Aplikace matematiky

Florian-Alexandru Potra

The rate of convergence of a modified Newton's process

Aplikace matematiky, Vol. 26 (1981), No. 1, 13–17

Persistent URL: http://dml.cz/dmlcz/103890

THE RATE OF CONVERGENCE OF A MODIFIED NEWTON'S PROCESS

F. A. POTRA

(Received October 18, 1978)

1. PRELIMINARIES

Let us first recall the definition of the rate of convergence given in [6].

Definition. *Let T be an interval of the form $T = \{t \in R;\ 0 < t \leqq t_0\}$ for a positive t_0. A rate of convergence on T is a function ω defined on T with the following properties:*

1° ω maps T into itself,

2° for each $t \in T$, the series

$$t + \omega(t) + \omega^2(t) + \ldots$$

is convergent, where $\omega^2(t)$ means $\omega(\omega(t))$, etc.

The sum of the above series will be denoted by $\sigma(t)$. The function σ satisfies evidently the functional equation

$$\sigma(t) = t + \sigma(\omega(t)) . \tag{1}$$

We can now restate the induction theorem [5].

Let (E, d) be a complete metric space, and let ω be a rate of convergence on the interval $T = \{t \in R;\ 0 < t \leqq t_0\}$. For each $t \in T$ let $Z(t)$ be a subset of E. We denote by $Z(0)$ the limit of the family $Z(\cdot)$, i.e. $Z(0) = \bigcap_{s>0} \left(\bigcup_{t \leqq s} Z(t)^- \right)$, and for any $M \subset E$ we denote by $U(M, r)$ the set $\{x \in E;\ \mathrm{d}(x, M) \leqq r\}$.

Theorem. *If*

$$Z(t) \subset U(Z(\omega(t)), t) \tag{2}$$

for each $t \in T$, then

$$Z(t) \subset U(Z(0), \sigma(t)) \tag{3}$$

for each $t \in T$.

The above theorem will be the basis for the proof of our main theorem concerning the convergence of the modified Newton process, which will be stated in the next section.

2. THE MODIFIED NEWETON PROCESS

Let us consider the operational equation

$$f(x) = 0 \tag{4}$$

where f is a nonlinear operator defined in a Banach space E with values in a Banach space F. Starting with an initial approximation x_0 we may obtain a sequence $(x_n)_{n=0}^{\infty}$ converging to a solution x^* of the equation (4) through the algorithm

$$x_{n+1} = x_n - (f'(x_n))^{-1} f(x_n), \quad n = 0, 1, 2, \ldots \tag{5}$$

where $f'(x)$ denotes the Fréchet derivative [1] of the operator f at the point x. This generalization of the classical Newton process was first studied by L. V. Kantorovič [2], who gave sufficient conditions for the convergence of this method together with estimates for the bounds of the errors $\|x_n - x^*\|$, $n = 1, 2, \ldots$. Using the method of nondiscrete mathematical induction, V. Pták [6] has found that the rate of convergence of Newton's process is given by

$$\omega(r) = \frac{r^2}{2(r^2 + d)^{1/2}}. \tag{6}$$

He has also obtained sharp estimates for the error bounds.

In applied numerical analysis the Newton process (5) is often replaced by a modified process given by the relations

$$x_{n+1} = x_n - (f(x_0))^{-1} f(x_n), \quad n = 0, 1, 2, \ldots \tag{7}$$

which, in contrast to (5), does not require the inversion of a linear operator at each step. This modification of the Newton method was first studied also by L. V. Kantorovič [3].

In the sequel we shall use the method of V. Pták to find out a rate of convergence for the process (7) and to obtain sharp error bound estimates.

One of the main ideas in the quoted paper of V. Pták [6] is that the study of convergence of the process (5) can be reduced in a certain sense from the general case to the case where f is a real quadratic polynomial. We have done the same thing for the process (7).

