BULL. AUSTRAL. MATH. SOC. 65J15,65B05,65L50,65M50
VOL. 37 (1988) [131–147]

NEWTON-LIKE METHODS UNDER MILD DIFFERENTIABILITY CONDITIONS WITH ERROR ANALYSIS

IOANNIS K. ARGYROS

We apply Newton-like methods to operator equations where the operator has Hölder continuous derivatives. Our results reduce to the ones obtained by Rockne when the ordinary Newton method is applied to find solutions of nonlinear operator equations.

The results are applied to a second order differential equation.

INTRODUCTION

Consider an equation

$$F(x) = 0 \tag{1}$$

where F is a nonlinear operator between two Banach spaces $E, \hat{E}$. A Newton-like method can be defined as any iterative method of the form

$$x_{n+1} = x_n - L_n^{-1}F(x),\ n = 0,1,2,\dots;\ x_0 \text{ pre-chosen} \tag{2}$$

for generating approximate solutions to (1). The $\{L_n\}$ denotes a sequence of invertible linear operators. This is plainly too general and what is really implicit in the title is that L_n should be a conscious approximation to $F'(x_n)$, since when $L_n = F'(x_n)$, the method reduces to the Newton-Kantorovich method. The convergence of (2) to a solution of (1) has been described already in [**2**], [**3**], [**6**] and the references there. The basic assumption made is that F is twice Fréchet-differentiable in some ball around the initial iterate. We relax this requirement to operators that are only once Fréchet-differentiable. An error analysis is also provided.

Our results can be compared with the ones obtained in [**2**], [**4**] and [**8**] when $L_n = F'(x_n)$, $n = 0,1,2,\dots$. But, even then they are proved to be stronger.

Received 9 April 1987
The author wishes to express his gratitude to Professor Florian Potra for his constructive suggestions for the last part of this paper.

1 Preliminaries

From now on we assume that F is once Fréchet-differentiable at a point $x \in E$ and note that $F'(x) \in L\left(E, \hat{E}\right)$, the space of bounded linear operators from E to $\hat{E}$.

DEFINITION 1: We say that the Fréchet-derivative $F'(x)$ *is Hölder continuous* over a domain D if for some $c > 0$, $p \in [0,1]$.

$$\|F'(x) - F'(y)\| \leqslant c\|x - y\|^p, \text{ for all } x, y \in D. \tag{3}$$

We then say that $F'(\bullet) \in H_D(c,p)$.

DEFINITION 2: Let t_0 and t' be non-negative real numbers and let g be a continuously differentiable real function on $[t_0, t_0 + t']$ and P be a continuously Fréchet-differentiable operator on

$$\overline{U}(x_0, t') = \{x \in E \mid \|x - x_0\| \leqslant t'\} \subset E$$

into $\hat{E}$. Then the equation

$$t = g(t)$$

will be said to *majorise* the equation

$$x = P(x) \text{ on } U(x_0, t')$$

if

$$\|P(x_0) - x_0\| \leqslant g(t_0) - t_0$$

and

$$\|P'(x)\| \leqslant g'(t) \text{ for } \|x - x_0\| \leqslant t - t_0 < t'.$$

We will need the following results, whose proofs can be found in [**2**] and [**8**] respectively.

LEMMA 1. *Let $\{x_k\}$, $k = 0,1,2,\ldots$ be a sequence in E and $\{t_k\}$, $k = 0,1,2,\ldots$ a sequence of non-negative real numbers such that*

$$\|x_{k+1} - x_k\| \leqslant t_{k+1} - t_k, \; k = 0,1,2,\ldots$$

and

$$t_k \to t^\star < \infty \text{ as } k \to \infty.$$

Then there exists a point $x^\star \in E$ such that

$$x_k \to x^\star \text{ as } k \to \infty$$

and

$$\|x^* - x_k\| \leqslant t^* - t_k, \; k = 0, 1, 2, \ldots.$$

LEMMA 2. *Let* $F : E \to E$ *and* $D \subseteq E$. *Assume* D *is open and that* $F'(\bullet)$ *exists for every* $x \in D$. *Let* D_0 *be a convex set with* $D_0 \subseteq D$ *such that* $F'(\bullet) \in H_{D_0}(c, p)$. *Then*

$$\|F(x) - F(y) - F'(x)(x - y)\| \leqslant \frac{c}{1+p}\|x - y\|^{p+1} \text{ for all } x,\; y \in D_0.$$

2 MAIN CONVERGENCE RESULTS

We can now prove the following:

