Aplikace matematiky

Götz Alefeld

Über die Existenz einer eindeutigen Lösung bei einer Klasse nichtlinearer Gleichungssysteme und deren Berechnung mit Iterationsverfahren

Aplikace matematiky, Vol. 17 (1972), No. 4, 267–294

Persistent URL: http://dml.cz/dmlcz/103418

ÜBER DIE EXISTENZ EINER EINDEUTIGEN LÖSUNG BEI EINER KLASSE NICHTLINEARER GLEICHUNGSSYSTEME UND DEREN BERECHNUNG MIT ITERATIONSVERFAHREN

GÖTZ ALEFELD

(Eingegangen am 30. April 1971)

EINLEITUNG

Die vorliegende Arbeit beschäftigt sich mit Aussagen über die Existenz einer eindeutigen Lösung bei einer Klasse nichtlinearer Gleichungssysteme und mit verschiedenen Verfahren zur Berechnung dieser. In Abschnitt 2 wird für ein geeignet definiertes Relaxationsverfahren in Gesamt- bzw. Einzelschritten die Konvergenz gegen die eindeutige Lösung bewiesen. In Abschnitt 3 werden Verallgemeinerungen dieser Relaxationsverfahren angegeben, bei welchen die Lösung fortwährend eingeschloßen werden kann. Diese Verfahren können so modifiziert werden, daß die einschließenden Folgen monoton gegen die Lösung konvergieren (Abschnitt 4). Es wird gezeigt, daß diese Verfahren unter den in Abschnitt 2 für die Existenz und Eindeutigkeit der Lösung angegebenen Bedingungen definiert sind und gegen die Lösung konvergieren. In Abschnitt 5 werden die in den Abschnitten 3 und 4 angegebenen Verfahren bezüglich ihrer Konvergenzgeschwindigkeit verglichen und das optimale Verfahren angegeben. Abschnitt 6 enthält hinreichende Kriterien für die Existenz genau einer Lösung und die Abschnitte 7 und 8 einige Möglichkeiten zur Berechnung dieser unter Verwendung der Ableitung, ohne jedoch die beim Newton-Verfahren entstehenden linearen Gleichungssysteme vollständig aufzulösen. Im Abschnitt 9 werden diese Verfahren so modifiziert, daß man monoton gegen die Lösung konvergierende Folgen erhält. Schließlich enthält Abschnitt 10 Aussagen über die Lösung einschließende und monoton überlinear konvergente Verfahren. Der letzte Abschnitt enthält numerische Beispiele. Es wird gezeigt, daß die in den Abschnitten 9 und 10 angegebenen Verfahren vorteilhaft zur Berechnung der Lösung der durch Diskretisierung einer Klasse von Randwertaufgaben entstehenden nichtlinearen Gleichungssysteme verwendet werden können.

Mit $\mathsf{x} = (x_i)$, $\mathsf{y} = (y_i)$, ... bezeichnen wir reelle n-dimensionale Vektoren: $\mathsf{x}, \mathsf{y}, \ldots \in V_n(\boldsymbol{R})$, entsprechend mit $\mathsf{X} = (x_{ij})$, $\mathsf{Y} = (y_{ij})$ reelle $n \times n$-Matrizen: $\mathsf{X}, \mathsf{Y}, \ldots \in M_n(\boldsymbol{R})$. Unter $I(\boldsymbol{R})$ verstehen wir die Menge der reellen abgeschlossenen und beschränkten Intervalle $X = [x_1, x_2]$, $Y = [y_1, y_2]$, ..., und unter $V_n(I(\boldsymbol{R}))$ die Menge der Vektoren $\mathbf{x} = (X_i)$, $\mathbf{y} = (Y_i)$, ... mit $X_i, Y_i, \ldots \in I(\boldsymbol{R})$. Entsprechend bezeichnet $M_n(I(\boldsymbol{R}))$ die Menge der Matrizen $\mathbf{X} = (X_{ij})$, $\mathbf{Y} = (Y_{ij})$, ... mit $X_{ij}, Y_{ij} \in I(\boldsymbol{R})$. In $I(\boldsymbol{R})$ sind vier Verknüpfungen definiert, welche sich auf die Verknüpfung der Intervallschranken zurückführen lassen. Verknüpfungen in $M_n(I(\boldsymbol{R}))$ sind wie üblich definiert.

In $I(\boldsymbol{R})$ wir durch

$$q(X, Y) := \max\left(|x_1 - y_1|, |x_2 - y_2|\right)$$

eine Metrik definiert. $(I(\boldsymbol{R}), q)$ ist vollständig.

$$|X| := q(X, 0) \quad \text{bzw.} \quad d(X) := x_2 - x_1$$

heißt Betrag bzw. Durchmesser des Intervalles X. Es wird Gebrauch von folgenden Rechenregeln gemacht:

$$\begin{aligned}
&q(A + C, B + C) = q(A, B)\,; && x \in \boldsymbol{R} \Rightarrow d(x) = 0\,;\\
&q(AB, AC) \leqq |A|\, q(B, C)\,; && d(A \pm B) = d(A) + d(B)\,;\\
&q(A + B, C + D) \leqq q(A, C) + q(B, D)\,; && d(AB) \leqq |A|\, d(B) + |B|\, d(A)\,;\\
& && d(A) = |A - A|\,;
\end{aligned}$$

$m(X) := \frac{1}{2}(x_1 + x_2)$ heißt Mittelpunkt des Intervalles X.

Mit Hilfe von d, m und $|.|$ definiert man

$$d(\mathbf{x}) := (d(X_i))\,, \quad m(\mathbf{x}) := (m(X_i))\,, \quad |\mathbf{x}| := (|X_i|) \quad \text{für} \quad \mathbf{x} \in V_n(I(\boldsymbol{R}))$$

bzw.

$$d(\mathbf{X}) := (d(X_{ij}))\,, \quad m(\mathbf{X}) := (m(X_{ij}))\,, \quad |\mathbf{X}| := (|X_{ij}|) \quad \text{für} \quad \mathbf{X} \in M_n(I(\boldsymbol{R}))\,.$$

Eine Abbildung

$$\mathsf{f} : \mathsf{D} \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})\,, \quad \mathsf{f} = (f_i(\mathsf{x}))\,,$$

heißt diagonal wenn $f_i(\mathsf{x}) = f_i(x_i)$, $i = 1(1)\,n$, gilt.

Wir setzen stets voraus, daß $V_n(\boldsymbol{R})$ durch

$$\mathsf{x} \leqq \mathsf{y} \Leftrightarrow x_i \leqq y_i\,, \quad i = 1(1)\,n\,,$$

halbgeordnet ist.

Eine Abbildung

$$\mathsf{f} : D \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R}), \quad \mathsf{f} = (f_i(\mathsf{x}))$$

heißt isoton in $D_0 \subset D$, wenn aus $\mathsf{x}, \mathsf{y} \in D_0$, $\mathsf{x} \leqq \mathsf{y}$, die Beziehung $\mathsf{f}(\mathsf{x}) \leqq \mathsf{f}(\mathsf{y})$ folgt.

Ist $\mathbf{x} \in V_n(I(\boldsymbol{R}))$ und $\mathsf{f} : D \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ stetig und $\mathbf{x} \subset D$, so bezeichnen wir mit $\mathbf{f}(\mathbf{x}) \in V_n(I(\boldsymbol{R}))$ einen Intervallvektor, für den gilt

$$\{\mathsf{f}(\mathsf{x}) \mid \mathsf{x} \in \mathbf{x}\} \subset \mathbf{f}(\mathbf{x}).$$

Im folgenden wird vorausgesetzt, daß alle auftretenden Abbildungen der Teilmengeneigenschaft genügen:

$$\mathbf{x} \subset \mathbf{y} \Rightarrow \mathbf{f}(\mathbf{x}) \subset \mathbf{f}(\mathbf{y}),$$

und daß mit $\lim\limits_{m \to \infty} \mathbf{x}^{(m)} = \mathsf{x}$ die Beziehung $\lim\limits_{m \to \infty} d(\mathbf{f}(\mathbf{x}_m)) = \mathsf{o}$ gilt (Stetigkeit der intervallmäßigen Auswertung).

Die meisten der in dieser Arbeit verwendeten Hilfsmittel der Intervallrechnung findet man in [1].

Mit $\varrho(\mathsf{A})$ bezeichnen wir den Spektralradius einer Matrix $\mathsf{A} \in M_n(\boldsymbol{R})$.

$\mathsf{A} = \mathsf{D} - \mathsf{B}$ (D Diagonalmatrix) heißt M-Matrix, wenn $a_{ij} \leqq 0$ für $i \neq j$ und $\mathsf{A}^{-1} \geqq \mathsf{O}$ gilt. Für eine M-Matrix A gilt $\varrho(\mathsf{D}^{-1}\mathsf{B}) < 1$. (Siehe etwa [4].)

Zum Beweis von einigen unten angegebenen Sätzen zitieren wir die folgenden Aussagen:

Satz 1. *Sei* $\mathsf{A} = \mathsf{D} - \mathsf{B}$, $\mathsf{B} = \mathsf{L} + \mathsf{R}$ (D *Diagonalmatrix*, L *bzw.* R *strenge untere bzw. obere Dreiecksmatrix*), D *nicht singulär. Ist* $\varrho(|\mathsf{D}^{-1}\,\mathsf{B}|) < 1$, *so gilt*

$$\varrho((\mathsf{E} - \omega|\mathsf{D}^{-1}\,\mathsf{L}|)^{-1}\{|1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\,\mathsf{R}|\} < 1 \quad \textit{für}$$
$$0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\,\mathsf{B}|)).$$

(Der Beweis ergibt sich unmittelbar aus Satz 1 in [2], wenn man dort $\mathsf{A} := \mathsf{E} - \mathsf{D}^{-1}\,\mathsf{B}$ setzt.)

Satz 2. *Unter den Voraussetzungen von Satz* 1 *gilt*

$$\varrho(|1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\,\mathsf{B}|) < 1 \quad \textit{für} \quad 0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\,\mathsf{B}|)).$$

Beweis. Die Matrix $\mathsf{P} = (\mathsf{E} - \omega|\mathsf{D}^{-1}\,\mathsf{L}|)^{-1}\{|1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\,\mathsf{R}|\}$ kann aufgefaßt werden als die zur Matrix $\mathsf{Q} = |1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\,\mathsf{B}|$ gehöreige Einzelschrittmatrix. Nach einem Satz von Stein und Rosenberg (siehe [4], Seite 68) und einer Verallgemeinerung davon (siehe [2]) folgt aus $\varrho(\mathsf{P}) < 1$ die Beziehung $\varrho(\mathsf{Q}) < 1$. Aus Satz 1 folgt daher die Behauptung.

Gegeben sei das nichtlineare Gleichungssystem

$$\mathsf{f}(\mathsf{x}) = \mathsf{o}$$

mit $\mathsf{f}(\mathsf{x}) = (f_i(x_1, \ldots, x_n))$.

Dann heißt die Iterationsvorschrift

$$\mathsf{x}^{(0)} \in V_n(\boldsymbol{R}) ;$$

Löse

$$f_i(x_1^{(k)}, \ldots, x_{i-1}^{(k)}, x_i, x_{i+1}^{(k)}, \ldots, x_n^{(k)}) = 0 , \quad i = 1(1)\, n , \quad k = 0, 1, 2, \ldots .$$

Setze

$$x_i^{(k+1)} = x_i^{(k)} + \omega(x_i - x_i^{(k)}) , \quad i = 1(1)\, n , \quad k = 0, 1, 2, \ldots , \quad (\omega > 0) .$$

Relaxationsverfahren in Gesamtschritten (RG).

Entsprechend heißt das Verfahren

$$\mathsf{x}^{(0)} \in V_n(\boldsymbol{R}) ;$$

Löse

$$f_i(x_{(k+1)}^{(k)}, \ldots, x_{i-1}^{(k+1)}, x_i, x_{i+1}^{(k)}, \ldots, x_n^{(k)}) = 0 , \quad i = 1(1)\, n , \quad k = 0, 1, 2, \ldots .$$

Setze

$$x_i^{(k+1)} = x_i^{(k)} + \omega(x_i - x_i^{(k)}) , \quad i = 1(1)\, n , \quad k = 0, 1, 2, \ldots , \quad (\omega > 0) .$$

Relaxationsverfahren in Einzelschritten (RE).

Im folgenden Satz wird für eine Klasse von Systemen $\mathsf{f}(\mathsf{x}) = \mathsf{o}$ die Existenz einer eindeutigen Lösung und die globale Konvergenz der Verfahren (RG) und (RE) gezeigt.

Satz 3. (*Globale Konvergenzaussage für die Verfahren* (RG) *und* (RE)). *Sei* $\mathsf{A} \in M_n(\boldsymbol{R})$ *und* $\mathsf{A} = \mathsf{D} - \mathsf{L} - \mathsf{R}$ *mit einer nichtsingulären Diagonalmatrix* $\mathsf{D} \geqq \mathsf{O}$, *einer strengen unteren Dreiecksmatrix* L *und einer strengen oberen Dreiecksmatrix* R, $\mathsf{B} = \mathsf{L} + \mathsf{R}$. *Es sei* $\varrho(|\mathsf{D}^{-1}\,\mathsf{B}|) < 1$ *und* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *eine stetige, diagonale und isotone Abbildung mit* $\mathsf{b}(\mathsf{x}) = (b_i(x_i))$. *Dann sind die Verfahren* (RG) *und* (RE) *für beliebiges* $\mathsf{x}^{(0)} \in V_n(\boldsymbol{R})$ *definiert und für* $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$ *konvergent gegen die eindeutige Lösung von* $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$.

