Časopis pro pěstování matematiky

Florian-Alexandru Potra; Vlastimil Pták

Nondiscrete induction and an inversion-free modification of Newton's method

Časopis pro pěstování matematiky, Vol. 108 (1983), No. 4, 333--341

Persistent URL: `http://dml.cz/dmlcz/118177`

ČASOPIS PRO PĚSTOVÁNÍ MATEMATIKY

Vydává Matematický ústav ČSAV, Praha

SVAZEK 108 * PRAHA 30. 11. 1983 * ČÍSLO 4

NONDISCRETE INDUCTION AND AN INVERSION-FREE MODIFICATION OF NEWTON'S METHOD

FLORIAN A. POTRA, Bucuresti, VLASTIMIL PTÁK, Praha

(Received August 3, 1982)

1. INTRODUCTION

The central notion of the method of nondiscrete induction is that of a rate of convergence. It is mainly this notion which makes it possible to establish, for a number of important iterative processes of numerical mathematics, error estimates sharp for the whole length of the process, not only asymptotically.

In a series of papers [4], [6] – [29], [31] devoted to the development of the method of nondiscrete induction, a number of rates of convergence have been described and discussed. In all these examples the corresponding rates of convergence were expressed in a natural manner as functions of the distances between the successive approximations.

It is the purpose of the present note to show that it is convenient to extend the notion of rate of convergence by admitting functions depending on parameters whose meaning is not necessarily that of a distance.

We intend to apply the method of nondiscrete mathematical induction to an iterative process of Newton type, the inverse of the derivative being replaced by an approximative inverse. This method was studied before by several authors [1], [2], [3], [5], [6], [30]. It is the purpose of the present note to establish the natural rate of convergence of this process; in the process of doing so we also obtain a very simple proof of the main theorem.

The "inversion – free" modification of the Newton method is based on the following observation. Suppose y is an invertible linear operator on a Banach space and suppose g is an approximate left inverse for y such that the error

$$1 - gy$$

is less than one in norm. Then g is invertible as well and we may write y in the form

$$y = g^{-1}(1 - (1 - gy)) \quad \text{whence} \quad y^{-1} = \left(\sum_{0}^{\infty} (1 - gy)^n\right) g \, .$$

If we take just the first two terms of this series we obtain an element

$$g' = g + (1 - gy)\,g$$

which is a better approximation of y^{-1}. Indeed,

$$1 - g'y = (1 - gy)^2\,.$$

This leads to the consideration of the following iterative process for a Fréchet differentiable mapping f from a Banach space X into a Banach space Y.

Given an approximation x and an approximate inverse G for $f'(x)$ we first construct G' by the formula

(1) $$G' = (2 - G f'(x))\,G$$

and then define the new approximation x' like in the Newton process by setting

(2) $$x' = x - G' f(x)$$

It will be convenient, for further reference, to collect the formulae in a lemma.

1.1. *Let X and Y be two Banach spaces.*

$$D, D' \in D(X, Y)\,, \quad G \in B(Y, X)\,.$$

If we set

$$G' = G + (1 - GD)\,G$$

then

$$1 - G'D = (1 - GD)^2$$

$$1 - G'D' = (1 - GD)^2 + G'(D - D')\,.$$

Proof. Immediate verification.

This lemma will be used in the following situation: we shall consider a Fréchet differentiable mapping f defined on an open subset U of X and taking its values in Y. Then D will be the derivative at a point x, $D = f'(x)$ and $D' = f'(x')$ where x' is defined as

$$x' = x - G' f(x)\,.$$

2. RATES OF CONVERGENCE

Let T be either the set of all positive numbers or a half-open interval of the form $\{t;\ 0 < t \leqq t_0\}$ for some t_0. Let E be an arbitrary set. A mapping w of $T \times E$ into itself will be called a rate of convergence on $T \times E$ with respect to the first component if the series

$$\sum_0^\infty P \circ w^{(n)} = s(t, e)$$

is convergent on the whole set $T \times E$. Here P stands for the projection on the first component. The sum $s(t, e)$ will be called the estimate function corresponding to w. It satisfies the following functional equation

$$s = t + s \circ w$$

on the whole set $T \times E$.

It is easy to state and prove the corresponding induction theorem.