PROPOSITION 1. *Let* $F'(\bullet) \in H_{D_0}(c, p)$, *where* D_0 *is the closure of an open convex set and* $D_0 \subset D$. *Assume that for every* n *with* $\{x_k\} \subset D_0$, $k = 0, 1, 2, \ldots, n$, *there exists an invertible operator* $L_n \in \left(E, \hat{E}\right)$ *and a positive real number* d_n *such that:*

$$(4) \qquad \|L_n^{-1}\| \leqslant d_n^{-1}.$$

For a *and* $b > 0$, *both independent of* n *with*

$$a \geqslant \frac{2}{p(p+1)} \text{ if } p \neq 0,$$

and

$$a \geqslant 1 \text{ if } p = 0,$$

assume:

$$(5) \qquad \|F'(x_n) - L_n\| \leqslant d_n + ap\left(\sum_{j=1}^{n}\|x_j - x_{j-1}\|\right)^p - b, \quad n = 0, 1, 2, \ldots$$

Set

$$(6) \qquad f(t) = \frac{ca}{p+1}t^{p+1} - bt + d_0\|L_0^{-1}F(x_0)\|, \quad t \in [0, \infty)$$

and

$$(7) \qquad t_{n+1} = t_n + \frac{f(t_n)}{d_n}; \quad t_0 = 0.$$

Then if $\{x_n\} \subset D_0$, *(7) majorizes iteration (2).*

PROOF: We will use induction on n and definitions 1 and 2. Note:

$$\|x_1 - x_0\| = \|L_0^{-1}F(x_0)\| = t_1 - t_0$$

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

and assume that:

$$\{x_k\} \subset D_0,\ k = 0, 1, 2, \dots, n$$

and

$$\|x_j - x_{j-1}\| \leqslant t_j - t_{j-1} \text{ for } j = 1, \dots, n.$$

The iterate x_{n+1} is well defined since $F(x_n)$ and L_n^{-1} are. We will use the obvious estimate

$$\sum_{j=1}^{n} \|x_j - x_{j-1}\| \leqslant t_n$$

to compute

$$\begin{aligned}(8)\qquad \|x_{n+1} - x_n\| &\leqslant \|L_n^{-1}\|\ \|F(x_n)\| \\ &\leqslant d_n^{-1}[\|F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1})\| \\ &\qquad + \|L_{n-1} - F'(x_{n-1})\|\|x_n - x_{n-1}\|] \\ &\leqslant d_n^{-1}[\frac{c}{p+1}\|x_n - x_{n-1}\|^{p+1} + \big(d_{n-1} + act_{n-1}^{p} - b\big)\|x_n - x_{n-1}\|] \\ &\leqslant d_n^{-1}[\frac{c}{p+1}(t_n - t_{n-1})^{p+1} + \big(act_{n-1}^{p} - b\big)(t_n - t_{n-1}) + d_{n-1}(t - t_{n-1})] \\ &\leqslant d_n^{-1}[\frac{1}{2}f''(t_{n-1})(t_n - t_{n-1})^2 + f'(t_{n-1})(t_n - t_{n-1}) + f(t_{n-1})].\end{aligned}$$

But,

$$\begin{aligned}f(t_n) &= f(t_{n-1}) + f'(t_{n-1})(t_n - t_{n-1}) \\ &\qquad + \frac{1}{2}f''(t_{n-1})(t_n - t_{n-1})^2 + \frac{1}{6}f'''(\tilde{t}_n)(t_n - t_{n-1})^3 \\ &\geqslant f(t_{n-1}) + f'(t_{n-1})(t_n - t_{n-1}) + \frac{1}{2}f''(t_{n-1})(t_n - t_{n-1})^2\end{aligned}$$

since

$$f'''(\tilde{t}) = p(p-1)ca(\tilde{t}_n)^{p-2} \geqslant 0 \quad \text{for some } \tilde{t}_n \in [t_{n-1}, t_n] \subset [0, \infty)$$

and

$$t_n \geqslant t_{n-1}$$

by the induction hypothesis.

Therefore (8) becomes

$$\|x_{n-1} - x_n\| \leqslant d_n^{-1} f(t_n) = t_{n+1} - t_n$$

and the induction is completed. ∎

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

PROPOSITION 2. *Let $F'(\bullet) \in H_{D_0}(c,p)$ where D_0 is the closure of an open convex set and $D_0 \subset D$. Assume:*

(1) *Inequality (5) holds for $n = 0$;*

(2) $\dfrac{\|F'(x_0) - L_0\|}{a_0} \leqslant \delta^1 < 1$; and

(3) *the function $\bar{f}(t)$ defined by*

$$\bar{f}(t) = \frac{c}{p+1} t^{p+1} + (\delta^1 - 1)d_0 t + d_0 \|L_0^{-1} F(x_0)\|, \ t \in [0, \infty) \tag{9}$$

has a minimum positive zero r_0' such that $U(x_0, r_0') \subset D_0$.

