Časopis pro pěstování matematiky

Vlastimil Pták

A modification of Newton's method

Časopis pro pěstování matematiky, Vol. 101 (1976), No. 2, 188--194

Persistent URL: http://dml.cz/dmlcz/117905

Časopis pro pěstování matematiky, roč. 101 (1976), Praha

A MODIFICATION OF NEWTON'S METHOD

VLASTIMIL PTÁK, Praha

(Received June 30, 1975)

INTRODUCTION

In a recent paper [2] the author obtained a simple theorem of the closed graph type, the so-called "Induction Theorem" which gives an abstract model for iterative existence proofs in analysis and numerical analysis. The induction theorem not only provides a heuristic method for the investigation of iterative constructions but also yields considerable simplifications of proofs.

The induction theorem forms the basis of the method of nondiscrete mathematical induction described in [8]. The method consists in reducing the given problem to a set of inequalities for several indeterminate functions one of which is to be a rate of convergence.

In the present remark we intend to apply the method of nondiscrete mathematical induction to a modification of Newton's method due to Jürgen Moser. The purpose of the remark is twofold. First of all, the method of nondiscrete induction provides, we believe, a deeper insight into the essence of the use of approximate solutions of the linearized equation (obtained, in concrete situations, by means of smoothing operators or additional viscosities or similar devices). At the same time, Moser's theorem provides a good example to illustrate the advantages of the nondiscrete method.

1. DEFINITIONS AND NOTATION

Let T be an interval of the form $T = \{t; 0 < t < t_0\}$. A function ω mapping T into itself will be called a *rate of convergence* if, for each $r \in T$, the series

$$r + \omega(r) + \omega(\omega(r)) + \omega(\omega(\omega(r))) + \ldots$$

is convergent.

If ω is a rate of convergence we denote by σ the sum of the above series. We observe that σ satisfies the functional equation

$$\sigma(\omega(r)) + r = \sigma(r).$$

(1,1) Lemma. *Let a be a number with $a > 1$. Then $t \mapsto t^a$ is a rate of convergence on the interval $0 < t < 1$. For sufficiently small t we have the estimate $\sigma(t) \leqq 2t$. More precisely, it suffices to have*

$$t \leqq 2^{-1/(a-1)}.$$

Proof. If

$$0 < t \leqq 2^{-1/(a-1)}$$

we have $t^a \leqq \frac{1}{2}t$ and it is easy to see, by induction, that

$$t^{a^n} \leqq (\tfrac{1}{2})^n t.$$

Hence $\sigma(t) \leqq 2t$.

Let (E, d) be a metric space.

An approximate set in E is a family of subsets of E $t \mapsto W(t)$, $t \in T$. We define the limit $W(0)$ of this family as follows

$$W(0) = \bigcap_{r>0} \Big(\bigcup_{s \leqq r} W(s) \Big)^- ;$$

hence $W(0)$ is the set of all limits of convergent sequences x_n such that $x_n \in W(r_n)$ for a suitable sequence $r_n \to 0$.

If $x \in E$ and $r > 0$ we define

$$U(x, r) = \{z \in E;\ d(z, x) < r\}$$

similarly, if $M \subset E$ and $r > 0$,

$$U(M, r) = \{z \in E;\ d(z, m) < r \text{ for some } m \in M\}.$$

(1,2) The Induction Theorem. *Let ω be a rate of convergence on T. Let (E, d) be a complete metric space; for each $t \in T$ let $W(t)$ be a subset of E. Suppose that, for each $t \in T$*

$$W(t) \subset U(W(\omega(t)), t);$$

then

$$W(t) \subset U(W(0), \sigma(t)).$$

The proof is simple and straightforward; the theorem is closely related to the closed graph theorem of Functional Analysis. The proof is given and the relation to the closed graph theorem explained in the author's remark [4]. Applications to existence theorems are given in [3], [2], [6], [7]. The general principles governing the application of the nondiscrete induction method are discussed in the author's lecture [8] where further examples are given.

Sometimes, it is more convenient to use the induction theorem in the following equivalent form.

