

Über das Divergenzverhalten des Intervall-Newton-Verfahrens

*Herrn Prof. Dr. J. W. Schmidt, Dresden, anlässlich seines
Geburstages am 4.8.1991 gewidmet*

G. Alefeld, Karlsruhe

Received January 18, 1991

Abstract — Zusammenfassung

On the divergence behaviour of the interval-Newton-method. If the starting interval contains no zero then the interval-Newton-method breaks down because the intersection becomes empty after a finite number of steps. It is shown that in a certain sense this divergence behaviour is quadratic.

AMS Subject Classification: 65 G10, 65 H10

Key words: Interval-Newton-Method, Speed of Divergence

Über das Divergenzverhalten des Intervall-Newton-Verfahrens. Es ist bekannt, daß das Intervall-Newton-Verfahren wegen leeren Durchschnitts abbricht, wenn das Startintervall keine Nullstelle enthält. Wir zeigen, daß in einem bestimmten Sinne diese Divergenz quadratisch erfolgt.

1. Einleitung

Eine Liste der verwendeten Bezeichnungen und Hilfsmittel findet man im Abschnitt 2. Ansonsten wird generell auf [3] verwiesen.

Für einen Intervallvektor $[x] = ([x]_i)$ und die intervallmäßige Auswertung $f'([x])$ einer Frechet-differenzierbaren Abbildung $f: [x] \subset D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ wird der Intervall-Newton-Operator durch

$$N[x] = m[x] - \text{IGA}(f'([x]), f(m[x])) \quad (1)$$

definiert. $N[x]$ ist wieder ein Intervallvektor. Mit Hilfe von $N[x]$ wird für einen gegebenen Intervallvektor $[x]^0$ das *Intervall-Newton-Verfahren* definiert als

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

wobei die Durchschnittsbildung elementweise für die entsprechenden Komponenten von $N[x]^k$ und $[x]^k$ erfolgt.

In [1] wurden über das Verfahren (2) u.a. die folgenden Aussagen bewiesen.

Theorem. *Es sei $[x]^0$ ein Intervallvektor und $f: [x]^0 \subset D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Abbildung, deren Frechet-Ableitung intervallmäßig auswertbar ist. Wir setzen voraus, daß*

$IGA(f'([x]^0))$ existiert (was damit äquivalent ist, daß der Gauss'sche Algorithmus mit $f'([x]^0)$ und jeder rechten Seite durchführbar ist). Es sei

$$A = |I - IGA(f'([x]^0)) \cdot f'([x]^0)|$$

und für den Spektralradius ρ von A gelte

$$\rho(A) < 1.$$

- (a) Falls f in $[x]^0$ eine Nullstelle x^* besitzt, so ist die nach (2) definierte Folge von Intervallvektoren wohldefiniert. Es gilt $x^* \in [x]^k$ und

$$\lim_{k \rightarrow \infty} [x]^k = x^*.$$

- (b) Besitzt f keine Nullstellen in $[x]^0$, so gibt es ein $k_0 \geq 0$, sodaß

$$N[x]^{k_0} \cap [x]^{k_0} = \emptyset, \quad (\text{leere Menge}), \quad (3)$$

d.h. das Verfahren (2) bricht wegen leeren Durchschnitts ab. $\square$

Das Verhalten in (b) bezeichnen wir als *Divergenz* des Intervall-Newton-Verfahrens.

Falls (a) gilt so läßt sich unter der Voraussetzung

$$d(f'([x])_{ij}) \leq c \|d[x]\|, \quad c \geq 0, \quad (4)$$

die quadratische Konvergenz der Folge (2) nachweisen: Es gilt

$$\|d[x]^{k+1}\| \leq \gamma \|d[x]^k\|^2.$$

In beiden Fällen (a) und (b) gilt das folgende

Lemma 1. (Siehe [1], Lemma 6). Falls die Voraussetzungen des Theorems gelten, so ist

$$m[x]^k \notin [x]^{k+1}. \quad (5)$$

$\square$

Dieses Lemma besagt, daß in jedem Schritt des Intervall-Newton-Verfahrens der Durchmesser mindestens einer Komponente von $[x]^k$ mehr als halbiert wird. Daraus kann jedoch über die Größenordnung der notwendigen Schritt k_0 bis (3) eintritt (falls keine Nullstelle von f in $[x]^0$ liegt), keine Folgerung gezogen werden.

Wir wollen in dieser Note zeigen, daß k_0 klein ist, wenn der Durchmesser von $[x]^0$ klein ist. In einem unten näher präzisierten Sinne können wir von *quadratischer Divergenz* sprechen. Für eine Gleichung mit einer Unbekannten, d.h. für das eindimensionale Newton-Intervall-Verfahren, wurde dieses Ergebnis bereits in [2] bewiesen.