Lemma. *If d_0, r_0 and K are any positive numbers satisfying the relation*

$$2Kr_0 \leqq d_0, \tag{8}$$

then:

I.

$$\omega(r) = \frac{1}{2\, d_0}(Kr^2 + 2d_0 r - 2r\sqrt{(d_0^2 - 2d_0 K(r_0 - r))}) \tag{9}$$

is a rate of convergence on the interval $T = \{r; 0 < r \leqq r_0\}$ *and the corresponding function* σ *is given by*

$$(10) \qquad \sigma(r) = \frac{1}{K}\left(\sqrt{(d_0^2 - 2d_0 K(r_0 - r))} - \sqrt{(d_0^2 - 2d_0 r_0 K)}\right);$$

ω *as well as* σ *are increasing functions on the interval* T.

II. *Given the real polynomial*

$$(11) \qquad \varphi(x) = \frac{1}{2}Kx^2 - \frac{1}{2K}d_0(d_0 - 2Kr_0)$$

then the algorithm

$$(12) \qquad x_0 = d_0/K$$

$$x_{n+1} = x_n - \varphi(x_n)/\varphi'(x_0)$$

yields a sequence $(x_n)_{n=0}^{\infty}$, *converging to the root* $x^* = (1/K)\sqrt{(d_0(d_0 - 2Kr_0))}$ *of the equation* $\varphi(x) = 0$, *and the following equalities are fulfilled*:

$$(13) \qquad \varphi'(x_0) = d_0\,,$$

$$(14) \qquad (\varphi'(x_0))^{-1}\,\varphi(x_0) = r_0\,,$$

$$(15) \qquad x^* - x_0 = \sigma(r_0)\,,$$

$$(16) \qquad |x_{n+1} - x_n| = \omega^n(r_0)\,, \quad n = 0, 1, 2, \ldots,$$

$$(17) \qquad |x_n - x^*| = \sigma(\omega^n(r_0))\,, \quad n = 0, 1, 2, \ldots.$$

Proof. The equalities (13)–(14) can be verified by simple calculation. It is clear that the process

$$(18) \qquad y_{n+1} = y_n - \varphi(y_n)/d_0\,, \quad n = 0, 1, 2, \ldots$$

with an arbitrary starting point y_0, $x^* < y_0 \leqq x_0$, yields a sequence $(y_n)_{n=0}^{\infty}$ which converges to the root x^*. Impossing the condition $\varphi(y_0)/d_0 = r$ with $0 < r \leqq r_0$ we get $y_0 = (1/K)\sqrt{(d_0^2 - 2K\,d_0(r_0 - r))}$. Setting now $\omega(r) = \varphi(y_1)/d_0$ and $\sigma(r) = y_0 - x^*$ we obtain (9) and (10). Finally, the equalities (15–(17) are obvious, since we have $y_0 = d_0/K = x_0$ for $r = r_0$.

We are now able to state our main result:

Let E and F be two Banach spaces, let x_0 be a point of E and f a mapping of $U = U(x_0, m)$ into F. Suppose that f is twice Fréchet differentiable at each $x \in U$ and that $f'(x_0)$ has a bounded inverse.

Theorem. *If there exist three positive constants K, d_0 and r_0 such that the following conditions are satisfied:*

$$\|f''(x)\| \leqq K \quad \text{for all} \quad x \in U\,, \tag{18}$$

$$\|(f'(x_0))^{-1}\|^{-1} \geqq d_0 > 0\,, \tag{19}$$

$$\|(f'(x_0))^{-1} f(x_0)\| \leqq r_0\,, \tag{20}$$

$$2Kr_0 \leqq d_0\,, \tag{21}$$

$$m \geqq \frac{1}{K}\left(d_0 - \sqrt{(d_0(d_0 - 2Kr_0))}\right), \tag{22}$$

then the modified Newton process (7) *makes sense, the sequence $(x_n)_{n=0}^{\infty}$ obtained is convergent to a solution x^* of the equation* (4), *and we have the following estimates:*

$$\|x^* - x_0\| \leqq \frac{1}{K}\left(d_0 - \sqrt{(d_0(d_0 - 2Kr_0))}\right), \tag{23}$$

$$\|x_{n+1} - x_n\| \leqq \omega^n(r_0)\,, \quad n = 0, 1, 2, \ldots, \tag{24}$$

$$\|x_n - x^*\| \leqq \sigma(\omega^n(r_0))\,, \quad n = 0, 1, 2, \ldots, \tag{25}$$

where the functions ω and σ are given by (9) *and* (10).