Beweis. (a) Das Verfahren (RG) nimmt unter den Voraussetzungen dieses Satzes die folgende Gestalt an:

$$\mathsf{x}^{(0)} \in V_n(\boldsymbol{R}) ;$$

Löse

$$a_{ii}x_i + b_i(x_i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j^{(k)} = 0 , \quad i = 1(1)\, n , \quad k = 0, 1, 2, \ldots .$$

Setze

$$x_i^{(k+1)} = x_i^{(k)} + \omega(x_i - x_i^{(k)})\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots .$$

Wir setzen

$$r_i(t) = a_{ii}t + b_i(t)\,, \quad i = 1(1)\,n\,, \quad t \in \boldsymbol{R}\,.$$

Wegen $a_{ii} > 0$ und der Isotonie von b_i ist die Abbildung $r_i : \boldsymbol{R} \to \boldsymbol{R}$ bijektiv. Damit ist die Folge $\mathsf{x}^{(k)}$ definiert, insbesondere ist die Abbildung $\mathsf{D} + \mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ mit $(\mathsf{D} + \mathsf{b})(\mathsf{x}) := \mathsf{Dx} + \mathsf{b}(\mathsf{x})$ invertierbar für alle $\mathsf{x} \in V_n(\boldsymbol{R})$. Wir setzen

$$\mathsf{g} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R}) \quad \text{mit} \quad \mathsf{g}(\mathsf{x}) = (1 - \omega)\,\mathsf{x} + \omega(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{Bx})\,.$$

Aus der Stetigkeit der $b_i(t)$ und der strengen Isotonie von $r_i(t)$ folgt die Stetigkeit von $(\mathsf{D} + \mathsf{b})^{-1}$ und damit die Stetigkeit von g.

x^* ist Fixpunkt von g genau dann, wenn $\mathsf{f}(\mathsf{x}) = \mathsf{o}$ gilt. Außerdem gilt für die Iterierten $\mathsf{x}^{(k)}$ von (RG)

$$\mathsf{x}^{(k+1)} = (1 - \omega)\,\mathsf{x}^{(k)} + \omega(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{Bx}^{(k)})\,.$$

Wir zeigen, daß g kontrahierend ist. Wegen der Isotonie der b_i gilt

$$|t_1 - t_2| \leqq \left|t_1 - t_2 + \frac{1}{a_{ii}}\,[b_i(t_1) - b_i(t_2)]\right| = \frac{1}{a_{ii}}\,|r_i(t_1) - r_i(t_2)|$$

für alle $t_1, t_2 \in \boldsymbol{R}$, $i = 1(1)\,n$. Sind $s_1, s_2 \in \boldsymbol{R}$ und $t_1 = r_i^{-1}(s_1)$, $t_2 = r_i^{-1}(s_2)$, so gilt daher

$$|r_i^{-1}(s_1) - r_i^{-1}(s_2)| \leqq \frac{1}{a_{ii}}\,|s_1 - s_2|\,, \quad i = 1(1)\,n\,,$$

oder

$$|(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{x}) - (\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{y})| \leqq \mathsf{D}^{-1}|\mathsf{x} - \mathsf{y}| \quad \text{für alle} \quad \mathsf{x}, \mathsf{y} \in V_n(\boldsymbol{R})\,.$$

Damit folgt

$$\begin{aligned}
&|\mathsf{g}(\mathsf{x}) - \mathsf{g}(\mathsf{y})| = \\
&= |(1 - \omega)\,\mathsf{x} + \omega(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{Bx}) - (1 - \omega)\,\mathsf{y} - \omega(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{By})| \leqq \\
&\leqq |1 - \omega|\,|\mathsf{x} - \mathsf{y}| + \omega|(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{Bx}) - (\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{By})| \leqq \\
&\leqq |1 - \omega|\,|\mathsf{x} - \mathsf{y}| + \omega\mathsf{D}^{-1}|\mathsf{Bx} - \mathsf{By}| \leqq \\
&\leqq \{|1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{B}|\}\,|\mathsf{x} - \mathsf{y}|\,.
\end{aligned}$$

Nach Satz 2 gilt $\varrho(|1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{B}|) < 1$ für $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$. Damit ist die Behauptung für das Verfahren (RG) gezeigt.

(b) Das Verfahren (RE) nimmt unter den Voraussetzungen dieses Satzes die folgende Gestalt an:

$$\mathsf{x}^{(0)} \in V_n(\boldsymbol{R})\,.$$

Löse

$$a_{ii}x_i + b_i(x_i) + \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} = 0\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots .$$

Setze

$$x_i^{(k+1)} = x_i^{(k)} + \omega(x_i - x_i^{(k)})\,, \quad (\omega > 0)\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots .$$

Wie oben setzen wir

$$r_i(t) = a_{ii}t + b_i(t)\,, \quad i = 1(1)\,n\,, \quad t \in \boldsymbol{R}\,.$$

Dann läßt sich diese Iterationsvorschrift in der Form

$$x_i^{(k+1)} = (1 - \omega)\, x_i^{(k)} + \omega r_i^{-1}\Big(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\Big)\,, \quad i = 1(1)\,n\,,$$

schreiben. Wir setzen

$$\mathsf{g} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R}) \quad \text{mit}$$
$$\mathsf{g}(\mathsf{x}) = (1 - \omega)\,\mathsf{x} + \omega(\mathsf{D} + \mathsf{b})^{-1}\,(\mathsf{L}\mathsf{g}(\mathsf{x}) + \mathsf{R}\mathsf{x})\,.$$

g ist stetig. Für zwei beliebige Vektoren $\mathsf{x}, \mathsf{y} \in V_n(\boldsymbol{R})$ gilt mit $\mathsf{g}(\mathsf{x}) = (g_i(\mathsf{x}))$, $\mathsf{g}(\mathsf{y}) = (g_i(\mathsf{y}))$

$$\begin{aligned}
|\mathsf{g}(\mathsf{x}) - \mathsf{g}(\mathsf{y})| &= \Big(\Big|(1 - \omega)\, x_i + \omega r_i^{-1}\Big(-\sum_{j=1}^{i-1} a_{ij}g_j(\mathsf{x}) - \sum_{j=i+1}^{n} a_{ij}x_j\Big) - \\
&\quad - (1 - \omega)\, y - \omega r_i^{-1}\Big(-\sum_{j=1}^{i-1} a_{ij}g_i(\mathsf{y}) - \sum_{j=i+1}^{n} a_{ij}y_i\Big)\Big|\Big) - \\
&\quad - \Bigg(|1 - \omega|\, |x_i - y_i| + \frac{\omega}{a_{ii}}\Bigg|-\sum_{j=1}^{i-1} a_{ij}g_j(\mathsf{x}) - \sum_{j=i+1}^{n} a_{ij}x_j + \\
&\quad + \sum_{j=1}^{i-1} a_{ij}g_j(\mathsf{y}) + \sum_{j=i+1}^{n} a_{ij}y_j\Bigg|\Bigg) \leqq \Bigg(|1 - \omega|\, |x_i - y_i| + \frac{\omega}{a_{ii}} \sum_{j=1}^{i-1} |a_{ij}|\, |g_j(\mathsf{x}) - g_j(\mathsf{y})| + \\
&\quad + \frac{\omega}{a_{ii}} \sum_{j=i+1}^{n} |a_{ij}|\, |x_j - y_j|\Bigg) = (|1 - \omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{R}|)\,|\mathsf{x} - \mathsf{y}| + \omega|\mathsf{D}^{-1}\mathsf{L}|\,\Big|\mathsf{g}(\mathsf{x}) - \mathsf{g}(\mathsf{y})\Big)\Big|\,.
\end{aligned}$$

Daraus folgt

$$|\mathsf{g}(\mathsf{x}) - \mathsf{g}(\mathsf{y})| \leqq \mathsf{P}|\mathsf{x} - \mathsf{y}|$$

mit

$$\mathsf{P} = (\mathsf{E} - \omega|\mathsf{D}^{-1}\mathsf{L}|)^{-1} \cdot \{|1 - \omega|\mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{R}|\}\,.$$

Nach Satz 1 gilt $\varrho(\mathsf{P}) < 1$ für $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$. Damit ist der Satz bewiesen.

Als unmittelbare Folge von Satz 3 erhalten wir

Korollar 1. $\mathsf{A} \in M_n(\boldsymbol{R})$ *sei eine M-Matrix und* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *erfülle die Voraussetzungen von Satz 3. Dann sind die Verfahren* (RG) *und* (RE) *für beliebiges* $\mathsf{x}^{(0)} \in V_n(\boldsymbol{R})$ *definiert und für* $0 < \omega < 2/(1 + \varrho(\mathsf{D}^{-1}\mathsf{B}))$ *konvergent gegen die eindeutige Lösung* x^* *von* $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$.

Korollar 2. *Ist* $\mathsf{A} = \mathsf{D} - \mathsf{B} \in M_n(\boldsymbol{R})$ *mit einer nichtsingulären Diagonalmatrix* $\mathsf{D} \geqq \mathsf{O}$, A *entweder streng diagonal-dominant oder irreduzibel diagonaldominant, und gelten für* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *die Voraussetzungen von Satz* 3, *so sind die Verfahren* (RG) *und* (RE) *für beliebiges* $\mathsf{x}^{(0)} \in V_n(\boldsymbol{R})$ *definiert, und für* $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$ *konvergent gegen die eindeutige Lösung* x^* *von* $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$.

Bemerkung. Die Aussagen von Korollar 1 sind für $0 < \omega \leqq 1$ in [6] bewiesen worden.

3. EINSCHLIESSENDE FOLGEN MIT DEM RELAXATIONSVERFAHREN IN GESAMT- BZW. EINZELSCHRITTEN

Wir zeigen nun, daß man unter den Voraussetzungen von Satz 3 Folgen von. Intervallvektoren angeben kann, deren Schranken gegen die Lösung x^* konvergieren. Gilt darüberhinaus $\mathsf{x}^* \in \mathbf{x}^{(0)}$, so gilt $\mathsf{x}^* \in \mathbf{x}^{(k)}$ für alle k. Dazu betrachten wir zunächst die folgende Verallgemeinerung des Verfahrens (RG):

$$\mathbf{x}^{(0)} = (X_i^{(0)}) \in V_n(I(\boldsymbol{R}))\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots,$$

$$V_i^{(k)} = [v_{i,1}^{(k)}, v_{i,2}^{(k)}] := -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij} X_j^{(k)}\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots.$$

Löse

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)}\;; \quad a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}^{(k)}\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots.$$

Setze

$$X_i^{(k+1)} = (1 - \omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}]\,, \quad (\omega > 0)\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots.$$

Wir bezeichnen dieses Verfahren als intervallmäßige Durchführung des Relaxationsverfahrens in Gesamtschritten (RGI).

In Analogie zum Verfahren (RGI) betrachten wir die Iterationsvorschrift:

$$\mathbf{x}^{(0)} = (X_i^{(0)}) \in V_n(I(\boldsymbol{R}))\;;$$

$$V_i^{(k)} = [v_{i,1}^{(k)}, v_{i,2}^{(k)}] := -\sum_{j=1}^{i-1} a_{ij} X_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} X_j^{(k)}\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots.$$

Löse

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)}\ ; \quad a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}^{(k)}\ , \quad i = 1(1)\,n\ , \quad k = 0, 1, 2, \ldots .$$

Setze

$$X_i^{(k+1)} = (1 - \omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}]\ , \quad (\omega > 0)\ , \quad i = 1(1)\,n\ , \quad k = 0, 1, 2, \ldots .$$

Dieses Iterationsverfahren bezeichnen wir als intervallmäßige Durchführung des Relaxationsverfahrens in Einzelschritten (REI). Es gilt der folgende

Satz 4. *Für* A *und* $\mathsf{b}(\mathsf{x})$ *mögen die Voraussetzungen von Satz* 3 *bestehen. Dann ist das Iterationsverfahren* (RGI) *für beliebiges* $\mathbf{x}^{(0)} \in V_n(I(\boldsymbol{R}))$ *definiert und die Folge* $\mathbf{x}^{(k)}$ *konvergiert für* $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$ *gegen die Lösung* x^* *von* $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$. *Gilt darüberhinaus* $\mathsf{x}^* \in \mathbf{x}^{(0)}$, *so gilt* $\mathsf{x}^* \in \mathbf{x}^{(k)}$ *für falle k. Dieselben Aussagen gelten für das Verfahren* (REI).

