

## Das symmetrische Einzelschrittverfahren bei linearen Gleichungen mit Intervallen als Koeffizienten

G. Alefeld, Berlin

Eingegangen am 18. August 1976

### Zusammenfassung — Abstract

**Das symmetrische Einzelschrittverfahren bei linearen Gleichungen mit Intervallen als Koeffizienten.** Für das symmetrische Einzelschrittverfahren bei linearen Gleichungen mit Intervallen als Koeffizienten werden notwendige und hinreichende Bedingungen für die Konvergenz angegeben. Für eine Modifikation dieses Verfahrens wird gezeigt, daß der Fixpunkt eines Intervallgleichungssystems ohne Erhöhung des Aufwands in jedem Iterationsschritt mindestens genau so gut wie mit den bisher betrachteten Verfahren eingeschlossen wird. Dieser Sachverhalt wird auch durch numerische Ergebnisse erläutert.

**The Symmetric Single-Step-Method for Systems of Simultaneous Linear Equations With Intervals as Coefficients.** For systems of linear equations whose coefficients are intervals we consider the symmetric single-step-method. Sufficient and necessary conditions for convergence are given. For a modification of this method we show that in each step the fixed point of a linear system with intervals as coefficients can be included at least as good as by all other known methods. This is demonstrated by some numerical examples.

### 1. Einleitung

Gegeben sei eine Intervallmatrix<sup>1</sup>  $\mathfrak{A}$  und ein Intervallvektor  $b$ . Die Gleichung

$$x = \mathfrak{A} x + b$$

besitzt genau dann einen eindeutigen Fixpunkt  $x^*$ , gegen den das Gesamtschrittverfahren (G)

$$x^{k+1} = \mathfrak{A} x^k + b, \quad k=0, 1, 2, \dots,$$

für jeden Intervallvektor  $x^0$  als Anfangsvektor konvergiert, wenn der Spektralradius der reellen Matrix  $|\mathfrak{A}|$  der Ungleichung  $\rho(|\mathfrak{A}|) < 1$  genügt<sup>2</sup>. (Siehe dazu etwa [3], Abschnitt 12, Satz 1.) Zur Berechnung von  $x^*$  kann man neben dem Gesamtschrittverfahren auch verschiedene andere Iterationsverfahren anwenden. (Siehe etwa die Abschnitte 12—14 in [3].) Wählt man als Startvektor einen Inter-

<sup>1</sup> Die verwendeten Bezeichnungen sind in Abschnitt 2 zusammengestellt.

<sup>2</sup> Der Intervallvektor  $x^*$  besitzt die wichtige Eigenschaft  $\{x = (\mathfrak{I} - \mathfrak{A})^{-1} b \mid \mathfrak{A} \in \mathfrak{A}, b \in b\} \subseteq x^*$ .

vallvektor  $x^0$ , der den Fixpunkt  $x^*$  enthält, so besitzen bei den bekannten Iterationsverfahren auch alle Iterierten diese vom numerischen Standpunkt aus sehr wünschenswerte Eigenschaft. Unter allen bisher in der Literatur betrachteten Verfahren liefert dann bei gleichem Startvektor das sogenannte *Einzelschrittverfahren mit Durchschnittsbildung nach jeder Komponente* (EDK) die im Sinne der Inklusion optimale gegen  $x^*$  konvergente Folge. (Siehe [3], Abschnitt 14, Satz 4.)

In dieser Note betrachten wir nun das *symmetrische Einzelschrittverfahren* (SE). Als Spezialfall aus Ergebnissen über ein allgemeineres Iterationsverfahren erhalten wir eine notwendige und hinreichende Konvergenzaussage. Als wichtigstes Ergebnis dieser Arbeit geben wir unter sinngemäßer Verwendung der in der Arbeit [4] von W. Niethammer enthaltenen Tatsache, daß sich das symmetrische Einzelschrittverfahren im wesentlichen mit dem gleichen Aufwand wie das (gewöhnliche) Einzelschrittverfahren durchführen läßt, eine Modifikation des symmetrischen Einzelschrittverfahrens an, welche bei gleichem Ausgangsvektor im Sinne der Inklusion mindestens genauso gut wie das bereits erwähnte Verfahren (EDK) konvergiert. Die angegebene Modifikation ist daher unter allen bisher betrachteten Verfahren als das optimale zur fortwährenden Einschließung des Fixpunktes  $x^*$  anzusehen.

## 2. Bezeichnungen

Wir verwenden die Bezeichnungen und die Terminologie wie in [3]:

- |   |   |
|---|---|
| $A, B, C, \dots$  | <i>Intervalle.</i> ([3], Abschnitt 1. Außerdem Abschnitt 5, Def. 1 und Def. 5.)   |
| $x, \eta, \zeta, \dots$   | Vektoren mit Intervallen als Komponenten, sogenannte <i>Intervallvektoren.</i> ([3], Abschnitt 10.)   |
| $\bar{x}, \bar{\eta}, \bar{\zeta}, \dots$                           | <i>Punktvektoren</i> , d. h. Vektoren mit reellen oder komplexen Zahlen als Komponenten. Davon abweichend bezeichnet $\mathbf{o}$ stets den Nullvektor.   |
| $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \dots$                   | Matrizen mit Intervallen als Elemente, sogenannte <i>Intervallmatrizen.</i> ([3], Abschnitt 10.)  |
| $\bar{\mathfrak{A}}, \bar{\mathfrak{B}}, \bar{\mathfrak{C}}, \dots$ | <i>Punktmatrizen</i> , d. h. Matrizen mit reellen oder komplexen Zahlen als Elemente. Davon abweichend bezeichnet $\bar{\mathfrak{I}}$ stets die Einheitsmatrix, $\bar{\mathfrak{O}}$ stets die Nullmatrix. |
| $q(A, B)$   | <i>Abstand</i> zweier Intervalle. ([3], Abschnitt 2, Def. 1. Außerdem Abschnitt 6, Def. 1 und Def. 4.)  |
| $q(x, \eta)$  | <i>Abstand</i> zweier Intervallvektoren. ([3], Abschnitt 10, Def. 7.)   |
| $d(A)$  | <i>Durchmesser</i> eines Intervalls. ([3], Abschnitt 2, Def. 8. Außerdem Abschnitt 6, Def. 3 und Def. 4.)   |
| $d(x)$  | <i>Durchmesser</i> eines Intervallvektors. ([3], Abschnitt 10, Def. 6.)   |



- |A| Betrag eines Intervalles. ([3], Abschnitt 2, Def. 6. Außerdem Abschnitt 6, Def. 2 und Def. 4.)
- | $\mathfrak{A}$ | Betrag einer Intervallmatrix. ([3], Abschnitt 10, Def. 6.)
- $\rho$  Spektralradius einer Punktmatrix.

Alle in dieser Arbeit benötigten Eigenschaften der eingeführten Begriffe aus der Intervallrechnung findet man in [3].

### 3. Das symmetrische Einzelschrittverfahren

Gegeben sei eine Intervallmatrix  $\mathfrak{A}$  und ein Intervallvektor  $\mathfrak{b}$ . Die Matrix  $\mathfrak{A}$  sei zerlegt in

$$\mathfrak{A} = \mathfrak{Q} + \mathfrak{U},$$

wobei  $\mathfrak{Q}$  eine strenge untere und  $\mathfrak{U}$  eine strenge obere Dreiecksmatrix ist<sup>3</sup>. Dann betrachten wir das Iterationsverfahren

$$\begin{cases} \mathfrak{x}^{k+\frac{1}{2}} = \mathfrak{Q} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{U} \mathfrak{x}^k + \mathfrak{b} \\ \mathfrak{x}^{k+1} = \mathfrak{Q} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{U} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{b} \end{cases} \\ k=0, 1, 2, \dots,$$

welches wir als *symmetrisches Einzelschrittverfahren* (SE) bezeichnen. Für Punktgleichungssysteme wurde dieses Verfahren erstmals von Aitken [1] betrachtet<sup>4</sup>. Wie W. Niethammer in [4] für Punktgleichungssysteme gezeigt hat, läßt sich das Verfahren (SE), abgesehen vom ersten Iterationsschritt, mit dem gleichen Aufwand an Multiplikationen wie das Einzelschrittverfahren durchführen. Wir diskutieren diesen Sachverhalt ausführlicher in Abschnitt 4, wo wir eine Modifikation des Verfahrens (SE) betrachten, für die der gleiche Sachverhalt zutrifft. Das Konvergenzverhalten des Verfahrens (SE) ergibt sich als Korollar aus den folgenden, allgemeineren Aussagen.

**Satz 1:** Die Intervallmatrix  $\mathfrak{A}$  sei zerlegt in die Summe  $\mathfrak{A} = \mathfrak{M} + \mathfrak{N}$  zweier Intervallmatrizen  $\mathfrak{M}$  und  $\mathfrak{N}$ . Es sei  $\rho(|\mathfrak{M}|) < 1$  und  $\rho(|\mathfrak{N}|) < 1$ . Dann gelten für jeden beliebigen Intervallvektor  $\mathfrak{b}$  die folgenden Aussagen:

(a) Für jeden Intervallvektor  $\mathfrak{x}^0$  existiert eine eindeutige Folge  $\{\mathfrak{x}^k\}_{k=0}^{\infty}$ , welche der Iterationsvorschrift

$$(V) \begin{cases} \mathfrak{x}^{k+\frac{1}{2}} = \mathfrak{M} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{N} \mathfrak{x}^k + \mathfrak{b} \\ \mathfrak{x}^{k+1} = \mathfrak{M} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{N} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{b} \end{cases} \\ k=0, 1, 2, \dots,$$

genügt.

<sup>3</sup> Verschwinden nicht alle Diagonalelemente von  $\mathfrak{A}$ , so betrachten wir stattdessen das Iterationsverfahren

$$\begin{cases} \mathfrak{x}^{k+\frac{1}{2}} = \mathfrak{Q} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{D} \mathfrak{x}^k + \mathfrak{U} \mathfrak{x}^k + \mathfrak{b} \\ \mathfrak{x}^{k+1} = \mathfrak{Q} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{D} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{U} \mathfrak{x}^{k+\frac{1}{2}} + \mathfrak{b} \end{cases} \quad k=0, 1, 2, \dots,$$

für das sich analoge Aussagen beweisen lassen.

<sup>4</sup> Eine ausführliche Darstellung, welche Bedeutung dieses Verfahren und die durch Überrelaxation beschleunigte Version bei Punktgleichungssystemen besitzen, findet man bei Young in [8] und [9].