Then (1) has a unique solution $x^* \in \overline{U}(x_0, r_0')$. If r_0' is the unique fixed point of the function

$$g(t) = t + \frac{\bar{f}(t)}{d_0}$$

on some interval $[r_0', r_1']$, $r_0' \leqslant r_1'$, then x^* is also unique in $D_0 \cap U(x_0, r_1')$.

Moreover

(1) *The iteration*

$$x_{n+1}' = x_n' - L_0^{-1} F(x_n');$$

converges to x^* for $\|x_0' - x_0\| < r_2 \leqslant r_1'$ and $U(x_0, r_2) \subset D_0$.

(2) *The following estimate is true:*

$$\|x_n' - x^*\| \leqslant |r_0' - t_n'|$$

where $\{t_n'\}$ is generated by

$$t_{n+1}' = t_n' + \frac{\bar{f}(t_n')}{d_0}.$$

PROOF: Define the nonlinear operator P on D_0 by

$$P(x) = x - L_0^{-1} F(x).$$

We will show that if $t' \in [r_0', r_1')$, then $g(t)$ majorises $P(x)$ on $\overline{U}(x_0, t') \cap D_0$.

We have

$$\|P(x_0) - x_0\| = \|L_0^{-1} F(x_0)\| = g(0) - 0.$$

Let x, t be such that $x \in \overline{U}(x_0, t') \cap D_0$ and $\|x - x_0\| \leqslant t < t'$. Then

$$\begin{aligned}
\|P'(x)\| = \|I - L_0^{-1} P'\| &= \|L_0^{-1}((L_0 - F'(x_0)) + (F'(x_0) - F'(x)))\| \\
&\leqslant \|L_0^{-1}\|(\|F'(x) - F'(x_0)\| + \|F'(x_0) - L_0\|) \\
&\leqslant \delta^1 + c\frac{t^p}{d_0} = g'(t).
\end{aligned}$$

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

By hypothesis r_0' is the unique fixed point of $g(t)$ in $[0, t']$ and $g(t) \leqslant t'$ with equality holding if and only if $t' = r_0'$.

The results now follow from the well-known classical theorem on the existence and uniqueness of solutions of equation (1) via majorizing sequences given in Kantorovich ([**5**, p. 697]). ∎

We remark that if $\{t_n\}$ converges to $t^\star$, then $t^\star$ is the least upper bound for $\sum_{j=1}^{n} \|x_j - x_{j-1}\|$, independent of n. Therefore, if we assume that $U(x_0, t^\star) \subset D_0$, using Lemma 1 we obtain that $\{x_n\}$ exists and converges to a solution $x^\star$ of (1).

Usually we do not wish to calculate the derivative of each L_n but instead use L_n in place of $L_{n+1}, \ldots, l_{n+q}$ and then calculate L_{n+q+1} and use it for $\bar{q}$ calculations. That is why, as in [**5**], we find it useful to define a nondecreasing sequence of non-negative real numbers $\{e_n\}$ such that

$$e_0 = 0$$

and

$$e_n = e_{n-1} \text{ or } e_n = n.$$

We then replace (2) by the iteration

$$x_{n+1} = x_n - L_{e_n}^{-1} F(x), \; n = 0, 1, 2, \ldots. \tag{10}$$

We can now prove the basic result.

THEOREM 1. *Assume:*

(1) *The hypotheses of Proposition 1 hold;*
(2) *the sequence $\{d_n\}$ is uniformly bounded above and*
(3) *hypothesis (iii) of Proposition 2 is true.*

Then (10) converges to a solution $x^\star$ of (1) according to

$$\|x_{n+1} - x^\star\| \leqslant r_0 - t_n - d_{e_n}^{-1}(f(t_n)); \; t_0 = 0, 2, \ldots. \tag{11}$$

Moreover if

(4) *the hypothesis on r_1' in Proposition 2 holds then $x^\star$ is the unique solution of (1) in $U(x_0, r_1') \cap D_0$.*

PROOF: Let us define $C_n = L_{e_n}$ and $c_n = d_{e_n}$, $n = 0, 1, 2, \ldots$. The proof will be a consequence of the following steps.

Step 1. We will show that $\{x_n\} \subset U(x_0, r_0) \subset D_0$ and that the rest of the hypotheses of Proposition 1 hold.

We easily note:

(1) (4) holds for C_n and c_n, $n = 0, 1, 2, \ldots$;
(2) (5) holds by the choice of a and $d_n \leqslant n$.

