

Über die Konvergenzordnung des Intervall-Newton-Verfahrens

G. Alefeld, Karlsruhe

Eingegangen am 13. Juli 1987

Herrn E. Martensen, Karlsruhe, anlässlich seines 60. Geburtstages gewidmet

Zusammenfassung — Abstract

Über die Konvergenzordnung des Intervall-Newton-Verfahrens. Es ist bekannt, daß das klassische Newton-Verfahren kubisch gegen eine einfache Nullstelle konvergiert, wenn die zweite Ableitung an der Nullstelle verschwindet. Wir zeigen zunächst, daß sich diese Eigenschaft nicht auf das Intervall-Newton-Verfahren überträgt. Verwendet man jedoch anstelle der intervallmäßigen Auswertung der Ableitung die Mittelwertform oder die zentrierte Form, so erhält man wieder kubische Konvergenz.

On the Order of Convergence of the Interval-Newton-Method. It is well known that the classical Newton method is cubically convergent to a simple zero if the second derivative vanishes at the zero. We first show that this property does not hold for the interval-Newton-method. If, however, the interval arithmetic evaluation of the derivative in this method is replaced by the mean-value form or by the centered form, respectively, then the method is again cubically convergent.

AMS Subject Classifications: 65G10, 65H05, 65H10.

Key words: Interval arithmetic, Newton's method, order of convergence.

1.

Wendet man das Newton-Verfahren

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}, \quad k=0, 1, 2, \dots, \quad (1)$$

zur Berechnung einer Nullstelle $x^* \in D$ der zweimal stetig differenzierbaren Funktion $f: D \subseteq \mathbb{R}^1 \rightarrow \mathbb{R}^1$ an, so konvergiert im Falle $f'(x^*) \neq 0$ die Folge $\{x^k\}_{k=0}^\infty$ lokal – d. h. für genügend nahe an x^* gelegene Startwerte x^0 – gegen x^* und es gilt

$$|x^{k+1} - x^*| \leq c |x^k - x^*|^2, \quad k=0, 1, 2, \dots, \quad (2)$$

was als *quadratische Konvergenz* bezeichnet wird.

Ist f dreimal stetig differenzierbar und gilt

$$f(x^*)=0, \quad f'(x^*) \neq 0, \quad f''(x^*)=0, \quad (3)$$

so gilt sogar

$$|x^{k+1} - x^*| \leq c |x^k - x^*|^3, \quad k=0, 1, 2, \dots,$$

was als *kubische Konvergenz* bezeichnet wird.

Siehe dazu etwa [4].

2.

Zur fortwährenden *Einschließung* einer im kompakten reellen Intervall $[x]^0 \subseteq D$ gelegenen Nullstelle x^* von $f: D \subseteq \mathbb{R}^1 \rightarrow \mathbb{R}^1$ mit einer für $[x] \subseteq [x]^0$ existierenden intervallmäßigen Auswertung $f'([x])$ der Ableitung kann man das intervallmäßige Newton-Verfahren verwenden:

$$[x]^{k+1} = N[x]^k \cap [x]^k, \quad k=0, 1, 2, \dots, \quad (5)$$

$N[x]$ ist der sogenannte *Intervall-Newton-Operator*:

$$N[x] = m[x] - \frac{f(m[x])}{f'([x])}. \quad (6)$$

Hierbei ist $m[x]$ für ein kompaktes reelles Intervall

$$[x] = [x_1, x_2] = \{x \mid x_1 \leq x \leq x_2\}$$

ein beliebiges Element aus $[x]$. Gewöhnlich wählt man den Mittelpunkt. $f(m[x])$ bezeichnet den Funktionswert von f an der Stelle $m[x]$ und $f'([x])$ bezeichnet die sogenannte intervallmäßige Auswertung der Ableitung, die man — vereinfacht ausgedrückt — dadurch erhält, daß man in dem formelmäßigen Ausdruck für $f'(x)$ die Variable x durch das Intervall $[x]$ ersetzt und $f'([x])$ nach den Regeln der Intervallrechnung berechnet. Die intervallmäßige Auswertung $f'([x])$ ergibt ein kompaktes reelles Intervall. Siehe dazu [2, Chapter 3]. Die Durchschnittsbildung in (5) ist wie üblich mengentheoretisch zu verstehen. Das Ergebnis $[x]^{k+1}$ kann allein durch die Schranken von $[x]^k$ und $N[x]^k$ ausgedrückt werden.