(b) Ist  $\rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|) < 1$ , so besitzt die Gleichung  $x = \mathfrak{A}x + b$  genau einen Fixpunkt  $x^*$ . Gilt außerdem

$$\mathfrak{A}x^* = (\mathfrak{M} + \mathfrak{N})x^* = \mathfrak{M}x^* + \mathfrak{N}x^{*5},$$

so konvergiert die nach (V) berechnete Folge  $\{x^k\}_{k=0}^\infty$  für alle Intervallvektoren  $x^0$  gegen  $x^*$ .

(c) Besitzt umgekehrt die Gleichung  $x = \mathfrak{A}x + b$  genau einen Fixpunkt  $x^*$ , und konvergiert das Verfahren (V) für alle  $x^0$  gegen  $x^*$ , so gilt

$$\mathfrak{A}x^* = (\mathfrak{M} + \mathfrak{N})x^* = \mathfrak{M}x^* + \mathfrak{N}x^{*5}$$

und

$$\rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|) < 1.$$

*Beweis:* Zu (a): Für einen beliebigen Intervallvektor  $z$  gilt unter Verwendung der Regeln (23) und (25) in [3], Seite 155, für alle Intervallvektoren  $x$  und  $\eta$

$$q(\mathfrak{M}x + \mathfrak{N}z + b, \mathfrak{M}\eta + \mathfrak{N}z + b) = q(\mathfrak{M}x, \mathfrak{M}\eta) \leq |\mathfrak{M}| q(x, \eta).$$

Da nach Voraussetzung  $\rho(|\mathfrak{M}|) < 1$  ist, heißt dies nach Satz 1 in Abschnitt 12 von [3], daß für jedes  $k \geq 0$  die Gleichung

$$x^{k+\frac{1}{2}} = \mathfrak{M}x^{k+\frac{1}{2}} + \mathfrak{N}x^k + b$$

genau einen Fixpunkt  $x^{k+\frac{1}{2}}$  besitzt. Entsprechend zeigt man für jedes  $k \geq 0$  die Existenz eines eindeutigen Fixpunktes  $x^{k+1}$  der Gleichung

$$x^{k+1} = \mathfrak{M}x^{k+\frac{1}{2}} + \mathfrak{N}x^{k+1} + b.$$

Damit ist die Existenz und Eindeutigkeit der Folge  $\{x^k\}_{k=0}^\infty$ , ausgehend von einem beliebigen Intervallvektor  $x^0$  nachgewiesen.

Zu (b): Wir zeigen, daß  $\rho(|\mathfrak{A}|) < 1$  gilt. Wegen  $\rho(|\mathfrak{M}|) < 1$ ,  $\rho(|\mathfrak{N}|) < 1$  existieren nach Theorem 3.8 in [7] die Inversen

$$(\mathfrak{I} - |\mathfrak{M}|)^{-1} \text{ und } (\mathfrak{I} - |\mathfrak{N}|)^{-1}$$

und diese sind nichtnegativ. Somit ist auch die reelle Matrix

$$(\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}| = (\mathfrak{I} - |\mathfrak{N}|)^{-1} (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{M}| |\mathfrak{N}|$$

nichtnegativ. Durch nochmalige Anwendung von Theorem 3.8 in [7] erhält man daher

$$\begin{aligned} \mathfrak{D} &\leq (\mathfrak{I} - (\mathfrak{I} - |\mathfrak{N}|)^{-1} (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{M}| |\mathfrak{N}|)^{-1} = \\ &= (\mathfrak{I} - (|\mathfrak{M}| + |\mathfrak{N}|))^{-1} (\mathfrak{I} - |\mathfrak{N}|) (\mathfrak{I} - |\mathfrak{M}|) \end{aligned}$$

und unter Verwendung von  $(\mathfrak{I} - |\mathfrak{M}|)^{-1} \geq \mathfrak{D}$ ,  $(\mathfrak{I} - |\mathfrak{N}|)^{-1} \geq \mathfrak{D}$  schließlich

$$(\mathfrak{I} - (|\mathfrak{M}| + |\mathfrak{N}|))^{-1} \geq \mathfrak{D}.$$

Daraus folgt nach Theorem 3.10 in [7] die Ungleichung

$$\rho(|\mathfrak{M}| + |\mathfrak{N}|) < 1,$$

<sup>5</sup> Das Distributivgesetz ist im allgemeinen bei Intervallmatrizen nicht erfüllt. (Siehe etwa [3], Seite 151, Formel (6).)



und wegen

$$|\mathfrak{A}| = |\mathfrak{M} + \mathfrak{N}| \leq |\mathfrak{M}| + |\mathfrak{N}|.$$

(siehe [3], Seite 153) folgt aufgrund des Satzes von Perron und Frobenius ([7], Theorem 2.7)

$$\rho(|\mathfrak{A}|) \leq \rho(|\mathfrak{M}| + |\mathfrak{N}|) < 1.$$

Nach Satz 1 in Abschnitt 12 von [3] besitzt daher die Gleichung  $x = \mathfrak{A}x + b$  genau einen Fixpunkt  $x^*$ .