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

We now estimate

$$\begin{aligned}\|C_n - F'(x_n)\| &= \|L_{e_n} - F'(x_n)\| \\ &= \|(L_{e_n} - F'(x_{e_n})) + F'(x_{e_n}) - F'(x_n)\| \\ &\leqslant \|L_{e_n} - F'(x_{e_n})\| + \|F'(x_{e_n}) - F'(x_n)\| \\ &\leqslant d_n + ap\left(\sum_{j=1}^{e_n} \|x_j - x_{j-1}\|\right)^p - b + c\|x_{e_n} - x_n\|^p \\ &\leqslant d_n + ap\left(\sum_{j=1}^{n} \|x_j - x_{j-1}\|\right)^p - b.\end{aligned}$$

Let f be defined by Proposition 1, r_0 be the smallest positive zero of f. By Proposition 1

$$\|x_{n+1} - x_n\| \leqslant g_n(t_n) - t_n,$$

where the function $g_n(t)$ is defined on $[0, \infty)$ by

$$g_n(t) = t + \frac{f(t)}{c_n},\ n = 0, 1, 2, \ldots.$$

Assume that

$$t_n < r_0.$$

Then via the mean value theorem we can find $z_n \in (t_n, r_0)$ such that

$$\begin{aligned}r_0 - t_{n+1} = g_n(r_0) - g_n(t_n) &= g_n'(z_n)(r_0 - t_n) \\ &= [1 + \frac{f'(z_n)}{c_n}](r_0 - t_n) \\ &= c_n^{-1}[c_n + caz_n^p - b](r_0 - t_n).\end{aligned}$$

Using (5) we easily get

$$0 \leqslant c_n^{-1}[c_n + caz_n^p - b](r_0 - t_n) < r_0 - t_{n+1} < c_n^{-1}[c_n + car_0^p - b](r_0 - t_n) \leqslant r_0 - t_n.$$

Therefore $\{t_n\}$ is bounded and convergent to some $t^\star \leqslant r_0$. The estimate,

$$0 = \lim_{n\to\infty} (t_{n+1} - t_n) = \lim_{n\to\infty} \frac{f(t_n)}{c_n} \geqslant \lim_{n\to\infty} \frac{f(t_n)}{e}$$

where e denotes the uniform upper bound on $\{d_n\}$, implies that

$$f(t^\star) = 0,$$

that is

$$t^\star = r_0$$

and (11) holds.

Step 2. We show that $x^* = \lim_{n\to\infty} x_n$ is a solution of F. We have

$$\begin{aligned}\|C_n\| &\leqslant \|F'(x_n)\| + c_n - b + ap\left(\sum_{j=1}^{n}\|x_j - x_{j-1}\|\right)^p \\ &\leqslant \|F'(x_0)\| + c\|x_0 - x_n\|^p + c_n - b + apr_0^p \\ &\leqslant \|F'(x_0)\| + (c+ap)r_0^p - b + e \equiv B.\end{aligned}$$

Therefore the inequality

$$\begin{aligned}\|F(x_n)\| &\leqslant \|C_n(x_{n+1} - x_n)\| \\ &\leqslant \|C_n\|\,\|x_{n+1} - x_n\| \\ &\leqslant B\|x_{n+1} - x_n\| \to 0 \text{ as } n \to \infty\end{aligned}$$

implies that $F(x^*) = 0$.

The uniqueness result will now hold if (ii) of Proposition 2 is satisfied and hypothesis (iii) and (iv) of the theorem hold.

For $n = 0$ in (5) we obtain

$$0 < b \leqslant d_0 - \|F'(x_0) - L_0\|$$

that is

$$0 < \frac{b}{d_0} < 1 - \delta^1,$$

so (ii) of Proposition 2 is satisfied. It can easily be checked that $r_0' \leqslant r_0 \leqslant r_1'$ and the proof of the theorem is now complete. ∎

We now state a theorem which seems to reduce to a minimum the assumptions necessary to apply the majorant technique.

THEOREM 2. *Let $F'(\bullet) \in H_{D_0}(c,p)$, where D_0 is the closure of an open convex set and $D_0 \subset D$. Assume:*

(1) *If $x_0 \in D$, let $L_0 \in L\left(E, \hat{E}\right)$ be an invertible operator with*

$$\|L_0^{-1}F(x_0)\| \leqslant \alpha$$

and

$$\|L_0^{-1}\| \leqslant \beta$$

(2) *there exist real numbers δ and γ such that if $\{x_n\} \subset D_0$, $k = 0, 1, 2, \ldots, n$ then*

$$\|L_{e_n} - F'(x_n)\| \leqslant \delta_n + \gamma\left(\sum_{j=1}^{n}\|x_j - x_{j-1}\|\right)^p;$$

$$\delta_n \leqslant \delta_0 = \delta,\ n = 0, 1, 2, \ldots,$$

(3) *the following estimate holds:*

$$3\beta\delta < 1,$$

(4) *there exists an interval* $[0, r_0]$ *such that for* $r \in [0, r_0]$

$$2\beta\delta + \beta(\gamma + c)r^p < 1,$$

and $U(x_0, r) \subset D_0$;

