ANNALES ACADEMIAE SCIENTIARUM FENNICAE

Series A

I. MATHEMATICA

450

EIN ÜBERQUADRATISCH KONVERGENTER ITERATIVER ALGORITHMUS

VON

PENTTI LAASONEN

HELSINKI 1969
SUOMALAINEN TIEDEAKATEMIA

doi:10.5186/aasfm.1969.450

Vorgelegt am 9 Mai 1969

KESKUSKIRJAPAINO
HELSINKI 1969

1. Einleitung

Zur iterativen Annäherung der Nullstelle x^* einer stetigen Funktion $f(x)$ einer Veränderlichen kann oft die gewöhnliche Iteration nach der Vorschrift

$$x_{n+1} = g(x_n) , \qquad n = 0 , 1 , 2 , \dots$$

angewandt werden, wo $g(x)$ eine solche Funktion ist, dass die beiden Gleichungen

$$(1.1) \qquad \begin{aligned} f(x) &= 0 \\ g(x) - x &= 0 \end{aligned}$$

äquivalent sind; z.B. genügt die einfache Wahl

$$g(x) = x - kf(x)$$

mit einem solchen konstanten Koeffizienten k, dass g in der Umgebung von x^* eine Lipschitz-Bedingung

$$|g(x) - g(y)| \leqq L|x - y| , \qquad L < 1$$

erfüllt. Die Konvergenz ist hierbei gewöhnlicherweise nur linear [3].

Die Regula falsi besteht aus den Schritten

$$x_{n+1} = x_n - k_n f(x_n) , \qquad n = 1 , 2 , \dots ,$$

mit

$$k_n = \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$$

und gibt eine Konvergenz von der Ordnung $\frac{1}{2}(1 + \sqrt{5}) \approx 1{,}618$ [2].

Das Verfahren von Steffensen wird nun aus abwechselnden Schritten der Regula falsi und der gewöhnlichen Iteration zusammengesetzt:

$$\begin{cases} x_{n+1} = g_n(x_n) , \\ \bar{x}_{n+1} = g(x_{n+1}) \end{cases} \qquad n = 0 , 1 , 2 , \dots$$

mit

$$g(x) = x - kf(x) ,$$

$$g_n(x) = x - k_n f(x) ,$$

$$k_n = \frac{\bar{x}_n - x_n}{f(\bar{x}_n) - f(x_n)} .$$

Es ist besonders zu bemerken, dass jetzt die Lipschitz-Konstante beliebig sein darf, wobei trotzdem die Konvergenz für einen Doppelschritt $x_n \to x_{n+1}$ quadratisch ist.

Es ist nun interessant zu beobachten, dass eine Abwandlung der Rechenvorschrift von ganz unwesentlichem Einfluss auf die praktische Handhabung überquadratische Konvergenz von der Ordnung $1 + \sqrt{2} = 2{,}414$ zustandebringt. Hierzu braucht man nur das Verfahren von Steffensen derart abzuändern, dass die beiden Teilschritte sich auf den veränderlichen Wert k_n stützen:

$$\begin{cases} x_{n+1} = g_n(x_n)\,, \\ \bar{x}_{n+1} = g_n(x_{n+1}) \end{cases} \qquad n = 0\,,\,1\,,\,2\,,\,\ldots \tag{1.2}$$

mit

$$g_n(x) = x - k_n f(x)\,,$$

$$k_n = \frac{\bar{x}_n - x_n}{f(\bar{x}_n) - f(x_n)}\,.$$

Im folgenden wird der diesem Verfahren zugehörige Satz in dermassen verallgemeinerter Form bewiesen, dass f eine Abbildung des Banachraumes R auf ihn selbst ist. In diesartigen Verallgemeinerungen itarativer Methoden haben Schmidt [5], [6] und Chen [1] lineare Operatoren benutzt, die den dividierten Differenzen entsprechen. Gleicherweise wird auch hier Existenz der Steigungen erster Ordnung vorausgesetzt: zu jedem Elementenpaar $x'\,,x''$ aus R gehört ein beschränkter linearer Operator $F(x'\,,x'')$ mit folgender definierender Eigenschaft:

$$F(x'\,,x'')\,(x' - x'') = f(x') - f(x'') \tag{1.3}$$

und von folgender Stetigkeit:

$$\|F(x'\,,x'') - F(y'\,,y'')\| \leqq K[\|x' - y'\| + \|x'' - y''\|]\,, \tag{1.4}$$

die sich auch in der Arbeit von Johnson und Scholz [4] als hinreichend gezeigt hat.

2. Konvergenz

Satz 1. Der stetige Operator f bildet den Banachraum R auf sich selbst ab. Es gebe zwei Punkte x_0 und $\bar{x}_0$ sowie zwei Konstanten a_0 und b_0 mit den folgenden Eigenschaften:

a) In der Umgebung

$$S_0 = \{x \mid \|x - x_0\| \leqq r_0\} \tag{2.1}$$

des Punktes x_0 mit dem Radius

(2.2) $$r_0 = 2\,a_0\,(1 + b_0)$$

sei die durch (1.3) definierte Steigung mit der Stetigkeitseigenschaft (1.4) für beliebige $x'\,, x''\,, y'\,, y'' \in S_0$ vorhanden.

b) Speziell sei

(2.3) $$\|[F(x_0\,, \bar{x}_0)]^{-1}\| \leqq b_0\,.$$

c) Die folgenden Beziehungen mögen gelten:

(2.4) $$\|f(x_0)\| \leqq a_0$$

(2.5) $$\|x_0 - \bar{x}_0\| \leqq a_0$$

(2.6) $$h_0 = 2\,K\,a_0 b_0 (1 + b_0)^2 \leqq \frac{2}{5}\,.$$

Unter diesen Umständen ist der Algorithmus

(2.7) $$C_n = [F(x_n\,, \bar{x}_n)]^{-1}\,,$$

(2.8) $$x_{n+1} = x_n - C_n f(x_n)\,,$$

(2.9) $$\bar{x}_{n+1} = x_{n+1} - C_n f(x_{n+1})$$

wohlbestimmt und gibt konvergente Folgen der Punkte (x_n) und $(\bar{x}_n)$ mit dem gemeinsamen Grenzpunkt x^*, der ein Nullpunkt der Abbildung $f(x)$ in S_0 ist.

Beweis. Für $n = 0$ erhält man aus (2.8) und (2.9) die Gleichungen

$$x_0 - x_1 = C_0 f(x_0)$$

$$\begin{aligned} x_1 - \bar{x}_1 &= x_0 - x_1 - C_0 (f(x_0) - (f(x_1)) \\ &= C_0 (F(x_0\,, \bar{x}_0) - F(x_0\,, x_1))\,(x_0 - x_1) \end{aligned}$$

und ferner die Schranken

$$\|x_0 - x_1\| \leqq a_0 b_0 < \frac{1}{2}\,r_0\,,$$

(2.10) $$\begin{aligned} \|x_1 - \bar{x}_1\| &\leqq K b_0\,[\|x_0 - \bar{x}_0\| + \|x_0 - x_1\|]\,\|x_0 - x_1\| \\ &\leqq K a_0^2 b_0^2 (1 + b_0) \leqq \frac{1}{2}\,a_0 h_0\,\frac{b_0}{1 + b_0} < \frac{1}{2}\,a_0 h_0 < \frac{1}{10}\,r_0 \end{aligned}$$

so dass sowohl x_1 als auch $\bar{x}_1$ zu S_0 gehört. Ferner erhält man

$$\begin{aligned} f(x_1) &= F(x_0\,, \bar{x}_0)\,(x_0 - x_1) - f(x_0) + f(x_1) \\ &= (F(x_0\,, \bar{x}_0) - F(x_0\,, x_1))\,(x_0 - x_1) \end{aligned}$$

$$(2.11)\qquad \|f(x_1)\| \leqq K\,\|\bar{x}_0 - x_1\|\,\|x_0 - x_1\| \leqq K(a_0 + a_0 b_0)\,a_0 b_0 < \frac{1}{2}\,a_0 h_0$$