Beweis. Wie im Beweis von Satz 3 unter (a) setzen wir

$$r_i(t) = a_{ii}t + b_i(t)\ , \quad i = 1(1)\,n\ , \quad t \in \boldsymbol{R}\ .$$

Wegen der strengen Isotonie von $r_i(t)$ folgt $x_{i,1} \leqq x_{i,2}$, $i = 1(1)\,n$. Denn wäre $x_{i,1} > x_{i,2}$, so würde

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)} > a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}^{(k)}\ ,$$

also ein Widerspruch folgen. Somit ist die Folge $\mathbf{x}^{(k)} = (X_i^{(k)})$ definiert und es gilt

$$\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)}) = (g_i(\mathbf{x}^{(k)}))$$

mit

$$g_i(\mathbf{x}^{(k)}) = (1 - \omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}] =$$

$$= (1 - \omega)\,X_i^{(k)} + \omega[r_i^{-1}((-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}X_j^{(k)})_1),\ r_i^{-1}((-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}X_j^{(k)})_2)]\ .$$

Die Abbildung $\mathbf{g} : V_n(I(\boldsymbol{R})) \to V_n(I(\boldsymbol{R}))$ ist stetig. Wir zeigen, daß $\mathbf{g}$ kontrahierend ist. Es gilt für $\mathbf{x} = (X_i)$, $\mathbf{y} = (Y_i) \in V_n(I(\boldsymbol{R}))$

$$q(g_i(\mathbf{x}), g_i(\mathbf{y})) = \max_{m=1,2} (|((1 - \omega)\,X_i)_m + \omega r_i^{-1}((-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}X_j)_m) -$$

$$- ((1 - \omega)\,Y_i)_m - \omega r_i^{-1}((-\sum_{\substack{j=1\\j\neq 1}}^{n} a_{ij}Y_j)_m)|) \leqq$$

$$\leqq \max_{m=1,2} (|((1 - \omega)\,X_i)_m - ((1 - \omega)\,Y_i)_m| + \omega|r_i^{-1}((-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}X_j)_m) -$$

$$- r_i^{-1}((-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}Y_j)_m)|) \leqq$$

$$\leqq q((1-\omega)X_i, (1-\omega)Y_i) + \frac{\omega}{a_{ii}} q(-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}X_j, -\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}Y_j) \leqq$$

$$\leqq |1-\omega|\, q(X_i, Y_i) + \frac{\omega}{a_{ii}} \sum_{\substack{j=1\\j\neq i}}^{n} |a_{ij}|\, q(X_j, Y_j),$$

also

$$(q(g_i(\mathbf{x}), g_i(\mathbf{y})) \leqq \{|1-\omega|\, \mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{B}|\}\,(q(X_i, Y_i)).$$

Nach Satz 2 folgt der erste Teil der Behauptung.

Es ist noch die letzte Behauptung des Satzes zu zeigen. Ist $\mathbf{x}^* \in \mathbf{x}^{(k)}$ so gilt

$$a_{ii}x_i^* + b_i(x_i^*) = -\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}x_j^* \in -\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}X_j^{(k)} = [v_{i,1}^{(k)}, v_{i,2}^{(k)}], \quad i = 1(1)\, n.$$

Wäre nun $x_i^* < x_{i,1}$, so würde

$$a_{ii}x_i^* + b_i(x_i^*) < a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)}$$

und entsprechend für $x_i^* > x_{i,2}$

$$v_{i,2}^{(k)} = a_{ii}x_{i,2} + b_i(x_{i,2}) < a_{ii}x_i^* + b_i(x_i^*),$$

also

$$a_{ii}x_i^* + b_i(x_i^*) \notin [v_{i,1}^{(k)} v_{i,2}^{(k)}]$$

folgen. Dies steht im Widerspruch zur oben angegebenen Beziehung. Somit gilt $x_i^* \in [x_{i,1}, x_{i,2}]$ und aufgrund der Teilmengeneigenschaft der Intervallrechnung

$$x_i^* = (1-\omega)\, x_i^* + \omega x_i^* \in (1-\omega)\, X_i^{(k)} + \omega[x_{i,1}, x_{i,2}] = X_i^{(k+1)}, \quad i = 1(1)\, n.$$

Das ist die Behauptung. Für das Verfahren (REI) wird der Beweis dieses Satzes ähnlich geführt.

Der Satz läßt sich unmittelbar auf Systeme der in Korollar 1 und 2 beschriebenen Form anwenden.

4. MONOTONE EINSCHLIESSENDE FOLGEN MIT DEN RELAXATIONSVERFAHREN IN GESAMT- UND EINZELSCHRITTEN

Die Schranken der Intervallvektoren $\mathbf{x}^{(k)}$ nach den Verfahren (RGI) bzw. (REI) konvergieren im allgemeinen nicht monoton gegen $\mathbf{x}^*$, selbst dann nicht, wenn $\mathbf{x}^* \in \mathbf{x}^{(k)}$ für alle k gilt. Wir geben jetzt Modifikationen der Verfahren (RGI) und (REI) an, bei welchen die Schranken monoton gegen die Lösung konvergieren. Als Verallgemeinerung des Verfahrens (RGI) betrachten wir die Iterationsvorschrift

$$\mathbf{x}^{(0)} = (X_i^{(0)}) \in V_n(I(\boldsymbol{R})):$$

$$V_i^{(k)} = [v_{i,1}^{(k)}, v_{i,2}^{(k)}] := -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij} X_j^{(k)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Löse

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)}; \quad a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}^{(k)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Setze

$$Y_i^{(k+1)} = (1-\omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}], \quad \omega > 0, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots;$$

$$X_i^{(k+1)} = Y_i^{(k+1)} \cap X_i^{(k)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Wir bezeichnen diese Iterationsvorschrift als intervallmäßige Durchführung des Relaxationsverfahrens in Gesamtschritten mit Durchschnittsbildung nach jedem Iterationsschritt (RGID).

In völliger Analogie dazu können wir das Verfahren (REI) verallgemeinern. Hier ergeben sich zwei verschiedene Möglichkeiten.

(a)
$$\mathbf{x}^{(0)} = (X_i^{(0)}) \in V_n(I(\boldsymbol{R}));$$

$$V_i^{(k)} = [v_{i,1}^{(k)}, v_{i,2}^{(k)}] := -\sum_{j=1}^{i-1} a_{ij} Y_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} X_j^{(k)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Löse

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)}; \quad a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}^{(k)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Setze

$$Y_i^{(k+1)} = (1-\omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}], \quad \omega > 0, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots;$$

$$X_i^{(k+1)} = X_i^{(k)} \cap Y_i^{(k+1)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Diese Iterationsvorschrift bezeichnen wir als intervallmäßige Durchführung des Relaxationsverfahrens in Einzelschritten mit Durchschnittsbildung nach jedem Iterationsschritt (REID).

(b)
$$\mathbf{x}^{(0)} = (X_i^{(0)}) \in V_n(I(\boldsymbol{R}));$$

$$V_i^{(k)} = [v_{i,1}^{(k)}, v_{i,2}^{(k)}] := -\sum_{j=1}^{i-1} a_{ij} X_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} X_j^{(k)}; \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Löse

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}^{(k)}; \quad a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}^{(k)}, \quad i = 1(1)\,n, \quad k = 0, 1, 2, \ldots.$$

Setze

$$Y_i^{(k+1)} = (1-\omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}]\,, \quad (\omega > 0)\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots\,;$$
$$X_i^{(k+1)} = Y_i^{(k+1)} \cap X_i^{(k)}\,, \quad i = 1(1)\,n\,, \quad k = 0, 1, 2, \ldots\,.$$

Diese Iterationsvorschrift bezeichnen wir als intervallmäßige Durchführung des Relaxationsverfahrens in Einzelschritten mit komponentenweiser Durchschnittsbildung (REIDK).

Bemerkung. Die intervallmäßige Durchführung des Relaxationsverfahrens in Gesamtschritten mit komponentenweiser Durchschnittsbildung (RGIDK) ist mit dem Verfahren (REIDK) identisch.

Satz 5. *Für* $\mathsf{A} \in M_n(\boldsymbol{R})$ *und* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *mögen die Voraussetzungen von Satz 3 bestehen. Die Lösung* x^* *von* $\mathsf{f}(\mathsf{x}) \equiv \mathsf{Ax} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$ *sei in* $\mathbf{x}^{(0)}$ *enthalten:* $\mathsf{x}^* \in \mathbf{x}^{(0)}$. *Dann sind die Verfahren* (RGID), (REID) *und* (REIDK) *definiert und für* $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$ *konvergent gegen* x^*.

Beweis. Wir führen den Beweis für das Verfahren (RGID). Wie im Beweis von Satz 4 gezeigt wurde, gilt $x_i^* \in (1-\omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}] =: Y_i^{(k+1)}$, falls $x_i^* \in X_i^{(k)}$ ist. Damit gilt dann auch $x_i^* \in Y_i^{(k+1)} \cap X_i^{(k)} = :X_i^{(k+1)}$, $i = 1(1)\,n$. Somit ist das Iterationsverfahren (RGID) definiert und für alle k gilt $\mathsf{x}^* \in \mathbf{x}^{(k)}$. Wegen der Durchschnittsbildung gilt für die Folge $\mathbf{x}^{(k)}$:

$$\mathbf{x}^{(0)} \supset \mathbf{x}^{(1)} \supset \mathbf{x}^{(2)} \supset \ldots \supset \mathbf{x}\,,$$

und damit für die Folge der Durchmesser

$$d(\mathbf{x}^{(0)}) \geqq d(\mathbf{x}^{(1)}) \geqq d(\mathbf{x}^{(2)}) \geqq \ldots \geqq \mathsf{o}\,.$$

Als monoton fallende und beschränkte Folge besitzt die Folge $d(\mathbf{x}^{(k)})$ einen Grenzwert $d(\mathbf{x})$. Für $k \to \infty$ gilt in (RGID):

$$V_i = [v_{i,1}, v_{i,2}] = -\sum_{\substack{j=1\\ j \neq i}}^{n} a_{ij} X_j\,, \quad i = 1(1)\,n$$

$$a_{ii}x_{i,1} + b_i(x_{i,1}) = v_{i,1}\,; \quad a_{ii}x_{i,2} + b_i(x_{i,2}) = v_{i,2}\,, \quad i = 1(1)\,n\,;$$
$$Y_i = (1-\omega)\,X_i + \omega[x_{i,1}, x_{i,2}]\,; \quad X_i = Y_i \cap X_i\,, \quad i = 1(1)\,n\,.$$

Daraus folgt

$$d(X_i) \leqq d(Y_i) = |1-\omega)\,d(X_i) + \omega|x_{i,2} - x_{i,1}| =$$
$$= |1-\omega|\,d(X_i) + \omega|r_i^{-1}((-\sum_{\substack{j=1\\ j \neq i}}^{n} a_{ij}X_j)_2) - r_i^{-1}((-\sum_{\substack{j=1\\ j \neq i}}^{n} a_{ij}X_j)_1)| \leqq$$

$$\leqq |1-\omega|\, d(X_i) + \frac{\omega}{a_{ii}} \Big|\Big(-\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}X_j\Big)_2 - \Big(-\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}X_j\Big)_1\Big| =$$

$$= |1-\omega|\, d(X_i) + \frac{\omega}{a_{ii}}\, d\Big(\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}X_j\Big) = |1-\omega|\, d(X_i) + \frac{\omega}{a_{ii}} \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\, d(X_j)\,,$$

oder

$$d(\mathbf{x}) \leqq \big(|1-\omega|\, \mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{B}|\big)\, d(\mathbf{x})\,.$$

Wegen $\varrho(|1-\omega|\,\mathsf{E} + \omega|\mathsf{D}^{-1}\mathsf{B}|) < 1$ für $0 < \omega < 2/(1 + \varrho(|\mathsf{D}^{-1}\mathsf{B}|))$ folgt $d(\mathbf{x}) = \mathsf{o}$, und wegen $\mathsf{x}^* \in \mathbf{x}^{(k)}$ für alle k ist $\lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathsf{x}^*$.

Die Beweise für die Verfahren (REID) und (REIDK) ergeben sich ähnlich.

5. VERGLEICH DER KONVERGENZGESCHWINDIGKEITEN

In den Abschnitten 3 und 4 haben wir mehrere Iterationsverfahren zur Einschließung der Lösung x^* von $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$ angegeben. In diesem Abschnitt sollen diese Verfahren bezüglich ihrer Konvergenzgeschwindigkeit miteinander verglichen werden. Dazu setzen wir in den folgenden Sätzen stets voraus, daß x^* im Startvektor enthalten ist. Unter dieser Voraussetzung sind alle Verfahren definiert. Als Ergebnis dieser Untersuchungen wird sich zeigen, daß das Verfahren (REIDK) für $\omega = 1$ in jedem Iterationsschritt die beste Einschließung für x^* liefert.

In einem ersten Satz vergleichen wir die Verfahren (RGI) und (RGID) bzw. (REI) und (REID) miteinander.

Satz 6. *Es seien $\mathbf{x}^{(k)} = (X_i^{(k)})$ und $\bar{\mathbf{x}}^{(k)} = (\bar{X}_i^{(k)})$ mit $\mathbf{x}^{(0)} = \bar{\mathbf{x}}^{(0)}$ die nach den Verfahren* (RGI) *und* (RGID) *berechneten Iterierten. Dann gilt für festes ω und alle $k : \bar{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$. Die gleiche Aussage gilt für die nach den Verfahren* (REI) *und* (REID) *berechneten Iterierten.*

Beweis. Wir beweisen diese Aussage für die Iterierten nach den Verfahren (RGI) und (RGID). Nach Voraussetzung gilt $\bar{\mathbf{x}}^{(0)} = \mathbf{x}^{(0)}$. Es sei nun $\bar{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$ für ein $k \geqq 0$. Dann folgt

$$\bar{V}_i^{(k)} = -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}\bar{X}_j^{(k)} \subset -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}X_j^{(k)} = V_i^{(k)}\,, \quad i = 1(1)\,n\,,$$

und damit $[\bar{x}_{i,1}, \bar{x}_{i,2}] \subset [x_{i,1}, x_{i,2}]$, $i = 1(1)\,n$. Dann folgt aufgrund der Teilmengeneigenschaft

$$\bar{Y}_i^{(k+1)} = (1-\omega)\,\bar{X}_i^{(k)} + \omega[\bar{x}_{i,1}, \bar{x}_{i,2}] \subset (1-\omega)\,X_i^{(k)} + \omega[x_{i,1}, x_{i,2}] = X_i^{(k+1)}\,,$$

also auch

$$\overline{X}_i^{(k+1)} = \overline{Y}_i^{(k+1)} \cap \overline{X}_i^{(k)} \subset X_i^{(k+1)}, \quad i = 1(1)\,n\,.$$

Das ist die Behauptung $\overline{\mathbf{x}}^{(k+1)} \subset \mathbf{x}^{(k+1)}$.