These estimates are sharp in the following sense: for each triple K, d_0, r_0 of positive numbers satisfying the inequality (21) *there exists a mapping f for which the estimates turn into equalities.*

Proof. First of all let us remark that the sharpness of the estimates is guaranteed by the preceding lemma. Let us consider now for each r, $0 < r \leqq r_0$, the set

$$W(r) = \{x \in E;\ \|(f'(x_0))^{-1} f(x)\| \leqq r,\ \|x - x_0\| \leqq \sigma(r_0) - \sigma(r)\}\,. \tag{26}$$

It is clear from the above definition and (20) that $W(r_0) = \{x_0\}$, so that according to the induction theorem and the lemma, the proof of our theorem will be complete if we can prove that $x' \in W(\omega(r))$ for each $x \in W(r)$, where we have denoted

$$x' = x - (f'(x_0))^{-1} f(x)\,. \tag{27}$$

The condition $\|x' - x_0\| \leqq \sigma(r_0) - \sigma(\omega(r))$ is obviously fulfilled because of the relations (1), (26) and the triangle inequality, so that all we still have to prove is that the following inequality is satisfied:

$$\|(f'(x_0))^{-1} f(x')\| \leqq \frac{1}{2d_0}\left(Kr^2 + 2d_0 r - 2r\sqrt{(d_0^2 - 2d_0 K(r_0 - r))}\right). \tag{28}$$

From (27), we infer that

$$\begin{aligned} f(x') &= f(x') - f(x) - f'(x_0)(x' - x) = \\ &= f(x') - f(x) - f'(x)(x' - x) + (f'(x) - f'(x_0))(x' - x)\,. \end{aligned}$$

According to the Taylor formula (see for instance [7], t. I, ch. III, § 7) and the conditions (18) and (19), the above equality yields

$$\|(f'(x_0))^{-1} f(x')\| \leqq \frac{K}{2d_0} \|x' - x\|^2 + \frac{K}{d_0} \|x - x_0\| \|x' - x\| \leqq$$

$$\leqq \frac{K}{2d_0} r^2 + \frac{K}{\mathrm{d}_0} (\sigma(r_0) - \sigma(r)) r .$$

Substituting in the above in equality the expression for $\sigma(r)$ given by (10) we get the inequality (28).

References

[1] *M. Frechet:* La notion de différentielle dans l'analyse générale. Ann. Ec. Norm. Sup., 42, (1925), 293—323.

[2] *Л. В. Канторович*: Функциональный анализ и прикладная математика. У. М. Н., 3, 89—185 (1948).

[3] *Л. В. Канторович*: О методе Ньютона. Труды Матем. Института Стеклова, 28, 104—144 (1949).

[4] *Л. В. Канторович, Г. П. Акилов:* Функциональный анализ. Издание второе переработанное, Москва (1977).

[5] *V. Pták:* Nondiscrete mathematical induction and iterative existence proofs. Linear algebra and its applications 13 (1976), 223—238.

[6] *V. Pták:* The rate of convergence of Newton's process. Num. Math., 25 (1976), 279—285.

[7] *L. Schwartz:* Analyse Mathématique, Herman, Paris, 1967.

Souhrn

RYCHLOST KONVERGENCE MODIFIKOVANÉHO NEWTONOVA PROCESU

F. A. Potra

V. Pták [6] použil metodu nediskretní matematické indukce [5] k nalezení rychlosti konvergence Newtonova procesu. Rychlost konvergence je přitom považována za funkci, nikoliv číslo. V článku se analogickou metodou dokazuje, že rychlost konvergence modifikovaného Newtonova procesu má tvar

$$\omega(r) = \frac{1}{2d_0} (Kr^2 + 2d_0 r - 2r \sqrt{(d_0^2 - 2d_0 K(r_0 - r))}) ,$$

kde K, d_0, r_0 jsou kladná čísla, závisející na počátečních podmínkách. Tento přístup umožňuje získat ostré odhady chyb modifikovaného Newtonova procesu.

Author's address: Dr. F. A. *Potra*, Department of Mathematics National Institute of Scientific and Technical Creation, Bdul Pacii 220, 77538 Bucharest, Romania.