(5) *there exists a nonempty interval* $[r_3, r_4] = [0, \bar{r}_0] \cap [r_0, r_1']$ *where* r_0 *is the small positive solution of (6).*

Then the following are true:

(1) *for* $\bar{a}$, $\bar{b}$ *such that*

$$\bar{a} \geqslant \max\left(\frac{2\gamma + c}{p}, \frac{2}{p(p+1)}\right)$$

and

$$0 < \bar{b} \leqslant \frac{1 - 3\beta\delta}{\beta},$$

$$\delta_n + \gamma\left(\sum_{j=1}^{n} \|x_j - x_{j-1}\|\right)^p \leqslant \bar{d}_n + \bar{a}p\left(\sum_{j=1}^{n} \|x_j - x_{j-1}\|\right)^p - \bar{b} \tag{12}$$

where

$$\bar{d}_n^{-1} = \beta[1 - \beta(\delta + \delta_n) - \beta(\gamma + c)\left(\sum_{j=1}^{n} \|x_j - x_{j-1}\|\right)^p]^{-1},\ n = 0, 1, 2, \ldots$$

and $\bar{d}_n \leqslant d_0$.

(2) *the sequence* $\{x_n\}$ *given by (10) exists in* $U(x_0, r)$, $r_3 \leqslant r \leqslant r_4$ *and converges to a unique solution* $x^\star$ *of (1) in* $U(x_0, r_3)$. *Moreover, the solution* $x^\star$ *is unique in* $U(x_0, r_4)$;

(3) *the following estimate holds if* $t_1 = \alpha$

$$\|x_{n+1} - x^\star\| \leqslant \bar{r}_0 - t_n - d_{e_n}^{-1}(f(t_n))$$

where f *is given by (6) with*

$$a = \bar{a}, \quad b = \bar{b} \text{ and } d_n = \bar{d}_n,\ n = 0, 1, 2, \ldots$$

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

PROOF: As in (5), assume $\{x_k\} \subset U(x_0, \bar{r}_0)$, $k = 0,1,2,\ldots,n$ and $\sum_{j=1}^{n} \|x_j - x_{j-1}\| < r$ with $r_3 \leqslant r \leqslant r_4$.

We have

$$\begin{aligned} \|L_n - L_0\| &\leqslant \|L_n - F'(x_n)\| + \|F'(x_n) - F'(x_0)\| + \|F'(x_0) - L_n\| \\ &\leqslant \delta_n + \gamma\left(\sum_{j=1}^{n} \|x_j - x_{j-1}\|\right)^p + c\|x_n - x_0\|^p + \delta \\ &\leqslant \delta_0 + \delta_n + \gamma(r)^p + cr^p \\ &\leqslant 2\delta + (\gamma + c)r^p < \frac{1}{\beta}. \end{aligned}$$

Therefore,

$$\|L_0^{-1}L_n - I\| \leqslant \beta(\delta + \delta_n) + \beta(\gamma + c)\left(\sum_{j=1}^{n} \|x_j - x_{j-1}\|\right)^p.$$

The Banach Lemma [2] can now be used to show that L_n^{-1} exists and is bounded in norm by the $(\bar{d}_n)^{-1}$, $n = 0,1,2,\ldots$ Moreover $\{\bar{d}_n\}$ is uniformly bounded by $\bar{d}_0$.

It is now easy to check that (12) is satisfied by the choice of $\bar{d}_n$, $\bar{a}$ and $\bar{b}$.

The rest of the theorem now follows from Theorem 1. ∎

3 Error analysis and applications

When we solve equation (1) numerically using iteration (10) we generate instead of the sequence $\{x_n\}$ the perturbed sequence $\{z_n\}$ given by

$$z_{n+1} = z_n + [L_{e_n} + \bar{L}_{z_n}]^{-1}[F(z_n) + a_{z_n}] - q_{z_n},\ n = 0,1,2,\ldots,$$

assuming $z_0 = x_0$ and $[L_{e_n} + \bar{L}_{z_n}]^{-1}$ exists for $n = 0,1,2,\ldots$.

The problem of estimating the bound on $\|x_n - z_n\|$ when $L_{e_n} = F'(x_n)$ and under certain assumptions, basically on the norm of the linear operator L_{z_n} and on the norm of the elements of $\bar{L}_{z_n}$, a_{z_n} and q_{z_n}, has been solved in [8].

Here we can easily prove the analogue of Lemma 2 and Theorem 3 in [8] for the more general iteration (10). However, we leave that to the motivated reader and we show that the order of convergence of (2) when $L_n = F'(x_n)$ to a solution x^* of (1) is $1 + p$.