Wegen

$$x^* = \mathfrak{A}x^* + b = \mathfrak{M}x^* + \mathfrak{N}x^* + b$$

folgt unter Verwendung der Regeln (23)–(25) in [3], Seite 155 aus (V)

$$\begin{aligned} q(x^{k+1}, x^*) &= q(\mathfrak{M}x^{k+\frac{1}{2}} + \mathfrak{N}x^{k+1} + b, \mathfrak{M}x^* + \mathfrak{N}x^* + b) \leq \\ &\leq q(\mathfrak{M}x^{k+\frac{1}{2}} + \mathfrak{N}x^{k+1}, \mathfrak{M}x^{k+\frac{1}{2}} + \mathfrak{N}x^*) + q(\mathfrak{M}x^{k+\frac{1}{2}} + \mathfrak{N}x^*, \mathfrak{M}x^* + \mathfrak{N}x^*) \\ &\leq |\mathfrak{N}| q(x^{k+1}, x^*) + |\mathfrak{M}| q(x^{k+\frac{1}{2}}, x^*), \end{aligned}$$

oder wegen  $(\mathfrak{T} - |\mathfrak{N}|)^{-1} \geq \mathfrak{D}$ ,

$$q(x^{k+1}, x^*) \leq (\mathfrak{T} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| q(x^{k+\frac{1}{2}}, x^*).$$

Entsprechend erhält man

$$q(x^{k+\frac{1}{2}}, x^*) \leq (\mathfrak{T} - |\mathfrak{M}|)^{-1} |\mathfrak{N}| q(x^k, x^*),$$

also insgesamt

$$\begin{aligned} q(x^{k+1}, x^*) &\leq (\mathfrak{T} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{T} - |\mathfrak{M}|)^{-1} |\mathfrak{N}| q(x^k, x^*) \\ &\leq \{(\mathfrak{T} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{T} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|\}^{k+1} q(x^0, x^*). \end{aligned}$$

Da der Spektralradius der in geschweiften Klammern stehenden Matrix kleiner als Eins ist, folgt die Behauptung  $\lim_{k \rightarrow \infty} x^k = x^*$ .

Zu (c): Die Gleichung  $x = \mathfrak{A}x + b$  besitze einen eindeutigen Fixpunkt  $x^*$ . Aus der Ungleichung

$$q(x^{k+\frac{1}{2}}, x^*) \leq (\mathfrak{T} - |\mathfrak{M}|)^{-1} |\mathfrak{N}| q(x^k, x^*),$$

die man wie beim Beweis von (b) herleiten kann, folgt, daß auch die Folge  $\{x^{k+\frac{1}{2}}\}_{k=0}^{\infty}$  für beliebiges  $x^0$  gegen  $x^*$  konvergiert. Somit folgt aus dem ersten Teil von (V) für  $k \rightarrow \infty$

$$x^* = \mathfrak{M}x^* + \mathfrak{N}x^* + b,$$

also

$$\begin{aligned} 0 &= q(x^*, x^*) = q(\mathfrak{M}x^* + \mathfrak{N}x^* + b, \mathfrak{A}x^* + b) \\ &= q(\mathfrak{M}x^* + \mathfrak{N}x^*, \mathfrak{A}x^*), \end{aligned}$$

d. h.

$$\mathfrak{A}x^* = (\mathfrak{M} + \mathfrak{N})x^* = \mathfrak{M}x^* + \mathfrak{N}x^*.$$

Es ist jetzt noch  $\rho((\mathfrak{T} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{T} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|) < 1$  zu zeigen. Dazu gehen wir ähnlich wie beim Beweis von Satz 1 in Abschnitt 12 von [3] vor. Aufgrund des

Satzes von Perron und Frobenius ([7], Theorem 2.7) besitzt die nichtnegative Matrix  $(\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|$  einen nichtnegativen Eigenvektor zum nichtnegativen Eigenwert

$$\tilde{\lambda} = \rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|).$$

Wir wählen nun  $x^0$  so, daß  $d(x^0)$  Eigenvektor zum Eigenwert  $\tilde{\lambda}$  ist, und daß mindestens eine Komponente von  $d(x^0)$  größer als die entsprechende Komponente von  $d(x^*)$  ist. Dann folgt aus (V) unter Verwendung von (12) und (18) aus Abschnitt 10 von [3]

$$d(x^{k+\frac{1}{2}}) \geq |\mathfrak{M}| d(x^{k+\frac{1}{2}}) + |\mathfrak{N}| d(x^k)$$

oder

$$d(x^{k+\frac{1}{2}}) \geq (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}| d(x^k),$$

sowie

$$d(x^{k+1}) \geq |\mathfrak{M}| d(x^{k+\frac{1}{2}}) + |\mathfrak{N}| d(x^{k+1})$$

oder

$$d(x^{k+1}) \geq (\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| d(x^{k+\frac{1}{2}}),$$

insgesamt also

$$\begin{aligned} d(x^{k+1}) &\geq (\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}| d(x^k) \\ &\geq \{(\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|\}^{k+1} d(x^0) \\ &= \tilde{\lambda}^{k+1} d(x^0). \end{aligned}$$

Aus der Konvergenz des Iterationsverfahrens (V) gegen  $x^*$  folgt die Konvergenz der Folge  $\{d(x^k)\}_{k=0}^{\infty}$  gegen  $d(x^*)$ . Die Annahme  $\tilde{\lambda} \geq 1$  führt nun auf die Ungleichung

$$d(x^{k+1}) \geq \tilde{\lambda}^{k+1} d(x^0) \geq d(x^0), \quad k \geq 0,$$

und für  $k \rightarrow \infty$  auf

$$d(x^*) \geq d(x^0).$$

Dies steht im Widerspruch zur Wahl von  $x^0$ . Also ist  $\tilde{\lambda} < 1$ . Damit ist der Satz bewiesen. ■