Der Beweis für die Iterierten nach den Verfahren (REI) und (REID) erfolgt ähnlich.

Im nächsten Satz werden die Verfahren (RGID) und (RGIDK) bzw. (REID) und (REIDK) miteinander verglichen.

Satz 7. *Es seien* $\mathbf{x}^{(k)} = (X_i^{(k)})$ *und* $\overline{\mathbf{x}}^{(k)} = (\overline{X}_i^{(k)})$ *mit* $\mathbf{x}^{(0)} = \overline{\mathbf{x}}^{(0)}$ *die nach den Verfahren* (RGID) *und* (RGIDK) (=(REIDK)) *berechneten Iterierten. Dann gilt für festes* ω *und alle* $k : \overline{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$. *Die gleiche Aussage gilt für die nach den Verfahren* (REID) *und* (REIDK) *berechneten Iterierten.*

Beweis. Wir beweisen diese Aussage für die Iterierten nach den Verfahren (REID) und (REIDK). Nach Voraussetzung gilt $\overline{\mathbf{x}}^{(0)} = \mathbf{x}^{(0)}$. Sei nun $\overline{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$ für ein $k \geqq 0$. Dann folgt

$$\overline{V}_1^{(k)} = -\sum_{j=2}^{n} a_{1j}\overline{X}_j^{(k)} \subset -\sum_{j=2}^{n} a_{1j}X_j^{(k)} = V_1^{(k)}$$

und damit $[\bar{x}_{1,1}, \bar{x}_{1,2}] \subset [x_{1,1}, x_{1,2}]$,

$$Y_1^{(k+1)} = (1-\omega)\,\overline{X}_1^{(k)} + \omega[\bar{x}_{1,1}, \bar{x}_{1,2}] \subset (1-\omega)\,X_1^{(k)} + \omega[x_{1,1}, x_{1,2}] = Y_1^{(k+1)},$$

also

$$\overline{X}_1^{(k+1)} = \overline{Y}_1^{(k+1)} \cap \overline{X}_1^{(k)} \subset Y_1^{(k+1)} \quad \text{und} \quad \overline{X}_1^{(k+1)} \subset Y_1^{(k+1)} \cap X_1^{(k)} = X_1^{(k+1)}.$$

Wegen $\overline{X}_1^{(k+1)} \subset Y_1^{(k+1)}$ folgt aufgrund der Teilmengeneigenschaft

$$\overline{V}_2^{(k)} = -a_{21}\overline{X}_1^{(k+1)} - \sum_{j=3}^{n} a_{2j}\overline{X}_j^{(k)} \subset -a_{21}Y_1^{(k+1)} - \sum_{j=3}^{n} a_{2j}X_j^{(k)} = V_2^{(k)}$$

und damit $[\bar{x}_{2,1}, \bar{x}_{2,2}] \subset [x_{2,1}, x_{2,2}]$,

$$\overline{Y}_2^{k+1} = (1-\omega)\,\overline{X}_2^{(k)} + \omega[\bar{x}_{2,1}, \bar{x}_{2,2}] \subset (1-\omega)\,X_2^{(k)} + \omega[x_{2,1}, x_{2,2}] = Y_2^{(k+1)},$$

$$\overline{X}_2^{(k+1)} = \overline{Y}_2^{(k+1)} \cap \overline{X}_2^{(k)} \subset Y_2^{(k+1)}, \quad \overline{X}_2^{(k+1)} \subset Y_2^{(k+1)} \cap X_2^{(k)} = X_2^{(k+1)}.$$

Durch die gleichen Überlegungen zeigt man

$$\overline{X}_i^{(k+1)} \subset X_i^{(k+1)}, \quad i = 3(1)\,n, \quad \text{also} \quad \overline{\mathbf{x}}^{(k+1)} \subset \mathbf{x}^{(k+1)}.$$

Die nächste Aussage zeigt, daß man beim Verfahren (REIDK) für $\omega = 1$ die beste Einschließungsfolge erhält.

Satz 8. *Es seien* $\mathbf{x}^{(k)} = (X_i^{(k)})$ *und* $\overline{\mathbf{x}}^{(k)} = (\overline{X}_i^{(k)})$ *mit* $\mathbf{x}^{(0)} = \overline{\mathbf{x}}^{(0)}$ *die Iterierten des Verfahrens* (REIDK) *für* ω *bzw.* $\overline{\omega}$ *mit* $0 < \omega < \overline{\omega} \leqq 1$. *Dann gilt* $\overline{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$ *für alle k. Die gleiche Aussage gilt für* $1 \leqq \overline{\omega} < \omega$.

Beweis. Wir zitieren zunächst einige Aussagen, welche beim Beweis benötigt werden. X, Y seien aus $I(\boldsymbol{R})$. Dann gilt:

(a) $0 < \omega \leqq 1,\ X \cap Y \neq \emptyset : \{(1-\omega)X + \omega Y\} \cap X = (1-\omega)X + \omega(Y \cap X);$

(b) $0 < \omega < \bar{\omega} \leqq 1,\ Y \subset X : (1-\bar{\omega})X + \bar{\omega}Y \subset (1-\omega)X + \omega Y;$

(c) $1 \leqq \bar{\omega} \leqq \omega,\ X \cap Y \neq \emptyset : (1-\bar{\omega})X + \bar{\omega}Y \subset (1-\omega)X + \omega Y;$

Die Aussagen (a) und (b) wurden von O. Mayer in [5] angegeben. Die Aussage (c) läßt sich einfach verifizieren.

Wir beweisen nun die Aussagen von Satz 8.

Es sei zunächst $0 < \omega \leqq \bar{\omega} \leqq 1$. Nach Voraussetzung gilt $\bar{\mathbf{x}}^{(0)} = \mathbf{x}^{(0)}$. Sei nun $\bar{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$ für ein $k \geqq 0$. Dann gilt

$$\bar{V}_1^{(k)} = -\sum_{j=2}^{n} a_{1j}\bar{X}_j^{(k)} \subset -\sum_{j=2}^{n} a_{1j}X_j^{(k)} = V_1^{(k)},$$

und damit $[\bar{x}_{1,1}, \bar{x}_{1,2}] \subset [x_{1,1}, x_{1,2}]$. Daraus folgt

$$\begin{aligned}
\bar{X}_1^{(k+1)} &= \{(1-\bar{\omega})\bar{X}_j^{(k)} + \bar{\omega}[\bar{x}_{1,1}, \bar{x}_{1,2}]\} \cap \bar{X}_1^{(k)} \\
&= (1-\bar{\omega})\bar{X}_1^{(k)} + \bar{\omega}\{[\bar{x}_{1,1}, \bar{x}_{1,2}] \cap \bar{X}_1^{(k)}\} && \text{(nach (a))} \\
&\subset (1-\bar{\omega})X_1^{(k)} + \bar{\omega}\{[x_{1,1}, x_{1,2}] \cap X_1^{(k)}\} && \text{(Teilmengeneigenschaft)} \\
&\subset (1-\omega)X_1^{(k)} + \omega\{[x_{1,1}, x_{1,2}] \cap X_1^{(k)}\} && \text{(nach (b))} \\
&= \{(1-\omega)X_1^{(k)} + \omega[x_{1,1}, x_{1,2}]\} \cap X_1^{(k)} && \text{(nach (a))} \\
&= X_1^{(k+1)}.
\end{aligned}$$

Damit folgt

$$\bar{V}_2^{(k)} = -a_{21}\bar{X}_1^{(k+1)} - \sum_{j=3}^{n} a_{2j}\bar{X}_j^{(k)} \subset -a_{21}X_1^{(k+1)} - \sum_{j=3}^{n} a_{2j}X_j^{(k)} = V_2^{(k)}.$$

Wie für die erste Komponente folgt dann

$$\bar{X}_2^{(k+1)} \subset X_2^{(k+1)}.$$

Entsprechend zeigt man $\bar{X}_i^{(k+1)} \subset X_i^{(k+1)}$, $i = 3(1)\,n$, also $\bar{\mathbf{x}}^{(k+1)} \subset \mathbf{x}^{(k+1)}$.

Ist $1 \leqq \bar{\omega} < \omega$, so folgt aus der Induktionsannahme $\bar{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)}$ die Beziehung

$$\bar{V}_1^{(k)} = -\sum_{j=2}^{n} a_{1j}\bar{X}_j^{(k)} \subset -\sum_{j=2}^{n} a_{1j}X_j^{(k)} = V_1^{(k)},$$

also $[\bar{x}_{1,1}, \bar{x}_{1,2}] \subset [x_{1,1}, x_{1,2}]$, und somit

$$\begin{aligned}
\bar{Y}_1^{(k+1)} &= (1-\bar{\omega})\bar{X}_1^{(k)} + \bar{\omega}[\bar{x}_{1,1}, \bar{x}_{1,2}] \\
&\subset (1-\bar{\omega})X_1^{(k)} + \bar{\omega}[x_{1,1}, x_{1,2}] && \text{(Teilmengeneigenschaft)} \\
&\subset (1-\omega)X_1^{(k)} + \omega[x_{1,1}, x_{1,2}] && \text{(nach (c))} \\
&= Y_1^{(k+1)},
\end{aligned}$$

also auch

$$\overline{X}_1^{(k+1)} = \overline{Y}_1^{(k+1)} \cap \overline{X}_1^{(k)} \subset Y_1^{(k+1)} \cap X_1^{(k)} = X_1^{(k+1)} ;$$

entsprechend zeigt man $\overline{X}_i^{(k+1)} \subset X_i^{(k+1)}$, $i = 2(1)\,n$, also $\overline{\mathbf{x}}^{(k+1)} \subset \mathbf{x}^{(k+1)}$. Damit ist der Satz bewiesen.

Wir bezeichnen das Verfahren (REIDK) für $\omega = 1$ als intervallmäßige Durchführung des Einzelschrittverfahrens mit komponentenweiser Durchschnittsbildung (EIDK). Die intervallmäßige Durchführung des Gesamtschrittverfahrens mit komponentenweiser Durchschnittsbildung (GIDK) ist mit dem Verfahren (EIDK) idetisch. Wir sind nun in der Lage das optimale Verfahren anzugeben.

Es sei $\boldsymbol{I}$ die Menge der Iterationsverfahren I mit

$$I \in \boldsymbol{I} = \{(\mathrm{RGI}), (\mathrm{REI}), (\mathrm{RGID}), (\mathrm{REID}), (\mathrm{REIDK}), (\mathrm{EIDK})\} .$$

Es bezeichne $\mathbf{x}^{(k)}$ die Iterierten nach Verfahren I. Dann wird durch

$$\bar{I} \leqq I \Leftrightarrow \overline{\mathbf{x}}^{(k)} \subset \mathbf{x}^{(k)} \quad \text{für alle } k ,$$

in $\boldsymbol{I}$ eine Halbordnung definiert. Wir zeigen, daß (EIDK) $\leqq I$ für alle $I \in \boldsymbol{I}$ gilt, woraus dann folgt, daß die Lösung $\mathbf{x}^*$ mit dem Verfahren (EIDK) in jedem Iterationsschritt am besten eingeschlossen wird.

Nach Satz 8 gilt

$$(\mathrm{EIDK}) \leqq (\mathrm{REIDK}) \quad \text{für alle} \quad \omega > 0 ,$$

und nach Satz 7 und Satz 6

$$(\mathrm{REIDK}) \leqq (\mathrm{RGID}) \leqq (\mathrm{RGI}) \quad (\omega > 0,\ \text{fest}) .$$

$$(\mathrm{REIDK}) \leqq (\mathrm{REID}) \leqq (\mathrm{REI}) \quad (\omega > 0,\ \text{fest}) .$$

Damit folgt für alle $\bar{\omega} > 0$

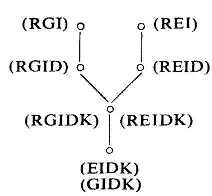
$$(\mathrm{EIDK}) \leqq (\mathrm{REIDK})|_{\omega=\bar{\omega}} \leqq (\mathrm{RGID})|_{\omega=\bar{\omega}} \leqq (\mathrm{RGI})|_{\omega=\bar{\omega}}$$

und

$$(\mathrm{EIDK}) \leqq (\mathrm{REIDK})|_{\omega=\bar{\omega}} \leqq (\mathrm{REID})|_{\omega=\bar{\omega}} \leqq (\mathrm{REI})|_{\omega=\bar{\omega}} .$$

Das ist die Behauptung.

Die bewiesenen Aussagen sind im nachfolgenden Ordnungsdiagramm veranschaulicht:



Der nächste Satz gibt eine nichtkonstruktive Aussage für die Existenz genau einer Lösung von $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$ unter der Voraussetzung, daß $\mathsf{f}(\mathsf{x})$ (Frechet-) differenzierbar ist.

Satz 9. *Es sei* $\mathsf{A} \in M_n(\boldsymbol{R})$, $\mathsf{A} = \mathsf{D} - \mathsf{B}$, *mit einer Diagonalmatrix* D *und* $\mathsf{b} : V_n(\boldsymbol{R}) \to \to V_n(\boldsymbol{R})$, $\mathsf{b} = (b_i(x_i))$, *sei eine stetig differenzierbare diagonale Abbildung mit* $|a_{ii} + b_i'(x_i)| \geqq \delta_i > 0$, $i = 1(1)\,n$, *für alle* $\mathsf{x} \in V_n(\boldsymbol{R})$. *Es sei* $\Delta = \operatorname{diag}(\delta_i)$ *und* $\varrho(|\Delta^{-1}\mathsf{B}|) < 1$. *Dann ist die Abbildung* $\mathsf{f}(\mathsf{x}) = \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x})$ *ein Homeomorphismus von* $V_n(\boldsymbol{R})$ *auf* $V_n(\boldsymbol{R})$.

