

Numerische Mathematik SS 2019

Dozent: Prof. Dr. ANDREAS FISCHER

7. April 2019

Inhaltsverzeichnis

| | | |
|------------|---|-----------|
| I | Das gewöhnliche Iterationsverfahren | 3 |
| 1 | Fixpunkte | 3 |
| 2 | Der Fixpunktsatz von Banach | 4 |
| 3 | Gewöhnliche Iterationsverfahren | 5 |
| 4 | Das NEWTON-Verfahren als Fixpunktiteration | 8 |
| II | Iterative Verfahren für lineare Gleichungssysteme | 10 |
| 1 | Fixpunktiteration | 10 |
| 1.1 | Das JACOBI-Verfahren | 12 |
| 1.2 | Das GAUSS-SEIDEL-Verfahren | 12 |
| 1.3 | SOR-Verfahren | 14 |
| 2 | KRYLOV-Raum-basierte Verfahren | 16 |
| 2.1 | KRYLOV-Räume | 16 |
| 2.2 | Basisalgorithmen zur Lösung von $Ax = b$ | 16 |
| 2.3 | Das CG-Verfahren | 16 |
| 2.4 | Fehlerverhalten des CG-Verfahrens | 16 |
| 2.5 | Vorkonditionierung | 16 |
| 2.6 | Ausblick und Anmerkungen | 16 |
| III | Numerische Behandlung von Anfangswertaufgaben | 17 |
| 1 | Aufgabe und Lösbarkeit | 17 |
| 2 | Einschrittverfahren | 19 |
| 2.1 | Grundlagen | 19 |
| 2.2 | Lokaler Diskretisierungsfehler und Konsistenz | 20 |
| 2.3 | Konvergenz von Einschrittverfahren | 21 |
| 2.4 | Stabilität gegenüber Rundungsfehlern | 21 |
| 2.5 | RUNGE-KUTTA-Verfahren | 21 |
| 3 | Mehrschrittverfahren | 22 |
| 3.1 | Grundlagen | 22 |
| 3.2 | Konsistenz- und Konvergenzordnung für lineare MSV | 22 |
| 4 | A-Stabilität | 23 |
| 5 | Einblick: Steife Probleme | 24 |
| 6 | Ausblick | 25 |
| | Anhang | 27 |
| | Index | 27 |

Vorwort

Vorwort

Literatur

- Bollhöfer/Mehrmann: Numerische Mathematik, Vieweg 2004
- Deuffhard/Hohmann: Numerische Mathematik1, de Gruyter 2008
- Deuffhard/Bornemann: Numerische Mathematik, de Gruyter 2008
- Deuffhard/Weiser: Numerische Mathematik 3, de Gruyter 2011
- Freund/Hoppe: Stoer/Bulirsch: Numerische Mathematik 1, Springer 2007
- Hämmerlin/Hoffmann: Numerische Mathematik, Springer 2013
- Knorrenschild, M: Numerische Mathematik, Fachbuchverlag 2005
- Plato, R: Numerische Mathematik kompakt, Vieweg 2009
- Preuß/Wenisch: Lehr- und Übungsbuch Numerische Mathematik, Fachbuchverlag 2001
- Quarteroni/Sacco/Saleri: Numerische Mathematik 1+2, Springer 2002
- Roos/Schwetlick: Numerische Mathematik, Teubner 1999
- Schaback/Wendland: Numerische Mathematik, Springer 2004
- Stoer/Bulirsch: Numerische Mathematik II, Springer 2005

Kapitel I

Das gewöhnliche Iterationsverfahren

1. Fixpunkte

Seien ein Vektorraum V , eine Menge $U \subseteq V$ und eine Abbildung $\Phi : U \rightarrow V$ gegeben. Dann heißt $x^* \in U$ Fixpunkt der Abbildung Φ , falls $\Phi(x^*) = x^*$ gilt. Die Aufgabe

$$\Phi(x) = x$$

(eigentlich die Aufgabe, diese Gleichung zu lösen) wird als Fixpunktaufgabe bezeichnet. Die Abbildung Φ heißt Fixpunktabbildung. Im Unterschied zur Fixpunktaufgabe heißt

$$F(x) = 0$$

Nullstellenaufgabe. Zu jeder Nullstellenaufgabe gibt es eine äquivalente Fixpunktaufgabe (z.B. $F(x) = 0 \Leftrightarrow \Phi(x) = x$ mit $\Phi(x) := F(x) + x$) und umgekehrt (z.B. $\Phi(x) = x \Leftrightarrow F(x) = 0$ mit $F(x) := \Phi(x) - x$).

