

Intervallanalytische Methoden bei nichtlinearen Gleichungen

von Götz Alefeld

1. Einleitung

Die Intervallrechnung stellt einen jungen Zweig der Numerischen Mathematik dar, dessen Ausbau nach wie vor in lebhafter Entwicklung begriffen ist. Während die meisten Regeln für Verknüpfungen reeller Intervalle in wesentlich allgemeinerem Zusammenhange bereits von R. C. Young [20] im Jahre 1931 bewiesen wurden, sind diese und ihre Eigenschaften erst zu Beginn des letzten Jahrzehnts in der Numerischen Mathematik systematisch verwendet worden. Es ist jedoch offensichtlich nicht allgemein bekannt, daß bereits in der 1951 in russischer Sprache erschienenen Enzyklopädie der Elementarmathematik, Band I, Arithmetik, die Gesetzmäßigkeiten der Intervallrechnung explizit verwendet wurden, um Einschließungen für Funktionswerte zu berechnen, wenn die Eingangsdaten mit Fehlern behaftet sind, d. h. in Intervallen liegen. Siehe dazu [7], S. 346 ff.. Die ersten, schon relativ weitreichenden Ergebnisse werden aufgrund dieser übersehenen Literaturstelle heute allgemein Sunaga [17] zugeschrieben. Unbestreitbar ist jedoch, daß mit den grundlegenden Ergebnissen von R. E. Moore und der zusammenfassenden Darstellung in seinem 1966 erschienenen Buch [13] das Interesse an der Intervallrechnung allgemein geweckt wurde. Eine weitere zusammenfassende Darstellung mit vielen neuen bedeutenden Ergebnissen wurde 1968 von U. Kulisch [11] verfaßt.

Daß die intensive Beschäftigung mit der Intervallrechnung gerade um 1960 einsetzte, war kein Zufall. Es war dies die Zeit, zu der bereits ein großer Personenkreis praktische Erfahrungen mit der numerischen Behandlung zahlreicher Probleme auf digitalen Rechenanlagen gemacht hatte. Die unangenehmen, damals noch weitgehend unerklärbaren Rundungsfehlereffekte hatten einen solchen Schock ausgelöst, daß man in der systematischen Verwendung der Intervallrechnung auf digitalen Rechenanlagen das Allheilmittel sah, um die Rundungsfehler zu kontrollieren. Es zeigte sich jedoch sehr bald, daß nur die durchdachte und durch sehr feinsinnige Überlegungen untermauerte Anwendung der Intervallrechnung hierbei zu praktisch verwertbaren Ergebnissen führt.

Die Tatsache, daß die systematische Ersetzung der reellen Größen eines im reellen Zahlenkörper hergeleiteten Algorithmus durch Intervalle bei der Durchführung auf einer Rechenanlage sehr häufig Ergebnisintervalle liefert, die so groß sind, daß sie praktisch wertlos sind, wurde von einigen Numeri-

schen Mathematikern als Anlaß genommen, um die Intervallrechnung als Ganzes als wertlos für die Numerische Mathematik anzusehen. Dabei wurde übersehen, daß die Ursache für diesen Sachverhalt nicht bei dem Hilfsmittel Intervallrechnung zu suchen ist, sondern in der in diesem Falle naiven Anwendung dieses Hilfsmittels begründet liegt. Es wurde weiter außer acht gelassen, daß die Beschäftigung mit der Erfassung von Rundungsfehlern durch die Intervallrechnung zu einer sauberen Beschreibung der beim numerischen Rechnen auf einer Rechenanlage vorliegenden Räume geführt hat. Eine Zusammenfassung dieser Ergebnisse findet man in dem Buch von U. Kulisch [12] sowie in Albrecht – Kulisch [1].

Unabhängig davon gibt es weitreichende Anwendungen der Intervallrechnung in der Numerischen Mathematik, die nichts mit der Erfassung von Rundungsfehlern zu tun haben. Es ist die Absicht dieses Aufsatzes, auf einige der von Rundungsfehlern unabhängigen Anwendungen der Intervallrechnung hinzuweisen. Dabei kann natürlich keinerlei Vollständigkeit angestrebt werden.

Die im folgenden verwendeten Bezeichnungen schließen sich eng an die in [3], S. 397–398, benutzten an und werden aus Platzgründen nicht nochmals aufgelistet. Lemmata und Sätze werden abschnittsweise fortlaufend numeriert. Entsprechend wird auf sie Bezug genommen.

2. Der Gaußsche Algorithmus bei Gleichungen mit Intervallen als Koeffizienten

Gegeben sei eine reelle Intervallmatrix $A = (A_{ij})$ und ein Intervallvektor $b = (B_j)$. Falls der Gaußsche Algorithmus mit der Intervallmatrix A durchführbar ist – dies ist der Fall, falls alle auftretenden Nennerintervalle nicht die Null enthalten, was eventuell durch Zeilen- oder Spaltenvertauschungen erreicht werden kann –, so liefert er einen Intervallvektor x mit der Eigenschaft

$$\{A^{-1}b \mid A \in A, b \in b\} \subseteq x,$$

d.h. der Intervallvektor x schließt die Lösungen der linearen Gleichungssysteme

$$Ax = b, \quad A \in A, b \in b,$$

ein. Einen Beweis für diesen Sachverhalt findet man z.B. in [3], S. 218 ff.

Die Aufgabe, einen Intervallvektor mit der angegebenen Eigenschaft zu bestimmen, ist als solche von Interesse. Sie liegt z.B. vor, wenn man die Koeffizienten eines linearen Gleichungssystems bei der Eingabe in eine Rechenmaschine runden muß, und daher die exakten Koeffizienten nur durch auf der Maschine darstellbare Schranken einschließen kann. Viel wichtiger im Rahmen dieses Aufsatzes sind Aufgabenstellungen, bei denen die Bestimmung eines Einschließungsvektors für die Lösungsmenge eines linearen Gleichungssystems mit Intervallen als Koeffizienten ohne jeden

Bezug auf fehlerhafte Eingangsdaten oder Rundungen während der Rechnung durchgeführt wird. (Siehe Abschnitte 3 und 4).