Beweis. Die Abbildung f ist stetig differenzierbar:

$$\mathsf{f}'(\mathsf{x}) = \mathsf{A} + \mathsf{b}'(\mathsf{x}) = \mathsf{D} + \mathsf{b}'(\mathsf{x}) - \mathsf{B}\,.$$

($\mathsf{f}'(\mathsf{x}) \in M_n(\boldsymbol{R})$ bezeichnet die Frechetableitung). Wegen $\delta_i \neq 0$, $i = 1(1)\,n$, existiert $(\mathsf{D} + \mathsf{b}'(\mathsf{x}))^{-1}$ für alle $\mathsf{x} \in V_n(\boldsymbol{R})$ und es ist

$$|(\mathsf{D} + \mathsf{b}'(\mathsf{x}))^{-1}\,\mathsf{B}| \leqq |\Delta^{-1}\mathsf{B}| \quad \text{für alle} \quad \mathsf{x} \in V_n(\boldsymbol{R})\,.$$

Damit folgt unter Verwendung der Neumannschen Reihe

$$\begin{aligned} |\mathsf{f}'(\mathsf{x})^{-1}| &= |\{\mathsf{E} - (\mathsf{D} + \mathsf{b}'(\mathsf{x}))^{-1}\,\mathsf{B}\}^{-1}\,(\mathsf{D} + \mathsf{b}'(\mathsf{x}))^{-1}| \leqq \\ &\leqq (\mathsf{E} - |\Delta^{-1}\mathsf{B}|)^{-1}\,\Delta^{-1} = (\Delta - |\mathsf{B}|)^{-1}\,. \end{aligned}$$

Unter Verwendung einer monotonen Matrixnorm folgt

$$\|\mathsf{f}'(\mathsf{x})^{-1}\| \leqq \|(\Delta - |\mathsf{B}|)^{-1}\|\,,$$

und mit Hilfe des Normäquivalenzsatzes

$$\|\mathsf{f}'(\mathsf{x})^{-1}\| \leqq c\|(\Delta - |\mathsf{B}|)^{-1}\|$$

für eine beliebige Norm mit einer von x unabhängigen Konstanten c. Aufgrund des Hadamardschen Satzes (siehe etwa [6], Seite 137) folgt die Behauptung.

Als unmittelbare Folgerungen aus Satz 9 erhalten wir

Korollar 3. *Sei* $\mathsf{A} \in M_n(\boldsymbol{R})$ *eine M-Matrix*, $\mathsf{A} = \mathsf{D} - \mathsf{B}$, *und* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *stetig differenzierbar und diagonal und* $\mathsf{b}'(\mathsf{x})$ *nicht negativ. Dann ist die Abbildung* $\mathsf{f}(\mathsf{x}) = \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x})$ *ein Homeomorphismus von* $V_n(\boldsymbol{R})$ *auf* $V_n(\boldsymbol{R})$.

Beweis. Wegen $a_{ii} > 0$, $i = 1(1)\,n$, und $b_i'(x_i) \geqq 0$ kann man in Satz 9 $\delta_i = a_{ii}$ wählen. Mit $\varrho(\mathsf{D}^{-1}\mathsf{B}) < 1$ folgt die Behauptung.

Korollar 4. *Sei* $\mathsf{A} \in M_n(\boldsymbol{R})$ *streng diagonaldominant oder irreduzibel diagonaldominant und* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *stetig differenzierbar und diagonal und* $\operatorname{sign}(b_i'(x_i)) = \operatorname{sign}(a_{ii})$ *für alle* x. *Dann ist die Abbildung* $\mathsf{f}(\mathsf{x}) = \mathsf{Ax} + \mathsf{b}(\mathsf{x})$ *ein Homeomorphismus von* $V_n(\boldsymbol{R})$ *auf* $V_n(\boldsymbol{R})$.

Beweis. Es gilt $|a_{ii} + b_i(x_i)| \geqq |a_{ii}|$, $i = 1(1)\,n$, für alle x. Man kann daher in Satz 9 $\delta_i = |a_{ii}|$, $i = 1(1)\,n$, wählen. Wegen $\varrho(|\mathsf{D}^{-1}\mathsf{B}|) < 1$ folgt die Behauptung.

Bemerkung. Der Beweis von Korollar 3 ist in [6], Seite 141 enthalten.

7. VERALLGEMEINERTE LINEARE ITERATIONSVERFAHREN

Ist ein lineares Gleichungssystem der Form $\mathsf{Ax} = \mathsf{b}$ mit $\mathsf{A} \in M_n(\boldsymbol{R})$, $\mathsf{b} \in V_n(\boldsymbol{R})$ gegeben, $\mathsf{A} = \mathsf{D} - \mathsf{L} - \mathsf{R}$, $\mathsf{B} = \mathsf{L} + \mathsf{R}$ die Zerlegung von A in eine Diagonalmatrix D und eine strenge untere bzw. strenge obere Dreiecksmatrix L bzw. R, so bezeichnet man bekanntlich die Iterationsvorschrift

$$\mathsf{x}^{(k+1)} = (1-\omega)\,\mathsf{x}^{(k)} + \omega\{\mathsf{D}^{-1}\mathsf{Bx}^{(k)} + \mathsf{D}^{-1}\mathsf{b}\}\,, \quad k = 0, 1, 2, \ldots, \quad (\omega > 0)\,,$$

als Relaxationsverfahren in Gesamtschritten (RG), die Iterationsvorschrift

$$\mathsf{x}^{(k+1)} = (\mathsf{D} - \omega\mathsf{L})^{-1}\{(1-\omega)\,\mathsf{D} + \omega\mathsf{R}\}\,\mathsf{x}^{(k)} + \omega(\mathsf{D} - \omega\mathsf{L})^{-1}\,\mathsf{b}\,,$$
$$k = 0, 1, 2, \ldots, \quad (\omega > 0)\,,$$

als Relaxationsverfahren in Einzelschritten (RE).

Mit $\mathsf{U} = (1/\omega)\,\mathsf{D}$, $\mathsf{C} = (1/\omega)\,[\mathsf{D} + \omega\mathsf{B}]$,

$$\mathsf{H} = \mathsf{U}^{-1}\mathsf{C} = \mathsf{D}^{-1}[(1-\omega)\,\mathsf{D} + \omega\mathsf{B}]$$

läßt sich beim Verfahren (RG) $\mathsf{x}^{(k)}$ durch $\mathsf{x}^{(0)}$ folgendermaßen darstellen:

$$\mathsf{x}^{(k+1)} = \mathsf{x}^{(k)} - \omega(\mathsf{E} + \mathsf{H} + \ldots + \mathsf{H}^k)\,\mathsf{D}^{-1}(\mathsf{Ax}^{(0)} - \mathsf{b})\,.$$

Entsprechend erhält man mit

$$\mathsf{U} = 1/\omega\,(\mathsf{D} - \omega\mathsf{L})\,, \quad \mathsf{C} = 1/\omega[(1-\omega)\,\mathsf{D} + \omega\mathsf{R}]\,,$$
$$\mathsf{H} = \mathsf{U}^{-1}\mathsf{C} = (\mathsf{D} - \omega\mathsf{L})^{-1}\,\{(1-\omega)\,\mathsf{D} + \omega\mathsf{R}\}$$

beim Verfahren (RE)

$$\mathsf{x}^{(k+1)} = \mathsf{x}^{(0)} - \omega(\mathsf{E} + \mathsf{H} + \ldots + \mathsf{H}^k)\,(\mathsf{D} - \omega\mathsf{L})^{-1}\,(\mathsf{Ax}^{(0)} - \mathsf{b})\,.$$

(Siehe [6], Seite 215.)

Ist eine Abbildung $\mathsf{f} : D \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ gegeben, und ist $\mathsf{x}^{(0)}$ eine Näherung für eine Nullstelle x^* von f, so genügt die Folge $\{\mathsf{x}^{(k)}\}$ der mit dem Newton-Verfahren berechneten Näherungen den Gleichungen

$$\mathsf{f}'(\mathsf{x}^{(k)}) \cdot \mathsf{x}^{(k+1)} = \mathsf{f}'(\mathsf{x}^{(k)}) \cdot \mathsf{x}^{(k)} - \mathsf{f}(\mathsf{x}^{(k)}), \quad k = 0, 1, 2, \ldots .$$

Statt nun $\mathsf{x}^{(k+1)}$ für jedes k exakt zu bestimmen, kann man mit einem Iterationsverfahren zur Auflösung linearer Gleichungssysteme durch Ausführung einiger Iterationsschritte eine Näherung für $\mathsf{x}^{(k+1)}$ bestimmen, die Ableitung neu berechnen usw.

Mit

$$\mathsf{f}'(\mathsf{x}^{(k)}) = \mathsf{D}'_k - \mathsf{L}'_k - \mathsf{R}'_k$$

erhält man bei Verwendung des Verfahrens (RG) und bei Ausführung von m Iterationsschritten zur Auflösung des durch Anwendung des Newton-Verfahrens entstehenden linearen Gleichungssystems die folgende Iterationsvorschrift

$$\mathsf{x}^{(k+1)} = \mathsf{x}^{(k)} - \omega(\mathsf{E} + \mathsf{H}_k + \ldots + \mathsf{H}_k^{m-1})\, \mathsf{D}_k'^{-1}\mathsf{f}(\mathsf{x}^{(k)}), \quad k = 0, 1, 2, \ldots, \quad (\omega > 0).$$

Wir bezeichnen diese Iterationsvorschrift als Newton-Relaxationsverfahren in Gesamtschritten (NRG).

Entsprechend erhält man bei Verwendung des Verfahrens (RE)

$$\mathsf{x}^{(k+1)} = \mathsf{x}^{(k)} - \omega(\mathsf{E} + \mathsf{H}_k + \ldots + \mathsf{H}_k^{m-1})\,(\mathsf{D}'_k - \omega \mathsf{L}'_k)^{-1}\, \mathsf{f}(\mathsf{x}^{(k)}),$$
$$k = 0, 1, 2, \ldots, \quad (\omega > 0).$$

Dieses Iterationsverfahren bezeichnen wir als Newton-Relaxationsverfahren in Einzelschritten (NRE).

8. LOKALE KONVERGENZAUSSAGEN

Für die im letzten Abschnitt angegebenen Verfahren (NRG) und (NRE) erhebt sich nun die Frage, wann die Folgen $\mathsf{x}^{(k)}$ gegen eine Nullstelle x^* von $\mathsf{f}(\mathsf{x}) = \mathsf{o}$ konvergieren.

Wir zitieren dazu den folgenden

Satz 10. *Es sei* $\mathsf{f} : \mathsf{D} \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *differenzierbar in einer offenen Umgebung* $\mathsf{S}_0 \subset \mathsf{D}$ *eines Punktes* $\mathsf{x}^* \in \mathsf{D}$, *für den* f' *stetig und* $\mathsf{f}(\mathsf{x}^*) = \mathsf{o}$ *ist. Es sei* $\mathsf{f}'(\mathsf{x}^*) = \mathsf{D}'(\mathsf{x}^*) - \mathsf{L}'(\mathsf{x}^*) - \mathsf{R}'(\mathsf{x}^*)$ *die Zerlegung von* $\mathsf{f}'(\mathsf{x}^*)$ *in eine Diagonalmatrix* $\mathsf{D}'(\mathsf{x}^*)$, *und eine strenge untere bzw. obere Dreiecksmatrix* $\mathsf{L}'(\mathsf{x}^*)$ *bzw.* $\mathsf{R}'(\mathsf{x}^*)$. *Ist* $\mathsf{D}'(\mathsf{x}^*)$ *nichtsingulär und* $\varrho(\mathsf{D}'(\mathsf{x}^*)^{-1}\{(1-\omega)\,\mathsf{D}'(\mathsf{x}^*) + \omega(\mathsf{L}'(\mathsf{x}^*) + \mathsf{R}'(\mathsf{x}^*))\}) < 1$, *dann gibt es eine offene Umgebung* $\mathsf{S} = \mathsf{S}(\mathsf{x}^*, \delta) \subset \mathsf{S}_0$, *so daß für alle* $\mathsf{x}^{(0)} \in \mathsf{S}$ *die mit dem Verfahren* (NRG) *berechnete Folge* $\mathsf{x}^{(k)}$ *in* S *bleibt und gegen* x^* *konvergiert.*

Die gleiche Aussage gilt für das Verfahren (NRE), *falls* $\varrho((\mathsf{D}'(\mathsf{x}^*) - \omega \mathsf{L}'(\mathsf{x}^*))^{-1} \cdot \{(1-\omega)\,\mathsf{D}'(\mathsf{x}^*) + \omega \mathsf{R}'(\mathsf{x}^*)\}) < 1$ *ist.* Für einen Beweis siehe [6], Seite 321.

Als Anwendung dieser Aussagen geben wir den folgenden

Satz 11. *Sei* $\mathsf{f} : \mathsf{D} \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *differenzierbar in einer offenen Umgebung* $S_0 \subset \mathsf{D}$ *eines Punktes* $\mathsf{x}^* \in \mathsf{D}$ *für den* f' *stetig und* $\mathsf{f}(\mathsf{x}^*) = \mathsf{o}$ *ist. Ist* $\varrho(|\mathsf{D}'(\mathsf{x}^*)^{-1} \cdot \mathsf{B}'(\mathsf{x}^*)|) < 1$, *dann gibt es für alle* $0 < \omega < 2/(1 + \varrho(|\mathsf{D}'(\mathsf{x}^*)^{-1}\, \mathsf{B}'(\mathsf{x}^*)|))$ *eine offene Umgebung* $S = S(\mathsf{x}^*, \delta) \subset S_0$, *so daß für alle* $\mathsf{x}^{(0)} \in S$ *die mit dem Verfahren* (NRG) *berechneten Folgen* $\mathsf{x}^{(k)}$ *in* S *bleiben und gegen* x^* *konvergieren. Die gleiche Aussage gilt für das Verfahren* (NRE).