2. Der Fixpunktsatz von Banach

Der folgende Satz gibt (unter gewissen Bedingungen) eine konstruktive Möglichkeit an, einen Fixpunkt näherungsweise zu ermitteln.

Satz 2.1 (Banach)

Seien $(V, \|\cdot\|)$ ein Banach-Raum, $U \subseteq V$ eine abgeschlossene Menge und $\Phi : U \rightarrow V$ eine Abbildung. Die Abbildung Φ sei selbstabbildend, d.h. es gilt

$$\Phi(U) \subseteq U.$$

Außerdem sei Φ kontraktiv, d.h. es gibt $\lambda \in [0, 1)$, so dass

$$\|\Phi(x) - \Phi(y)\| \leq \lambda \|x - y\|, \text{ für alle } x, y \in U.$$

Dann besitzt Φ genau einen Fixpunkt $x^* \in U$. Weiterhin konvergiert die durch

$$x^{k+1} := \Phi(x^k) \tag{1}$$

erzeugte Folge $\{x^k\}$ für jeden Startwert $x^0 \in U$ gegen x^* und es gilt für alle $k \in \mathbb{N}$

$$\|x^{k+1} - x^*\| \leq \frac{\lambda}{1 - \lambda} \|x^{k+1} - x^k\| \quad \text{a posteriori Fehlerabschätzung,} \tag{2}$$

$$\|x^{k+1} - x^*\| \leq \frac{\lambda^{k+1}}{1 - \lambda} \|x^1 - x^0\| \quad \text{a priori Fehlerabschätzung,} \tag{3}$$

$$\|x^{k+1} - x^*\| \leq \frac{\lambda}{1 - \lambda} \|x^k - x^*\| \quad \text{Q-lineare Konvergenz mit Ordnung } \lambda. \tag{4}$$

Beweis. Verlesung zur Analysis. □

Die in Satz 2.1 vorkommende Zahl $\lambda \in [0, 1)$ wird Kontraktionskonstante genannt.

3. Gewöhnliche Iterationsverfahren

Durch Gleichung (1) erklärte Verfahren heißt gewöhnliches Iterationsverfahren oder Fixpunktiteration. Kritisch ist dabei, ob die Voraussetzungen (Φ ist selbstabbildend und kontraktiv) erfüllt werden können. Dies wird in diesem Abschnitt im Fall $V = \mathbb{R}^n$ mit einer beliebigen aber festen Vektornorm $\|\cdot\|$ untersucht. Die zugeordnete Matrixnorm wurde mit $\|\cdot\|_*$ bezeichnet.

Lemma 3.1

Sei $S \subseteq \mathbb{R}^n$ offen und konvex und $\Phi : D \rightarrow \mathbb{R}^n$ stetig differenzierbar. Falls $L > 0$ existiert mit

$$\|\Phi'(x)\|_* \leq L \text{ für alle } x \in D, \quad (1)$$

dann ist Φ Lipschitz-stetig in D mit der Lipschitz-Konstante L , d.h. es gilt

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \text{ für alle } x \in D. \quad (2)$$

Die Umkehrung dieser Aussage ist ebenfalls richtig.

Beweis. 1. Sei Gleichung (1) erfüllt. Mit Satz 5.1 aus der Vorlesung ENM folgt

$$\|\Phi(x) - \Phi(y)\|_* = \left\| \int_0^1 \Phi'(y + t(x - y))(x - y) dt \right\| \leq \|x - y\| \sup_{t \in [0,1]} \|\Phi'(y + t(x - y))\|_* \quad (3)$$

für alle $x, y \in D$. Also liefert Gleichung (1) unter Beachtung der Konvexität von D die Behauptung.

2. Sei nun Gleichung (2) erfüllt. Angenommen es gibt $\hat{y} \in D$ mit

$$\|\Phi'(\hat{y})\|_* > L. \quad (4)$$

Unter Berücksichtigung der Definition der zugeordneten Matrixnorm $\|\cdot\|_*$ folgt, dass $d \in \mathbb{R}^n$ existiert mit $\|d\| = 1$ und $\|\Phi'(\hat{y})d\| = \|\Phi'(\hat{y})\|_*$. Wendet man nun ENM mit $x := \hat{y} + sd$ und $y := \hat{y}$ an, so folgt für alle $s > 0$ hinreichend klein