Während für Punktgleichungssysteme, d.h. für Gleichungssysteme mit einer Punktmatrix \mathring{A} als Koeffizientenmatrix, bekannt ist, daß bei rundungsfehlerfreier Rechnung die Durchführbarkeit des Gaußschen Algorithmus – möglicherweise unter Verwendung notwendiger Zeilenvertauschungen – gesichert ist, konnte für Intervallmatrizen bisher eine analoge Aussage nicht bewiesen werden. Aufgrund der speziellen Gesetzmäßigkeiten der Intervallrechnung ist sogar nicht auszuschließen, daß die Ausgangsmatrix A keine singuläre Matrix enthält und der Gaußsche Algorithmus nicht durchführbar ist. Eine Begründung dafür ist in [3], S. 221 ff. dargelegt.

Allerdings ist bisher kein Beispiel bekannt, bei dem die Ausgangsmatrix A keine singuläre Matrix enthält und der Gaußsche Algorithmus abbricht*. Das Auffinden eines solchen Beispiels wird durch zwei Tatsachen erschwert: Erstens ist es relativ schwierig nachzuweisen, daß eine Intervallmatrix A keine singuläre Punktmatrix \mathring{A} enthält. Zweitens muß die Ordnung n des Gleichungssystems größer als zwei sein, womit die systematische explizite Diskussion mit Bleistift und Papier recht kompliziert wird. Daß die Anzahl n der Unbekannten mindestens gleich drei sein muß, ergibt sich aus dem folgenden Satz.

Satz 2.1: *Ist $1 \leq n \leq 2$ und enthält die $n \times n$ -Intervallmatrix $A = (A_{ij})$ keine singuläre Matrix \mathring{A} , so ist der Gaußsche Algorithmus durchführbar.*

Beweis: Im Falle $n = 1$ ist nach Voraussetzung $0 \notin A_{11} = A$, womit alles gezeigt ist. Für $n = 2$ können nicht beide Intervalle A_{11} und A_{21} die Null enthalten, da sonst A eine singuläre Punktmatrix \mathring{A} enthält. Es gelte o.E.d.A. $0 \notin A_{11}$. Anderenfalls vertausche man die beiden Zeilen von A . Dann liefert der Gaußsche Algorithmus (siehe [3], S. 218 ff.) für das Element A'_{22} das Intervall

$$A'_{22} = A_{22} - \frac{1}{A_{11}} A_{21} A_{12}.$$

A'_{22} kann aufgefaßt werden als intervallmäßige Auswertung ([3], S. 28 ff.) der rationalen Funktion a'_{22} der vier Variablen a_{11} , a_{21} , a_{12} und a_{22} , definiert durch

$$a'_{22}(a_{11}, a_{12}, a_{21}, a_{22}) = a_{22} - \frac{1}{a_{11}} a_{21} a_{12}.$$

Nach Voraussetzung ist für jedes $A \in \mathcal{A}$

$$\det(A) = a_{11} a_{22} - a_{21} a_{12} \neq 0,$$

* Inzwischen wurde von L. Platzöder eine 4×4 Matrix angegeben, für die dies der Fall ist.

also auch

$$a'_{22}(a_{11}, a_{12}, a_{21}, a_{22}) = \frac{1}{a_{11}} \det(A) \neq 0.$$

Da im Funktionsausdruck für a'_{22} jede Variable genau einmal auftritt, liefert die intervallmäßige Auswertung den Wertebereich der Funktion a'_{22} für $a_{11} \in A_{11}$, $a_{12} \in A_{12}$, $a_{21} \in A_{21}$, $a_{22} \in A_{22}$. Somit gilt $0 \notin A'_{22}$ und der Gaußsche Algorithmus ist durchführbar. ■

Der angegebene Beweis läßt sich für $n \geq 3$ nicht übertragen. Abgesehen von den in Satz 2.1 beschriebenen trivialen Fällen sind einige weitere Klassen von Intervallmatrizen bekannt, für die die Durchführbarkeit nachgewiesen werden kann. Wir zitieren zunächst die Ergebnisse aus [2].

Satz 2.2: *Es sei $A = (A_{ij})$ eine reelle Intervallmatrix. Aus A werde die Punktmatrix $B = (b_{ij})$ mit*

$$b_{ij} = \begin{cases} |m(A_{ii})| - \frac{1}{2} d(A_{ii}) & i = j \\ -|A_{ij}| & \text{sonst} \end{cases}$$

gebildet. Falls B eine M-Matrix ist, so ist der Gaußsche Algorithmus mit der Intervallmatrix A durchführbar (und zwar sogar ohne Spalten und Zeilenvertauschungen). ■

Eine reelle Matrix B heißt dabei bekanntlich M-Matrix, falls außerhalb der Diagonale nur nichtpositive Elemente stehen, B nichtsingulär ist, und alle Elemente von B^{-1} nichtnegativ sind (Siehe z.B. [19], S. 85).

Die Voraussetzungen von Satz 2.2 sind beispielsweise erfüllt, wenn die Intervallmatrix streng diagonaldominant ist, d.h. wenn

$$|m(A_{ii})| - \frac{1}{2} d(A_{ii}) > \sum_{\substack{j=1 \\ j \neq i}}^n |A_{ij}|, \quad 1 \leq i \leq n$$

gilt. Die hier definierte Eigenschaft „streng diagonaldominant“ für Intervallmatrizen ist eine unmittelbare Verallgemeinerung der entsprechenden Definition für Punktmatrizen und geht im Spezialfall, daß A eine Punktmatrix ist, in diese über.

Intervallmatrizen, welche die Voraussetzungen von Satz 2.2 erfüllen, treten bei der numerischen Auflösung nichtlinearer Gleichungssysteme auf, welche durch Diskretisierung von Differentialgleichungen entstehen. (Siehe Abschnitt 4). Aus diesem Grunde ist das Ergebnis von Satz 2.2 von erheblichem praktischen Interesse.

Wir erwähnen noch, daß für eine Teilmenge der in Satz 2.2 enthaltenen Menge von Intervallmatrizen die Durchführbarkeit des Gaußschen Algorithmus in etwas anderem Zusammenhange in [5] gezeigt wurde.