We then show that iteration (10) under appropriate choice of the L_{e_n}'s converges to x^* with order $1 + 2p$.

This improves the results in [8] where the order of convergence is not given. If the second Fréchet derivative of F is bounded and $p = 1$, our results coincide with those in [2].

We then compare the numerical efficiency of (10) with the iteration

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \tag{N}$$

using the example of Rockne given in [8].

PROPOSITION 3. *Let $L_{e_n} = F'(x_n)$ in (10).*

Then under the hypotheses of Theorem 2, the solution $x^\star$ of (1) obtained via iteration (10) is such that

$$\|x_{n+1} - x^\star\| \leqslant k\|x_n - x^\star\|^{1+p},\ n = 0, 1, 2, \dots \tag{13}$$

where

$$k = \frac{c\bar{d}_0}{(p+1)^2}.$$

PROOF: We have

$$\begin{aligned} x_{n+1} - x^\star &= x_n - x^\star - F'(x_n)F(x_n) \\ &= (F'(x_n))^{-1}[F'(x_n)(x_n - x^\star) - (F(x_n) - F(x^\star))]. \end{aligned}$$

By taking the norms in the above identity we obtain

$$\begin{aligned} \|x_{n+1} - x^\star\| &\leqslant \|F'(x_n)^{-1}\| \left\| \int_0^1 (F'(x_n) - F(x_n + t(x^\star - x_n)))dt \right\| \|x_n - x^\star\| \\ &\leqslant \frac{c\bar{d}_0}{p+1}\|x_n - x^\star\|^{p+1} \int_0^1 t^p dt \\ &\leqslant k\|x_n - x^\star\|^{p+1},\ n = 0, 1, 2, \dots \end{aligned}$$

∎

PROPOSITION 4. *Consider the iteration (10) for the solution (1) given in the form*

$$\begin{aligned} y_n &= x_n - F'(x_n)^{-1}F(x_n),\ n = 0, 1, 2, \dots \\ x_{n+1} &= y_n - F'(x_n)^{-1}F(y_n),\ n = 0, 1, 2, \dots \end{aligned} \tag{14}$$

with x_0 pre-chosen.

Then under the hypotheses of Theorem 2 the solution $x^\star$ of (1) obtained via iteration (14) is such that

$$\|x_{n+1} - x^\star\| \leqslant k_1\|x_n - x^\star\|^{1+2p},\ n = 0, 1, 2, \dots \tag{15}$$

where

$$k_1 = \frac{2^p \left(c\bar{d}_0\right)^2}{(p+1)^3}.$$

Proof: We have

$$\begin{aligned} x_{n+1} - x^\star &= y_n - x^\star - F'(x_n)^{-1} F(y_n) \\ &= F'(x_n)^{-1}[F'(x_n)(y_n - x^\star) - (F(y_n) - F(x^\star))] \end{aligned}$$

By taking the norms of the above identity we obtain

$$\begin{aligned} \|x_{n+1} - x^\star\| &\leqslant \|F'(x_n)^{-1}\| \left\| \int_0^1 (F'(x_n) - F'(x^\star + t(y_n - x^\star)))dt \right\| \cdot \|y_n - x^\star\| \\ &\leqslant \frac{c\bar{d}_0}{p+1} \|(x_n - x^\star) + t(y_n - x^\star)\|^p \|y_n - x^\star\| \\ &\leqslant \frac{c\bar{d}_0}{p+1} (\|x_n - x^\star\| + \|y_n - x^\star\|)^p \|y_n - x^\star\| \\ &\leqslant \frac{2^p c\bar{d}_0}{p+1} \|x_n - x^\star\|^p \|y_n - x^\star\| \text{ (since } \|y_n - x^\star\| \leqslant \|x_n - x^\star\|). \end{aligned} \tag{16}$$

Similarly,

$$\begin{aligned} y_n - x^\star &= x_n - x^\star - F'(x_n)^{-1} F(x_n) \\ &= F'(x_n)^{-1}[F'(x_n)(x_n - x^\star) - (F(x_n) - F(x^\star))]. \end{aligned}$$

Therefore,

$$\begin{aligned} \|y_n - x^\star\| &\leqslant \|F'(x_n)^{-1}\| \left\| \int_0^1 (F'(x_n) - F'(x^\star + t(x_n - x^\star)))dt \right\| \cdot \|x_n - x^\star\| \\ &\leqslant \frac{c\bar{d}_0}{p+1} \left\| \int_0^1 (1-t)^p dt \right\| \cdot \|x_n - x^\star\|^{p+1} \\ &\leqslant \frac{c\bar{d}_0}{(p+1)^2} \|x_n - x^\star\|^{p+1}. \end{aligned} \tag{17}$$

Finally, by (16) and (17), we obtain

$$\begin{aligned} \|x_{n+1} - x^\star\| &\leqslant \frac{2^p c\bar{d}_0}{p+1} \|x_n - x^\star\|^p \frac{c\bar{d}_0}{(p+1)^2} \|x_n - x^\star\| \\ &\leqslant k_1 \|x_n - x^\star\|^{1+2p}. \end{aligned}$$

∎

Example 1: Consider the function G defined on $[0, b]$ by

$$G(t) = At^{1+p} + Bt$$

where, A, $B \in \mathbf{R}$, $\bar{p} \in [0,1]$ and $b > 0$.