(1.3) *Let (E, d) be a complete metric space, let ω be a function which maps $T = (0, t_0)$ into itself and such that $\omega^{(n)}(t)$ tends to zero for all $t \in T$. Let φ be a positive increasing function defined on T such that*

$$\sigma_\varphi(t) = \sum \varphi(\omega^{(n)}(t)) < \infty$$

for each $t \in T$. Then $\varphi \circ \omega \circ \varphi^{-1}$ is a rate of convergence. Given a family $W(t)$ of subsets of E such that

$$W(t) \subset U(W(\omega(t)), \varphi(t))$$

for each $t \in T$, then

$$W(t) \subset U(W(0), \sigma_\varphi(t))$$

for each $t \in T$.

Proof. Set $Z(t) = W(\varphi^{-1}(t))$ and apply the induction theorem to the family $Z(\cdot)$ and the rate of convergence $\tilde{\omega} = \varphi \circ \omega \circ \varphi^{-1}$.

2. MOSER'S THEOREM

In this section we state and prove a slightly improved version of Moser's theorem. There are some formal simplifications in the statement of the theorem and the proof is considerably simpler.

(2,1) Theorem. *Suppose $E_1 \subset E_\varrho \subset E_0$ are three vector spaces over the complex field, each equipped with a norm (indexed by the same number). Suppose that the norm of E_ϱ satisfies the inequality*

$$|u|_\varrho \leqq c|u|_0^{1-\sigma} |u|_1^\sigma \quad \text{for all} \quad u \in E_1$$

and a fixed $0 < \sigma < 1$ and that the space $(E_\varrho, |\cdot|_\varrho)$ is complete.

Further, let $F_1 \subset F_0$ be two vector spaces over the complex field, each equipped with a norm (indexed by the same number).

Let R be a positive number and let

$$D_1 = \{u \in E_1;\ |u|_\varrho \leqq R\}.$$

Let D be the closure of D_1 in the space $(E_\varrho, |\cdot|_\varrho)$. Let f be a continuous mapping of $(D, |\cdot|_\varrho)$ into $(F_0, |\cdot|_0)$ such that f maps D_1 into F_1. We shall make the following assumptions about f.

1° *(growth) there exist two positive numbers M and $S \geqq 1$ such that*

$$|f(u)|_1 \leqq M \max(S, |u|_1) \quad \text{for all} \quad u \in D_1 .$$

2° *(approximation by a differential) there exists a mapping g of $D_1 \times E_1$ into F_1 and a number β, $0 < \beta < 1$ such that*

$$|f(u + v) - f(u) - g(u, v)|_0 \leqq M|v|_0^{2-\beta} |v|_1^\beta$$

whenever both u and $u + v$ belong to D_1.

3° *(solvability of the linearized equation) there exist two positive numbers λ and μ with the following properties*

$$\textit{if} \quad u \in D_1 \quad \textit{and} \quad g \in f(D_1) \quad \textit{are such that}$$

$|g|_0 \leqq m^{-\lambda}$ where $m = \max((1/M)|g|_1, |u|_1, S)$ and if Q is any number greater than 1, there exists at least one $v \in E_1$ for which

$$|g(u, v) - g|_0 \leqq MmQ^{-\mu},$$
$$|v|_1 \leqq MmQ,$$
$$|v|_0 \leqq M|g(u, v)|_0 .$$

4° *suppose that $\mu > \lambda$ and*

$$\frac{\mu + 1}{\mu - \lambda} < \min\left(2 - \beta \cdot \frac{\lambda + (\lambda + 1)(\mu + 1)}{\lambda(\mu + \beta)}, \ \lambda \frac{1 - \sigma}{\sigma}\right).$$

Then there exists a number $\delta > 0$ such that $|f(0)|_0 < \delta$ implies the existence of an element $u \in D$ for which $f(u) = 0$.

Proof. Suppose that $u \in D_1$ and $|f(u)|_0 \leqq m^{-\lambda}$ where

(1) $$m \geqq \max(|u|_1, S).$$

According to 1°, we have $\max(|u|_1, S) \geqq |f(u)|_1/M$ so that $m \geqq \max(|f(u)|_1/M, |u|_1, S)$.