Es sei hier weiter erwähnt, daß auch für die von Gargantini und Henrici

[6] eingeführte Arithmetik für Kreisscheiben der komplexen Zahlenebene die Aussagen von Satz 2.2 sinngemäß bestehen (siehe [2]). Dagegen ist Satz 2.1 in diesem Falle nicht richtig.

Schließlich hat K. Reichmann [16] eine Klasse von Hessenbergmatrizen mit Intervallen als Koeffizienten angegeben, für die der Gaußsche Algorithmus stets durchführbar ist. (Eine Matrix heißt obere Hessenbergmatrix, falls unterhalb der ersten Nebendiagonale alle Elemente verschwinden). Unter geeigneten Voraussetzungen über die auf der Rechenmaschine vorliegende Rundung gilt dies sogar bei der Durchführung mit gerundeter Intervallrechnung.

Es wäre wünschenswert, wenn man weitere Klassen von Intervallmatrizen angeben könnte, für die die Durchführbarkeit des Gaußschen Algorithmus gezeigt werden kann.

3. Existenzaussagen für Lösungen nichtlinearer Gleichungssysteme

Wir wollen in diesem Abschnitt darlegen, wie unter Verwendung intervallarithmetischer Hilfsmittel praktisch nachprüfbar, d.h. zum Beispiel auf einer digitalen Rechenanlage programmierbare Aussagen für die Existenz bzw. Nichtexistenz einer Lösung eines nichtlinearen Gleichungssystems gewonnen werden können. Alle Aussagen beruhen im wesentlichen auf dem nachfolgenden Lemma 3.1, welches eine einfache Anwendung des Brouwerschen Fixpunktsatzes darstellt. Der direkte Nachweis, daß die Voraussetzungen des Brouwerschen Fixpunktsatzes vorliegen, nämlich daß eine kompakte konvexe Menge des $V_n(\mathbb{R})$ durch eine stetige Abbildung in sich abgebildet wird, ist im allgemeinen mit erheblichen Schwierigkeiten verbunden. Dagegen kann mit Hilfe intervallarithmetischer Operationen das Vorliegen dieser Eigenschaft zumindest für Intervallvektoren unmittelbar nachgeprüft werden.

Lemma 3.1: *Es sei $f : x \subseteq D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ eine stetige Abbildung. Desgleichen sei $Y : x \subseteq D \subseteq V_n(\mathbb{R}) \rightarrow M_{nn}(\mathbb{R})$ stetig und für alle $x \in x$ existiere die Inverse $Y(x)^{-1}$. Für die Abbildung $p : x \rightarrow V_n(\mathbb{R})$ mit*

$$p(x) = x - Y(x) f(x).$$

gelte $p(x) \in x$ für $x \in x$. Dann besitzt f eine Nullstelle in x . ■

Der einfache *Beweis* ergibt sich aus dem Brouwerschen Fixpunktsatz: p bildet die kompakte und konvexe Menge $\{x | x \in x\} \subseteq V_n(\mathbb{R})$ in sich ab. p besitzt daher einen Fixpunkt in x . Da $Y(x)$ nichtsingulär ist, besitzt f eine Nullstelle in x .

Lemma 3.1 und der Beweis sind triviale Verallgemeinerungen der entsprechenden für konstantes Y bewiesenen Aussagen von R. E. Moore [14].

Lemma 3.2: Es sei $f: X \subseteq D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ in D (im Sinne von Frechet) differenzierbar. Dann gilt für ein festes $y \in X$ und für alle $x \in X$

$$f(y) - f(x) = Z(x)(y - x)$$

mit der Matrix

$$Z(x) = \begin{pmatrix} f'_1(x + t_1(y - x)) \\ \vdots \\ f'_n(x + t_n(y - x)) \end{pmatrix} \in M_{nn}(\mathbb{R}),$$

wobei $t_i \in (0, 1)$, $i = 1(1)n$. ■

Den Beweis findet man (unter etwas schwächeren Voraussetzungen) in [15], S. 68/69.

Mit Hilfe dieser beiden Lemmata lassen sich eine Reihe von Aussagen beweisen, die die Existenz einer Nullstelle einer Abbildung f garantieren, oder eine Menge liefern, in der keine Nullstelle von f liegt. Wichtig ist dabei, daß diese Aussagen durch Ausführung von intervallarithmetischen Operationen konkret nachprüfbar sind.

Satz 3.3: Es sei $f: D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ in D differenzierbar. Die Ableitung sei für $x \in D$ intervallmäßig auswertbar (siehe [3], S. 28 ff.). Die intervallmäßige Auswertung $f'(x)$ sei zerlegt in

$$f'(x) = D(x) + B(x).$$

(Dabei bezeichne $D(x)$ den in der Hauptdiagonale stehenden Anteil von $f'(x)$ und $B(x)$ den verbleibenden Rest, sowie $D(x)$ und $B(x)$ die entsprechenden intervallmäßigen Auswertungen). Es sei

$$0 \notin \frac{\partial}{\partial x_i} f_i(x), \quad i = 1(1)n, \quad \text{und} \quad \bar{D}(x) := \text{diag} \left(1 / \frac{\partial}{\partial x_i} f_i(x) \right).$$

Gilt dann für ein $y \in X$

$$y - \bar{D}(x) \{B(x)(y - x) + f(y)\} \subseteq X$$

so besitzt f in X eine Nullstelle. Ist dagegen

$$X \cap \{y - \bar{D}(x) \{B(x)(y - x) + f(y)\}\} = \emptyset,$$

so besitzt f in X keine Nullstelle.