2.1. *Let X be a complete metric space, Z a set. For each $[t, e] \in T \times E$ let $M(t, e)$ be a subset of $X \times Z$.*

Suppose that

$$PM(r, e) \subset U(PM(w(r, e)), r)$$

for each $[r, e] \in T \times E$. Then

$$PM(r, e) \subset U(\lim P\tilde{M}(\cdot), s(r, e))$$

for each $[r, e] \in T \times E$. Here P is the projection on the first component and

$$\tilde{M}(r) = \bigcup_{e \in E} M(r, e), \quad \lim W(\cdot) = \bigcap_{s>0} \Big(\bigcup_{r<s} W(r)\Big)^{-}.$$

In the case when a concrete algorithm is given we have the following corollary:

2.2. *Suppose D is a subset of $X \times Z$ such that*

$$M(r, e) \subset D \quad \textit{for all} \quad [r, e] \in T \times E .$$

Suppose F is a mapping of D into itself which satisfies the following condition

$$[x, z] \in M(r, e) \quad \textit{implies} \quad F(x, z) \in M(w(r, e)), \quad d(x, PF(x, z)) \leqq r .$$

Then the algorithm

$$(x_{n+1}, z_{n+1}) = F(x_n, z_n)$$

starting at a point $[x_0, z_0] \in M(r_0, e_0)$ is meaningful, the sequence x_n converges to a point $x_* \in X$ and

$$d(x_n, x_*) \leqq s(w^n(r_0, e_0))$$

for every $n = 0, 1, 2, \ldots$.

It will be convenient to have another characterization of rates of convergence: convergence of the series may be replaced, roughly speaking, by the existence of an estimate function.

2.3. *Suppose w is a mapping of $T \times E$ into itself and let h be a nonnegative function defined on $T \times E$ such that*

$$h(r, e) = r + h(w(r, e))$$

for each $[r, e] \in T \times E$. *Then* w *is a rate of convergence on* $T \times E$ *with respect to the first component. If*

$$\lim_{n \to \infty} h(w^{(n)}(t, e)) = a$$

exists for all $[t, e] \in T \times E$
then

$$s(t, e) = h(t, e) - a .$$

Proof. It is easy to see that

$$h(t, e) = t + w(t, e) + w^{(2)}(t, e) + \ldots + w^{(n)}(t, e) + h(w^{(n+1)}(t, e))$$

so that

$$t + \ldots + w^{(n)}(t, e) \leqq h(t, e)$$

for all n.

3. THE ALGORITHM

This section is devoted to the study of the natural rate of convergence of the algorithm (1)–(2).

3.1. *Let* Q *be the set of all pairs* (r, e) *such that* $r > 0$ *and* $0 \leqq e < 1$. *Let* $a \geqq 0$. *Consider the mapping* w *which assigns to each pair* $(r, e) \in Q$ *the pair*

$$r' = \tfrac{1}{2} r \left((1 + e^2) + (1 - e^2) \frac{r}{x} \right) \left(2e^2 + (1 - e^2) \frac{r}{x} \right)$$

$$e' = e^2 + (1 - e^2) \frac{r}{x}$$

where

$$x = x(r, e) = \frac{r + (r^2 + a^2(1 - e^2)^2)^{1/2}}{1 - e^2} .$$

Then w *is a mapping of* Q *into itself; if* P *stands for the first coordinate then the series*

$$\sum_0^\infty P \circ w^{(n)}$$

converges for each $(r, e) \in Q$ *and its sum equals* $x - a$. *The function* x *satisfies the relation*

$$x(r', e') = x(r, e) - r .$$

The function $g(r, e) = (1 - e^2)/x(r, e)$ *satisfies*

$$g(r', e') = (1 + e') \, g(r, e) \tag{3}$$

In terms of g the pair r', e' is given by

$$r' = (1 + e')(e^2 r + \tfrac{1}{2}g(r, e)\, r^2) \tag{4}$$

$$e' = e^2 + g(r, e)\, r\,. \tag{5}$$

Proof. Consider the iterative procedure (1)–(2) applied to the function $f(x) = x^2 - a^2$. Suppose we are given an $r > 0$ and an e, $0 \leqq e < 1$. We intend to show that there exists a point $x_0 \geqq a$ und a number $g_0 > 0$ such that the first two steps of the iteration satisfy the following relations

$$\begin{aligned} x_0 - x_1 &= r \\ 1 - g_0\, 2x_0 &= e \\ x_1 - x_2 &= r' \\ 1 - g_1\, 2x_1 &= e' \end{aligned}$$