According to 4°, there exists, for each $Q > 1$, at least one $v \in E_1$ such that

$$|g(u, v) + f(u)|_0 \leqq MmQ^{-\mu} \quad \text{and} \quad |v|_1 \leqq MmQ .$$

Suppose further that

(2) $$u + v \in D_1$$

and let us estimate $|f(u + v)|_0$ and $|v|_0$. We have, by 2°

$$|f(u + v)|_0 \leqq |f(u + v) - f(u) - g(u, v)|_0 + |g(u, v) + f(u)|_0 =$$
$$\leqq M|v|_0^{2-\beta} |v|_1^\beta + MmQ^{-\mu} \leqq M|v|_0^{2-\beta} (MmQ)^\beta + MmQ^{-\mu} .$$

According to 3°

$$|v|_0 \leqq M|g(u, v)|_0 \leqq M(|g(u, v) + f(u)|_0 + |f(u)|_0) \leqq$$
$$\leqq M(MmQ^{-\mu} + m^{-\lambda}).$$

Let us assume further that

$$MmQ^{-\mu} \leqq m^{-\lambda}. \tag{3}$$

Under these assumptions it follows that $|v|_0 \leqq 2Mm^{-\lambda}$ whence

$$|v|_\varrho \leqq c\, 2^{1-\sigma} M m^{-\lambda(1-\sigma)+\sigma} Q^\sigma$$

$$|f(u + v)|_0 \leqq pQ^\beta + qQ^{-\mu}$$

where $p = M^3\, 2^{2-\beta} m^{-\lambda(2-\beta)+\beta}$ and $q = Mm$.

We shall write $1/r$ for m. We intend to show that there exists a number $a > 1$ such that

$$qQ^{-\mu} \leqq \tfrac{1}{2} r^{a\lambda} \tag{5}$$

$$pQ^\beta \leqq \tfrac{1}{2} r^{a\lambda} \tag{6}$$

for a suitable $Q > 1$. First of all let us note that condition (5) implies condition (3). If such a number a exists it is possible to expect that $t \to t^a$ will be a suitable rate of convergence. Since $|u|_1 \leqq m = 1/r$ it is natural to impose further the following condition

$$Q \leqq \frac{1}{2M} r^{1-a} \tag{7}$$

which will ensure the estimate $|v|_1 \leqq \frac{1}{2}(1/r)^a$ whence

$$|u + v|_1 \leqq |u|_1 + |v|_1 \leqq \frac{1}{r} + \frac{1}{2}\left(\frac{1}{r}\right)^a$$

and this will not exceed $(1/r)^a$ as soon as

$$r \leqq \left(\frac{1}{2}\right)^{1/(a-1)}. \tag{8}$$

Summing up: our task reduces to finding an $a > 1$ for which there exists a Q satisfying the following conditions

$$1 < Q$$

(5) $\quad (2M)^{1/\mu}\, r^{-(1+a\lambda)/\mu} = (2qr^{-a\lambda})^{1/\mu} \leqq Q$

(7) $$Q \leqq \frac{1}{2M} r^{1-a}$$

(6) $$Q \leqq \left(\frac{1}{2p} r^{a\lambda}\right)^{1/\beta} = $$

$$= \left(2^{3-\beta}M^3\right)^{-1/\beta} r^{-(\lambda(2-\beta)-\beta-a\lambda)/\beta}.$$

If there exists an a such that

(57) $$\frac{1 + a\lambda}{\mu} < a - 1$$

and

(56) $$\frac{1 + a\lambda}{\mu} < \frac{1}{\beta}(\lambda(2 - \beta) - \beta - a\lambda)$$

then there exists an $r(a)$ with the following property: for each $0 < r \leqq r(a)$ there exists a $Q > 1$ (depending on r) satisfying (5), (6) and (7). Since $\mu > \lambda$, condition (57) is equivalent to

(57) $$a > \frac{\mu + 1}{\mu - \lambda}$$

and condition (56) to

(56) $$a < 2 - \beta \frac{\lambda + (\lambda + 1)(\mu + 1)}{\lambda(\mu + \beta)}.$$

If condition 4° is satisfied, it follows that it is possible to choose an a which satisfies both (56) and (57).