Beweis: Es sei $y \in X$ fest und $x \in X$ beliebig. Wir betrachten die Abbildung

$$p: \{x \mid x \in X\} \subset V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$$

mit

$$p(x) = x - \bar{D}(x)^{-1} f(x).$$

Dabei ist $\bar{D}(x)$ über die nach Lemma 3.2 bestehende Gleichung $f(y) - f(x) = Z(x)(y - x)$ als Diagonalanteil von $Z(x)$ definiert. Wir setzen außerdem $\bar{B}(x) := \bar{D}(x) - Z(x)$. Aufgrund der Teilmengeneigenschaft ist $\bar{D}(x) \in D(x)$. Somit existiert $\bar{D}(x)^{-1}$ für jedes $x \in X$ und für $p(x)$

gilt aufgrund der Teilmengeneigenschaft

$$\begin{aligned} p(x) &= x - \bar{D}(x)^{-1} f(x) \\ &= y - \bar{D}(x)^{-1} f(y) + x - y - \bar{D}(x)^{-1} (f(x) - f(y)) \\ &= y - \bar{D}(x)^{-1} f(y) + \bar{D}(x)^{-1} \bar{B}(x) (x - y) \\ &= y - \bar{D}(x)^{-1} \{ \bar{B}(x) (y - x) + f(y) \} \\ &\in y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \} \\ &\subseteq x. \end{aligned}$$

Somit ist für alle $x \in X$ die Beziehung $p(x) \in X$ erfüllt. Nach Lemma 3.1 folgt der erste Teil der Behauptung.

Um den zweiten Teil zu beweisen, nehmen wir an, daß f eine Nullstelle x^* in X besitzt. Dann gilt für die oben definierte Abbildung p wie vorher für beliebiges $y \in X$

$$\begin{aligned} p(x^*) &= x^* - \bar{D}(x^*)^{-1} f(x^*) \\ &\in y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \} \end{aligned}$$

Wegen $x^* \in X$ müßte somit

$$X \cap \{ y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \} \} \neq \emptyset$$

sein. Widerspruch. ■

Wir wollen noch darauf hinweisen, daß die Existenzaussage aus dem bewiesenen Satz noch etwas präzisiert werden kann:

Gilt

$$y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \} \subseteq X,$$

so besitzt f in $y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \}$ eine Nullstelle.

Unter der genannten Voraussetzung existiert – wie bewiesen – in X eine Nullstelle. Wir zeigen, daß jede Nullstelle x^* von f in X auch in $y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \}$ liegt. Dazu gehen wir von der in Lemma 3.2 angegebenen Gleichung mit einem $y \in X$ und mit $x := x^*$ aus, und lösen diese Gleichung wie folgt nach x^* auf:

$$x^* = y - \bar{D}(x^*)^{-1} \{ \bar{B}(x^*) (y - x^*) + f(y) \}.$$

Dann gilt aufgrund der Teilmengeneigenschaft

$$x^* \in y - \bar{D}(x) \{ \bar{B}(x) (y - x) + f(y) \}$$

Die gleiche Bemerkung gilt für die nachfolgenden Sätze.

Im wesentlichen analog wie Satz 3.3 kann die nächste Aussage bewiesen werden.

Satz 3.4: Es sei $f: D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ in D differenzierbar. Die Ableitung sei für $x \subseteq D$ intervallmäßig auswertbar. Der Gaußsche Algorithmus sei mit der intervallmäßigen Auswertung $f'(x)$ durchführbar (siehe Abschnitt 2) und liefere mit der rechten Seite $f(y)$ für ein $y \in x$ den Intervallvektor z . Gilt dann

$$y - z \subseteq x,$$

so besitzt f in x eine Nullstelle. Gilt dagegen

$$x \cap \{y - z\} = \emptyset,$$

so enthält x keine Nullstelle von f .

Beweis: Da der Gaußsche Algorithmus mit der Intervallmatrix $f'(x)$ durchführbar ist, sind alle Punktmatrizen aus $f'(x)$, insbesondere die in Lemma 3.2 definierte Matrix $Z(x)$ nichtsingulär. Wir betrachten die Abbildung

$$p: x \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$$

mit

$$p(x) = x - Z(x)^{-1} f(x).$$

Es gilt

$$\begin{aligned} p(x) &= x - Z(x)^{-1} f(y) + Z(x)^{-1} (f(y) - f(x)) \\ &= y - Z(x)^{-1} f(y) \\ &\in y - z. \end{aligned}$$

Gilt somit $y - z \subseteq x$, so gilt $p(x) \in x$ für $x \in x$ und nach Lemma 3.1 besitzt f in x eine Nullstelle. Der Rest des Beweises kann analog wie im vorangehenden Satz erbracht werden. ■

Der Beweis des vorangehenden Satzes zeigt, daß man auf die Anwendung des Gaußschen Algorithmus verzichten kann, wenn man eine Intervallmatrix V mit der Eigenschaft

$$Z(x)^{-1} \in V \text{ für } x \in x$$

kennt. Man erhält dann mit der im vorangehenden Satz definierten Abbildung p

$$p(x) = x - Z(x)^{-1} f(x) \in y - Vf(y).$$

Somit gilt

Satz 3.5: Es sei $f: D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ in D differenzierbar. Die Ableitung sei für $x \subseteq D$ intervallmäßig auswertbar. Mit der oben definierten Intervallmatrix V gelte für ein festes $y \in x$

$$y - Vf(y) \subseteq x.$$

Dann besitzt f in x eine Nullstelle. Gilt dagegen

$$\{y - Vf(y)\} \cap x = \emptyset,$$

so gibt es in x keine Nullstelle von f . ■

Im allgemeinen wird man die Matrix V mit der in diesem Satz geforderten Eigenschaft nur durch n -malige Anwendung des Gaußschen Algorithmus auf die Intervallmatrix $f'(x)$ mit den Spalten der Einheitsmatrix als rechten Seiten bestimmen können. Eine entsprechende Aussage wie in den vorangehenden Sätzen, bei der jedoch die Durchführung des Gaußschen Algorithmus mit einer Intervallmatrix nicht notwendig ist, wurde von R. E. Moore [14] bewiesen:

Satz 3.6: *Es sei $f: D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ in D differenzierbar. Die Ableitung sei für $x \in D$ intervallmäßig auswertbar. Gilt dann für ein $y \in x$ und eine nichtsinguläre Punktmatrix $Y \in M_{nn}(\mathbb{R})$ und mit der Einheitsmatrix I*

$$y - Yf(y) + (I - Yf'(x))(x - y) \subseteq x,$$

so besitzt f in x eine Nullstelle. Ist dagegen

$$x \cap \{y - Yf(y) + (I - Yf'(x))(x - y)\} = \emptyset,$$

so gibt es keine Nullstelle in x . ■

Es ist naheliegend, für Y die Wahl $m(f'(x))^{-1}$ zu treffen. Dies erfordert dann die Invertierung einer Punktmatrix und zur Nachprüfung der Inklusionsbeziehungen noch einige Multiplikationen von Matrizen mit Matrizen bzw. Vektoren. Diese Multiplikationen kann man sich ersparen, indem man ein lineares Gleichungssystem mit einer Punktmatrix als Koeffizientenmatrix und einem Intervallvektor als rechte Seite auflöst. Dies ist im folgenden Satz präzisiert.