Beweis. Der Beweis folgt unmittelbar aus dem vorhergehenden Satz, den Sätzen 1 und 2, und der Tatsache daß für eine Matrix $\mathsf{P} \in M_n(\boldsymbol{R})$ aus $\mathsf{O} \leqq |\mathsf{P}| \leqq \mathsf{Q}$ und $\varrho(\mathsf{Q}) < 1$ die Beziehung $\varrho(\mathsf{P}) < 1$ folgt.

Bemerkung. Im Falle, daß $\mathsf{f}'(\mathsf{x}^*)$ eine M-Matrix ist, wurde diese Aussage für das Verfahren (NRE) und für $0 < \omega \leqq 1$ in [6] angegeben.

Als eine unmittelbare Folgerung erhalten wir

Satz 12. *Es sei* $\mathsf{A} \in M_n(\boldsymbol{R})$, $\mathsf{A} = \mathsf{D} - \mathsf{B}$, *und* $\mathsf{b} : V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$, $\mathsf{b} = (b_i(x_i))$, *sei eine stetig differenzierbare diagonale Abbildung mit* $|a_{ii} + b_i'(x_i)| \geqq \delta_i > 0$, $i = 1(1)\, n$, *für alle* $\mathsf{x} \in V_n(\boldsymbol{R})$. *Es sei* $\Delta = \operatorname{diag}(\delta_i)$ *und* $\varrho(|\Delta^{-1}\mathsf{B}|) < 1$. *Dann gibt es für alle* $0 < \omega < 2/(1 + \varrho(|(\mathsf{D} + \mathsf{b}'(\mathsf{x}^*))^{-1}\, \mathsf{B}|))$ *eine offene Umgebung* $S = S(\mathsf{x}^*, \delta)$, *so daß für alle* $\mathsf{x}^{(0)} \in S$ *die mit dem Verfahren* (NRG) *berechnete Folge* $\{\mathsf{x}^{(k)}\}$ *in* S *bleibt und gegen* x^* *konvergiert. Dabei ist* x^* *die eindeutige Lösung von* $\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$. *Die gleiche Aussage besteht für das Verfahren* (NRE).

Der Beweis ergibt sich aus dem vorhergehenden Satz und Satz 9 unter Berücksichtigung von

$$\varrho(|(\mathsf{D} + \mathsf{b}'(\mathsf{x}^*))^{-1}\, \mathsf{B}|) < \varrho(|\Delta^{-1}\mathsf{B}|) < 1\,.$$

Satz 12 ist insbesondere anwendbar, wenn die Voraussetzungen der Korollare 3 und 4 bestehen.

9. MONOTONE KONVERGENZ MIT DEM NEWTON-RELAXATIONSVERFAHREN IN GESAMT- UND EINZELSCHRITTEN

Im folgenden setzen wir voraus, daß die Abbildung $\mathsf{f} : \mathsf{D} \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ in D stetig differenzierbar ist, und daß der Intervallvektor $\mathbf{x}^{(0)}$ in D enthalten ist: $\mathbf{x}^{(0)} \subset \mathsf{D}$. f besitze in $\mathbf{x}^{(0)}$ genau eine Nullstelle x^*. Aufgrund des Mittelwertsatzes der Differentialrechnung gilt dann für die Funktionen

$$f_i : \mathbf{x}^{(0)} \subset V_n(\boldsymbol{R}) \to V_1(\boldsymbol{R}) = \boldsymbol{R} \quad \text{für ein} \quad \mathsf{x} \in \mathbf{x}^{(0)}$$

$$f_i(\mathsf{x}) - f_i(\mathsf{x}^*) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j} f_i(\mathsf{x} + \theta_i(\mathsf{x}^* - \mathsf{x}))\,(x_j - x_j^*)\,, \quad 0 \leqq \theta_i \leqq 1\,, \quad i = 1(1)\, n\,,$$

oder $\mathsf{f}(\mathsf{x}) - \mathsf{f}(\mathsf{x}^*) = \mathsf{f}' \cdot (\mathsf{x} - \mathsf{x}^*)$ mit

$$\mathsf{f}' := \left(\frac{\partial}{\partial x_j} f_i(\mathsf{x} + \theta_i(\mathsf{x}^* - \mathsf{x}))\right) \in M_n(\boldsymbol{R})\,.$$

(Mit f' bezeichnen wir von jetzt ab die Matrix $\mathsf{f}'(\mathsf{x}) := ((\partial/\partial x_j) f_i(\mathsf{x} + \theta_i(\mathsf{x}^* - \mathsf{x})))$ und entsprechend mit $\mathbf{f}'(\mathbf{x})$ eine Obermenge von $\{\mathsf{f}'(\mathsf{x}) \mid \mathsf{x} \in \mathsf{x}\}$ aus $M_n(I(\boldsymbol{R}))$.) Wegen $\mathsf{f}(\mathsf{x}^*) = \mathsf{o}$ gilt dann $\mathsf{f}' \cdot \mathsf{x}^* = \mathsf{f}' \cdot \mathsf{x} - \mathsf{f}(\mathsf{x})$. Mit $\mathsf{f}' = \mathsf{D}' - \mathsf{B}'$ und $\mathsf{B}' = \mathsf{L}' + \mathsf{R}'$ (D' Diagonalmatrix, L' bzw. R' strenge untere bzw. obere Dreiecksmatrix) erhalten wir

$$\mathsf{D}' \cdot \mathsf{x}^* = \mathsf{B}'\mathsf{x}^* + (\mathsf{D}' - \mathsf{B}')\,\mathsf{x} - \mathsf{f}(\mathsf{x})\,.$$

Ist D' nicht singulär, so ergibt sich

$$\mathsf{x}^* = \mathsf{D}'^{-1}\{\mathsf{B}'\mathsf{x}^* - \mathsf{B}'\mathsf{x} - \mathsf{f}(\mathsf{x})\} + \mathsf{x} = \mathsf{D}'^{-1}\{\mathsf{B}'(\mathsf{x}^* - \mathsf{x}) - \mathsf{f}(\mathsf{x})\} + \mathsf{x}\,,$$

oder

$$\mathsf{x}^* = \omega[\mathsf{x} - \mathsf{D}'^{-1}\{\mathsf{B}'(\mathsf{x} - \mathsf{x}^*) + \mathsf{f}(\mathsf{x})] + (1 - \omega)\,\mathsf{x}^*\,, \quad \omega > 0\,.$$

Wegen $\mathsf{x} + \theta_i(\mathsf{x}^* - \mathsf{x}) \in \mathbf{x}^{(0)}$ gilt $(\partial/\partial x_j) f_i(\mathsf{x} + \theta_i(\mathsf{x}^* - \mathsf{x})) \in (\partial/\partial x_j) f_i(\mathbf{x}^{(0)})$, $i, j = 1(1)\,n$, und somit speziell für $\mathsf{x} := m(\mathbf{x}^{(0)})$

$$\mathsf{x}^* \in \omega[m(\mathbf{x}^{(0)}) - \mathbf{D}'(\mathbf{x}^{(0)})^{-1}\{\mathbf{B}'(\mathbf{x}^{(0)})(m(\mathbf{x}^{(0)}) - \mathbf{x}^{(0)}) + \mathsf{f}(m(\mathbf{x}^{(0)}))\}] + \\ + (1 - \omega)\,\mathbf{x}^{(0)} =: \mathbf{y}^{(1)}\,.$$

Wegen $\mathsf{x}^* \in \mathbf{x}^{(0)}$ gilt auch $\mathsf{x}^* \in \mathbf{x}^{(1)} := \mathbf{x}^{(0)} \cap \mathbf{y}^{(1)}$. Durch wiederholte Anwendung dieser Schlußweise ergibt sich der

Beweis für den folgenden

Satz 13. *Sei* $\mathsf{f} : \mathbf{x}^{(0)} \subset V_n(\boldsymbol{R}) \to V_n(\boldsymbol{R})$ *stetig differenzierbar und* $\mathbf{f}'(\mathbf{x}^{(0)}) = \mathbf{D}'(\mathbf{x}^{(0)}) - \mathbf{L}'(\mathbf{x}^{(0)}) - \mathbf{R}'(\mathbf{x}^{(0)}) \in M_n(I(\boldsymbol{R}))$. *Besitzt* f *in* $\mathbf{x}^{(0)}$ *genau eine Nullstelle* x^*, *und ist* $\mathsf{O} \notin \mathbf{D}'(\mathbf{x}^{(0)})$, *so gilt* $\mathsf{x}^* \in \mathbf{x}^{(k)}$ *für alle* k. *Dabei ist die Folge* $\{\mathbf{x}^{(k)}\}$ *durch die Iterationsvorschrift*

$$\mathbf{y}^{(k+1)} = (1 - \omega)\,\mathbf{x}^{(k)} + \omega[m(\mathbf{x}^{(k)}) - \mathbf{D}'(\mathbf{x}^{(k)})^{-1}\{\mathbf{B}'(\mathbf{x}^{(k)})(m(\mathbf{x}^{(k)}) - \mathbf{x}^{(k)}) + \\ + \mathsf{f}(m(\mathbf{x}^{(k)}))\}]\,;$$

$$\mathbf{x}^{(k+1)} = \mathbf{y}^{(k+1)} \cap \mathbf{x}^{(k)}\,, \quad k = 0, 1, 2, \ldots\,;$$

definiert.

Wir bezeichnen die angegebene Iterationsvorschrift als intervallmäßige Durchführung des Newton-Relaxationsverfahrens in Gesamtschritten mit Durchschnittsbildung und Berechnung der Ableitung nach jedem Iterationsschritt (NRGID).

Das Iterationsverfahren

$$\mathbf{y}^{(k+1)} = (1 - \omega)\,\mathbf{x}^{(k)} + \omega[m(\mathbf{x}^{(k)}) - \mathbf{D}'(\mathbf{x}^{(k)})^{-1}\{\mathbf{L}'(\mathbf{x}^{(k)})(m(\mathbf{y}^{(k+1)}) - \\ - \mathbf{y}^{(k+1)}) + \mathbf{R}'(\mathbf{x}^{(k)})(m(\mathbf{x}^{(k)}) - \mathbf{x}^{(k)})) + \mathsf{f}(m(\mathbf{x}^{(k)}))\}]\,;$$

$$\mathbf{x}^{(k+1)} = \mathbf{y}^{(k+1)} \cap \mathbf{x}^{(k)}\,, \quad k = 0, 1, 2, \ldots\,, \quad (\omega > 0)\,,$$

bezeichnen wir als intervallmäßige Durchführung des Newton-Relaxationsverfahrens in Einzelschritten mit Durchschnittsbildung und Berechnung der Ableitung nach jedem Iterationsschritt (NREID).

Die Iterationsvorschrift

$$Y_i^{(k+1)} = \omega \left\{ m(X_i^{(k)}) - \frac{1}{a'_{ii}(\mathbf{x}^{(k)})} \left[\sum_{j=1}^{i-1} a'_{ij}(\mathbf{x}^{(k)}) \left(m(X_j^{(k+1)}) - X_j^{(k+1)} \right) + \right.\right.$$
$$\left.\left. + \sum_{j=i+1}^{n} a'_{ij}(\mathbf{x}^{(k)}) \left(m(X_j^{(k)}) - X_j^{(k)} \right) + f_i(m(\mathbf{x}^{(k)})) \right] \right\} + (1-\omega) X_i^{(k)} ;$$
$$X_i^{(k+1)} = Y_i^{(k+1)} \cap X_i^{(k+1)}, \quad i = 1(1)\,n, \quad (\omega > 0), \quad k = 0, 1, 2, \ldots,$$

bezeichnen wir als intervallmäßige Durchführung des Newton-Relaxationsverfahrens in Einzelschritten mit komponentenweiser Durchschnittsbildung und Berechnung der Ableitung nach jedem Iterationsschritt (NREIDK).

Bemerkung. Die intervallmäßige Durchführung des Relaxationsverfahrens in Gesamtschritten mit komponentenweiser Durchschnittsbildung und Berechnung der Ableitung nach jedem Iterationsschritt (NRGIDK) ist mit dem Verfahren (NREIDK) identisch.

Die Aussagen von Satz 13 bestehen sinngemäß für die Verfahren (NREID) und (NREIDK).

Es erhebt sich nun die Frage, wann in den Verfahren (NRGID), (NREID) und (NREIDK) die Folgen $\mathbf{x}^{(k)}$ gegen die Nullstelle $\mathsf{x}^* \in \mathbf{x}^{(0)}$ von $\mathsf{f}(\mathsf{x}) = \mathsf{o}$ konvergieren. Eine Antwort darauf gibt der folgende

Satz 14. *Sei* $\mathsf{f} : \mathbf{x}^{(0)} \subset V_n(R) \to V_n(R)$, $\mathsf{f} = (f_i(\mathsf{x})$. *Die* f_i *mögen stetige partielle Ableitungen besitzen und es sei*

$$\mathsf{f}'(\mathbf{x}^{(0)}) = (a'_{ij}(\mathbf{x}^{(0)})) \quad \textit{mit} \quad \mathsf{f}'(\mathbf{x}^{(0)}) = \mathbf{D}'(\mathbf{x}^{(0)}) - \mathbf{L}'(\mathbf{x}^{(0)}) - \mathbf{R}'(\mathbf{x}^{(0)}),$$
$$\mathbf{B}'(\mathbf{x}^{(0)}) = \mathbf{L}'(\mathbf{x}^{(0)}) + \mathbf{R}'(\mathbf{x}^{(0)}).$$

Ist $0 \notin a'_{ii}(\mathbf{x}^{(0)})$, $i = 1(1)\,n$, *und* $\varrho(|\mathbf{D}'(\mathbf{x}^{(0)})^{-1}| \cdot |\mathbf{B}'(\mathbf{x}^{(0)})|) < 1$, *so konvergieren in den Verfahren* (NRGID), (NREID) *und* (NREIDK) *die Folgen* $\mathbf{x}^{(k)}$ *für* $0 < \omega < 2/(1 + \varrho(|\mathbf{D}'(\mathbf{x}^{(0)})^{-1}| \cdot |\mathbf{B}'(\mathbf{x}^{(0)})|))$ *gegen* $\mathbf{x}^*$.