Genügt der Durchmesser d — zur Definition siehe unten unter 3. — der intervallmäßigen Auswertung der Ableitung einer Abschätzung

$$d(f'([x])) \leq \gamma d[x], \quad \gamma \geq 0, \quad [x] \subseteq [x]^0, \quad (7)$$

was z. B. bei rationalen Funktionen stets der Fall ist, so gilt für die Iterierten des intervallmäßigen Newton-Verfahrens (5)

$$d[x]^{k+1} \leq c (d[x]^k)^2, \quad c \geq 0, \quad k=0, 1, 2, \dots, \quad (8)$$

d. h. die Durchmesser $d[x]^k$ konvergieren quadratisch gegen Null. Siehe dazu [2, Chapter 7].

(Die Durchführbarkeit des Verfahrens (5) und die Konvergenz gegen x^* ist gesichert, wenn f in $[x]^0 \subseteq D$ eine Nullstelle x^* besitzt und $0 \notin f'([x]^0)$ gilt. Es gilt dann sogar $x^* \in [x]^k$, $k \geq 0$. Siehe dazu [2, Chapter 7].)

3.

Wir demonstrieren zunächst an einem einfachen *Beispiel*, daß die in 1. unter der Voraussetzung (3) beschriebene kubische Konvergenz des Newton-Verfahrens sich nicht ohne weiteres auf das intervallmäßige Newton-Verfahren (5) überträgt:

Das Polynom

$$f(x) = \frac{1}{6}x^3 - \frac{1}{2}x^2 + \frac{3}{2}x - \frac{7}{6} \quad (9)$$

besitzt an der Stelle $x^* = 1$ eine Nullstelle, für welche (3) gilt, wie man unter Verwendung von

$$f'(x) = \frac{1}{2}x^2 - x + \frac{3}{2}$$

$$f''(x) = x - 1$$

nachprüfen kann.

Für die nachfolgenden Rechnungen benötigen wir einige Regeln für das Rechnen mit Durchmessern und Beträgen von Intervallen. Wir wiederholen zunächst die Definitionen (siehe etwa [2, Chapter 2]):

Für $[x] = [x_1, x_2]$ heißt

- a) $d[x] = x_2 - x_1$ *Durchmesser*
und
- b) $|[x]| = \max\{|x_1|, |x_2|\}$ *Betrag* des Intervalles $[x]$.

Es gelten die folgenden Regeln:

- $\alpha)$ $[x] \subseteq [y] \Rightarrow |[x]| \leq |[y]|$ und $d[x] \leq d[y]$;
- $\beta)$ $d([x] \pm [y]) = d[x] + d[y]$;
- $\gamma)$ $0 \in [x], 0 \in [y] \Rightarrow d([x] \cdot [y]) \leq d[x] d[y]$;
- $\delta)$ $\alpha \in \mathbb{R} \Rightarrow d(\alpha[x]) = |\alpha| d[x]$;
- $\kappa)$ $0 \notin [x] \Rightarrow d\left(\frac{1}{[x]}\right) \leq \left|\frac{1}{[x]}\right|^2 d[x]$.

$\alpha)$, $\beta)$ und $\delta)$ findet man in [2, Chapter 2]. $\gamma)$ wurde in [3] bewiesen. $\kappa)$ kann man durch elementare Rechnung beweisen.

Es sei nun $[x]^0 = [1, 1 + \varepsilon]$, $\varepsilon > 0$. Es ist dann $d[x]^0 = \varepsilon$ und $m[x]^0 = 1 + \frac{\varepsilon}{2}$. Das Intervall $[x]^0$ enthält die Nullstelle $x^* = 1$ des Polynoms (9).