Suppose first that we have an x_0 which satisfies the first two equations. Then, from the second equation

$$g_0 = \frac{1 - e}{2x_0} \quad \text{whence} \quad g_1 = (1 + e)\, g_0 = \frac{1 - e^2}{2x_0}$$

and

$$x_0 - x_1 = g_1(x_0^2 - a^2) = \frac{1 - e^2}{2x_0}(x_0^2 - a^2)$$

so that

$$r = \frac{1 - e^2}{2x_0}(x_0^2 - a^2)\,. \tag{6}$$

This is a condition on x_0 necessary for the first two equations. It is easy to see that the choice

$$x = \frac{r + (r + a^2(1 - e^2)^2)^{1/2}}{1 - e^2} \tag{7}$$

will satisfy the first two equations.

We intend to show now that the other two equations will be satisfied as well.

Let us show that

$$\begin{aligned} e' &= 1 - g_1\, 2x_1 \\ r' &= x_1 - x_2\,. \end{aligned}$$

To see that, note first that $g_1 = (1 + e)\, g_0 = (1 - e^2)/2x_0$.

Hence

$$1 - g_1\, 2x_1 = e^2 + g_1\, 2r = e^2 + \frac{1 - e^2}{2x_0}\, 2r = e'\,. \tag{8}$$

Furthermore, using (6), we obtain

$$(9) \qquad x_1 - x_2 = g_2(x_1^2 - a^2) = (1 + e')\, g_1(x_1^2 - a^2) =$$

$$= (1 + e')\,\frac{1 - e^2}{2x_0}\,[(x_0 - r)^2 - a^2] =$$

$$= (1 + e')\left[\frac{1 - e^2}{2x_0}(x_0^2 - a^2) + \frac{1 - e^2}{2x_0}(-2x_0 r + r^2)\right] =$$

$$= (1 + e')\left[r - (1 - e^2)\, r + \frac{1 - e^2}{2x_0} r^2\right] = \tfrac{1}{2} r(1 + e')\left(2e^2 + \frac{1 - e^2}{x_0} r\right).$$

Since

$$1 + e' = (1 + e^2) + (1 - e^2)\frac{r}{x_0}$$

by (8) this proves that $x_1 - x_2 = r'$.

To complete the proof, let us show now that

$$x(r', e') = x(r, e) - r$$

and

$$g(r', e') = (1 + e')\, g(r, e)\,.$$

To verify the first formula, it suffices, by (6), to prove the equation

$$r' = \frac{1 - e'^2}{2(x - r)}[(x - r)^2 - a^2]\,.$$

However, we know from (9) that

$$r' = (1 + e')\,\frac{1 - e^2}{2x}\,[(x - r)^2 - a^2]$$

so that it will be sufficient to prove

$$\frac{1 - e'^2}{2(x - r)} = (1 + e')\,\frac{1 - e^2}{2x}$$

in other words

$$(10) \qquad (1 - e')\, x = (1 - e^2)(x - r)\,.$$

Now, by (8), we have

$$1 - e' = 1 - e^2 - (1 - e^2)\frac{r}{x}$$

whence our assertion. The formula for g follows from (10) upon multiplying both sides by $(1 + e')/x(x - r)$.

Now we are able to prove the following

3.2. Theorem. *Let X and Y be two Banach spaces and let f be a mapping defined in a neighbourhood U of a point $x_0 \in X$ and taking its values in Y. Suppose that f is Fréchet differentiable in U and that the following conditions are satisfied*