2. Hilfsmittel

Intervalle werden wir mit $[x] = [x_1, x_2]$, $[y] = [y_1, y_2]$, ..., Intervallvektoren mit $[x] = ([x]_i)$, $[y] = ([y]_i)$, ... und Intervallmatrizen mit $[X] = ([x]_{ij})$, $[Y] = ([y]_{ij})$, ... bezeichnen. Dabei sind die $[x]_i$, $[y]_i$, $[x]_{ij}$, $[y]_{ij}$, ... Intervalle.

$m[x] = (x_1 + x_2)/2$ bezeichnet den Mittelpunkt des Intervalles $[x] = [x_1, x_2]$. Der Abstand zweier Intervalle $[x]$ und $[y]$ ist durch $q([x], [y]) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$ definiert. $q([x], 0)$ heißt Betrag des Intervalles $[x]$, i.Z. $|[x]|$. Mit $d[x] = x_2 - x_1$ wird der Durchmesser des Intervalles $[x]$ bezeichnet. Für Intervallvektoren und Intervallmatrizen sind diese Begriffe über die Elemente erklärt.

Wendet man auf eine Intervallmatrix $[A]$ ("Koeffizientenmatrix") und einen Intervallvektor $[b]$ ("Rechte Seite") die Formeln des Gauss'schen Algorithmus an (Siehe dazu etwa [3], Kapitel 15), so läßt sich der berechnete Intervallvektor $[x] = \text{IGA}([A], [b])$ darstellen als ein Produkt aus Intervallmatrizen und dem Intervallvektor $[b]$. Dabei ist wegen der fehlenden Assoziativität durch geeignete Klammerung auf die Reihenfolge der Operationen zu achten. Wir geben hier die explizite Darstellung von $\text{IGA}([A], [b])$ nicht an. Man findet sie z.B. in [1] und [5]. Für das entsprechend geklammerte Produkt, welches allein aus den auftretenden Matrizen besteht, schreiben wir $\text{IGA}([A])$. Die Darstellung von $\text{IGA}([A])$ findet man z.B. in [1], Formel (4).

Für eine Frechet-differenzierbare Funktion $f: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist die Steigung $\delta f(x^1, x^2)$ eine (n, n) -Matrix, für die

$$f(x^1) - f(x^2) = \delta f(x^1, x^2)(x^1 - x^2), \quad x^1, x^2 \in D \quad (6)$$

gilt. Die Elemente von $\delta f(x^1, x^2)$ können durch geeignete Differenzenquotienten bzw. partielle Ableitungen der Komponenten von f dargestellt werden. Siehe dazu [4]. Es gilt für die intervallarithmetische Auswertung der Ableitung

$$\delta f(x^1, x^2) \in f'([x]), \quad x^1, x^2 \in [x]. \quad (7)$$

3. Quadratische Divergenz

Da für eine Intervallmatrix $[A]$ und einen Vektor b

$$\text{IGA}([A], b) \subseteq \text{IGA}([A]) \cdot b$$

gilt (siehe dazu [5], S. 33), erhält man für einen Intervallvektor $[x]$, für den $N[x]$ berechnet werden kann,

$$\begin{aligned} N[x] &= m[x] - \text{IGA}(f'([x]), f(m[x])) \\ &\subseteq m[x] - \text{IGA}(f'([x])) \cdot f(m[x]) =: \tilde{N}[x]. \end{aligned}$$

Fassen wir die unteren und oberen Schranken der Komponenten von $N[x]$ bzw. $\tilde{N}[x]$ jeweils zu reellen Vektoren n^1 und n^2 bzw. $\tilde{n}^1$ und $\tilde{n}^2$ zusammen, so ist die obige Inklusion $N[x] \subseteq \tilde{N}[x]$ äquivalent mit

$$\tilde{n}^1 \leq n^1 \leq n^2 \leq \tilde{n}^2. \quad (8)$$

Für eine Intervallmatrix $[A]$ und einen reellen Vektor b gilt (siehe dazu [3], (1) in Kapitel 10)

$$[A]b = \{Ab | A \in [A]\}.$$

Daher folgt aus der Definition von $\tilde{N}[x]$

$$\begin{cases} \tilde{n}^1 = m[x] - A^2 f(m[x]) \\ \tilde{n}^2 = m[x] - A^1 f(m[x]) \end{cases} \quad (9)$$

mit reellen Matrizen A^1 und $A^2 \in \text{IGA}(f'([x]))$. Mit (8) und (9) folgt nun unter Verwendung der Steigung (6), falls wir die rechten Endpunkte der Komponenten von $[x]$ zum Vektor x^2 zusammenfassen,

$$\begin{aligned} x^2 - n^1 &\leq x^2 - \tilde{n}^1 \\ &= x^2 - m[x] + A^2 f(m[x]) \\ &= x^2 - m[x] + A^2 \{f(x^2) + \delta f(m[x], x^2)(m[x] - x^2)\} \\ &= (I - A^2 \delta f(m[x], x^2))(x^2 - m[x]) + A^2 f(x^2). \end{aligned} \quad (10)$$