Let $\|\cdot\|$ denote the max norm on $\mathbf{R}$. Then

$$\|G'(t)\| = \max_{t\in[0,b]} \left|A(1+\bar{p})\bar{p}t^{\bar{p}-1}\right| = \infty,$$

which implies that Newton's method cannot be used to find a solution of the equation

$$G(t) = 0. \tag{18}$$

However, it can easily be seen that $G'(t)$ is Hölder continuous on $[0,b]$ with

$$c = A(1+\bar{p}) \qquad \text{and} \qquad p = \bar{p}.$$

Therefore, under the assumptions of Theorem 2, iteration (14) can be used to find a solution t^* of (18).

Definition 3: Define the *efficiency* E of an iteration $\{x_n\}$ for solving (1) in the sense of [5] by

$$E = \frac{\ln k}{T},$$

where k is the order of convergence of $\{x_n\}$ and T denotes the "time per step", that is, the number of function evaluations required to compute each iterate x_n for $n = 0, 1, 2, \ldots$.

Let E_1, E_2 denote the efficiencies of iterations (N) and (14) respectively. Take $p = \frac{1}{2}$. Then

$$E_1 = \frac{\ln\left(\frac{3}{2}\right)}{2} < E_2 = \frac{\ln 2}{3}.$$

A more interesting nontrivial application of Theorem 2 is given by the following example.

Example 2: Consider the differential equation

$$\begin{gathered} x'' + x^{1+p} = 0, \; p \in [0,1] \\ x(0) = x(1) = 0. \end{gathered} \tag{20}$$

We divide the interval $[0,1]$ into n subintervals and we set $h = \frac{1}{n}$. Let $\{v_k\}$ be the points of the subdivisions with

$$0 = v_0 < v_1 < \cdots < v_n = 1.$$

A standard approximation for the second derivative is given by

$$x_i'' = \frac{x_{i-1} - 2x_i + x_{i+1}}{h^2}, x_i = x(v_i), \quad i = 1, 2, \ldots, n-1.$$

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

Take $x_0 = x_n = 0$ and define the operator $F : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ by

$$F(x) = H(x) + h^2\varphi(x) \tag{21}$$

$$H = \begin{bmatrix} 2 & -1 & \dots & 0 \\ -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ 0 & \dots & -1 & 2 \end{bmatrix},$$

$$\varphi(x) = \begin{bmatrix} x_1^{1+p} \\ x_2^{1+p} \\ \vdots \\ x_{n-1}^{1+p} \end{bmatrix},$$

and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_{n-1} \end{bmatrix}.$$

Then

$$F'(x) = H + h^2(p+1)\begin{bmatrix} x_1^p & & & 0 \\ & x_2^p & & \\ & & \ddots & \\ 0 & & & x_{n-1}^p \end{bmatrix}.$$

Newton's method cannot be applied to the equation

$$F(x) = 0. \tag{22}$$

We may not be able to evaluate the second Fréchet-derivative since it would involve the evaluation of quantities of the form x_i^{-p} and they may not exist.

Let $x \in \mathbb{R}^{n-1}$, $H \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ and define the norms of x and H by

$$\|x\| = \max_{1 \leqslant j \leqslant n-1} |x_j|$$

$$\|H\| = \max_{1 \leqslant j \leqslant n-1} \sum_{k=1}^{n-1} |h_{jk}|.$$

For all x, $z \in \mathbb{R}^{n-1}$ for which $|x_i| > 0$, $|z_i| > 0$, $i = 1, 2, \dots, n-1$ we obtain

$$\begin{aligned} \|F'(x) - F'(z)\| &= \|\operatorname{diag}\{(1+p)h^2(x_j^p - z_j^p)\}\| \\ &= (1+p)h^2 \max_{1 \leqslant j \leqslant n-1} |x_j^p - z_j^p| \leqslant (1+p)h^2[\max|x_j - z_j|]^p \\ &= (1+p)h^2\|x - z\|^p. \end{aligned}$$

https://doi.org/10.1017/S0004972700004238 Published online by Cambridge University Press

For $z_0 \in \mathbf{R}^{n-1}$ iteration (14) can be written as a system of linear equations

$$\begin{aligned} F'(z_n)(z_n - \bar{z}_n) &= F(z_n) \\ F'(z_n)(\bar{z}_n - z_{n+1}) &= F(\bar{z}_n) \end{aligned} \quad n = 0, 1, 2, \dots .$$

By (15), since $p = \frac{1}{2}$ iteration (14) will converge under the assumptions of Theorem 2 to a solution x^* of equation (22) with the order of convergence being 2.