Wir setzen nun in Satz 1 speziell

$$\mathfrak{M} := \mathfrak{Q}, \quad \mathfrak{N} := \mathfrak{U}.$$

Dann ist  $\rho(|\mathfrak{M}|) = \rho(|\mathfrak{N}|) = 0$ . Da aufgrund der speziellen Gestalt der Matrizen  $\mathfrak{Q}$  und  $\mathfrak{U}$  sogar für alle Intervallvektoren die Gleichheit

$$\mathfrak{U} x = \mathfrak{Q} x + \mathfrak{U} x$$

gilt, erhält man aus Satz 1 jetzt unmittelbar das

**Korollar 1:** Das symmetrische Einzelschrittverfahren (SE) konvergiert genau dann für jeden Intervallvektor  $x^0$  gegen den eindeutigen Fixpunkt  $x^*$  der Gleichung  $x = \mathfrak{U} x + b$ , wenn der Spektralradius der reellen Matrix

$$(\mathfrak{I} - |\mathfrak{U}|)^{-1} |\mathfrak{Q}| (\mathfrak{I} - |\mathfrak{Q}|)^{-1} |\mathfrak{U}|$$

kleiner als Eins ist. ■



Beim Beweis von (b) in Satz 1 haben wir bereits gezeigt, daß aus

$$\rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|) < 1$$

stets die Ungleichung  $\rho(|\mathfrak{A}|) < 1$  folgt. Unter der zusätzlichen Voraussetzung

$$|\mathfrak{A}| = |\mathfrak{M}| + |\mathfrak{N}|$$

zeigen wir jetzt, daß mit den bereits in Satz 1 gemachten Voraussetzungen  $\rho(|\mathfrak{M}|) < 1$  und  $\rho(|\mathfrak{N}|) < 1$  auch die Umkehrung gilt. Sei also

$$\rho(|\mathfrak{A}|) = \rho(|\mathfrak{M}| + |\mathfrak{N}|) < 1.$$

Dann existiert nach Theorem 3.8 in [7] die Inverse von  $\mathfrak{I} - |\mathfrak{A}|$  und es gilt  $(\mathfrak{I} - |\mathfrak{A}|)^{-1} \geq \mathfrak{D}$ . Wir betrachten nun die Zerlegung von  $\mathfrak{I} - |\mathfrak{A}|$  in

$$\mathfrak{I} - |\mathfrak{A}| = (\mathfrak{I} - |\mathfrak{M}|) \cdot (\mathfrak{I} - |\mathfrak{N}|) - |\mathfrak{M}| |\mathfrak{N}|.$$

Wegen  $\rho(|\mathfrak{M}|) < 1$ ,  $\rho(|\mathfrak{N}|) < 1$  existieren die Inversen von  $\mathfrak{I} - |\mathfrak{M}|$  und  $\mathfrak{I} - |\mathfrak{N}|$  und diese sind nichtnegativ. Daher ist auch das Produkt  $(\mathfrak{I} - |\mathfrak{N}|)^{-1} (\mathfrak{I} - |\mathfrak{M}|)^{-1}$  nichtnegativ. Da außerdem  $|\mathfrak{M}| |\mathfrak{N}|$  nichtnegativ ist, ist die Zerlegung von  $\mathfrak{I} - |\mathfrak{A}|$  regulär ([7], Definition 3.5), so daß nach Theorem 3.13 in [7] die Beziehung

$$\rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{M}| |\mathfrak{N}|) = \rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|) < 1$$

folgt.

Zusammenfassend erhält man unter Verwendung von (b) aus Satz 1 die folgende Aussage.

**Satz 2:** Die Intervallmatrix  $\mathfrak{A}$  sei zerlegt in die Summe  $\mathfrak{A} = \mathfrak{M} + \mathfrak{N}$  zweier Intervallmatrizen, für welche

$$|\mathfrak{A}| = |\mathfrak{M}| + |\mathfrak{N}|$$

und

$$\rho(|\mathfrak{M}|) < 1, \rho(|\mathfrak{N}|) < 1$$

gilt. Dann ist

$$\rho((\mathfrak{I} - |\mathfrak{N}|)^{-1} |\mathfrak{M}| (\mathfrak{I} - |\mathfrak{M}|)^{-1} |\mathfrak{N}|) < 1$$

genau dann, wenn

$$\rho(|\mathfrak{A}|) < 1$$

gilt. ■

Die Gleichheit  $|\mathfrak{A}| = |\mathfrak{M}| + |\mathfrak{N}|$  besteht z. B. dann, wenn man  $\mathfrak{A} = (A_{ij})$  so in  $\mathfrak{A} = \mathfrak{M} + \mathfrak{N}$  mit  $\mathfrak{M} = (M_{ij})$ ,  $\mathfrak{N} = (N_{ij})$  zerlegt, daß für alle  $1 \leq i, j \leq n$  gilt:

Mindestens einer der beiden Summanden  $M_{ij}$  bzw.  $N_{ij}$  ist gleich Null. Diese Bedingung ist z. B. bei der Zerlegung, welche auf das symmetrische Einzelschrittverfahren führt, erfüllt. Aus Satz 2 erhält man daher unmittelbar