Wir setzen nun voraus, daß die intervallmäßige Auswertung $f'([x])$ der Ableitung von f einer Lipschitzbedingung (4) genügt. Dann kann man

$$\|d(\text{IGA}(f'([x])))\| \leq \kappa \|d[x]\| \quad (11)$$

folgern. (Der Beweis kann ähnlich wie der von Lemma 4 und 5 in [3], Kapitel 19, geführt werden). Wegen $\delta f(m[x], x^2) \in f'([x])$ ist $\delta f(m[x], x^2)$ nichtsingulär und es gilt $\delta f(m[x], x^2)^{-1} \in \text{IGA}(f'([x]))$. Somit kann man schreiben

$$\begin{aligned} &|(I - A^2 \delta f(m[x], x^2))(x^2 - m[x])| \\ &= |(\delta f(m[x], x^2)^{-1} - A^2) \cdot \delta f(m[x], x^2) \cdot (x^2 - m[x])| \\ &\leq d(\text{IGA}(f'([x]))) \cdot |\delta f(m[x], x^2)| \cdot d[x]. \end{aligned}$$

Wegen (11) können wir (9) daher schreiben als

$$x^2 - n^1 \leq O(\|d[x]\|^2) + A^2 f(x^2), \quad (12)$$

wobei $O(\|d[x]\|^2)$ einen Vektor aus $\mathbb{R}^n$ bezeichnet, dessen Komponenten alle die Größenordnung $O(\|d[x]\|^2)$ haben.

Vollständig analog erhält man, wenn man die linken Endpunkte der Komponenten von $[x]$ zum Vektor x^1 zusammengefaßt,

$$n^2 - x^1 \leq O(\|d[x]\|^2) - A^1 f(x^1). \quad (13)$$

Somit haben wir bewiesen

Lemma 2. Es sei $[x]$ ein Intervallvektor und $f: [x] \subset D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Abbildung, deren Frechet-Ableitung eine intervallmäßige Auswertung $f'([x])$ besitzt, für die (4) gilt. Dann gelten (12) und (13). $\square$

Lemma 3. Für hinreichend kleinen Durchmesser $d[x]$ von $[x]$ sind die beiden Matrizen A^1 und A^2 aus (9) (und damit aus (12) und (13)) nichtsingulär.

Beweis. Es genügt zu zeigen, daß für eine Intervallmatrix $[A]$, für die $\text{IGA}([A])$ existiert, $\text{IGA}([A])$ keine singuläre Matrix enthält, falls $d([A])$ hinreichend klein ist. Dies sieht man so:

Nach Definition von IGA ($[A]$) gilt $[A] \cdot \text{IGA}([A]) = I + [R]$, I Einheitsmatrix, mit einer "Restmatrix" $[R]$, für die $\lim d[R] \rightarrow 0$ für $d[A] \rightarrow 0$ gilt. Somit sind für hinreichend kleinen Durchmesser $d[A]$ alle reellen Matrizen aus $I + [R]$ invertierbar. Wegen

$$\{XY | X \in [A], Y \in \text{IGA}([A])\} \subseteq [A] \cdot \text{IGA}([A]) = I + [R]$$

sind dann auch alle reellen Matrizen $Y \in \text{IGA}([A])$ für hinreichend kleinen Durchmesser $d[A]$ invertierbar. $\square$

Wir zeigen als nächstes

Lemma 4. *Unter der Voraussetzung (4) und $f(x) \neq 0$ für $x \in [x]$ gibt es für hinreichend kleinen Durchmesser $d[x]$ mindestens ein $i_0 \in \{1, 2, \dots, n\}$ sodaß*

$$(A^1 f(x^1))_{i_0} \neq 0 \quad (\alpha)$$

und

$$\text{sign}(A^1 f(x^1))_{i_0} = \text{sign}(A^2 f(x^2))_{i_0}. \quad (\beta)$$

Beweis. Wegen $f(x^1) \neq 0$ und da A^1 nach Lemma 3 nichtsingulär ist, gilt $A^1 f(x^1) \neq 0$. Somit gilt (α) . Wegen $A^1, A^2 \in \text{IGA}(f'([x]))$ gilt