Die Differenz zwischen den zwei linearen Operatoren $F(x_0\,,\bar{x}_0)$ und $F(x_1\,,\bar{x}_1)$ kann folgendermassen abgeschätzt werden:

$$\|F(x_0\,,\bar{x}_0) - F(x_1\,,\bar{x}_1)\| \leqq K[2\,\|x_0 - x_1\| + \|x_0 - \bar{x}_0\| + \|x_1 - \bar{x}_1\|]$$

$$< K\left[2\,a_0 b_0 + a_0 + \frac{1}{5}\,a_0\right] = K a_0 b_0\left(\frac{6}{5} + 2\,b_0\right)\cdot b_0^{-1} < h_0 b_0^{-1}\,.$$

Da die Norm der Inversen von $F(x_0\,,\bar{x}_0)$ durch b_0 beschränkt ist, sichert eine Variante des wohlbekannten Satzes von Banach die Existenz der Inversen von $F(x_1\,,\bar{x}_1)$ (z.B. [2], s. 79), und zwar mit der oberen Schranke

$$(2.12)\qquad \|(F(x_1\,,\bar{x}_1))^{-1}\| \leqq \frac{b_0}{1 - h_0}\,.$$

Werden jetzt die neuen Konstanten durch die Definitionen

$$(2.13)\qquad a_n = \frac{1}{2}\,a_{n-1}\,h_{n-1}$$

$$(2.14)\qquad b_n = \frac{b_{n-1}}{1 - h_{n-1}}$$

$$(2.15)\qquad r_n = 2\,a_n(1 + b_n)$$

$$(2.16)\qquad h_n = 2\,K a_n b_n (1 + b_n)^2$$

und der Kugelbereich S_n durch die Definition

$$(2.17)\qquad S_n = \{x \mid \|x - x_n\| \leqq r_n\}$$

eingeführt, so folgt aus (2.12), (2.11), (2.10), dass die den Ungleichungen (2.3—5) entsprechenden Abschätzungen

$$(2.18)\qquad \|C_n\| \leqq b_n$$

$$(2.19)\qquad \|f(x_n)\| \leqq a_n$$

$$(2.20)\qquad \|x_n - \bar{x}_n\| \leqq a_n$$

auch für $n = 1$ gelten. Ferner folgen aus (2.13—16) die Abschätzungen

$$(2.21)\qquad h_1 = K a_0 h_0\,\frac{b_0}{1 - h_0}\left(1 + \frac{b_0}{1 - h_0}\right)^2 < \frac{1}{2}\,\frac{h_0^2}{(1 - h_0)^3} < \frac{2}{5}$$

$$r_1 = a_0 h_0\left(1 + \frac{b_0}{1 - h_0}\right) \leqq \frac{2}{3}\,a_0(1 + b_0) < \frac{1}{2}\,r_0\,.$$

Da auch $\|x_0 - x_1\| < \frac{1}{2} r_0$ ist, folgen daraus die Beziehungen

$$h_n < \frac{2}{5} \tag{2.22}$$

$$S_n \subset S_{n-1} \tag{2.23}$$

für $n = 1$. Vollständige Induktion ist demnach möglich, um die allgemeine Gültigkeit der Eigenschaften (2.18—22) für jedes positive n zu sichern, wenn der Algorithmus (2.7—9) und die Folge der Abkürzungen unbegrenzt fortgesetzt werden.