Es ist

$$f(m[x]^0) = \frac{\varepsilon}{2} + \frac{1}{6} \left(\frac{\varepsilon}{2}\right)^3$$

und

$$\begin{aligned} f'([x]^0) &= \frac{1}{2} [1, 1+\varepsilon]^2 - [1, 1+\varepsilon] + \frac{3}{2} \\ &= \left[1-\varepsilon, 1+\varepsilon + \frac{1}{2}\varepsilon^2 \right]. \end{aligned}$$

Für $\varepsilon < 1$ gilt $0 \notin f'([x]^0)$.

Damit erhält man für den Intervall-Newton-Operator (6)

$$\begin{aligned} N[x]^0 &= m[x]^0 - \frac{f(m[x]^0)}{f'([x]^0)} \\ &= 1 + \frac{\varepsilon}{2} - \left(\frac{\varepsilon}{2} + \frac{1}{6} \left(\frac{\varepsilon}{2} \right)^3 \right) \cdot \left[\frac{1}{1+\varepsilon+\frac{1}{2}\varepsilon^2}, \frac{1}{1-\varepsilon} \right] \\ &= \left[1 + \frac{\varepsilon}{2} - \frac{\frac{\varepsilon}{2} + \frac{1}{6} \left(\frac{\varepsilon}{2} \right)^3}{1-\varepsilon}, 1 + \frac{\varepsilon}{2} - \left[\frac{\frac{\varepsilon}{2} + \frac{1}{6} \left(\frac{\varepsilon}{2} \right)^3}{1+\varepsilon+\frac{1}{2}\varepsilon^2} \right] \right]. \end{aligned}$$

Wegen $1 = x^* \in N[x]^0$ ist die linke Schranke von $N[x]^0$ nicht größer als die linke Schranke von $[x]^0$ und wegen $m[x]^0 \notin N[x]^0$ ist die rechte Schranke von $N[x]^0$ sicherlich kleiner als die rechte Schranke von $[x]^0$.

Somit gilt in (5) für $k=0$

$$[x]^1 = N[x]^0 \cap [x]^0 = \left[1, 1 + \frac{\varepsilon}{2} - \frac{\frac{\varepsilon}{2} + \frac{1}{6} \left(\frac{\varepsilon}{2} \right)^3}{1+\varepsilon+\frac{1}{2}\varepsilon^2} \right].$$

Daraus erhält man

$$\begin{aligned} d[x]^1 &= \frac{\varepsilon}{2} - \frac{\frac{\varepsilon}{2} + \frac{1}{6} \left(\frac{\varepsilon}{2} \right)^3}{1+\varepsilon+\frac{1}{2}\varepsilon^2} \\ &= \frac{1}{2} \frac{1 + \frac{11}{24}\varepsilon}{1+\varepsilon+\frac{1}{2}\varepsilon^2} (d[x]^0)^2 \\ &= \frac{1}{2} \frac{1 + \frac{11}{24}d[x]^0}{1+d[x]^0+\frac{1}{2}(d[x]^0)^2} (d[x]^0)^2. \end{aligned}$$

Durch wiederholte Anwendung dieser Schlußweise erhält man

$$d[x]^{k+1} = \frac{1}{2} \frac{1 + \frac{11}{24}d[x]^k}{1+d[x]^k+\frac{1}{2}(d[x]^k)^2} (d[x]^k)^2.$$

Daraus folgt

$$\lim_{k \rightarrow \infty} \frac{d[x]^{k+1}}{(d[x]^k)^2} = \frac{1}{2},$$

d. h. die Folge der Durchmesser konvergiert quadratisch gegen Null.

4.

Es ist bekannt, daß die in 2. genannten Eigenschaften des Verfahrens (5) auch gelten, wenn man anstelle von $f'([x])$ ein Intervall $[a]_{[x]}$ verwendet, für welches

$$\{f'(x) \mid x \in [x]\} \subseteq [a]_{[x]} \quad \text{mit} \quad 0 \notin [a]_{[x]} \quad (10)$$

und

$$d([a]_{[x]}) \leq \gamma d[x], \quad \gamma \geq 0, \quad [x] \subseteq [x]^0 \quad (11)$$

gelten. Siehe dazu [2, Chapter 7].

