notes the first zero of the derivative of $J_s(x)$, $N_s(x)$ is negative and increasing. Hence for $j'/v \geq x \geq 1$, a simple bound for $N_s(vx)$ is

$$\left| N_s(vx) \right| \leq \left| N_s(v) \right| \leq 1.73 \left| J_s(v) \right| .$$

References


University of Michigan

NEWTON'S METHOD IN BANACH SPACES

ROBERT G. BARTLE

In this note we show that if $f$ is a mapping between Banach spaces which is in class $C'$ in the sense of Hildebrandt and Graves [4], then the equation $f(x) = 0$ may be solved by an iterative process:

$$x_{n+1} = x_n - \left[ f'(x_n) \right]^{-1} f(x_n), \quad n = 0, 1, 2, \ldots,$$

provided that the initial guess $x_0$ and the arbitrarily selected points $z_n$ are sufficiently close to the solution desired. Here the derivative is taken in the sense of Fréchet. If we let $z_n = x_n$, $n = 0, 1, 2, \ldots$, we obtain the usual Newton process; if $z_n = x_0$, $n = 0, 1, 2, \ldots$, we obtain what is sometimes called the modified Newton process. Naturally, in any application, the computer would determine the $z_n$ so as to minimize effort.

This result is closely related to recent theorems of Kantorovič [5; 6; 7] and Mysovskih [8; 9], although these authors assumed the existence and boundedness of the second Fréchet derivative of $f$. In turn for this assumption, they were able to establish more rapid convergence. Under the assumption of analyticity, Stein [10] eliminated explicit mention of the second derivative, which is desirable

Presented to the Society, December 29, 1954; received by the editors October 21, 1954 and, in revised form, December 9, 1954.

1 This paper was written while the author was on Contract nonr 609(04) with the Office of Naval Research.

2 Numbers in brackets refer to the list of references at the end.
since it plays no role in the iteration. Fenyo [1] treated the Newton process in the case when the first derivative satisfies a Lipschitz condition at $x_0$, and the modified process without this restriction.

Let $\mathcal{X}$ and $\mathcal{Y}$ be (real or complex) Banach spaces and let $f$ map an open set of $\mathcal{X}$ into $\mathcal{Y}$. We recall that the Fréchet derivative $f'(x_0)$ of $f$ at a point $x_0$ in the domain of $f$ is the bounded linear operator of $\mathcal{X}$ into $\mathcal{Y}$ such that

$$\lim_{|h| \to 0} \frac{|f(x_0 + h) - f(x_0) - f'(x_0)(h)|}{|h|} = 0,$$

provided such an operator exists. If $\mathcal{G}$ is an open set in $\mathcal{X}$, we say, following Hildebrandt and Graves [4], that $f$ is in class $C'(\mathcal{G})$ if $f'(x)$ exists for $x \in \mathcal{G}$, and if the mapping $x \mapsto f'(x)$ is a continuous map of $\mathcal{G}$ into the space $\mathfrak{B}(\mathcal{X}, \mathcal{Y})$ of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ equipped with the uniform operator topology.

If $f$ is in class $C'(\mathcal{G})$, let $\delta(x_0, \epsilon)$ denote the modulus of continuity of $f'$ at $x_0$. This means that if $|x - x_0| \leq \delta(x_0, \epsilon)$ then $|f'(x) - f'(x_0)| \leq \epsilon$. Further, let $\mathcal{S}(x_0, \alpha)$ denote the sphere $\{x \in \mathfrak{X}: |x - x_0| < \alpha\}$.

**Lemma 1.** If $f$ is in class $C'(\mathcal{S}(x_0, \alpha))$ and if $x_1$ and $x_2$ are such that $|x_i - x_0| \leq \delta(x_0, \epsilon)$, then

$$|f(x_1) - f(x_2) - f'(x_0)(x_1 - x_2)| \leq \epsilon |x_1 - x_2|.$$

**Proof.** It follows from the general form of Taylor’s theorem, due to Graves [2, p. 173], that

$$f(x_1) - f(x_2) = \int_0^1 f'(tx_1 + (1-t)x_2)(x_1 - x_2) dt,$$

from which the statement follows.