**Korollar 2:** Das symmetrische Einzelschrittverfahren konvergiert genau dann für jeden Intervallvektor  $x^0$  gegen den eindeutigen Fixpunkt  $x^*$  der Gleichung  $x = \mathfrak{A}x + b$ , wenn  $\rho(|\mathfrak{A}|) < 1$  gilt, d. h. wenn das Gesamtschrittverfahren (und damit auch das Einzelschrittverfahren) für jeden Intervallvektor  $x^0$  gegen  $x^*$  konvergiert. ([3], Abschnitt 12, Satz 1 und Satz 4.) ■

#### 4. Das symmetrische Einzelschrittverfahren mit Durchschnittsbildung

In diesem Abschnitt setzen wir voraus, daß alle auftretenden Intervalle reelle kompakte Intervalle oder achsenparallele kompakte Rechtecke der komplexen Zahlenebene sind. Die Gesamtheit dieser Intervalle ist (bei nichtleerem Durchschnitt) bezüglich Durchschnittsbildung abgeschlossen.

Beginnt man nun das Verfahren (SE) unter der Voraussetzung  $\rho(|\mathfrak{A}|) < 1$  mit einem Intervallvektor  $\mathfrak{x}^0 = (X_i^0)$ , welcher  $\mathfrak{x}^* = (X_i^*)$  enthält, so gilt

$$X_1^* = \sum_{j=2}^n A_{1j} X_j^* \subseteq \sum_{j=2}^n A_{1j} X_j^0 = X_1^{\frac{1}{2}}.$$

Es ist daher naheliegend, den Durchschnitt von  $X_1^{\frac{1}{2}}$  mit  $X_1^0$  zu bilden, und die so eventuell nochmals verbesserte Einschließung für  $X_1^*$  bei der Berechnung aller weiteren Komponenten bereits zu verwenden.

Führt man auch bei allen anderen Komponenten die Durchschnittsbildung mit der letzten berechneten Näherung durch, so führt dies auf die folgende Modifikation des Verfahrens (SE):

$$\begin{cases} X_i^{k+\frac{1}{2}} = \left\{ \sum_{j=1}^{i-1} A_{ij} X_j^{k+\frac{1}{2}} + \sum_{j=i+1}^n A_{ij} X_j^k + B_i \right\} \cap X_i^k, & i=1(1)n. \\ X_i^{k+1} = \left\{ \sum_{j=1}^{i-1} A_{ij} X_j^{k+\frac{1}{2}} + \sum_{j=i+1}^n A_{ij} X_j^{k+1} + B_i \right\} \cap X_i^{k+\frac{1}{2}}, & i=n(-1)1. \end{cases}$$

$k=0, 1, 2, \dots, (\mathfrak{x}^* \subseteq \mathfrak{x}^0).$

Wir bezeichnen diese Iterationsvorschrift als *symmetrisches Einzelschrittverfahren mit Durchschnittsbildung nach jeder Komponente* (SEDK)<sup>6</sup>. Durch vollständige Induktion läßt sich  $\mathfrak{x}^* \subseteq \mathfrak{x}^k$ ,  $k=0, 1, 2, \dots$ , zeigen. Ähnlich wie in [3], Abschnitt 14, Satz 1, beweist man außerdem, daß unter den Voraussetzungen von Korollar 1 auch die nach Verfahren (SEDK) berechnete Folge  $\{\mathfrak{x}^k\}_{k=0}^{\infty}$  gegen  $\mathfrak{x}^*$  konvergiert.

Wir betrachten außerdem das Einzelschrittverfahren mit Durchschnittsbildung nach jeder Komponente (EDK) ([3], Abschnitt 14)

$$\left\{ Z_i^{k+1} = \left\{ \sum_{j=1}^{i-1} A_{ij} Z_j^{k+1} + \sum_{j=i+1}^n A_{ij} Z_j^k + B_i \right\} \cap Z_i^k, \quad i=1(1)n \right.$$

$k=0, 1, 2, \dots, (\mathfrak{x}^* \subseteq \mathfrak{z}^0).$

Das Verfahren (SEDK) kann so durchgeführt werden, daß es — abgesehen vom ersten Schritt — pro Schritt die gleiche Anzahl von Intervallmultiplikationen wie das Verfahren (EDK), nämlich  $n^2 - n$ , und damit im wesentlichen den gleichen Aufwand benötigt. (Siehe dazu W. Niethammer [4].) Dazu geht man folgendermaßen vor:

<sup>6</sup> Es ist offensichtlich, wie dieses Verfahren zu modifizieren ist, falls nicht alle Diagonalelemente von  $A$  verschwinden.



