Computing 14, 141—147 (1975)

Eine Anwendung des Brouwerschen Fixpunktsatzes in der Intervallarithmetik der Matrizen

P. Thieler, Bonn

Eingegangen am 11. Mai 1974

Zusammenfassung — Abstract

Eine Anwendung des Brouwerschen Fixpunktsatzes in der Intervallarithmetik der Matrizen. Mittels des Satzes von Brouwer wird ein Satz bewiesen, nach dem man entscheiden kann, ob eine gegebene Intervallmatrix $[S]$ die Einschließungseigenschaft bezüglich der unbekannten Inversen A^{-1} einer bekannten Matrix A hat. Es wird vorgeführt, in welchen Fällen der Satz anwendbar ist und wie $[S]$ in diesen Fällen gewählt werden kann.

An Application of the Brouwer Fixed Point Theorem in Matrix Interval Arithmetic. A theorem is proved by means of Brouwer's theorem which allows to decide whether a given interval matrix $[S]$ has the inclusion property relative to the unknown inverse A^{-1} of a known matrix A. It is demonstrated in which cases the theorem may be used and how to choose $[S]$ in these cases.

1. Bezeichnungen und allgemeine Voraussetzungen

Es sei $\mathbb{R}$ der Körper der reellen Zahlen mit den Elementen $a, b, \ldots$. Unter $M_n(\mathbb{R})$ wird der Raum der $n \times n$-Matrizen mit Gliedern aus $\mathbb{R}$ verstanden, seine Elemente werden mit $A, B, \ldots$ bezeichnet. Insbesondere ist E die Einheitsmatrix und N die Nullmatrix. — $\|\cdot\|$ bezeichnet eine Norm auf $M_n(\mathbb{R})$.

$I(\mathbb{R})$ ist die Menge der abgeschlossenen Intervalle aus $\mathbb{R}$. Die Elemente aus $I(\mathbb{R})$ werden wie üblich $[a_1, a_2]$, $[b_1, b_2], \ldots$ geschrieben (für die Ecken wird vorausgesetzt $a_1 \leq a_2$, $b_1 \leq b_2, \ldots$). Insbesondere werden Intervalle der Form $[a, a]$, $[b, b], \ldots$ durch $\underset{\cdot}{a}, \underset{\cdot}{b}, \ldots$ gekennzeichnet (Punktintervalle). $M_n(I(\mathbb{R}))$ steht für die Menge aller $n \times n$-Matrizen mit Gliedern aus $I(\mathbb{R})$, die Elemente von $M_n(I(\mathbb{R}))$ heißen Intervallmatrizen und werden mit $[A], [B], \ldots$ bezeichnet. Sind alle Glieder einer Intervallmatrix Punktintervalle, so wird $\underset{\cdot}{A}, \underset{\cdot}{B}, \ldots$ geschrieben (Punktmatrizen). Eine Matrix $C \in M_n(\mathbb{R})$ wird durch $\underset{\cdot}{C}$ in $M_n(I(\mathbb{R}))$ dargestellt, indem man jedem Glied c der Matrix C das Intervall $[c, c] = \underset{\cdot}{c}$ als entsprechendes Glied von $\underset{\cdot}{C}$ zuordnet. Sind alle Glieder einer Intervallmatrix $[D]$ gleich dem Intervall $[d_1, d_2]$, so wird dies durch $[D] = [[d_1, d_2]]$ ausgedrückt.

In $I(\mathbb{R})$ bzw. $M_n(I(\mathbb{R}))$ seien die intervallarithmetischen Operationen $+$, $-$, $\cdot$, $/$ bzw. $+$, $-$, $*$ wie üblich erklärt (vgl. [1], [2], [3]). Insbesondere stehe $\cdot$ auch für die Multiplikation einer Intervallmatrix mit einem Intervall (intervall-arithmetische Skalarmultiplikation). Für Verknüpfungen in $\mathbb{R}$ bzw. $M_n(\mathbb{R})$ werden die gleichen Zeichen in hergebrachter Weise benutzt.