Satz 3.7: *Es sei $f: D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ in D differenzierbar. Die Ableitung sei für $x \in D$ intervallmäßig auswertbar. Die reelle Punktmatrix \tilde{Y} sei nichtsingulär. Der Gaußsche Algorithmus werde auf die Punktmatrix \tilde{Y} und die rechte Seite $f(y) - (\tilde{Y} - f'(x))(x - y)$ mit einem $y \in x$ angewandt und liefere einen Intervallvektor z . Es gelte $y - z \in x$. Dann besitzt f in x eine Nullstelle. Ist dagegen $x \cap (y - z) = \emptyset$, so gibt es keine Nullstelle in x .*

Beweis. Wir betrachten die Abbildung $p: x \subseteq D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$ mit

$$p(x) = x - \tilde{Y}^{-1} f(x)$$

Dann gilt

$$\begin{aligned} p(x) &= y - \tilde{Y}^{-1} f(y) + x - y - \tilde{Y}^{-1} (f(x) - f(y)) \\ &= y - \tilde{Y}^{-1} f(y) + \tilde{Y}^{-1} (\tilde{Y} - Z(x))(x - y) \end{aligned}$$

oder

$$\tilde{Y}(p(x) - y) = -f(y) + (\tilde{Y} - Z(x))(x - y).$$

Wegen $x \in x$ und $Z(x) \in f'(x)$ gilt

$$-f(y) + (\tilde{Y} - Z(x))(x - y) \in -f(y) + (\tilde{Y} - f'(x))(x - y).$$

Wendet man daher auf das lineare Gleichungssystem mit \tilde{Y} als Koeffizientenmatrix und $f(y) - (\tilde{Y} - f'(x))(x - y)$ als rechter Seite den Gaußschen

Algorithmus an, so erhält man einen Intervallvektor z , für welchen $p(\bar{x}) - y \in -z$ für $\bar{x} \in x$ gilt. Ist daher $y - z \subseteq x$, so ist $p(\bar{x}) \in x$ für $\bar{x} \in x$ und der Beweis kann wie in den vorangehenden Sätzen zu Ende geführt werden. ■

Die Aussage dieses Satzes ist nicht nur wegen der Aufwandsersparnis im Vergleich zum Moore'schen Ergebnis aus Satz 3.6 von Interesse. Gegenüber dem analogen Resultat aus Satz 3.4 besteht der Hauptvorteil darin, daß die Durchführbarkeit des Gaußschen Algorithmus bei einer nichtsingulären Punktmatrix gesichert ist.

In den vorangehenden Sätzen spielt die intervallmäßige Auswertung $f'(x)$ der Ableitung der Abbildung f eine wesentliche Rolle. Es ist bemerkenswert (und nicht nur von theoretischem Interesse), daß man in allen diesen Sätzen in der intervallmäßigen Auswertung der Ableitung gewisse Intervalle durch reelle Zahlen ersetzen kann. Um dies zu konkretisieren, bemerken wir zunächst, daß in dem beim Beweis aller dieser Aussagen verwendeten Lemma 3.2 die Matrix $Z(\bar{x})$ durch eine Matrix $\bar{Z}(\bar{x})$ – wie nachfolgend angegeben – ersetzt werden kann.

Lemma 3.2': *Es sei $f : D \subseteq V_n(\mathbb{R})$ in D differenzierbar. Dann gilt für festes $y \in x \subseteq D$ und $\bar{x} \in x \subseteq D$*

$$f(y) - f(\bar{x}) = \bar{Z}(\bar{x}) (y - \bar{x})$$

wobei die j -te Spalte $\bar{Z}(\bar{x})_j$ der Matrix $\bar{Z}(\bar{x})$ sich darstellt als

$$\bar{Z}(\bar{x})_j = \begin{cases} \frac{1}{y_j - x_j} [f(y + \sum_{k=1}^{j-1} (x_k - y_k) e^{(k)}) - f(y + \sum_{k=1}^j (x_k - y_k) e^{(k)})] & \text{für } y_j \neq x_j \\ f'(y + \sum_{k=1}^j (x_k - y_k) e^{(k)}) e^{(j)} & \text{für } y_j = x_j. \end{cases}$$

($e^{(k)}$ bezeichnet den k -ten Einheitsvektor). ■

Der Beweis erfolgt durch elementares Nachrechnen. Die Matrix $\bar{Z}(\bar{x})$ wird bei J. W. Schmidt [18] als Steigung bezeichnet. Sie tritt bei der Übertragung der Regula-Falsi auf Gleichungssysteme auf.

Bei der Verwendung der in Lemma 3.2' angegebenen Gleichung zur Herleitung entsprechender Aussagen wie in den Sätzen 3.3 bis 3.7 kommt es nur darauf an, die reelle Matrix $\bar{Z}(\bar{x})$ für ein festes $y \in x$ und für alle $\bar{x} \in x$ durch eine Intervallmatrix einzuschließen. Dazu betrachten wir zunächst für $y_j \neq x_j$ das Element $\bar{z}_{ij}(\bar{x})$ der Matrix $\bar{Z}(\bar{x})$. Durch Anwendung des Mittelwertsatzes erhält man

$$\bar{z}_{ij}(\bar{x}) = \frac{1}{y_j - x_j} [f_i(x_1, \dots, x_{j-1}, y_j, \dots, y_n) - f_i(x_1, \dots, x_j, y_{j+1}, \dots, y_n)]$$

$$= \frac{\partial}{\partial x_j} f_i(x_1, \dots, x_{j-1}, y_j + \theta_{ij}(x_j - y_j), y_{j+1}, \dots, y_n)$$

mit $\theta_{ij} \in (0, 1)$. Wegen $y_j + \theta_{ij}(x_j - y_j) \in X_j$ gilt dann aufgrund der Teilmengeneigenschaft

$$\tilde{z}_{ij}(x) \in \frac{\partial}{\partial x_j} f_i(X_1, \dots, X_j, y_{j+1}, \dots, y_n),$$

für ein festes $y = (y_i) \in x$ und alle $x \in x = (X_i)$. Dieselbe Relation erhält man im Falle $y_j = x_j$.