The order of convergence of (N), which was used in [8] to solve the same problem, is $\frac{3}{2}$.

Let $\tilde{E}_1$, $\tilde{E}_2$ denote the efficiencies of iterations (N) and (14) respectively, then it easily follows from the discussion made after the definition that

$$\tilde{E}_1 < \tilde{E}_2.$$

To show further the advantages of (14) when compared to (N), set, as in [8], $n = 10$ and choose the initial approximation to be $130 \sin \pi x$. We then get

$$z_0 = \begin{bmatrix} 4.01524E+01 \\ 7.63785E+01 \\ 1.05135E+02 \\ 1.23611E+02 \\ 1.29999E+02 \\ 1.23675E+02 \\ 1.05257E+02 \\ 7.65462E+01 \\ 4.03495E+01 \end{bmatrix}$$

After 2 iterations we obtain

$$z_2 = \begin{bmatrix} 3.35741E+01 \\ 6.52027E+01 \\ 9.15665E+01 \\ 1.09168E+02 \\ 1.15364E+02 \\ 1.09167E+02 \\ 9.15665E+01 \\ 6.52027E+01 \\ 3.35742E+01 \end{bmatrix}$$

We choose z_2 as our x_0 for Theorem 2. Since $L_{e_n} = F'(x_n)$, with $e_n = n - 1$, conditions (ii) and (iii) in Theorem 2 are satisfied for

$$\delta_n = \delta_0 = \gamma = \delta = 0, \; n = 0, 1, 2, \dots .$$

We also have,

$$\alpha = \|L_0^{-1}F(x_0)\| = \|F'(x_0)^{-1}F(x_0)\| = 9.15312E-05$$
$$\beta = \|L_0^{-1}\| = \|F'(x_0)^{-1}\| = 2.55883E+01$$
$$c = (p+1)h^2 = \frac{3}{2}h^2 = 0.015$$

Condition (iv) will be satisfied if we choose $r > 0$ such that

$$\beta(\gamma + c)r^p < 1$$

or

$$0 < r < 6.7879398 \equiv \bar{r}_0.$$

Moreover, equation (6) for

$$\bar{a} = \max\left(\frac{2\gamma + c}{p}, \frac{2}{p(p+1)}\right) = 2.666666$$
$$\bar{d}_0 = \bar{b} = \frac{1 - 3\beta\delta}{\beta} = \frac{1}{\beta} = 3.90803E+02$$

becomes

$$(2.666666E-02)t^{\frac{3}{2}} - (3.90808E-02)t + 3.577066E-06 = 0$$

with solution

$$r_0 = 9.21E-05 = r_3.$$

Hence, by Theorem 2, the iteration (14) remains in $U(x_0, \bar{r}_0)$ and converges quadratically to a solution $x^\star$ of equation (22).

Finally, note that x_0 was found using two iterations in (14) instead of four that were required using (N). Moreover, the ball S used in Rockne [8] is such that $U(x_0, r_3) \subset S$.

References

[1] I.K. Argyros, 'On the approximation of some nonlinear equations', *Aequationes Math.* **32** (1987), 87–95.

[2] J.E. Dennis, 'Towards a unified convergence theory for Newton-like methods', in *Nonlinear Function Analysis and Applications*, L.B. Rall ed. (Academic Press, 1970).

[3] L.V. Kantorovich and G.P. Akilov, *Functional Analysis in Normed Spaces* (Pergamon Press, Oxford, 1964).

[4] P. Lancaster, 'Error analyiss for the Newton-Raphson method', *Numer. Math.* **9** (1968), 55–68.

[5] F.A. Potra and V. Ptak, 'Nondiscrete induction and iterative processes' (Pitman Publ).

[6] W.C. Rheinboldt, 'A unified convergence theory for a class of iterative processes', *SIAM J. Numer. Anal.* 5 **1** (1968), 371–391.

[7] W.C. Rheinboldt, *Numerical Analysis of Parametrized Nonlinear Equations* (John Wiley, 1986).

[8] J. Rockne, 'Newton's method under mild differentiability conditions with error analysis', *Numer. Math.* **18** (1972), 401–412.

Department of Mathematics
New Mexico State University
Las Cruces, NM 88003
United States of America