Zwischen Elementen aus $\mathbb{R}$ und $I(\mathbb{R})$ ist die Relation $\in$ in üblicher Weise zu verstehen. In $I(\mathbb{R})$ ist $\subset$ erklärt durch

$$[a_1, a_2] \subset [b_1, b_2] \Leftrightarrow b_1 \leq a_1 \wedge a_2 \leq b_2 .$$

Für Elemente aus $M_n(\mathbb{R})$ und $M_n(I(\mathbb{R}))$ sind $\in$ und $\subset$ komponentenweise zu lesen.

Die Matrix $A \in M_n(\mathbb{R})$ sei regulär, d. h. es gibt eine Matrix $A^{-1} \in M_n(\mathbb{R})$ mit der Eigenschaft $A^{-1} * A = E$. Zu A und A^{-1} seien definiert:

$$S_A := E - A, \quad S_E := A^{-1} - E .$$

2. Eine Anwendung des Brouwerschen Fixpunktsatzes

Bei vielen intervallarithmetischen Iterationsverfahren zur Invertierung von Matrizen $A \in M_n(\mathbb{R})$ muß eine Start-Intervallmatrix mit der Einschließungseigenschaft zur Verfügung gestellt werden, d. h. diese muß die gesuchte Lösung bereits enthalten (vgl. [4], 3. 4. 2.). — Ist umgekehrt irgendeine Näherungslösung für A^{-1} bekannt und gewinnt man aus dieser Näherung und einer Fehlerabschätzung eine Intervallmatrix, für die die Einschließungseigenschaft garantiert werden kann, so kann die Einschließung so grob sein, daß eine Verfeinerung vonnöten ist. Diese muß so ausgeführt werden, daß die Einschließungseigenschaft nicht verletzt wird.

In beiden Fällen können folgende Kriterien von Nutzen sein, nach denen man beurteilen kann, ob eine gegebene Intervallmatrix die Einschließungseigenschaft hat:

Satz 1: *Sei* $[S] \in M_n(I(\mathbb{R}))$ *beliebig. Es gilt:*

(*i*) $$([S] + \underset{.}{E}) * \underset{.}{S}_A \subset [S] \Rightarrow A^{-1} \in \underset{.}{E} + [S] .$$

(*ii*) $$\{S_A \text{ regulär}\} \wedge [S] * \underset{.}{S}_A^{-1} - \underset{.}{E} \subset [S] \Rightarrow A^{-1} \in \underset{.}{E} + [S] .$$

Beweis:
$M_n(\mathbb{R})$ ist mit der natürlichen Topologie ein lokalkonvexer topologischer linearer Raum. $[S]$ ist eine konvexe kompakte Teilmenge des $M_n(\mathbb{R})$. Die Abbildungen

$$\begin{array}{llll} T_1: M_n(\mathbb{R}) \to M_n(\mathbb{R}) & \text{bzw.} & T_2: M_n(\mathbb{R}) \to M_n(\mathbb{R}) \\ \quad X \quad \to (X+E) * S_A & & \quad X \quad \to X * S_A^{-1} - E \end{array}$$

sind stetig. Es gilt wegen der Inklusionsmonotonie der Intervallarithmetik (vgl. [1], [2]):

$$S \in [S] \Rightarrow (S+E) * S_A \in ([S] + \underset{.}{E}) * \underset{.}{S}_A \vee S * S_A^{-1} - E \in [S] * \underset{.}{S}_A^{-1} - \underset{.}{E} .$$

Nach den Voraussetzungen gilt also:

$$S \in [S] \Rightarrow T_1(S) \in [S] \vee T_2(S) \in [S],$$

d. h. T_1 bzw. T_2 sind Abbildungen von $[S]$ in sich. Nach dem Brouwerschen Fixpunktsatz (in der Verallgemeinerung von Schauder/Tychonoff, siehe [5]) haben daher die beiden Abbildungen je mindestens einen Fixpunkt in $[S]$ mit

$$X_1 = (X_1 + E) * S_A \quad \text{bzw.} \quad X_2 = X_2 * S_A^{-1} - E.$$

Man überzeugt sich leicht, daß

$$S_E * (E - S_A) = S_A = X_1 * (E - S_A) = X_2 * (E - S_A)$$

richtig ist. Man hat demnach $X_1 = X_2 = S_E = A^{-1} - E \in [S]$, und es folgen die Behauptungen. //