Es gilt somit $\tilde{Z}(x) \in \tilde{f}'(x)$ für $x \in x$, wobei die Elemente der Intervallmatrix $\tilde{f}'(x)$ definiert sind durch

$$(\tilde{f}'(x))_{ij} = \frac{\partial}{\partial x_j} f_i(X_1, \dots, X_j, y_{j+1}, \dots, y_n),$$

$$i, j = 1(1)n.$$

Aufgrund der Teilmengeneigenschaft gilt $\tilde{f}'(x) \subseteq f'(x)$.

Somit gelten die Aussagen der Sätze 3.3 bis 3.7 mit der Intervallmatrix $\tilde{f}'(x)$ an Stelle von $f'(x)$ sicherlich dann, wenn sie mit $f'(x)$ gelten.

Die Tatsache, daß man die Matrix $f'(x)$ durch $\tilde{f}'(x)$ ersetzen kann, wurde von E. Hansen in [8] verwendet. Um die gleiche Zeit wurde auch von J. Herzberger [9] darauf hingewiesen, daß in den Anwendungen gewöhnlich die Einschließung der Steigung (an Stelle der Ableitung) ausreichend ist. Siehe dazu auch [10].

4. Einige Iterationsverfahren zur Auflösung nichtlinearer Gleichungssysteme

Bei der Nachprüfung der Voraussetzungen der im vorangehenden Abschnitt angegebenen Sätze können gewöhnlich drei Fälle auftreten, die wir exemplarisch für Satz 3.4 explizit formulieren:

- (I) Es ist $y - z \subseteq x$.
- (II) Es ist $(y - z) \cap x = \phi$.
- (III) Es ist $(y - z) \cap x \neq \phi$ und $y - z \not\subseteq x$.

Im Falle (II) gibt es in x keine Nullstelle, im Falle (I) existiert in x eine Nullstelle und im Falle (III) kann eine Nullstelle von f in x existieren. Es ist daher naheliegend, in den Fällen (I) und (III) mit $(y - z) \cap x$ an Stelle von x nochmals die Voraussetzungen von Satz 3.4 zu überprüfen und damit (im Falle (I)) eine eventuell bessere Einschließung der Nullstelle oder (im Falle III)) die Existenz oder Nichtexistenz einer Nullstelle nachzuweisen.

Bezeichnen wir den Ausgangsvektor mit $x^{(0)}$ und wählen für y den reellen Vektor $m(x^{(0)})$ – allgemein $x^{(k)}$ und $m(x^{(k)})$ – so führt dies auf die folgende Iterationsvorschrift (wir formulieren diese für $\tilde{f}'(x)$). Beim prakti-

schen Rechnen sollte man aus den genannten Gründen die in Abschnitt 3 angegebene Matrix $\underline{f}'(\underline{x})$ verwenden).

Verfahren I:

- $$\left\{ \begin{array}{l} \text{(a) Anwendung des Gaußschen Algorithmus auf die Intervallmatrix} \\ \underline{f}'(\underline{x}^{(k)}) \text{ mit der rechten Seite } \underline{f}(m(\underline{x}^{(k)})) \text{ liefert } w^{(k)} \\ \text{(b) } \underline{x}^{(k+1)} = \{m(\underline{x}^{(k)}) - w^{(k)}\} \cap \underline{x}^{(k)} \\ k = 0, 1, 2, \dots \end{array} \right.$$

Enthält $\underline{x}^{(0)}$ eine Nullstelle von \underline{f} , so ist diese (siehe die Bemerkungen nach Satz 3.3) auch in $m(\underline{x}^{(0)}) - w^{(0)}$ und damit in $\underline{x}^{(1)}$, also in allen $\underline{x}^{(k)}$ enthalten. Aufgrund der Durchschnittsbildung konvergiert die Folge

$\{\underline{x}^{(k)}\}_{k=0}^{\infty}$ gegen einen Intervallvektor \underline{x}^* , der die Lösung \underline{x}^* von $\underline{f}(\underline{x}) = 0$ enthält.

Die Konvergenz der Folge $\underline{x}^{(k)}$ gegen \underline{x}^* kann allein unter der Voraussetzung, daß das Verfahren durchführbar ist (was identisch damit ist, daß der Gaußsche Algorithmus durchführbar ist) nicht bewiesen werden. Wir ändern daher die Iterationsvorschrift etwas ab und zeigen unter Verwendung einiger zusätzlicher Voraussetzungen die Konvergenz.

Verfahren I':

- $$\begin{array}{l} \text{(a) Anwendung des Gaußschen Algorithmus auf die Intervallmatrix} \\ \underline{f}'(\underline{x}^{(k)}) = \underline{D}(\underline{x}^{(k)}) - \underline{B}(\underline{x}^{(k)}), (\underline{D}(\underline{x}^{(k)})) \text{ Diagonalanteil der intervall-} \\ \text{mäßigen Auswertung von } \underline{f}', \underline{B}(\underline{x}^{(k)}) \text{ der außerhalb der Diagonale} \\ \text{verbleibende Rest) mit der rechten Seite } \underline{f}(m(\underline{x}^{(k)})) \text{ liefert } w^{(k)} \\ \text{(b) } \tilde{\underline{x}}^{(k+1)} = \{m(\underline{x}^{(k)}) - w^{(k)}\} \cap \underline{x}^{(k)}. \\ \text{(c) } \underline{x}^{(k+1)} = \{m(\underline{x}^{(k)}) - \underline{\tilde{D}}(\underline{x}^{(k)}) \{ \underline{B}(\underline{x}^{(k)}) (m(\underline{x}^{(k)}) - \tilde{\underline{x}}^{(k)}) + \\ + \underline{f}(m(\underline{x}^{(k)})) \} \} \cap \tilde{\underline{x}}^{(k)} \end{array}$$

wobei $\underline{\tilde{D}}(\underline{x}^{(k)}) = \text{diag}(1/\frac{\partial}{\partial x_i} f_i(\underline{x}^{(k)}))$.