Wir betrachten zwei Möglichkeiten für die Wahl von $[a]_{[x]}$:

A) Gilt $0 \notin f'([x]^0)$, so kann man für $[a]_{[x]}$ im Falle der Existenz der intervallmäßigen Auswertung $f''([x])$ der zweiten Ableitung den Durchschnitt von $f'([x]^0)$ mit der sogenannten *Mittelwertform* $f'_M([x])$ der Ableitung wählen. Dabei ist (siehe [2, Chapter 3])

$$f'_M([x]) = f'(m[x]) + f''([x]) \cdot ([x] - m[x]).$$

Für $x \in [x] \subseteq [x]^0$ gilt nämlich mit einem $\eta \in [x]$ aufgrund des Mittelwertsatzes und der Inklusionsisotonie der Intervallrechnung

$$\begin{aligned} f'(x) &= f'(m[x]) + f''(\eta)(x - m[x]) \\ &\in f'(m[x]) + f''([x]) \cdot ([x] - m[x]) = f'_M([x]) \end{aligned}$$

und daher

$$f'(x) \in f'_M([x]) \cap f'([x]^0) = [a]_{[x]}.$$

Eine Abschätzung der Form (11) besteht, falls

$$d(f''[x]) \leq \sigma d[x], \quad \sigma \geq 0, \quad [x] \subseteq [x]^0, \quad (12)$$

gilt.

B) Gilt $0 \notin f'([x]^0)$, so kann man für $[a]_{[x]}$ den Durchschnitt der *intervallmäßigen Auswertung* $f'_Z([x])$ der zentrierten Form von $f'(x)$ mit $f'([x]^0)$ wählen. Dabei ist

$$f'_Z([x]) = f'(m[x]) + h'([x], m[x]) \cdot ([x] - m[x]).$$

Siehe dazu [2, Chapter 3]. Nach Definition der zentrierten Form¹ gilt für $x \in [x]$

$$\begin{aligned} f'(x) &= f'(m[x]) + h'(x, m[x]) \cdot (x - m[x]) \\ &\in f'(m[x]) + h'([x], m[x]) \cdot ([x] - m[x]) = f'_Z([x]) \end{aligned}$$

und daher

$$f'(x) \in f'_Z([x]) \cap f'([x]^0) = [a]_{[x]}.$$

(11) gilt für dieses $[a]_{[x]}$ sicherlich für rationale Funktionen.

¹ Die Funktion $h'(x, y)$ ist definiert als

$$h'(x, y) = \begin{cases} \frac{h'(x) - h'(y)}{x - y}, & x \neq y \\ h''(x), & x = y. \end{cases}$$

5.

Wir zeigen jetzt — dies ist das Hauptergebnis dieser Note — daß man bei Verwendung der Mittelwertform der Ableitung unter den Voraussetzungen (3) kubische Konvergenz der Durchmesser gegen Null erhält. Zumindest für Polynome ist dies auch bei Verwendung der intervallmäßigen Auswertung $f'_Z([x])$ der zentrierten Form der Ableitung richtig, wie anschließend gezeigt wird.

Satz 1: Es sei $f: D \subseteq \mathbb{R}^1 \rightarrow \mathbb{R}^1$ zweimal stetig differenzierbar in D und die intervallmäßige Auswertung der Ableitung enthalte für ein $[x]^0 \subseteq D$ nicht die Null: $0 \notin f'([x]^0)$. Gilt (3) für ein $x^* \in [x]^0$, so gilt unter der Voraussetzung, daß die zweite Ableitung von f für $[x] \subseteq [x]^0$ intervallmäßig auswertbar ist mit

$$d(f''([x])) \leq \alpha d[x], \quad \alpha \geq 0, \quad (13)$$

für die Iterierten des Verfahrens

$$\left\{ \begin{array}{l} m[x]^k \in [x]^k \\ [x]^{k+1} = N_M[x]^k \cap [x]^k \\ \text{mit} \\ N_M[x]^k = m[x]^k - \frac{f(m[x]^k)}{f'_M([x]^k) \cap f'([x]^0)} \\ \text{und} \\ f'_M([x]^k) = f'(m[x]^k) + f''([x]^k) \cdot ([x]^k - m[x]^k) \\ k = 0, 1, 2, \dots, \end{array} \right.$$

die Ungleichung

$$d[x]^{k+1} \leq \xi (d[x]^k)^3,$$

d. h. die Durchmesser konvergieren (mindestens) kubisch gegen Null.