Die Argumentation zur Beweisführung ändert sich offenbar in den wesentlichen Teilen nicht, wenn S_A bzw. S_A^{-1} nicht als Punktmatrizen vorliegen, sondern nur Einschließungen $[S_A]$ bzw. $[S_A^{-1}]$ bekannt sind. Es läßt sich daher folgendes Korollar aussprechen:

Satz 2: *Sei* $[S] \in M_n(I(\mathbb{R}))$ *beliebig.*

(*i*) $$S_A \in [S_A] \wedge ([S] + E) * [S_A] \subset [S] \Rightarrow A^{-1} \in \underset{\cdot}{E} + [S].$$

(*ii*) $$\{S_A \; reg.\} \wedge S_A^{-1} \in [S_A^{-1}] \wedge [S] * [S_A^{-1}] - \underset{\cdot}{E} \subset [S] \Rightarrow A^{-1} \in \underset{\cdot}{E} + [S].$$

Es ist zu bemerken, daß man es in der Praxis aufgrund von Darstellungs- und Rundungsfehlern (problembedingt und/oder durch Einsatz von Rechenmaschinen) fast immer nur mit Intervallmatrizen $[S_A]$ bzw. $[S_A^{-1}]$ zu tun hat. Verknüpft man diese in der in Satz 2 angegebenen Weise maschinell und genügen die verwendeten Maschinenverknüpfungen den üblichen Anforderungen an eine Maschinenintervallarithmetik (vgl. [6], [2]), so ist bekanntlich das Maschinenergebnis $[Z_1]$ bzw. $[Z_2]$ gröber als das wahre Ergebnis:

$$([S] + \underset{\cdot}{E}) * [S_A] \subset [Z_1] \quad \text{bzw.} \quad [S] * [S_A^{-1}] - \underset{\cdot}{E} \subset [Z_2].$$

Beobachtet man aber

$$[Z_1] \subset [S] \quad \text{bzw.} \quad [Z_2] \subset [S],$$

so kann man noch immer die Einschließungseigenschaft von $\underset{\cdot}{E} + [S]$ garantieren. In diesem Sinne sind die Kriterien aus Satz 1 maschinenfreundlich zu nennen.

Es erhebt sich die Frage, wann die Bedingungen aus 1 (i) bzw. 1 (ii) beobachtbar sind. — Bearbeitet man die Aufgabe $X * A = E$ in normaler Maschinenarithmetik, so erhält man i. a. nur eine Näherungslösung X' mit $X' * A =: G \neq E$. Gilt nun (wie man es in der Praxis für viele Iterationsverfahren fordert) $\| E - A \| = \| S_A \| < 1$, so ist bekanntlich

$$A^{-1} = \sum_{n=0}^{\infty} S_A^{\,n}$$

konvergent (Neumannsche Reihe), und es ist

$$\| A^{-1} \| \leq 1/(1 - \| S_A \|).$$

Man hat außerdem

$$X - X' = A^{-1} - G * A^{-1} = (E - G) * A^{-1} = (E - X' * A) * A^{-1},$$

das heißt

$$\| A^{-1} - X' \| \leq \| E - X' * A \| / (1 - \| S_A \|) \leq (1 + \| X' * A \|) / (1 - \| S_A \|) = : a, a > 0,$$

beziehungsweise $A^{-1} \in \underset{\cdot}{X}' + [D]$, wobei $[D] := [[-a, a]]$.

Satz 3: *Sei X' eine Näherungslösung für die Aufgabe $X * A = E$, sei $\| \cdot \|$ die Spaltensummennorm und $\| S_A \| < 1$. $[D]$ sei wie oben definiert. Dann gilt: Die Intervallmatrix*

$$[S] := \underset{\cdot}{X}' + [D] - \underset{\cdot}{E}$$

erfüllt die Voraussetzungen aus Satz 1, *Aussage* (*i*).