$k = 0, 1, 2, \dots$

Teilabschnitt (c) erfordert bei einer vollbesetzten Matrix ungefähr n^2 (Intervall-)Operationen, ist also vom Aufwand her gegenüber Teilschritt (a), der ungefähr $n^3/3$ (Intervall-)Operationen benötigt, für größeres n vernachlässigbar.

Für das angegebene Verfahren I' gelten die folgenden Aussagen.

Satz 4.1: Es sei $\underline{x}^{(0)} \subseteq D \subseteq V_n(\mathbb{R})$ ein Intervallvektor und $\underline{x}^* \in \underline{x}^{(0)}$ eine Nullstelle der Abbildung $\underline{f} : D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$. Die intervallmäßige Auswertung $\underline{f}'(\underline{x}^{(0)})$ der Frechet-Ableitung erfülle die Voraussetzungen von Satz 2.2. Dann gilt

- (1) Das angegebene Verfahren ist für alle $k \geq 0$ durchführbar und es gilt $\lim_{k \rightarrow \infty} x^{(k)} = x^*$.
- (2) Jede Iterierte $x^{(k)}$, $k \geq 0$, enthält die Nullstelle x^* .
- (3) Es gilt

$$\|d(x^{(k+1)})\| \leq c \|d(x^{(k)})\|^2$$

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

Wir können auf den Beweis hier nicht eingehen und verweisen stattdessen auf [3], S. 288 ff., wo unter etwas spezielleren Voraussetzungen an $f'(x^{(0)})$ die gleichen Aussagen bewiesen wurden.

Der angegebene Satz ist von erheblichem praktischen Interesse. Betrachten wir etwa die nichtlineare Randwertaufgabe

$$y'' = g(t, y), \quad 0 \leq t \leq 1$$

mit $y(0) = \alpha$, $y(1) = \beta$, stetigem g und $g_y \geq 0$, so führt die Diskretisierung mit dem gewöhnlichen Differenzenverfahren (und bei genügend kleinem h auch mit dem verbesserten Differenzenverfahren) auf ein nichtlineares Gleichungssystem, welches – wie das stetige Problem – genau eine Lösung besitzt. Diese Lösung läßt sich sehr einfach durch einen Intervallvektor $x^{(0)}$ einschließen, während die tatsächliche Berechnung auf die bekannten Schwierigkeiten führen kann, nämlich auf die Tatsache, daß die meisten bekannten Verfahren nur lokal konvergent sind oder die Konvergenzgeschwindigkeit nur linear ist. Interessant ist nun, daß man unabhängig von der Güte der Einschließung, d.h. unabhängig von $x^{(0)}$, die Voraussetzungen von Satz 4.1 verifizieren kann. Das angegebene Verfahren konvergiert daher in dem beschriebenen Sinne global gegen die Lösung des diskreten Problems. Entsprechende Aussagen für klassische Verfahren wie z.B. für das Newton-Verfahren und seine zahlreichen Varianten sind allgemein nicht richtig und können gewöhnlich nur bewiesen werden, wenn man für das diskretisierte Problem Konvexität o.ä. nachweisen kann.

Die analogen Bemerkungen gelten für das durch Diskretisierung der nichtlinearen elliptischen Randwertaufgabe

$$\begin{aligned} \Delta u &\equiv u_{ss} + u_{tt} = g(s, t, u), & (s, t) \in \Omega \\ u(s, t) &= \varphi(s, t), & (s, t) \in \partial\Omega \end{aligned}$$

mit stetigem g , $g_u \geq 0$ und einfach zusammenhängendem, beschränktem Gebiet Ω entstehende nichtlineare Gleichungssystem. Auch in diesem Falle läßt sich die quadratische Konvergenz des oben angegebenen Verfahrens gegen die eindeutige Lösung x^* des diskreten Problems für beliebige Ausgangsvektoren $x^{(0)}$ mit $x^* \in x^{(0)}$ nachweisen. Die praktische Durchführung stößt allerdings auf einige ernsthafte Schwierigkeiten. Da der Diskretisierungsfehler klein gehalten werden soll, erhält man sehr rasch große nichtlineare Gleichungssysteme. Diese Systeme haben spezielle Gestalt. Die intervallmäßige Auswertung der Ableitung $f'(x^{(k)})$ ist eine

Blocktridiagonalmatrix. Die nichtverschwindenden Blöcke sind selbst wieder nur dünn besetzt und die Anwendung des Gaußschen Algorithmus erzeugt nichtverschwindende Elemente für Indexpaare, bei denen anfangs eine Null steht. Die benötigten Speicherplätze wachsen bei obigem Verfahren während der Rechnung i. allg. rasch an.

Dieser Sachverhalt ist der Hauptgrund dafür, daß man Verfahren sucht, bei denen die Anzahl der nichtverschwindenden Elemente während der Rechnung konstant bleibt. Ein solches Verfahren gewinnt man durch wiederholte Anwendung von Satz 3.3.

Verfahren II:

$$\begin{cases} \tilde{x}^{(k+1)} = m(x^{(k)}) - \tilde{D}(x^{(k)}) \{B(x^{(k)}) (m(x^{(k)}) - x^{(k)}) + f(m(x^{(k)}))\} \\ x^{(k+1)} = \tilde{x}^{(k+1)} \cap x^{(k)}, \\ k = 0, 1, 2, \dots, \end{cases}$$

$$(f'(x^{(k)})) = D(x^{(k)}) - B(x^{(k)}), \tilde{D}(x^{(k)}) := \text{diag} \left(1 / \frac{\partial}{\partial x_i} f_i(x^{(k)}) \right)$$

Verfahren II kann als eine Variante des bekannten Newton-Gesamtschrittverfahrens (siehe z.B. [15], S. 214 ff.) angesehen werden. Es geht in dieses über, wenn man die bei der Linearisierung von f entstehenden Fehler vernachlässigt. Für Verfahren II lassen sich die folgenden Aussagen beweisen ([3], S. 327 ff.).