Beweis: Die Durchführbarkeit und Konvergenz gegen x^* kann wie üblich gezeigt werden. (Siehe dazu [2, Chapter 7].) Wegen $f''(x^*) = 0$ und wegen $x^* \in [x]^k$ gilt $0 \in f''([x]^k)$. Außerdem ist $0 \in [x]^k - m[x]^k$. Damit folgt unter Verwendung von γ) und (13)

$$\begin{aligned} & d(f'_M([x]^k) \cap f'([x]^0)) \\ & \leq d(f'_M([x]^k)) = \\ & = d(f'(m[x]^k) + f''([x]^k) \cdot ([x]^k - m[x]^k)) \\ & \leq d(f''([x]^k)) d([x]^k - m[x]^k) \\ & \leq \alpha (d[x]^k)^2. \end{aligned}$$

Wegen $x^*, m[x]^k \in [x]^k$ gilt $|x^* - m[x]^k| \leq d[x]^k$ und daher mit einem $\xi_k \in [x]^k$ aufgrund des Mittelwertsatzes

$$|f(m([x]^k))| = |f'(\xi_k)| \cdot |m[x]^k - x^*| \leq |f'([x]^k)| d[x]^k.$$

Somit erhält man mit δ) und κ)

$$\begin{aligned} d[x]^{k+1} &\leq d(N_M[x]^k) \\ &= |f(m[x]^k)| d\left(\frac{1}{f'_M([x]^k) \cap f'([x]^0)}\right) \\ &\leq \xi (d[x]^k)^3. \end{aligned}$$

□

In [1] wurde gezeigt, daß für Polynome

$$h'([x], m[x]) \subseteq f''([x])$$

gilt. Daher erhält man aus Satz 1 unmittelbar

Satz 2: Für Polynome gilt die Aussage des vorangehenden Satzes, wenn die Mittelwertform $f'_M([x])$ durch die intervallmäßige Auswertung $f'_Z([x])$ der zentrierten Form von $f'(x)$ ersetzt wird. □

6.

Man könnte vermuten, daß die in den Sätzen 1 und 2 nachgewiesene kubische Konvergenz auch besteht, wenn man den dritten Teil von (3), also $f'''(x^*)=0$, nicht voraussetzt. Diese Vermutung erscheint deswegen naheliegend, da ja bekannt ist, daß der Abstand zwischen der Mittelwertform der Ableitung und dem Wertebereich der Ableitung mit dem Quadrat des Durchmessers des Intervalles $[x]$ gegen Null geht. Siehe dazu etwa [2, Chapter 3, Theorem 6]. Aus dieser Eigenschaft folgt jedoch nicht, daß der Durchmesser der Mittelwertform der Ableitung mit dem Quadrat des Durchmessers des Intervalles $[x]$ gegen Null geht. $d(f'_M([x]))$ geht i. A. nur linear gegen Null. Somit ist die obige Vermutung falsch. Ähnliches gilt für die intervallmäßige Auswertung der zentrierten Form der Ableitung.

Wir erwähnen schließlich, ohne auf Einzelheiten einzugehen, daß alle Aussagen dieser Note sinngemäß für die n -dimensionalen Analoga von (5) gelten. Siehe dazu [2, Chapter 19].

Literatur

- [1] Alefeld, G.: Bounding the slope of polynomial operators and some applications. *Computing* 26, 227–237 (1981).
- [2] Alefeld, G., Herzberger, J.: *Introduction to Interval Computations*. New York: Academic Press 1983.
- [3] Cornelius, H., Lohner, R.: Computing the range of values of real functions with accuracy higher than second order. *Computing* 33, 331–347 (1984).
- [4] Ortega, J. M., Rheinboldt, W. C.: *Iterative Solution of Nonlinear Equations in Several Variables*. New York: Academic Press 1970.

G. Alefeld
Institut für Angewandte Mathematik
Universität Karlsruhe
Kaiserstrasse 12
D-7500 Karlsruhe 1
Bundesrepublik Deutschland

Printed in Austria

Druck: Paul Gerin, A-1021 Wien