$$\|\Phi(\hat{y} + sd) - \Phi(\hat{y})\| \leq L\|sd\| = sL \quad (5)$$

und

$$\begin{aligned} \|\Phi(\hat{y} + sd) - \Phi(\hat{y})\| &= \left\| \int_0^1 \Phi'(\hat{y} + tsd)(sd) dt \right\| \\ &= \left\| \int_0^1 \Phi'(\hat{y} + tsd)(sd) dt + \int_0^1 \Phi'(\hat{y})(sd)(sd) dt - \int_0^1 \Phi'(\hat{y})(sd)(sd) dt \right\| \\ &\geq s \left\| \Phi'(\hat{y})d \right\| - s\|d\| \sup_{t \in [0,1]} \|\Phi'(\hat{y} + tsd) - \Phi'(\hat{y})\|_* \\ &= s(\|\Phi'(\hat{y})\|_* - \sup_{t \in [0,1]} \|\Phi'(\hat{y} + tsd) - \Phi'(\hat{y})\|_*) \\ &> sL, \end{aligned}$$

wobei sich die letzte Ungleichung wegen Gleichung (4) und der Stetigkeit von Φ' ergibt. Offenbar hat man damit einen Widerspruch, so dass die Annahme falsch ist. \square

■ **Beispiel 3.2**

Die Nullstellenaufgabe $\cos x - 2x = 0$ sei zu lösen. Eine mögliche Formulierung als Fixpunktaufgabe ist

$$\Phi(x) = x \text{ mit } \Phi(x) := -x + \cos x$$

Offenbar ist $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ selbstabbildend. Weiter ergibt sich

$$\Phi'(x) = -1 - \sin x$$

Für $x \in D := (0, 1)$ gilt daher $|\Phi'(x)| > 1$. Mit Lemma 3.1 folgt $|\Phi(x) - \Phi(y)| \geq |x - y|$ für mindestens ein Paar $(x, y) \in D \times D$. Somit ist Φ in D nicht kontrahierend. Definiert man Φ aber durch $\Phi(x) := \frac{1}{2} \cos x$, so ist die Fixpunktaufgabe $\frac{1}{2} \cos x = x$ wiederum zur Nullstellenaufgabe äquivalent und es folgt

$$\Phi'(x) = -\frac{1}{2} \sin x.$$

Damit hat man $|\Phi'(x)| \leq \frac{1}{2}$ für alle $x \in \mathbb{R}$. Also ist die zuletzt definierte Abbildung Φ kontrahierend auf \mathbb{R} (und dort natürlich selbstabbildend), so dass die Voraussetzungen des Banachschen Fixpunktsatzes erfüllt sind. Die Fixpunktiteration mit $\Phi(x) = \frac{1}{2} \cos x$ und $x^0 := 1$ ergibt:

$$x^1 = 0.270 \dots$$

$$x^2 = 0.481 \dots$$

$$x^3 = 0.433 \dots$$

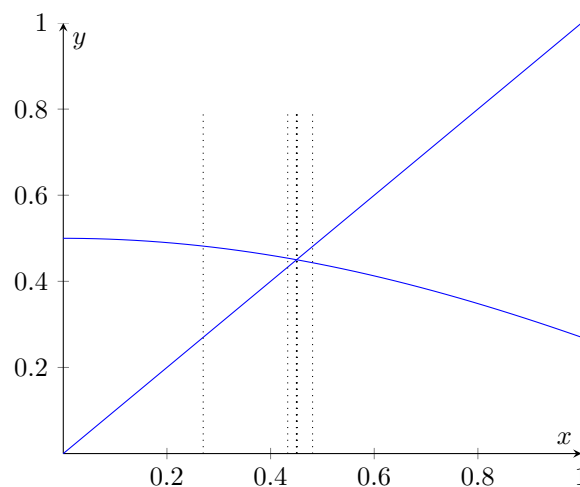
$$x^4 = 0.4517 \dots$$

$$x^5 = 0.4498 \dots$$

$$x^6 = 0.45025 \dots$$

$$x^7 = 0.450167 \dots$$

$$x^8 = 0.450187 \dots$$



Nehmen wir an, die Voraussetzungen des Banachschen Fixpunktsatzes seien gegeben. Dann hängt die Konvergenzgeschwindigkeit der Fixpunktiteration offenbar von der Kontraktionskonstanten $\lambda \in [0, 1)$ ab. Je kleiner λ ist, desto schneller ist die Konvergenzgeschwindigkeit. Unter Umständen kann die Umformulierung einer Fixpunktaufgabe mit Hilfe einer anderen Fixpunktabbildung helfen, die Konvergenzgeschwindigkeit zu verbessern (ggf. auf Kosten der Größe der Menge U , in der die Voraussetzungen des Banachschen Fixpunktsatzes erfüllt sind.) Ein Beispiel zur Konstruktion einer Fixpunktabbildung mit lokal beliebig kleiner Kontraktionskonstante gibt Abschnitt 1.4. In Abschnitt 2.1 wird gezeigt, wie Fixpunktabbildungen zur iterativen Lösung von linearen Gleichungssystemen eingesetzt werden können. Im Weiteren bezeichne $B(x^*, r) :=$ die abgeschlossene Kugel um x^* mit Radius r (bzgl. einer passenden Norm).