Beweis. Wir führen den Beweis für das Verfahren (NREID). Die Folge $\mathbf{x}^{(k)}$ konvergiert gegen $\mathbf{x}$. Wegen der Stetigkeit der Iterationsvorschrift folgt für $k \to \infty$

$$\mathbf{y} = \omega\{m(\mathbf{x}) - \mathbf{D}'(\mathbf{x})^{-1} [\mathbf{L}'(\mathbf{x})(m(\mathbf{y}) - \mathbf{y}) + \mathbf{R}'(\mathbf{x})(m(\mathbf{x}) - \mathbf{x}) + \mathsf{f}(m(\mathbf{x}))]\} +$$
$$+ (1-\omega)\,\mathbf{x} ;$$
$$\mathbf{y} = \mathbf{x} \cap \mathbf{y} .$$

Wir verwenden nun die folgenden Beziehungen:

(a) $A, B \in I(R)$, $0 < \omega \leqq 1$, $B \subset \omega A + (1-\omega) B \Rightarrow B \subset A$.

(b) $A, B \in I(R)$, $1 \leqq \omega \leqq 2$, $B \subset \omega A + (1-\omega) B \Rightarrow m(B) \in A$.

(c) $A, B \in I(R)$, $0 \notin A$, $0 \in B \Rightarrow d(AB) = |A|\, d(B)$.

Der Beweis von (a) ergibt sich folgendermaßen. Für $0 < \omega \leqq 1$ ist $B \subset \omega A + (1 - \omega) B$ äquivalent mit

$$\omega a_1 + (1 - \omega) b_1 \leqq b_1 , \quad b_2 \leqq \omega a_2 + (1 - \omega) b_2 ,$$

also $a_1 \leqq b_1 \leqq b_2 \leqq a_2$. d. h. $B \subset A$.

Wäre in (b) $m(B) \notin A$, d. h. $\frac{1}{2}(b_1 + b_2) < a_1$ oder $\frac{1}{2}(b_1 + b_2) > a_2$ so folgte wegen $B \subset \omega A + (1 - \omega) B$

$$\tfrac{1}{2}(b_1 + b_2)\,\omega + (1 - \omega) b_2 < a_1\omega + (1 - \omega) b_2 \leqq b_1$$

oder

$$b_2 \leqq \omega a_2 + (1 - \omega) b_1 < \tfrac{1}{2}(b_1 + b_2)\,\omega + (1 - \omega) b_1 ,$$

und in beiden Fällen $\omega > 2$. Der Beweis von (c) ist trivial.

Wir setzen nun den Beweis des Satzes fort.

Wegen $\mathbf{y} = \mathbf{x} \cap \mathbf{y}$ gilt $\mathbf{x} \subset \mathbf{y}$, d. h.

$$\mathbf{x} \subset (1 - \omega)\,\mathbf{x} + \omega\{m(\mathbf{x}) - \mathbf{D}'(\mathbf{x})^{-1} [\mathbf{L}'(\mathbf{x})\,(m(\mathbf{y}) - \mathbf{y}) + \mathbf{R}'(\mathbf{x}) \cdot \\ \cdot (m(\mathbf{x}) - \mathbf{x}) + \mathsf{f}(m(\mathbf{x}))]\} ,$$

und wegen (a) und (b) besteht die Beziehung

$$m(\mathbf{x}) \in m(\mathbf{x}) - \mathbf{D}'(\mathbf{x})^{-1} \{\mathbf{L}'(\mathbf{x})\,(m(\mathbf{y}) - \mathbf{y}) + \mathbf{R}'(\mathbf{x})\,(m(\mathbf{x}) - \mathbf{x}) + \mathsf{f}(m(\mathbf{x}))\} ,$$

also $\mathsf{o} \in -\mathbf{D}'(\mathbf{x})^{-1} \{\mathbf{L}'(\mathbf{x})\,(m(\mathbf{y}) - \mathbf{y}) + \mathbf{R}'(\mathbf{x})\,(m(\mathbf{x}) - \mathbf{x}) + \mathsf{f}(m(\mathbf{x}))\}$. Wegen $\mathsf{o} \notin \mathbf{D}'(\mathbf{x})^{-1}$ erhalten wir schließlich

$$\mathsf{o} \in \{\mathbf{L}'(\mathbf{x})\,(m(\mathbf{y}) - \mathbf{y}) + \mathbf{R}'(\mathbf{x})\,(m(\mathbf{x}) - \mathbf{x}) + \mathsf{f}(m(\mathbf{x}))\} .$$

Damit folgt unter Verwendung von (c)

$$d(\mathbf{x}) \leqq d(\mathbf{y}) = |1 - \omega|\, d(\mathbf{x}) + \omega|\mathbf{D}'(\mathbf{x})^{-1}|\, \{|\mathbf{L}'(\mathbf{x})|\, d(\mathbf{y}) + |\mathbf{R}'(\mathbf{x})|\, d(\mathbf{x})\} \leqq \\ \leqq |1 - \omega|\, d(\mathbf{y}) + \omega|\mathbf{D}'(\mathbf{x})^{-1}|\, \{|\mathbf{L}'(\mathbf{x})|\, d(\mathbf{y}) + \mathbf{R}'(\mathbf{x})\, d(\mathbf{y})\} ,$$

oder

$$(\mathsf{E} - \mathsf{P})\, d(\mathbf{y}) \leqq \mathsf{o} ,$$

mit

$$\mathsf{P} = (\mathsf{E} - \omega|\mathbf{D}'(\mathbf{x})^{-1}|\, |\mathbf{L}'(\mathbf{x})|)^{-1} \{|1 - \omega|\, \mathsf{E} + \omega|\mathbf{D}'(\mathbf{x})^{-1}|\, |\mathbf{R}'(\mathbf{x})|\} .$$

Wegen $\mathbf{x} \subset \mathbf{x}^{(0)}$ und Satz 1 ist

$$\varrho(\mathsf{P}) < 1 \quad \text{für} \quad 0 < \omega < 2/(1 + \varrho(|\mathbf{D}'(\mathbf{x}^{(0)})^{-1}| \,.\, |\mathbf{B}'(\mathbf{x}^{(0)})|)) ,$$

also $(\mathsf{E} - \mathsf{P})^{-1} \geqq \mathsf{O}$. Somit folgt $d(\mathbf{y}) \leqq \mathsf{o}$, also $d(\mathbf{y}) = \mathsf{o}$, und somit $d(\mathbf{x}) = \mathsf{o}$. Wegen $\mathsf{x}^* \in \mathbf{x}^{(k)}$ für alle k gilt $\lim\limits_{k \to \infty} \mathbf{x}^{(k)} = \mathsf{x}^*$.

Der Beweis für die Verfahren (NRGID) und (NREIDK) wird ähnlich geführt.

10. ÜBERLINEARE KONVERGENZ MIT DEN NEWTON-RELAXATIONSVERFAHREN

Die im letzten Abschnitt angegebenen Verfahren lassen verschiedene Modifikationen zu. So kann es in Fällen, in denen die Berechnung der Ableitungen sehr aufwendig ist, angebracht sein, nicht nach jedem Iterationsschritt diese neu zu berechnen. Führt man m_k Iterationsschritte zur Berechnung von $\mathbf{x}^{(k+1)}$ durch, so geht das Verfahren (NRGID) über in

$$\begin{aligned}
&\mathbf{x}^{(0)} \in V_n(I(R)), \quad \mathbf{x}^{(0,0)} = \mathbf{x}^{(0)} ;\\
&\mathbf{y}^{(k+1,i)} = (1-\omega)\,\mathbf{x}^{(k,i-1)} + \omega\big[m(\mathbf{x}^{(k,i-1)}) -\\
&\quad - \mathbf{D}'(\mathbf{x}^{(k)})^{-1}\{\mathbf{B}'(\mathbf{x}^{(k)})\,(m(\mathbf{x}^{(k,i-1)}) - \mathbf{x}^{(k,i-1)}) + \mathbf{f}(m(\mathbf{x}^{(k,i-1)}))\}\big] ;\\
&\mathbf{x}^{(k,i)} \quad = \mathbf{y}^{(k+1,i)} \cap \mathbf{x}^{(k,i-1)}, \quad i = 1(1)\, m_k ;\\
&\mathbf{x}^{(k+1)} \quad = \mathbf{x}^{(k,m_k)} ;\\
&\mathbf{x}^{(k+1,0)} = \mathbf{x}^{(k+1)}, \quad k = 0, 1, 2, \ldots, \quad (m_k \geqq 1) .
\end{aligned}$$

(Entsprechende Modifikationen lassen sich bei den Verfahren (NREID) und (NREIDK) vornehmen. Auf die explizite Formulierung wollen wir hier nicht eingehen.)

Das angegebene Iterationsverfahren bezeichnen wir mit (NRGID)*. Ähnlich wie in Satz 13 beweist man $x^* \in \mathbf{x}^{(k)}$ für alle k, falls $x^* \in \mathbf{x}^{(0)}$ gilt, und $\lim \mathbf{x}^{(k)} = x^*$ unter den in Satz 14 angegebenen Bedingungen.

Während in den Verfahren (NRGID), (NREID) und (NREIDK) die Folgen $d(\mathbf{x}^{(k)})$ im allgemeinen nur linear gegen $\mathbf{o}$ konvergieren, läßt sich unter bestimmten Voraussetzungen über die m_k in den Verfahren (NRGID)*, (NREID)* und (NREIDK)* überlineare Konvergenz nachweisen.

Satz 15. *Es seien die Voraussetzungen von Satz* 13 *und* 14 *erfüllt. Gilt in* (NRGID)* $m_k \to \infty$ *für* $k \to \infty$, *so gilt*

$$\|d(\mathbf{x}^{(k+1)})\| \leqq c_k \|d(\mathbf{x}^{(k)})\| \quad \textit{mit} \quad \lim_{k\to\infty} c_k = 0 .$$

Beweis. Es gilt

$$\begin{aligned}
d(\mathbf{x}^{(k,1)}) &\leqq d(\mathbf{y}^{(k+1,1)}) = |1-\omega|\, d(\mathbf{x}^{(k,0)}) +\\
&+ \omega d\big[\mathbf{D}(\mathbf{x}^{(k)})^{-1}\{\mathbf{B}'(\mathbf{x}^{(k)})\,(m(\mathbf{x}^{(k,0)}) - \mathbf{x}^{(k,0)}) + \mathbf{f}(m(\mathbf{x}^{(k,0)}))\}\big] \leqq\\
&\leqq |1-\omega|\, d(\mathbf{x}^{(k,0)}) + \omega d(\mathbf{D}'(\mathbf{x}^{(k)})^{-1}) \,.\, |\mathbf{f}(m(\mathbf{x}^{(k,0)}))| +\\
&+ \omega|\mathbf{D}'(\mathbf{x}^{(k)})^{-1}| \,.\, |\mathbf{B}'(\mathbf{x}^{(k)})|\, d(\mathbf{x}^{(k,0)}) \leqq |1-\omega|\, d(\mathbf{x}^{(k)}) +\\
&+ \omega d(\mathbf{D}'(\mathbf{x}'^{(k)})^{-1})\, |\mathbf{f}'(\mathbf{x}^{(k)})|\, d(\mathbf{x}^{(k)}) + \omega \mathbf{D}'(\mathbf{x}^{(k)})^{-1}| \,.\, |\mathbf{B}'(\mathbf{x}^{(k)})|\, d(\mathbf{x}^{(k)}) =\\
&= \{\omega d(\mathbf{D}'(\mathbf{x}^{(k)})^{-1})\, |\mathbf{f}'(\mathbf{x}^{(k)})| + |1-\omega| \mathbf{E} + \omega|\mathbf{D}'(\mathbf{x}^{(k)})^{-1}| \,.\, |\mathbf{B}'(\mathbf{x}^{(k)})|\}\, d(\mathbf{x}^{(k)}) ,
\end{aligned}$$

und durch wiederholte Anwendung dieser Schlußweise

$$d(\mathbf{x}^{(k+1)}) = d(\mathbf{x}^{(k,m_k)}) \leqq \{\omega d(\mathbf{D}'(\mathbf{x}^{(k)})^{-1})\, |\mathbf{f}'(\mathbf{x}^{(k)})| +$$
$$+ |1-\omega|\, \mathsf{E} + \omega|\mathbf{D}'(\mathbf{x}^{(k)})^{-1}| \,.\, |\mathbf{B}'(\mathbf{x}^{(k)}|\}^{m_k}\, d(\mathbf{x}^{(k)})\,.$$

Wegen $d(\mathbf{D}'(\mathbf{x}^{(k)})^{-1}) \to \mathsf{O}$ für $k \to \infty$ und $m_k \to \infty$ für $k \to \infty$ folgt die Behauptung bei Verwendung einer monotonen Norm mit

$$c_k = \|\{\omega d(\mathbf{D}'(\mathbf{x}^{(k)})^{-1})\, |\mathbf{f}'(\mathbf{x}^{(k)})| + |1-\omega|\, \mathsf{E} + \omega|\mathbf{D}'(\mathbf{x}^{(k)})^{-1}| \,.\, |\mathbf{B}'(\mathbf{x}^{(k)})|\}^{m_k}\|\,.$$

Die entsprechende Aussage gilt für die Verfahren (NREID)* und (NREIDK)*.