1° *there exists a constant k such that*

$$|f'(x) - f'(y)| \leqq k|x - y|$$

for all $x, y \in U$

2° *there exists a bounded linear operator*

$$G \in B(Y, X)$$

and three numbers e_0, g_0, r_0 such that

$$|1 - Gf'(x_0)| \leqq e_0 \tag{11}$$

$$k|(2 - Gf'(x_0))\,G| \leqq g_0 \tag{12}$$

$$|(2 - Gf'(x_0))\,Gf(x_0)| \leqq r_0 \tag{13}$$

3° *the numbers e_0, g_0, r_0 satisfy the inequality*

$$(1 - e_0^2)^2 \geqq 2r_0 g_0$$

4° *the domain U contains the closed spherical neighbourhood of x_0 with radius*

$$m = \frac{1 - e_0^2}{g_0} - \frac{2r_0}{1 - e_0^2} - \frac{1}{g_0}\left((1 - e_0^2)^2 - 2g_0 r_0\right)^{1/2}.$$

Then the process (1)–(2) starting at the point x_0 is meaningful and converges to a point x_* for which

$$Gf(x_*) = 0 \quad \text{and} \quad |x_* - x_0| \leqq m\,.$$

If a stands for the nonnegative square root of

$$\frac{1}{g_0^2}\left((1 - e_0^2)^2 - 2g_0 r_0\right)$$

then

$$|x_* - x_n| \geqq s(w^n(r_0, e_0))\,, \quad n = 0, 1, 2, \ldots$$

where w is the rate of convergence defined in Lemma 3.1.

Proof. Let w and g be the mappings defined in Lemma 3.1. For $(r, e) \in Q$ we shall define a set $M(r, e) \subset X \times B(Y, X)$ as follows: the pair $[x, G]$ belongs to $M(r, e)$ if and only if the following three conditions are satisfied

$$|1 - Gf'(x)| \leqq e$$

$$|(2 - Gf'(x))G| \leqq \frac{1}{k}\, g(r, e)$$

$$|(2 - Gf'(x))\, Gf(x)| \leqq r$$

Let us prove that $[x, G] \in M(r, e)$ implies $[x', G'] \in M \circ w(r, e)$. Write, for brevity, D for $f'(x)$ and D' for $f'(x')$. Let us show first that $|1 - G'D'| \leqq e'$. It follows from Lemma 1.1 that

$$|1 - G'D'| \leqq |(1 - GD)|^2 + |G'|\,|D - D'| \leqq \tag{14}$$

$$\leqq e^2 + \frac{1}{k}\, g(r, e)\, kr = e' ,$$

the last equation being a consequence of (5). Next, we have to show that $G'' = G' + (1 - G'D')\, G'$ satisfies $|G''| \leqq (1/k)\, g(e', r')$.

We have $G'' = (2 - G'D')\, G'$ so that, using (3)

$$G'' \leqq (1 + |1 - G'D'|)\, |G'| = (1 + e')\frac{1}{k}\, g(r, e) = \frac{1}{k}\, g(r', e') .$$

The last inequality to be proved is $|G''f(x')| \leqq r'$. Since

$$G''f(x') = G''(f(x') - f(x) - D(x' - x)) + G''(f(x) + D(x' - x)) .$$

Using the relations $G'' = (2 - G'D')\, G'$ and $G'f(x) = -(x' - x)$ we may rewrite the second term of the above relation in the form

$$\begin{aligned} G''(f(x) + D(x' - x)) &= -(2 - G'D')(1 - G'D)(x' - x) = \\ &= -(2 - G'D')(1 - GD)^2 (x' - x) \end{aligned} \tag{15}$$

so that, using (14),

$$|G''f(x')| \leqq |G''|\tfrac{1}{2}kr^2 + (1 + e')\, e^2 r .$$

Since $|G''| = |(2 - G'D')\, G'| \leqq (1 + e')(1/k)\, g$, this yields $|G''f(x')| \leqq (1 + e') \cdot (\frac{1}{2}gr^2 + e^2 r) = r'$ (see (4)). Conditions (11), (12), (13) ensure that

$$[x_0, G] \in M(r_0, e_0) .$$

It suffices to observe that the parameter a has been chosen so as to have

$$g(r_0, e_0) = g_0 .$$

The proof is complete.

References

[1] *W. Burmeister:* Inversionsfreie Verfahren zur Lösung nichtlinearer Operatorgleichungen, ZAMM *52* (1972), 101–110.

[2] *O. H. Hald:* On a Newton-Moser type method, Numer. Math. *23* (1975), 411–426.