$$|A^2 - A^1| \leq d(\text{IGA}(f'([x])))$$

und daher

$$A^2 = A^1 + T \quad (14)$$

mit einer Matrix T , für die $\|T\| = O(\|d[x]\|)$ wegen (4) gilt. Aus (14) folgt daher

$$A^2 f(x^2) = A^1 f(x^2) + T f(x^2) \quad (15)$$

mit $T f(x^2) \rightarrow 0$ für $d[x] \rightarrow 0$. Außerdem gilt nach (6)

$$f(x^2) = f(x^1) + \delta f(x^2, x^1)(x^2 - x^1)$$

und daher nach Multiplikation mit A^1

$$A^1 f(x^2) = A^1 f(x^1) + A^1 \delta f(x^2, x^1)(x^2 - x^1).$$

Setzt man dies in (15) ein, so erhält man

$$A^2 f(x^2) = A^1 f(x^1) + A^1 \delta f(x^2, x^1)(x^2 - x^1) + T f(x^2) \quad (16)$$

mit $A^1 \delta f(x^2, x^1)(x^2 - x^1) \rightarrow 0$ und $T f(x^2) \rightarrow 0$ für $d[x] \rightarrow 0$. Nach (α) gibt es ein i_0 mit $(A^1 f(x^1))_{i_0} \neq 0$. Daher gilt wegen (16) für hinreichend kleinen Durchmesser $d[x]$

$$\text{sign}(A^2 f(x^2))_{i_0} = \text{sign}(A^1 f(x^1))_{i_0}.$$

Damit ist (β) gezeigt. $\square$

Es sei nun $\text{sign}(A^1 f(x^1))_{i_0} = 1$ und $d[x]$ hinreichend klein. Dann gilt nach (13)

$$(n^2 - x^1)_{i_0} \leq (O(\|d[x]\|^2) - A^1 f(x^1))_{i_0} < 0. \quad (17)$$

Falls $\text{sign}(A^1 f(x^1))_{i_0} = -1$ gilt, so ist nach (β) für hinreichend kleinen Durchmesser $d[x]$

$$\text{sign}(A^2 f(x^2))_{i_0} = -1$$

und daher nach (12)

$$(x^2 - n^1)_{i_0} \leq (O(\|d[x]\|^2) + A^2 f(x^2))_{i_0} < 0. \quad (18)$$

Man zeigt nun leicht

Lemma 5. *Es gilt $N[x] \cap [x] = \emptyset$ genau dann, wenn für (mindestens) ein $i_0 \in \{1, 2, \dots, n\}$ gilt*

$$(\tfrac{1}{2}d[x] + \tfrac{1}{2}dN[x])_{i_0} < |m[x] - m(N[x])|_{i_0}.$$

Dies ist wiederum genau dann der Fall, wenn für (mindestens) ein $i_0 \in \{1, 2, \dots, n\}$

$$(n^2 - x^1)_{i_0} < 0 \quad \text{oder} \quad (x^2 - n^1)_{i_0} < 0$$

gilt. □

Die Ungleichungen (17) und (18) zeigen, daß wegen des Termes $O(\|d[x]\|^2)$ auf der rechten Seite eine dieser beiden Ungleichungen "schnell" eintritt, wenn $d[x]$ klein ist. Wegen des Termes $O(\|d[x]\|^2)$ können wir von *quadratischer Divergenz* sprechen.

Literatur

- [1] Alefeld, G.: On the convergence of some interval-arithmetic modifications of Newton's method. In: Stepleman R. et al. (ed) Scientific Computing. IMACS/North-Holland Publishing Company 223–230 (1983).
- [2] Alefeld, G.: Über die Divergenzgeschwindigkeit des Intervall-Newton-Verfahrens. *Fasciuli Mathematici* 19, 9–13 (1989).
- [3] Alefeld, G., Herzberger, J.: Introduction to interval computations. New York London: Academic Press (1983).
- [4] Schmidt, J. W.: Die Regula Falsi für Operatoren in Banachräumen. *Z. Angew. Math. Mech.* 41, 61–63 (1961).
- [5] Schwandt, H.: Schnelle fast global konvergente Verfahren für die Fünf-Punkt-Diskretisierung der Poissongleichung mit Dirichletschen Randbedingungen auf Rechteckgebieten. Dissertation. Fachbereich Mathematik der TU Berlin (1981).

G. Alefeld
Institut für Angewandte Mathematik
Universität Karlsruhe
Kaiserstraße 12
D-W-7500 Karlsruhe
Bundesrepublik Deutschland