**Lemma 2.** Let $f$ be in class $C'(\mathcal{S}(x_0, \alpha))$ and suppose that $f'(x_0)$ has a bounded inverse. For any $\lambda > |[f'(x_0)]^{-1}|$, there is a number $\beta \leq \min \{1, \alpha\}$ such that

(1) if $|x - x_0| \leq \beta$, then the operator $f'(x)$ has an inverse and $|[f'(x)]^{-1}| < \lambda$;

(2) if $|x_i - x_0| \leq \beta$, $i = 1, 2, 3$, then

$$|f(x_1) - f(x_2) - f'(x_3)(x_1 - x_2)| \leq (1/2\lambda) |x_1 - x_2|.$$

**Proof.** Choose $\delta_1$ such that if $|x - x_0| \leq \delta_1$, then $|f'(x) - f'(x_0)| \leq (4\lambda)^{-1}$. This implies that if $|x - x_0| \leq \delta_1$, then $f'(x)$ has an inverse.

---

8 Since no confusion can result, we use a single vertical bar to denote the norm of an element of the spaces $\mathcal{X}$ and $\mathcal{Y}$ and the norm of an operator between these two spaces.
operator. Since the function \(|T^{-1}|\) is an upper semi-continuous function for \(T\) in the open set of invertible operators in \(\mathcal{B}(X, Y)\) (see [3, p. 112]), there is an \(\alpha_1 \leq \min \\{\alpha, \delta\}\) such that if \(|x-x_0| \leq \alpha_1\), then \(|[f'(x)]^{-1}| < \lambda\). This proves (1). If \(\beta = \min \{1, \alpha_1, \delta(x_0, 1/4\lambda)\}\), it follows from Lemma 1 and the above that if \(|x_i-x_0| \leq \beta\), then (2) holds.

**Theorem.** Let \(f: \mathcal{S}(x_0, \alpha) \to Y\) be in class \(C'(\mathcal{S}(x_0, \alpha))\), and suppose that \(f'(x_0)\) has an inverse with \(|[f'(x_0)]^{-1}| < \lambda < \infty\). Let \(|f(x_0)| < \beta/2\lambda\), where \(\beta\) is as in Lemma 2, and let \(x_n, n=0, 1, 2, \cdots\), be completely arbitrary points with \(|x_n-x_0| \leq \beta\). Then the sequence \(\{x_n\}\) obtained by the iterative process

\[
x_{n+1} = x_n - [f'(x_n)]^{-1}f(x_n), \quad n = 0, 1, 2, \cdots,
\]

converges to a solution \(x\) of the equation \(f(x) = 0\). Further, \(|x-x_0| \leq \beta\) and is the only solution of the equation in this neighborhood of \(x_0\). The rapidity of the convergence is given by \(|x_n-x| < 2^{-n}\beta, n=0, 1, 2, \cdots\).

**Proof.** By definition, \(|x_1-x_0| \leq \lambda|f(x_0)| < \beta/2\). Further

\[
f(x_1) = f(x_0) - f'(x_0)(x_1 - x_0),
\]

and by (**) of Lemma 2, we have that

\[
|f(x_1)| \leq (1/2\lambda) |x_1 - x_0|.
\]

By induction, suppose that \(x_1, \cdots, x_n\) have been chosen such that for \(i=1, \cdots, n\) we have

(a) \(|x_i - x_0| < \beta,\)

(b) \(|x_i - x_{i-1}| \leq \lambda|f(x_{i-1})|,\)

(c) \(|f(x_i)| \leq (1/2\lambda) |x_i - x_{i-1}|.\)