Der Ungleichung (2.21) entsprechend erhält man eine allgemeine Abschätzung, präziser als (2.22),

$$h_{n+1} < \frac{1}{2} \frac{h_n^2}{(1 - h_n)^3} < \frac{25}{27} h_n ,$$

die zu erkennen gibt, dass die Folge (h_n) schneller als geometrisch gegen Null strebt. Hieraus folgt, dass das Produkt $\sum_0^\infty (1 - h_n)$ konvergiert und daher die Folge der Konstanten

$$b_n = b_0 \Big/ \prod_0^{n-1} (1 - h_\nu) < b \tag{2.24}$$

gleichmässig begrenzt ist. Die Folgen der Kanstanten a_n und r_n streben somit gegen Null und diejenigen der Bereiche S_n und der Punkte x_n und $\bar{x}_n$ gegen einen Grenzpunkt x^*. Aus der Stetigkeit folgt weiter nach (2.19), dass x^* ein Nullpunkt von f ist.

3. Ordnung der Konvergenz

Satz 2. Die Ordnung der Konvergenz der Folgen (x_n) und $(\bar{x}_n)$ gegen x^* im vorigen Satz ist $1 + \sqrt{2}$ in dem Sinne, dass es für die Fehler solche gegen Null strebende Schranken

$$\|x_n - x^*\| \leqq u_n , \quad \|\bar{x}_n - x^*\| \leqq v_n$$

gibt, die sukzessiv die folgenden Grössenverhältnisse haben:

$$u_{n+1} = 0(v_n^{1+\sqrt{1/2}}) = 0(u_n^{1+\sqrt{2}}) ; \tag{3.1}$$

$$v_{n+1} = 0(u_{n+1}^{\sqrt{2}}) \quad = 0(v_n^{1+\sqrt{2}}) . \tag{3.2}$$

Beweis. Aus (2.8) und (2.9) folgen die Gleichungen

$$\begin{aligned} x_{n+1} - x^* &= x_n - x^* - C_n(f(x) - f(x^*)) \\ &= C_n(F(x_n, \bar{x}_n) - F(x_n, x^*))(x_n - x^*) \\ \bar{x}_{n+1} - x^* &= x_{n+1} - x^* - C_n(f(x_{n+1}) - f(x^*)) \\ &= C_n(F(x_n, \bar{x}_n) - F(x_{n+1}, x^*))(x_{n+1} - x^*) \end{aligned}$$

und nach (2.24) die Abschätzungen

$$\|x_{n+1} - x^*\| \leqq Kb\, \|\bar{x}_n - x^*\|\, \|x_n - x^*\| \, . \tag{3.3}$$

$$\|\bar{x}_{n+1} - x^*\| \leqq Kb[\|x_n - x^*\| + \|x_{n+1} - x^*\| + \|\bar{x}_n - x^*\|]\, \|x_{n+1} - x^*\| \, . \tag{3.4}$$

Für genügend grosses $n(\geqq n_0)$ ist nach (3.3)

$$\|x_{n+1} - x^*\| \leqq \|x_n - x^*\|$$

und nach (3.4)

$$\|\bar{x}_n - x^*\| \leqq \|x_n - x^*\|$$

so dass (3.4) auf die Form

$$\|\bar{x}_{n+1} - x^*\| \leqq 3\, Kb\, \|x_n - x^*\|\, \|x_{n+1} - x^*\| \tag{3.5}$$

vereinfacht werden kann. Wenn jetzt in den Abschätzungen (3.3) und (3.5) für $n \geqq n_0$ die Ungleichungen durch Gleichungen ersetzt werden, hat man eine Folge von rekursiven Definitionen der oberen Schranken für $\|x_n - x^*\|$ und $\|\bar{x}_n - x^*\|$. Werden diese mit u_n und v_n bezeichnet, so erhält man die gekoppelten Differenzengleichungen

$$\begin{cases} u_{n+1} = Kb\, u_n\, v_n \\ v_{n+1} = 3\, Kb\, u_n\, u_{n+1} \end{cases} \qquad n = n_0, n_{0+1}, \ldots .$$