Angenommen, für ein  $k > 0$  sind die Summen

$$\sum_{j=i+1}^n A_{ij} X_j^k, \quad i=1(1)n-1$$

bekannt. Dann erfordert die Berechnung von  $x^{k+\frac{1}{2}}$  nach (SEDK) bei einer vollbesetzten Matrix  $\frac{1}{2}(n^2-n)$  Multiplikationen. Merkt man sich bei dieser Berechnung die  $n-1$  Summen

$$\sum_{j=1}^{i-1} A_{ij} X_j^{k+\frac{1}{2}}, \quad i=n(-1)2,$$

so erfordert die Berechnung von  $x^{k+1}$  aus  $x^{k+\frac{1}{2}}$  nochmals  $\frac{1}{2}(n^2-n)$  Multiplikationen. Insgesamt benötigt man also wie beim Verfahren (EDK)  $n^2-n$  Multiplikationen zur Berechnung einer neuen Näherung  $x^{k+1}$  aus  $x^k$ . Werden nun bei der Berechnung von  $x^{k+1}$  aus  $x^{k+\frac{1}{2}}$  die Summen

$$\sum_{j=i+1}^n A_{ij} X_j^{k+1}, \quad i=1(1)n-1,$$

wiederum abgespeichert, so lassen sich alle Überlegungen mit einem um Eins erhöhten Index durchführen.

In [3], Abschnitt 14, Satz 4, wurde gezeigt, daß bei gleichem Startvektor  $z^0$  mit  $x^* \subseteq z^0$  unter allen dort betrachteten Verfahren die Iterierten des Verfahrens (EDK) in jedem Schritt  $x^*$  am engsten einschließen. Hier zeigen wir nun

**Satz 3:** Es seien  $\{x^k\}_{k=0}^{\infty}$  und  $\{z^k\}_{k=0}^{\infty}$  die nach den Verfahren (SEDK) und (EDK) unter der Voraussetzung

$$x^* \subseteq x^0 \subseteq z^0$$

berechneten Folgen. Dann gilt für alle  $k \geq 0$

$$x^* \subseteq x^k \subseteq z^k.$$

*Beweis:* Es gelte

$$x^* \subseteq x^k \subseteq z^k$$

für ein  $k \geq 0$ . Dies ist nach Voraussetzung für  $k=0$  der Fall. Dann folgt aus dem ersten Teil von (SEDK) und aus (EDK) durch vollständige Induktion nach den Komponentenindizes

$$x^* \subseteq x^{k+\frac{1}{2}} \subseteq z^{k+1},$$

und mit Hilfe des zweiten Teiles von (SEDK) wiederum durch vollständige Induktion nach den Komponentenindizes

$$x^* \subseteq x^{k+1} \subseteq x^{k+\frac{1}{2}},$$

insgesamt also

$$x^* \subseteq x^{k+1} \subseteq z^{k+1}.$$

Damit ist der Beweis abgeschlossen. ■

Zusammenfassend erhält man also das wichtige Ergebnis, daß bei gleichem Startvektor  $x^0$  (mit  $x^* \subseteq x^0$ ) das Verfahren (SEDK) in jedem Iterationsschritt eine

Einschließung von  $x^*$  liefert, die mindestens genau so eng ist, wie die mit dem Verfahren (EDK) berechnete. Dabei wird bei beiden Verfahren pro Schritt im wesentlichen der gleiche Aufwand benötigt. Nach den Ergebnissen in [3], Abschnitt 14, ist daher unter allen bisher in der Literatur betrachteten Verfahren zur fortwährenden Einschließung von  $x^*$  das Verfahren (SEDK) in jedem Fall vorzuziehen.

## 5. Numerische Beispiele

Zur Illustration von Satz 3 haben wir verschiedene in der Literatur bereits betrachtete Beispiele auf der Rechenanlage UNIVAC 1108 (Mantissenlänge: 8 Dezimalstellen) am Rechenzentrum der Universität Karlsruhe gerechnet. Aufgrund der Durchschnittsbildung sind die oberen bzw. unteren Schranken der nach (EDK) und (SEDK) berechneten Intervallvektoren monoton fallend bzw. wachsend. Wegen der Endlichkeit der Menge der Maschinenzahlen werden daher auf der Maschine nach endlich vielen Schritten feste Intervallvektoren erreicht:  $x^{k^*} = x^{k^*+1} = \dots$

Bei den Beispielen ist jeweils der Intervallvektor  $x^0 = (X_i^0)$  sowie die Anzahl  $k^*$  der bis zum Stillstand der Iteration benötigten Iterationsschritte angegeben. Diese Zahlen zeigen, daß bei den angegebenen Beispielen das Verfahren (EDK) ungefähr 25% mehr Iterationsschritte als das Verfahren (SEDK) benötigt. Bei den ersten beiden Beispielen sind die Ausgangsdaten und der Fixpunkt echte Intervalle. Wir geben hier auch  $x^{k^*}$  an. Bei den anderen Beispielen geben wir jeweils für verschiedene Werte des Iterationsindex  $k$  den maximalen Durchmesser der Komponenten von  $x^k$ , also  $d^k := \max_{1 \leq i \leq n} \{d(X_i^k)\}$ , an.

Alle hier betrachteten Beispiele wurden jeweils durch Auflösen nach den Diagonalelementen auf die Form  $x = \mathfrak{A} x + b$  gebracht, so daß  $\mathfrak{A}$  verschwindende Diagonalelemente besitzt.

Für die Programmierung der Beispiele danke ich Herrn Kurt Speck sehr herzlichst.