Bei der praktischen Anwendung dieses Satzes kann man etwa $m_0 = 1$, $m_{k+1} = m_k + 1$, $k = 0, 1, 2, \ldots$, wählen.

11. BEMERKUNGEN UND NUMERISCHE BEISPIELE

Die in den Sätzen 10, 11 und 12 angegebenen lokalen Konvergenzaussagen sind praktisch kaum anwendbar, da man einen „genügend nahe" an der Lösung liegenden Startwert benötigt. Anderseits lassen sich diese lokalen Konvergenzaussagen nicht ohne weitere Einschränkungen zu globalen Konvergenzaussagen erweitern. Dies zeigt das folgende

Beispiel. (Siehe [6], Seite 464, E 13.4–5)

$$\mathsf{f}(\mathsf{x}) \equiv \mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) \quad \text{mit} \quad \mathsf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \quad \mathsf{b}(\mathsf{x}) = \begin{pmatrix} 2x_1 + 2\sin x_1 \\ x_2 + \sin x_2 \end{pmatrix}.$$

Für f sind die Voraussetzungen von Korollar 3 erfüllt. Also besitzt $\mathsf{f}(\mathsf{x}) = \mathsf{o}$ genau eine Lösung $\mathsf{x}^*(\mathsf{x}^* = \mathsf{o})$. Wählt man im Verfahren (NRE) (mit $m = 1$, $\omega = 1$) den Startvektor

$$\mathsf{x}^{(0)} = \begin{pmatrix} \pi \\ \pi \end{pmatrix}, \quad \text{so gilt} \quad \mathsf{x}^{(2k)} = \mathsf{x}^{(0)}\,, \quad k = 1, 2, \ldots.$$

Dagegen ist es in praktischen Beispielen oft leichter möglich einen Intervallvektor zu bestimmen, in welchem die Lösung liegt, so daß man die in den Abschnitten 9 und 10 angegebenen Verfahren anwenden kann. Als ein wichtiges Beispiel dazu betrachten wir die durch Diskretisierung einer Klasse von nichtlinearen Randwertaufgaben entstehenden Gleichungssysteme.

Gegeben sei die Randwertaufgabe

$$u'' = g(t, u)\,, \quad 0 \leqq t \leqq 1\,, \quad u(0) = \alpha\,, \quad u(1) = \beta\,.$$

Ist g zweimal stetig differenzierbar auf der Menge $S = \{(t, y) \mid 0 \leqq t \leqq 1, -\infty < < y < \infty\}$ und $g_y(t, y) \geqq 0$ für alle $(t, y) \in S$, so besitzt die Randwertaufgabe genau eine Lösung (siehe etwa [6], Seite 10).

Zur Berechnung von numerischen Werten führen wir die Teilpunkte

$$t_j = j \cdot h\,, \quad h = \frac{1}{n+1}\,, \quad j = 0, 1, \ldots, n+1\,,$$

ein, und ersetzen die Ableitung $u''(t_j)$ durch den zentralen Differenzenquotienten zweiter Ordnung

$$u''(t_j) \approx \frac{1}{h^2}\left[u(t_{j+1}) - 2u(t_j) + u(t_{j-1})\right].$$

Dies führt unter Vernachlässigung des Diskretisierungsfehlers auf das nichtlineare Gleichungssystem

$$\mathsf{A}\mathsf{x} + \mathsf{b}(\mathsf{x}) = \mathsf{o}$$

mit

$$\mathsf{A} = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}, \quad \mathsf{b}(\mathsf{x}) = h^2 \begin{pmatrix} g(t_1 x_1) - \dfrac{\alpha}{h^2} \\ g(t_2, x_2) \\ \vdots \\ g(t_{n-1}, x_{n-1}) \\ g(t_n, x_n) - \dfrac{\beta}{h^2} \end{pmatrix}.$$

A ist eine M-Matrix, $\mathsf{b}(\mathsf{x})$ ist diagonal und isoton; daher existiert nach Korollar 3 genau eine Lösung x^*. x_j^* wird aufgefaßt als Näherung für $u(t_j)$, $j = 1(1)\,n$.

Für die Lösung x^* gilt

$$-\mathsf{A}^{-1}|\mathsf{b}(\mathsf{o})| \leqq \mathsf{x}^* \leqq \mathsf{A}^{-1}|\mathsf{b}(\mathsf{o})|$$

(siehe [6], Seite 460).

Daher gilt mit $z_i := (\mathsf{A}^{-1}|\mathsf{b}(\mathsf{o})|)_i$ die Beziehung $\mathsf{x}^* \in ([-z_i, z_i])$. Will man sich die zur Bestimmung der z_i nötige Gleichungsauflösung ersparen, so erhält man unter Berücksichtigung von

$$\|\mathsf{A}^{-1}\| := \max_i \sum_{k=1}^{n} |\tilde{a}_{ik}| \leqq \frac{1}{8h^2}$$

(siehe [4], Seite 171) die Abschätzung $\max_i |z_i| \leqq (1/8h^2) \max_i (|\mathsf{b}(\mathsf{o})|)_i =: c$, und damit

$$\mathsf{x}^* \in \mathbf{x}^{(0)} = (X_i) \quad \text{mit} \quad X_i = [-c, c]\,, \quad i = 1(1)\,n\,.$$

Daß die Voraussetzungen von Satz 14 für jeden Intervallvektor $\mathbf{x}^{(0)}$ mit $\mathsf{x}^* \in \mathbf{x}^{(0)}$ erfüllt sind, ist leicht einzusehen.

Beispiele

Wir wählen n Teilpunkte und lösen das entstehende Gleichungssystem

(a) mit dem Verfahren (NREIDK) (Satz 14);

(b) mit dem Verfahren (NREIDK)* (Abschnitt 10) mit

$$m_0 = 1\,, \quad m_{k+1} = m_k + 1\,, \quad k = 0, 1, 2, \ldots .$$

In (a) und (b) wählen wir $\omega = 1$. Der Startvektor $\mathbf{x}^{(0)}$, welcher die Lösung enthält, wird mit Hilfe der oben angegebenen Normabschätzung bestimmt. Es wird solange iteriert, bis auf der Maschine zwei aufeinanderfolgende Näherungen übereinstimmen.

(1) $$u'' = 2(u - \tfrac{1}{2}t + 1)^3\,; \quad u(0) = u(1) = 0\,;$$

Ergebnisse für $n = 5$:

$\mathbf{x}^{(0)} = (X_i^{(0)})$ mit $X_i^{(0)} = [-c, c]$, $c = 0{,}192563657408$;

$X_1 = [-0{,}058708825843, -0{,}058708825841]$,

$X_2 = [-0{,}082332340109, -0{,}082332340106]$,

$X_3 = [-0{,}082424385956, -0{,}082424385953]$,

$X_4 = [-0{,}065988105177, -0{,}065988105174]$,

$X_5 = [-0{,}037511064647, -0{,}037511064645]$.

Anzahl der nötigen Iterationsschritte k bis zum Stillstand der Iteration auf der Maschine ($\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$):

(a) $k = 80$; (b) $k = 14$;

Ergebnisse für $n = 10$:

$$\mathbf{x}^{(0)} = (X_i^{(0)}) \quad \text{mit} \quad X_i^{(0)} = [-c, c]\,, \quad c = 0{,}217435199099\,;$$

$X_1 = [-0{,}037708266844, -0{,}037708266841]$,

$X_2 = [-0{,}062677945842, -0{,}062677945837]$,

$X_3 = [-0{,}077624767614, -0{,}077624767608]$,

$X_4 = [-0{,}084544991894, -0{,}084544991887]$,

$X_5 = [-0{,}084938601160, -0{,}084938601153]$,

$X_6 = [-0{,}079954348094, -0{,}079954348087]$,

$X_7 = [-0{,}070486788435, -0{,}070486788428]$,

$X_8 = [-0{,}057242857192, -0{,}057242857187]$,

$X_9 = [-0{,}040788578269, -0{,}040788578265]$,

$X_{10} = [-0{,}021582491261, -0{,}021582491258]$.

Anzahl der nötigen Iterationsschritte k bis zum Stillstand der Iteration:

$$\text{(a)}\quad k = 264\,; \qquad \text{(b)}\quad k = 24\,;$$

$$u'' = e^u\,; \quad u(0) = u(1) = 0\,; \tag{2}$$

Ergebnisse für $n = 5$:

$$\mathbf{x}^{(0)} = (X_i^{(0)}) \quad \text{mit} \quad X_i^{(0)} = [-c, c]\,, \quad c = 0{,}125\,;$$

$X_1 = [-0{,}063573023781,\ -0{,}063573023778]$,
$X_2 = [-0{,}101079225592,\ -0{,}101079225589]$,
$X_3 = [-0{,}113478165706,\ -0{,}113478165703]$,
$X_4 = [-0{,}101079225592,\ -0{,}101079225589]$,
$X_5 = [-0{,}063573023781,\ -0{,}063573023778]$.

Anzahl der nötigen Iterationsschritte k bis zum Stillstand der Iteration auf der Maschine:

$$\text{(a)}\quad k = 90\,; \qquad \text{(b)}\quad k = 14\,;$$

Ergebnisse für $n = 10$:

$$\mathbf{x}^{(0)} = (X_i^{(0)}) \quad \text{mit} \quad X_i^{(0)} = [-c, c]\,, \quad c = 0{,}125\,;$$

$X_1 = [-0{,}038047082285,\ -0{,}038047082282]$,
$X_2 = [-0{,}068138233865,\ -0{,}068138233859]$,
$X_3 = [-0{,}090509291758,\ -0{,}090509291751]$,
$X_4 = [-0{,}105331045137,\ -0{,}105331045129]$,
$X_5 = [-0{,}112714562782,\ -0{,}112714562773]$,
$X_6 = [-0{,}112714562782,\ -0{,}112714562773]$,
$X_7 = [-0{,}105331045137,\ -0{,}105331045129]$,
$X_8 = [-0{,}090509291758,\ -0{,}080509291751]$,
$X_9 = [-0{,}068138233865,\ -0{,}068138233859]$,
$X_{10} = [-0{,}038047082285,\ -0{,}038047082282]$.

Anzahl k der nötigen Iterationsschritte bis zum Stillstand der Iteration

$$\text{(a)}\quad k = 299\,; \qquad \text{(b)}\quad k = 25\,;$$

Die angegebenen Beispiele wurden auf der elektronischen Rechenanlage Elektrologica $X8$ am Rechenzentrum der Universität Karlsruhe gerechnet.

Diese Arbeit wurde von der Fakultät für Mathematik an der Universität Karlsruhe als Habilitationsschrift genehmigt.

Literaturverzeichnis

[1] *Kulisch, U.:* Grundzüge der Intervallrechnung. Überblicke 2, Bibliographisches Institut Mannheim (1968).

[2] *Apostolatos, N.* und *Kulisch, U.:* Über die Konvergenz des Relaxationsverfahrens bei nichtnegativen und diagonaldominanten Matrizen. Computing Vol. 2, Fasc. 1 (1967).

[3] *Kulisch, U.:* Über positive Zerlegungen von Matrizen. Numerische Mathematik 11, 444—443 (1968).

[4] *Varga, R. S.:* Matrix Iterative Analysis. Engewood Cliffs, N. J.: Prentice-Hall, Inc. 1966.

[5] *Mayer, O.:* Über intervallmäßige Iterationsverfahren bei linearen Gleichungssystemen und allgemeinen Intervallgleichungssystemen. ZAMM, 51, 177—124 (1971).

[6] *Ortega, J. M.* und *Rheinboldt, W. C.:* Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York and London (1970).

[7] *Alefeld, G.* und *Herzberger, J.:* Über das Newton-Verfahren bei nichtlinearen Gleichungssystemen. ZAMM, Heft 12 (1970).

[8] *Alefeld, G.* und *Herzberger, J.:* Iterative Einschließung der Lösung nichtlinearer Gleichungssysteme ohne Invertierung von Intervallmatrizen. Numerische Mathematik, 19, 56—64 (1972)

Souhrn

O EXISTENCI JEDNOZNAČNÉHO ŘEŠENÍ JISTÉ TŘÍDY NELINEÁRNÍCH SYSTÉMŮ ROVNIC A O JEHO VÝPOČTU ITERAČNÍ METODOU

Götz Alefeld

Práce se zabývá různými iteračními metodami řešení soustav nelineárních rovnic spočívajících v aplikaci a vzájemné kombinaci relaxační a Newtonovy metody. Jsou zkoumány různé modifikace těchto metod odpovídající Jacobiho a Gauss-Seidelově metodě. Článek se zabývá lokální konvergencí těchto metod a vzájemným porovnáním rychlosti konvergence. Jsou též uvedeny podmínky, za kterých je konvergence monotonní. Metody jsou vyšetřovány i pro případ intervalové aritmetiky. Některé z metod v článku zkoumaných jsou vhodné při řešení jisté třídy okrajových úloh. V článku jsou uvedeny i některé numerické výsledky.

Anschrift des Verfassers: Priv. Doz. Dr. *Götz Alefeld,* Universität Karlsruhe, Institut für Angewandte Mathematik, Englerstrasse 2, 75 Karlsruhe.