Beweis:

Man überprüft leicht, daß folgende Matrizenbeziehungen gelten:

$$[[-\| X' * A \| - a \cdot \| S_A \|, \| X' * A \| + a \cdot \| S_A \|]] \subset [[-a, a]] - \underset{\cdot}{E},$$

$$[[-\| X' * A \|, \| X' * A \|]] + [[-a \cdot \| S_A \|, a \cdot \| S_A \|]] \subset [D] - \underset{\cdot}{E}.$$

Nun gilt

$$-\underset{\cdot}{X}' + \underset{\cdot}{X}' * \underset{\cdot}{S}_A = -\underset{\cdot}{X}' * (\underset{\cdot}{E} - \underset{\cdot}{S}_A) = -\underset{\cdot}{X}' * \underset{\cdot}{A} \subset [[-\| X' * A \|, \| X' * A \|]],$$

und außerdem, weil $\| \cdot \|$ die Spaltensummennorm ist,

$$[D] * \underset{\cdot}{S}_A \subset [[-a \cdot \| S_A \|, a \cdot \| S_A \|]].$$

Damit folgt

$$-\underset{\cdot}{X}' + \underset{\cdot}{X}' * \underset{\cdot}{S}_A + [D] * \underset{\cdot}{S}_A \subset [D] - \underset{\cdot}{E}, \qquad (\underset{\cdot}{X}' + [D]) * \underset{\cdot}{S}_A \subset \underset{\cdot}{X}' + [D] - \underset{\cdot}{E},$$

$([S] + \underset{\cdot}{E}) * \underset{\cdot}{S}_A \subset [S]$, was zu zeigen war. Im Beweis wird mehrfach vom Distributivgesetz für Punktfaktoren in $I(\mathbb{R})$ und $M_n(I(\mathbb{R}))$ Gebrauch gemacht (vgl. [1], [2], [3]) und von der Tatsache, daß die Beziehung $=$ bzw. $\subset$ nicht verletzt wird durch Addition oder Subtraktion von Punktgrößen auf beiden Seiten. //

Man sieht, daß es in 1 (i) weniger darauf ankommt, ob $[S]$ schmal ist (hier wird $[S]$ ja sehr grob gewählt), vielmehr ist die Bedingung $\| S_A \| < 1$ wesentlich. — Es ist nun unter Umständen möglich, $[S]$ aus Satz 3 unter Beachtung von 1 (i) zu verfeinern. Dabei ist zu fallendem $r \in [0, 1]$ für $[S(r)] := \underset{\cdot}{X}' - \underset{\cdot}{E} + [D] \cdot \underset{\cdot}{r}$ die Beziehung

$$([S(r)] + \underset{\cdot}{E}) * \underset{\cdot}{S}_A \subset [S(r)]$$

zu prüfen, welche äquivalent ist zu der Beziehung

$$\underset{\cdot}{E} - \underset{\cdot}{X}' * \underset{\cdot}{A} + [D] * \underset{\cdot}{S}_A \cdot \underset{\cdot}{r} \subset [D] \cdot \underset{\cdot}{r}.$$

(Man verifiziert das unter Ausnutzung der gleichen Tatsachen, die im Beweis zu Satz 3 verwendet werden.) Die letzte Form erlaubt die einmalige Berechnung von $\underset{\cdot}{E} - \underset{\cdot}{X}' * \underset{\cdot}{A}$ und $[D] * \underset{\cdot}{S}_A$, ehe man r variiert. Wenn für $r_0 < 1$ die Matrix

$$[S(r_0)] = \underset{\cdot}{X}' + [D] \cdot \underset{\cdot}{r}_0 - \underset{\cdot}{E}$$

der Voraussetzung aus 1 (i) genügt, so hat

$$\underset{\cdot}{X}' + [D] \cdot \underset{\cdot}{r}_0 \subset \underset{\cdot}{X}' + [D]$$

noch immer die Einschließungseigenschaft. Auch hier gelten die Bemerkungen in Anschluß an Satz 2 sinngemäß.

Dieses Verfahren ist offenbar auch dann anwendbar, wenn $[D]$ anders als in Satz 3 angesetzt wird. Man betrachte hierzu das Beispiel im Abschnitt 3.

Es soll nun noch demonstriert werden, wie 1 (ii) zu einer Intervallmatrix mit Einschließungseigenschaft verhelfen kann.

