+1} (2\ q_0)^{2^{k+1}-1}\, \eta, \end{aligned}$$

since

$$\frac{L}{2\,(1 - L_0\ t_{k+1})} \left(\left(\frac{\delta}{2}\right)^{k-1} (2\ q_0)^{-1}\, \eta \right) \le 1 \qquad (k \ge 1). \tag{A.3}$$

Indeed, we can show instead of (A.3):

$$t_{k+1} \le \frac{1}{L_0} \left(1 - \left(\frac{\delta}{2}\right)^{k-1} \frac{L}{4\ \overline{\overline{L}}} \right),$$

which is true, since by (2.7) and the induction hypotheses:

$$\begin{aligned} t_{k+1} &\le t_k + \frac{\delta}{2}\ (t_k - t_{k-1}) \\ &\le t_1 + \frac{\delta}{2}\ (t_1 - t_0) + \cdots + \frac{\delta}{2}\ (t_k - t_{k-1}) \\ &\le \eta + \left(\frac{\delta}{2}\right) \eta + \cdots + \left(\frac{\delta}{2}\right)^k \eta \\ &= \frac{1 - \left(\frac{\delta}{2}\right)^{k+1}}{1 - \frac{\delta}{2}}\ \eta \\ &\le \frac{1}{L_0} \left(1 - \left(\frac{\delta}{2}\right)^{k-1} \frac{L}{4\ \overline{\overline{L}}} \right). \end{aligned}$$

That completes the induction for estimate (2.8).

Using estimate (2.8) for $j \ge k$, we obtain in turn for $2\ q_0 < 1$:

$$\begin{aligned} t_{j+1} - t_k &= (t_{j+1} - t_j) + (t_j - t_{j-1}) + \cdots + (t_{k+1} - t_k) \\ &\le \left(\left(\frac{\delta}{2}\right)^j (2\,q_0)^{2^j - 1} + \left(\frac{\delta}{2}\right)^{j-1} (2\,q_0)^{2^{j-1}-1} + \cdots + \left(\frac{\delta}{2}\right)^k (2\,q_0)^{2^k - 1} \right) \eta \\ &\le \left(1 + (2\,q_0)^{2^k} + \left((2\,q_0)^{2^k} \right)^2 + \cdots \right) \left(\frac{\delta}{2}\right)^k (2\,q_0)^{2^k-1}\ \eta \\ &= \left(\frac{\delta}{2}\right)^k \frac{(2\,q_0)^{2^k - 1}\ \eta}{1 - (2\,q_0)^{2^k}}. \end{aligned} \tag{A.4}$$

Estimate (2.9) follows from (A.4) by letting $j \longrightarrow \infty$.

That completes the proof of Lemma 2.1. $\square$

References

1. I. K. Argyros, *On the Newton-Kantorovich hypothesis for solving equations*, J. Comput. Appl. Math., **169** (2004), 315–332. MR2072881 (2005c:65047)
2. I. K. Argyros, *A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space*, J. Math. Anal. Appl., **298** (2004), 374–397. MR2086964
3. I. K. Argyros, *Convergence and applications of Newton-type iterations*, Springer-Verlag, 2008, New York. MR2428779
4. A. Ben-Israel, Y. Levin, *Maple programs for directional Newton methods*, are available at ftp://rutcor.rutgers.edu/pub/bisrael/Newton–Dir.mws.
5. Y. Levin, A. Ben-Israel, *Directional Newton methods in n variables*, Mathematics of Computation, A.M.S., **71** (2002), 251–262. MR1862998 (2002h:65083)
6. G. Lukács, *The generalized inverse matrix and the surface-surface intersection problem.* Theory and practice of geometric modeling (Blaubeuren, 1988), 167–185, Springer, Berlin, 1989. MR1042329 (91f:65041)
7. J. M. Ortega, W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970. MR0273810 (42:8686)
8. V. Pereyra, *Iterative methods for solving nonlinear least square problems*, SIAM J. Numer. Anal., **4** (1967), 27–36. MR0216732 (35:7561)
9. B. T. Polyak, *Introduction to optimization.* Translated from the Russian. With a foreword by D.P. Bertsekas. Translations Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York, 1987. MR1099605 (92b:49001)
10. A. Ralston, P. Rabinowitz, *A first course in numerical analysis*, 2nd Edition, McGraw-Hill, 1978. MR0494814 (58:13599)
11. J. Stoer, K. Bulirsch, *Introduction to numerical analysis*, Springer-Verlag, 1980. MR0557543 (83d:65002)
12. H. F. Walker, L. T. Watson, *Least-change secant update methods*, SIAM J. Numer. Anal., **27** (1990), 1227–1262. MR1061128 (91g:65121)

Department of Mathematics Sciences, Cameron University, Lawton, Oklahoma 73505
E-mail address: `iargyros@cameron.edu`