I. Beispiel 1 aus [2], Seite 356.

$$x^0 = \begin{pmatrix} [ 0.9; 1.2] \\ [ 0.4; 0.7] \\ [ 0 ; 0.2] \\ [-0.4; -0.1] \end{pmatrix}$$

(EDK) :  $k^* = 37$

(SEDK) :  $k^* = 31$

$$x^{k^*} = \begin{pmatrix} [ 1.032 860 1 ; 1.059 757 9 ] \\ [ 0.550 754 40 ; 0.574 813 98] \\ [ 0.099 483 623; 0.122 513 79] \\ [-0.243 545 82 ; -0.212 698 41] \end{pmatrix}$$



II. Beispiel 2 aus [2], S. 357.

$$\bar{x}^0 = \begin{pmatrix} [0.8 \ ; \ 1.0 \ ] \\ [0.65; 0.85] \\ [0.55; 0.7 \ ] \end{pmatrix}$$

(EDK) :  $k^* = 18$

(SEDK) :  $k^* = 15$

$$\bar{x}^{k^*} = \begin{pmatrix} [0.896 \ 368 \ 17; 0.896 \ 479 \ 91] \\ [0.765 \ 057 \ 55; 0.765 \ 202 \ 25] \\ [0.614 \ 247 \ 34; 0.614 \ 521 \ 84] \end{pmatrix}$$

III. Beispiel 1 aus [6], Seite 332.

$\bar{x}^0 = (X_i^0)$  mit  $X_i^0 = [0; 1], i = 1 (1) 8$ .

(EDK) :  $k^* = 27$ ,

(SEDK) :  $k^* = 22$ .

$d^k \backslash k$	0	5	10	15	20
(EDK)	1	$4.0_{10} - 2$	$8.4_{10} - 4$	$1.8_{10} - 5$	$3.8_{10} - 7$
(SEDK)	1	$7.0_{10} - 3$	$5.4_{10} - 5$	$4.1_{10} - 7$	$3.7_{10} - 9$

IV. Beispiel 2 aus [6], Seite 333.

$\bar{x}^0 = (X_i^0)$  mit  $X_i^0 = [0; 0.5], i = 1 (1) 4$ .

(EDK) :  $k^* = 56$ ,

(SEDK) :  $k^* = 44$ .

$d^k \backslash k$	0	5	10	20	30	40
(EDK)	$5.0_{10} - 1$	$1.1_{10} - 1$	$2.9_{10} - 2$	$6.7_{10} - 4$	$2.7_{10} - 5$	$1.5_{10} - 6$
(SEDK)	$5.0_{10} - 1$	$3.7_{10} - 2$	$4.7_{10} - 3$	$7.4_{10} - 5$	$1.2_{10} - 6$	$1.8_{10} - 8$

V. Beispiel A in [5], Seite 85.

$\bar{x}^0 = (X_i^0)$  mit  $X_i^0 = [-0.036 \ 016; 0.674 \ 056], i = 1 (1) 3$ .

(EDK) :  $k^* = 52$ ,

(SEDK) :  $k^* = 42$ .

$d^k \backslash k$	0	5	10	20	40
(EDK)	$7.1_{10} - 1$	$1.2_{10} - 1$	$2.0_{10} - 2$	$5.3_{10} - 4$	$3.7_{10} - 7$
(SEDK)	$7.1_{10} - 1$	$5.5_{10} - 2$	$5.7_{10} - 3$	$6.2_{10} - 5$	$3.7_{10} - 9$

## VI. Wie V. aber mit

$$X_i^0 = [0.059\ 459; 0.643\ 243], \quad i = 1(1)3.$$

$$(\text{EDK}) : k^* = 51,$$

$$(\text{SEDK}) : k^* = 41.$$

$d^k$ \ $k$	0	5	10	20	40
(EDK)	$5.8_{10} - 1$	$1.0_{10} - 1$	$1.7_{10} - 2$	$4.3_{10} - 4$	$3.0_{10} - 7$
(SEDK)	$5.8_{10} - 1$	$4.5_{10} - 2$	$4.7_{10} - 3$	$5.1_{10} - 5$	$3.7_{10} - 9$

## Literatur

- [1] Aitken, A. C.: Studies in practical mathematics V. On the iterative solution of a system of linear equations. Proc. Roy. Soc. Edinburgh Sec. A 63, 52—60 (1960).
- [2] Albrecht, J.: Monotone Iterationsfolgen und ihre Verwendung zur Lösung linearer Gleichungssysteme. Num. Mathematik 3, 345—358 (1961).
- [3] Alefeld, G., Herzberger, J.: Einführung in die Intervallrechnung. Mannheim: Bibliographisches Institut 1974.
- [4] Niethammer, W.: Relaxation bei komplexen Matrizen. Math. Zeitschrift 86, 34—40 (1964).
- [5] Schmidt, J. W.: Ausgangsvektoren für monotone Iterationen bei linearen Gleichungssystemen. Num. Mathematik 6, 78—88 (1964).
- [6] Schröder, J.: Computing Error Bounds in Solving Linear Systems. Math. Comp. 16, 323—337 (1962).
- [7] Varga, R. S.: Matrix Iterative Analysis. Englewood Cliffs, N.J.: Prentice-Hall 1962.
- [8] Young, D. M.: Iterative Solution of Large Linear Systems. New York: Academic Press 1971.
- [9] Young, D. M.: On the Accelerated SSOR Method for solving Large Linear Systems. Report CNA-92. Center for Numerical Analysis. The University of Texas, 1974.

Prof. Dr. G. Alefeld  
 Fachbereich 3 / Mathematik  
 Technische Universität Berlin  
 Straße des 17. Juni 135  
 D-1000 Berlin 12