Then, since \(x_{n+1} = x_n - [f'(x_n)]^{-1}f(x_n)\), it follows that

\[
|x_{n+1} - x_n| \leq \lambda|f(x_n)|,
\]

which is (b_{n+1}). Thus \(|x_{n+1} - x_n| < (1/2) |x_n - x_{n-1}|\). Iterating (b) and (c), we conclude readily that

\[
|x_{n+1} - x_0| \leq \left\{ \sum_{i=0}^{n} 2^{-i} \right\} |x_1 - x_0| < \left\{ 1 - 2^{-(n+1)} \right\} \beta.
\]

This proves (a_{n+1}). Since

\[
f(x_{n+1}) = f(x_n) - f'(x_n)(x_{n+1} - x_n),
\]
the validity of \( (a_{n+1}) \) permits the use of \( (**) \) to conclude that
\[
| f(x_{n+1}) | \leq (1/2\lambda) \left| x_{n+1} - x_n \right|,
\]
which is \( (c_{n+1}) \). Thus the inductive steps may be continued. Further, for any integers \( n \) and \( p \),
\[
\left| x_{n+p} - x_n \right| \leq \sum_{i=1}^{p} \left| x_{n+i} - x_{n+i-1} \right|
\]
\[(\dagger)\]
\[
\leq \lambda \left| f(x_0) \right| 2^{-n} \left\{ \sum_{i=0}^{p-1} 2^{-i} \right\} < 2^{-n}\beta,
\]
and hence \( \{x_n\} \) is a Cauchy sequence converging to an element \( \hat{x} \in \mathbb{X} \). In view of \( (a_i) \), we have \( |\hat{x} - x_0| \leq \beta \). From \( (c_i) \) it follows that \( f(\hat{x}) = 0 \). To see that \( \hat{x} \) is the unique solution of \( f(x) = 0 \) in this neighborhood, let \( \hat{x} \) be another solution with \( |\hat{x} - x_0| \leq \beta \). We have
\[
|\hat{x} - \hat{x}| = |f'(x_0)^{-1}f'(x_0)(\hat{x} - \hat{x})| \leq \lambda \left| f'(x_0)(\hat{x} - \hat{x}) \right|.
\]
By \( (*) \) we have \( |f'(x_0)(\hat{x} - \hat{x})| \leq (1/2\lambda)|\hat{x} - \hat{x}| \), and hence \( |\hat{x} - \hat{x}| \leq (1/2)|\hat{x} - \hat{x}| \) which is a contradiction unless \( \hat{x} = \hat{x} \). Finally the inequality concerning the speed of convergence follows from \( (\dagger) \) upon letting \( p \) approach infinity.

We wish to point out to the reader the similarity between what we have done and Theorems 1 and 2 of Graves [3].

Since the modified Newton process (i.e., when \( z_n = x_0 \)) avoids the necessity of computing a new inverse at each step, it appears to be particularly convenient. By analysing the proof, we may see that Lemma 2 is never used unless the point \( z_n \) differs from \( x_0 \), so in the modified process we require only the first lemma. This permits better estimates:

**Corollary.** In the modified Newton process where \( z_n = x_0, n = 0, 1, 2, \cdots \), the number \( \beta \) may be chosen to be \( \min \{ 1, \alpha, \delta(x_0, 1/2\lambda) \} \).

**Remark.** Since the calculation of the inverse operators \( [f'(z_n)]^{-1} \) is generally difficult, it is worth observing that the above proof applies directly to assure the convergence of an iteration
\[
x_{n+1} = x_n - T_n^{-1} f(x_n), \quad n = 0, 1, 2, \cdots,
\]
for any sequence \( \{T_n\} \) of bounded operators such that
\[
|T_n - f'(x_0)| < 1/4\lambda, \quad |T_n^{-1}| < \lambda.
\]
For application of the Newton process, we refer the reader to Kantorovič [6; 7].

References