Übergang zu Logarithmen liefert ein Paar von linearen Differenzengleichungen, die folgende Lösung haben:

$$\begin{cases} \ln u_n = K_1(1 + \sqrt{2})^n + K_2(1 - \sqrt{2})^n - \ln(Kb) - \frac{1}{2}\ln 3 \, , \\ \ln v_n = \sqrt{2}\, K_1(1 + \sqrt{2})^n - \sqrt{2}\, K_2(1 - \sqrt{2})^n - \ln(Kb) \, . \end{cases}$$

Hieraus ergeben sich die Resultate

$$\lim_{n\to\infty}\left(\ln u_{n+1} - \left(1 + \sqrt{\frac{1}{2}}\right)\ln v_n\right) = \sqrt{\frac{1}{2}}\ln(Kb) - \sqrt{\frac{1}{2}}\ln 3 \, ,$$

$$\lim_{n\to\infty}(\ln v_{n+1} - \sqrt{2}\ln u_{n+1}) = (\sqrt{2} - 1)\ln(Kb) + \sqrt{\frac{1}{2}}\ln 3 \, ,$$

aus denen unmittelbar die Verhältnisse (3.1) und (3.2) folgen.

4. Bestimmung von Fixpunkten

Die obigen Sätze gestatten natürlich eine einfache Anwendung auf das iterative Aufsuchen eines Fixpunktes einer stetigen Abbildung $g(x)$ eines Banachraumes R auf sich selbst. Die nötigen Modifikationen in den Annahmen des Satzes 1 sind:

Die beschränkte lineare Steigung $G(x', x'')$ wird durch

$$G(x', x'')(x' - x'') = g(x') - g(x'')$$

definiert und von derselben wird die durch

$$\|G(x', x'') - G(y', y'')\| \leqq K[\|x' - y'\| + \|x'' - y''\|]$$

ausgedrückte Stetigkeit gefordert; die Abschätzung (2.3) muss durch

$$\|[I - G(x_0, \bar{x}_0)]^{-1}\| \leqq b_0$$

ersetzt werden, und (2.4) durch

$$\|x_0 - g(x_0)\| \leqq a_0$$

Der Algorithmus (2.7—9) findet sein Gegenstück in der Form

$$\begin{aligned} C_n &= [I - G(x_n, \bar{x}_n)]^{-1}, \\ x_{n+1} &= [I - C_n]\, x_n + C_n\, g(x), \\ \bar{x}_{n+1} &= [I - C_n]\, x_{n+1} + C_n\, g(x_{n+1}). \end{aligned}$$

Dieser Algorithmus gibt dann Folgen von Punkten (x_n) und $(\bar{x}_n)$, die gegen einen Fixpunkt x^* von $g(x)$ streben:

$$x^* = g(x^*);$$

die Ordnung der Konvergenz in dem oben gebrauchten Sinne ist abermals $1 + \sqrt{2}$.

Technische Universität
Otaniemi, Finnland

Literatur

[1] Chen, K.-W.: Generalization of Steffensen's method for operator equations in Banach space. - Comment. Math. Univ. Carolinae *5*, p. 47–77 (1964)

[2] Collatz, L.: Funktionalanalysis und numerische Mathematik. - Berlin, Springer, 1964.

[3] Henrici, P.: Elements of numerical analysis. - New York, Wiley, 1964.

[4] Johnson, L. W. and Scholz, D. R.: On Steffensen's method. - SIAM J. Numer. Anal. *5*, p. 296–302 (1968)

[5] Schmidt, J. W.: Eine Übertragung der Regula Falsi auf Gleichungen in Banachräumen. - ZAMM, *34*, s. 1–8, 97–110 (1963)

[6] —»— Konvergenzgeschwindigkeit der Regula falsi und des Steffensen-Verfahrens im Banachraum. - ZAMM, *46*, s. 146–148 (1966)

Gedruckt Oktober 1969