Satz 4: *Sei* $\|\cdot\|$ *die Spaltensummennorm und sei* S_A *regulär mit* $\| S_A^{-1} \| < 1$. *Es gilt*:

$$b \geq 1/(1 - \| S_A^{-1} \|) \Rightarrow A^{-1} \in \underset{\cdot}{E} + [[-b, b]].$$

Beweis:

Man setzt zunächst $[S] := [[-b, b]]$ mit $0 \leq b \in \mathbb{R}$. Wegen der speziellen Wahl der Norm hat man wieder:

$$[S] * \underset{\cdot}{S}_A^{-1} \subset [[-b \cdot \| S_A^{-1} \|, \; b \cdot \| S_A^{-1} \|]] =: [F].$$

Gelingt es nun, $[F] - \underset{\cdot}{E} \subset [[-b, b]]$ herzuleiten, so ist $[S]$ für die Anwendung von 1 (ii) geeignet, man kann also $A^{-1} \in \underset{\cdot}{E} + [S]$ garantieren.

Aus der letzten Matrixbeziehung entnimmt man vier Bestimmungsungleichungen für b und $\| S_A^{-1} \|$:

(1) $-b \cdot \| S_A^{-1} \| - 1 \geq -b$ (2) $-b \cdot \| S_A^{-1} \| \geq -b$

(3) $b \cdot \| S_A^{-1} \| \leq b$ (4) $b \cdot \| S_A^{-1} \| - 1 \leq b.$

Nach (1) kann b nicht Null sein. Unter dieser Voraussetzung sind (2) und (3) äquivalent zu $\| S_A^{-1} \| \leq 1$. Dies wiederum zieht nach sich, daß (4) für jedes $0 \leq b \in \mathbb{R}$ richtig ist. Es geht also darum, unter Beachtung von $\| S_A^{-1} \| \leq 1$ und $0 < b \in \mathbb{R}$ ein b zu finden, das (1) genügt. Dies ist offenbar für $b \geq 1/(1 - \| S_A^{-1} \|)$ bei $\| S_A^{-1} \| < 1$ der Fall. //

Aufmerksam geworden durch Satz 4 findet man übrigens, daß die gleiche Aussage für jede beliebige Matrixnorm getroffen werden kann. Man braucht dazu lediglich die Submultiplikativität der Norm.

Die Abschätzung aus Satz 4 mit $b = 1/(1 - \| S_A^{-1} \|)$ ist scharf in dem Sinne, daß Glieder von A^{-1} jeweils einer Ecke der entsprechenden Glieder von $E + [[-b, b]]$ gleich sein können (man setze etwa $A = -E$).

3. Beispiel

$\|\cdot\|$ Zeilensummennorm; $A = \begin{pmatrix} 0.8 & -0.6 \\ 0.4 & 0.9 \end{pmatrix}$; $S_A = E - A = \begin{pmatrix} 0.2 & 0.6 \\ -0.4 & 0.1 \end{pmatrix}$; $\| S_A \| = 0.8$;

$A^{-1}=\begin{pmatrix} 0.9375 & 0.625 \\ -0.416 & 0.833 \end{pmatrix}$. Mit der Näherungslösung X'_i $(i=1,2,3)$ wird

$$\| A^{-1}-X'_i \| \leq \| E-X'_i * A \| / (1-\| S_A \|)=:a_i$$

berechnet, die Matrix $[D_i]:=[[-a_i, a_i]]$ gebildet und damit die herkömmliche Abschätzung

$$A^{-1} \in \underset{\cdot}{X}'_i+[D_i]=:[H_i]$$

angegeben. Anschließend wird geprüft, für welche $r>0$ die Beziehung

$$\underset{\cdot}{E}-\underset{\cdot}{X}'_i * \underset{\cdot}{A}+[D_i] * \underset{\cdot}{S}_A \cdot \underset{\cdot}{r} \subset [D_i] \cdot \underset{\cdot}{r}$$

erfüllt ist. Es wird r_i möglichst klein aus dem erlaubten Bereich von r gewählt und die verbesserte Abschätzung

$$A^{-1} \in \underset{\cdot}{X}'_i+[D_i] \cdot \underset{\cdot}{r}_i=:[V_i]$$

berechnet.