[3] *H.-J. Kornstaedt:* Funktionalungleichungen und Iterationsverfahren, Aequationes Math. *13* (1975), 21–45.

[4] *J. Křížková, P. Vrbová:* A remark on a factorization theorem, Comm. Math. Univ. Carol. *15* (1974), 611–614.

[5] *J. Moser:* Stable and random motions in dynamical systems with special emphasis on celestial mechanics, Princeton University Press 1973.
[6] *H. Petzeltová:* Remark on a Newton-Moser type method, CMUC (1980) 719—725.
[7] *H. Petzeltová, P. Vrbová:* An overrelaxed modification of Newton's method, Revue Roumaine des Mathématiques *22* (1977), 959—963.
[8] *H. Petzeltová, P. Vrbová:* A remark on small divisors problems, Czech. Math. J. *103* (1978), 1—12.
[9] *F. A. Potra:* On a modified secant method, Math. Rev. Anal. Numer. Theor. Approximation, Anal. Numer. Theor. Approximation, *8*, 2 (1979), 203—214.
[10] *F. A. Potra:* An application of the Induction method of V. Pták to the study of Regula Falsi, Aplikace Matematiky *26* (1981), 111—120.
[11] *F. A. Potra:* The rate of convergence of a modified Newton's process, Aplikace matematiky *26* (1981), 13—17.
[12] *F. A. Potra:* An error analysis for the secant method, Numer. Math., *38* (1982), 427—445.
[13] *F. A. Potra, V. Pták:* Nondiscrete induction and a double step secant method, Math. Scand. *46* (1980), 236—250.
[14] *F. A. Potra, V. Pták:* On a class of modified Newton processes, Numer. Funct. Anal. and Optimiz. *2* (1980), 107—120.
[15] *F. A. Potra, V. Pták:* Sharp error bounds for Newton's process, Numer. Math. *34* (1980), 63—72.
[16] *F. A. Potra, V. Pták:* A generalization of Regula Falsi, Numer. Math. *36* (1981), 333—346.
[17] *V. Pták:* Some metric aspects of the open mapping theorem, Math. Annalen *165* (1966), 95—104.
[18] *V. Pták:* A quantitative refinement of the closed graph theorem, Czech. Math. J. *99* (1974), 503—506.
[19] *V. Pták:* A theorem of the closed graph type, Manuscripta Math. *13* (1974), 109—130.
[20] *V. Pták:* Deux théorèmes de factorization. Comptes Rendus, Acad. Sci. Paris *278* (1974), 1091—1094.
[21] *V. Pták:* Concerning the rate of convergence of Newton's process, Comm. Math. Univ. Carolinae *16* (1975), 699—705.
[22] *V. Pták:* A modification of Newton's method, Čas. pěst. mat. *101* (1976), 188—194.
[23] *V. Pták:* Nondiscrete mathematical induction and iterative existence proofs. Linear Algebra and its Applications *13* (1976), 223—236.
[24] *V. Pták:* The rate of convergence of Newton's process, Numer. Math. *25* (1976), 279—285.
[25] *V. Pták:* Nondiscrete mathematical induction, in: General Topology and its Relations to Modern Analysis and Algebra IV., 166—178, Lecture Notes in Mathematics *609*, Springer Verlag, 1977.
[26] *V. Pták:* What should be a rate of convergence, RAIRO, Analyse Numérique *11* (1977), 279—286.
[27] *V. Pták:* Stability of exactness, Comm. Math. (Poznań) *21* (1978), 343—348.
[28] *V. Pták:* A rate of convergence, Numer. Funct. Anal. and Optimiz. *1* (1979), 255—271.
[29] *V. Pták:* Factorization in Banach algebras, Studia Math. *65* (1979), 279—285.
[30] *S. Ulm:* Iteration methods with succesive approximation of the inverse operator (Russian) Esti NSV Tead. Acad. Toimetised Füüs-Mat. *16* (1967), 403—411.
[31] *J. Zemánek:* A remark on transitivity of operator algebras. Čas. pěst. mat. *100* (1975), 176—178.

Authors' addresses: F. A. Potra, Department of Mathematics, INCREST, Bd. Păcii 220, Bucureşti, România, V. Pták, 115 67 Praha 1, Žitná 25 (Matematický ústav ČSAV). Czechoslovakia.