It follows that

$$|v|_\varrho \leqq c\, 2^{1-\sigma} M \left(\frac{1}{2M}\right)^{\sigma} r^{\lambda(1-\sigma)-\sigma a};$$

this will tend to zero with r if $\lambda(1 - \sigma) - \sigma a > 0$. Choose a positive ε such that $\omega = \lambda(1 - \sigma) - \sigma a - \varepsilon > 0$. If $r \leqq (c\, 2^{1-\sigma}M(2M)^{-\sigma})^{-1/\varepsilon}$, we shall have $|v|_\varrho \leqq r^{\omega}$. The new condition to be imposed on a is the following

(9) $$a < \lambda \frac{1 - \sigma}{\sigma}.$$

It follows from condition 4° that $(\mu + 1)/(\mu - \lambda) < (1 - \sigma)/\sigma$ so that a may be chosen so as to satisfy (56), (57) and (9). Once such a is fixed, for any $r \leqq r(a)$ there exists a $Q > 1$ satisfying (5), (6) and (7). Now let

$$r_0 = \min\left(S^{-1}, r(a), 2^{-1/(a-1)}, (c\, 2^{1-2\sigma}M^{1-\sigma})^{-1/\varepsilon}\right)$$

and, for each $r \leqq r_0$, set

$$W(r) = \{u \in D_1; |u|_\varrho \leqq R - \sigma(r^\omega), |u|_1 \leqq 1/r, |f(u)|_0 \leqq r^\lambda\}.$$

Here, of course, σ is the function corresponding to the rate of convergence $t \to t^a$, hence $\sigma(t) = t + t^a + t^{a^2} + \ldots$. The preceding discussion shows that

$$W(r) \subset U(W(r^a), r^\omega)$$

for $r \leqq r_0$, the neighbourhood being taken in the norm $|\cdot|_\varrho$.

To complete the proof it will be sufficient to show that there exists an $r \leqq r_0$ such that $0 \in W(r)$. If $r \leqq r_0$ then $0 \in W(r)$ is equivalent to $|f(0)|_0 \leqq r^\lambda$ and $\sigma(r^\omega) \leqq R$. If

$$0 < t < \left(\frac{1}{2}\right)^{1/(a-1)},$$

we have $\sigma(t) \leqq 2t$ by lemma (1,1). It follows that the inequality $\sigma(r^\omega) \leqq R$ will be satisfied if $r^\omega \leqq (\frac{1}{2})^{1(a-1)}$ and $2r^\omega \leqq R$, in other words, if $r \leqq \min((\frac{1}{2})^{1/\omega(a-1)}, (\frac{1}{2}R)^{1/\omega})$. Now set $\delta = (\min(r_0, (\frac{1}{2})^{1/\omega(a-1)}, (\frac{1}{2}R)^{1/\omega}))^\lambda$ and suppose that $|f(0)|_0 \leqq \delta$. If $r = (|f(0)|_0)^{1/\lambda}$, we have $|f(0)|_0 \leqq r^\lambda$ and $\sigma(r^\omega) \leqq R$ so that $0 \in W(r)$. It follows from the induction theorem that $W(0)$ is nonvoid, in other words, there exists a $u \in D$ for which $f(u) = 0$.

The proof is complete.

The author wishes to thank M. ŠTĚDRÝ for several stimulating conversations concerning Moser's theorem and its applications.

References

[1] *J. Moser:* A rapidly convergent iteration method and nonlinear partial differential equations, Ann. Scuola Norm. Sup. Pisa *20* (1966), 265—315.

[2] *V. Pták:* A theorem of the closed graph type, Manuscripta Math. *13* (1974), 109—130.

[3] *V. Pták:* Deux théoremes de factorisation, Comptes Rendus Ac. Sci. Paris *278* (1974), 1091—1094.

[4] *V. Pták:* A quantitative refinement of the closed graph theorem, Czech. Math. J. *99* (1974), 503—506.

[5] *V. Pták:* On the closed graph theorem, Czech. Math. J. *84* (1959), 523—527.

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

[7] *J. Zemánek:* A remark on transitivity of operator algebras, Čas. pěst. mat. *100* (1975), 176—178.

[8] *V. Pták:* Nondiscręte mathematical induction and iterative existence proofs, Lin. Algebra and its Appl., in print.

Author's address: 115 67 Praha 1, Žitná 25 (Matematický ústav ČSAV).