$(i=1)$ Bei $X'_1:=N$ ist $a_1=5$, und es geht um die Frage, ob die a-priori-Abschätzung $A^{-1} \in [[-5, 5]]=[H_1]$ verbessert werden kann.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\begin{pmatrix} [-3, 3] & [-3.5, 3.5] \\ [-3, 3] & [-3.5, 3.5] \end{pmatrix} \cdot \underset{\cdot}{r} \subset \begin{pmatrix} [-5, 5] & [-5, 5] \\ [-5, 5] & [-5, 5] \end{pmatrix} \cdot \underset{\cdot}{r}$$

ist erfüllt für $r \geq 2/3$. Mit $r_1:=0.7$ wird $[V_1]=[[-3.5, 3.5]] \subset [H_1]$.

$(i=2)$ $\qquad X'_2:=\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $a_2=4.5$, $[H_2]=\begin{pmatrix} [-3.5, 5.5] & [-3.5, 5.5] \\ [-5.5, 3.5] & [-3.5, 5.5] \end{pmatrix}$.

$$\begin{pmatrix} -0.2 & -0.3 \\ 0.4 & -0.5 \end{pmatrix}+\begin{pmatrix} [-2.7, 2.7] & [-3.15, 3.15] \\ [-2.7, 2.7] & [-3.15, 3.15] \end{pmatrix} \cdot \underset{\cdot}{r} \subset \begin{pmatrix} [-4.5, 4.5] & [-4.5, 4.5] \\ [-4.5, 4.5] & [-4.5, 4.5] \end{pmatrix} \cdot \underset{\cdot}{r}$$

ist für $r \geq 10/23$ richtig. $r_2:=0.5>0.4347\ldots=10/23$.

$$[V_2]=\begin{pmatrix} [-1.25, 3.25] & [-1.25, 3.25] \\ [-3.25, 1.25] & [-1.25, 3.25] \end{pmatrix} \subset [H_2].$$

$(i=3)$ $\qquad X'_3:=\begin{pmatrix} 0.9 & 0.6 \\ -0.4 & 0.8 \end{pmatrix}$, $a_3=0.2$, $[H_3]=\begin{pmatrix} [0.7, 1.1] & [0.4, 0.8] \\ [-0.6, -0.2] & [0.6, 1.0] \end{pmatrix}$.

$$\begin{pmatrix} 0.04 & 0 \\ 0 & 0.04 \end{pmatrix}+\begin{pmatrix} [-0.12, 0.12] & [-0.14, 0.14] \\ [-0.12, 0.12] & [-0.14, 0.14] \end{pmatrix} \cdot \underset{\cdot}{r} \subset [[-0.2, 0.2]] \cdot \underset{\cdot}{r}$$

verlangt $r \geq 2/3$. Wählt man $r_3=2/3$, dann ist

$$[D_3] \cdot r_3 \subset [[-0.14, 0.14]], \text{ d. h. } [V_3] \subset \begin{pmatrix} [0.76, 1.04] & [0.46, 0.74] \\ [-0.54, -0.26] & [0.66, 0.94] \end{pmatrix} \subset [H_3].$$

Diese Arbeit ist mit Unterstützung des von der Deutschen Forschungsgemeinschaft getragenen Sonderforschungsbereiches 72 an der Universität Bonn entstanden.

Literatur

[1] Moore, R. E.: Interval Analysis. Englewood Cliffs: Prentice-Hall 1966.
[2] Kulisch, U.: Grundzüge der Intervallrechnung, in: Überblicke Mathematik, Band 2 (Laugwitz, D., Hrsg.), S. 51—98. Mannheim: Bibliographisches Institut 1969.
[3] Apostolatos, N., Kulisch, U.: Grundzüge einer Intervallrechnung für Matrizen und einige Anwendungen. Elektron. Rechenanlagen *10*, 73—83 (1968).
[4] Geul, A.: Algebraische Eigenschaften der Intervallarithmetik, die quasilinearen Räume der Intervalle, Intervallvektoren und Intervallmatrizen und ihre Anwendung auf Probleme der linearen Algebra. Diplomarbeit, Stuttgart, 1972.
[5] Dunford, N., Schwartz, J. T.: Linear Operators, Vol. 1, 4. Aufl., S. 456. New York: Interscience 1967.
[6] Apostolatos, N., Kulisch, U.: Grundlagen einer Maschinenintervallarithmetik. Computing *2*, 89—104 (1967).

Dr. P. Thieler
Institut für Angewandte Mathematik
und Informatik
Universität Bonn
Wegelerstraße 6
D-5300 Bonn 1
Bundesrepublik Deutschland