Satz 4.2: Sei $x^{(0)} \subseteq D \subseteq V_n(\mathbb{R})$ ein Intervallvektor und $x^* \in x^{(0)}$ Nullstelle der Abbildung $f: D \subseteq V_n(\mathbb{R}) \rightarrow V_n(\mathbb{R})$. Die intervallmäßige Auswertung $f'(x^{(0)})$ erfülle die Voraussetzungen von Satz 2.2. Dann gilt

- (1) Das angegebene Verfahren II ist für alle $k \geq 0$ durchführbar und es gilt $\lim_{k \rightarrow \infty} x^{(k)} = x^*$.
- (2) Jede Iterierte $x^{(k)}$, $k \geq 0$, enthält die Nullstelle x^* . ■

Die praktische Bedeutung dieses Satzes liegt weniger in der fortwährenden Einschließung von x^* durch die Folge $\{x^{(k)}\}_{k=0}^{\infty}$, sondern – wie bei Satz 4.1 – in der Konvergenzaussage (1), die unabhängig von der Güte der Einschließung von x^* durch $x^{(0)}$ besteht. Verfahren II ist daher unter den genannten Voraussetzungen global konvergent. Für das oben erwähnte Newton-Gesamtschrittverfahren erhält man dagegen globale Konvergenz nur unter wesentlich spezielleren Voraussetzungen an f (siehe z.B. [15], S. 456 ff.).

Die in Satz 4.2 benötigten Voraussetzungen lassen sich für das aus der oben angegebenen partiellen Randwertaufgabe mit Hilfe der üblichen Ersetzung der partiellen Ableitungen durch Differenzenquotienten entstehende nichtlineare System nachweisen. Es soll hier jedoch nicht verschwiegen werden, daß die lineare Konvergenz der Durchmesser gegen Null mit

einem Konvergenzfaktor, der mit wachsender Anzahl der Gleichungen gegen Eins geht, konvergenzbeschleunigende Maßnahmen notwendig macht, um das Verfahren praktisch sinnvoll anwenden zu können. Dieser Fragenkreis ist gegenwärtig Gegenstand eingehender Untersuchungen.

5. Ausblick

Dieser Beitrag sollte aufzeigen, daß mit der Intervallrechnung ein Hilfsmittel zur Verfügung steht, welches unabhängig von der ursprünglich nur beabsichtigten Rundungsfehlererfassung, bestens geeignet ist, um konstruktive Existenzaussagen zu gewinnen und Iterationsverfahren zu formulieren, welche Eigenschaften besitzen, die mit klassischen Mitteln nicht gewonnen werden können. Es kann nicht das Ziel sein, die Verfahren der klassischen Numerischen Mathematik durch sogenannte „Intervallalgorithmen“ – was man auch immer darunter verstehen mag – zu ersetzen. Jedoch sollte man bei Problemen, bei denen die Verwendung der Intervallarithmetik offensichtliche Vorteile bringt, bereit sein, diese Erkenntnisse auch praktisch zu verwenden.

Literatur

- [1] R. Albrecht, U. Kulisch: Grundlagen der Computer-Arithmetik. Computing Supplementum 1 (1977)
- [2] G. Alefeld: Über die Durchführbarkeit des Gaußschen Algorithmus bei Gleichungen mit Intervallen als Koeffizienten. In [1], S. 15–19.
- [3] G. Alefeld, J. Herzberger: Einführung in die Intervallrechnung. Bibliographisches Institut Mannheim (1974)
- [4] N. Apostolatos, U. Kulisch: Approximation der erweiterten Intervallarithmetik durch die einfache Maschinenintervallarithmetik. Computing 2, 181–194 (1967)
- [5] W. Barth, E. Nuding: Optimale Lösung von Intervallgleichungssystemen. Computing 12, 117–125 (1974)
- [6] I. Gargantini, P. Henrici: Circular Arithmetic and the Determination of Polynomial Zeros. Numer. Math. 18, 305–320 (1972)
- [7] H. Grell, K. Maruhn, W. Rinow (Herausgeber): Enzyklopädie der Elementarmathematik, Band I, Arithmetik. Deutscher Verlag der Wissenschaften, Berlin (1954)
- [8] E. Hansen: On solving systems of equations using interval arithmetic. Math. Comp. 22, 374–384 (1968)
- [9] J. Herzberger: Persönliche Mitteilung (1968)
- [10] J. Herzberger: Einige Klassen von global konvergenten Iterationsverfahren zur Nullstelleneinschließung. Erscheint in Beiträge zur Numerischen Mathematik
- [11] U. Kulisch: Grundzüge der Intervallrechnung. In „Überblicke Mathematik“ 2, Bibliographisches Institut Mannheim (1969)
- [12] U. Kulisch: Grundlagen des numerischen Rechnens. Bibliographisches Institut Mannheim (1976)
- [13] R. E. Moore: Interval Analysis. Prentice-Hall, Inc., Englewood Cliffs, N.J. (1966)

- [14] R. E. Moore: A test for existence of solutions to non-linear systems. SIAM J. Numer. Anal. 14, 611–615 (1977)
- [15] J. M. Ortega, W. C. Rheinboldt: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York (1970)
- [16] K. Reichmann: Ein hinreichendes Kriterium für die Durchführbarkeit des Intervall-Gauß-Algorithmus bei Intervall-Hessenbergmatrizen ohne Pivotsuche. GAMM-Vortrag. Brüssel (1978)
- [17] T. Sunaga: Theory of an interval algebra and its application to numerical analysis. RAAG Memoirs 2, 29–46 (1958)
- [18] J. W. Schmidt: Die Regula-Falsi für Operatoren in Banachräumen. Z. Angew. Math. Mech. 41, 61–63. (1961)
- [19] R. S. Varga: Matrix Iterative Analysis. Englewood Cliffs, N. J., Printice Hall Inc. (1962)
- [20] R. C. Young: The algebra of many-valued quantities. Math. Ann. 104, 260–290 (1931)

Anschrift des Verfassers:

Götz Alefeld
Fachbereich 3 / Mathematik
Technische Universität Berlin
Straße des 17. Juni 135
1000 Berlin 12