Satz 3.3 (Ostrowski)

Seien $D \subseteq \mathbb{R}^n$ offen und $\Phi : D \rightarrow \mathbb{R}^n$ stetig differenzierbar. Die Abbildung Φ besitze einen Fixpunkt $x^* \in D$ mit $\|\Phi'(x^*)\|_* < 1$. Dann existiert $r > 0$, so dass das gewöhnliche Iterationsverfahren für jeden Startpunkt $x^0 \in B(x^*, r)$ gegen x^* konvergiert.

Beweis. Da Φ stetig differenzierbar ist und $\|\Phi'(x^*)\|_* < 1$, gibt es $\lambda \in [0, 1]$ und $r > 0$, sodass

$$\|\Phi'(x)\|_* \leq \lambda \quad \text{für alle } x \in B(x^*, r).$$

Nach Lemma 3.1 gilt daher

$$\|\Phi(x) - \Phi(y)\| \leq \lambda \|x - y\| \quad \text{für alle } x, y \in B(x^*, r). \quad (6)$$

Insbesondere folgt hieraus

$$\|\Phi(x) - \Phi(x^*)\| = \|\Phi(x) - x^*\| \leq \lambda \|x - x^*\| \quad \text{für alle } x \in B(x^*, r) \quad (7)$$

und damit $\Phi(x) \in B(x^*, r)$ für alle $x \in B(x^*, r)$. Also ist Φ bzgl. $B(x^*, r)$ selbstabbildend und kontraktiv. Daher liefert Satz 2.1 die gewünschte Aussage. \square

4. Das Newton-Verfahren als Fixpunktiteration

Sei $D \subseteq \mathbb{R}^n$ offen und $F : D \rightarrow \mathbb{R}^n$ stetig differenzierbar. Die Nullstellenaufgabe

$$F(x) = 0$$

wird nun in eine äquivalente Fixpunktaufgabe überführt. Dazu nehmen wir an, dass x^* eine reguläre Nullstelle von F ist. Wegen der vorausgesetzten Stetigkeit von F' gibt es $r > 0$ hinreichend klein, so dass $F'(x)$ für $x \in B(x^*, r)$ regulär ist. Damit erhält man

$$F(x) = 0 \Leftrightarrow 0 = -F'(x)^{-1}F(x) \Leftrightarrow x = x - F'(x)^{-1}F(x).$$

für $x \in B(x^*, r)$. Definiert man $\Phi : B(x^*, r) \rightarrow \mathbb{R}^n$ durch

$$\Phi(x) := x - F'(x)^{-1}F(x). \quad (1)$$

so kann das Newton-Verfahren als Fixpunktverfahren mit Φ als Fixpunktabbildung interpretiert werden. Ob Φ selbstabbildend und kontrahierend ist, müsste noch untersucht werden. Hier soll nur die Kontraktionseigenschaft in $B(x^*, r)$ für $r > 1$ hinreichend klein betrachtet werden. Die Eigenschaft der Selbstabbildung ergibt sich dann wie im Beweis zu Satz 3.3.

Lemma 4.1

Sei $D \subseteq \mathbb{R}^n$ offen und $F : D \rightarrow \mathbb{R}^n$ stetig differenzierbar. Weiter sei $x^* \in D$ eine reguläre Nullstelle von F . Dann ist Φ in x^* differenzierbar mit $\Phi'(x^*) = 0$.

Beweis. Wie zuvor gezeigt wurde, ist die durch Gleichung (1) definierte Abbildung Φ in $B(x^*, r) \subset D$ hinreichend kleines $r > 0$ wohldefiniert. Falls

$$\lim_{x \rightarrow x^*} \frac{\|\Phi(x) - \Phi(x^*) - G(x - x^*)\|}{\|x - x^*\|} \quad (2)$$

mit $G = 0 \in \mathbb{R}^{n \times n}$ gilt, folgt die Behauptung des Lemmas aus der Definition der Fréchet-Differenzierbarkeit. Unter Beachtung von $\Phi(x^*) = x^*$ ergibt sich

$$\Phi(x) - \Phi(x^*) = x - F'(x)^{-1}F(x) - x^* = -F'(x)^{-1}(F'(x)(x^* - x) + F(x))$$

und mit Satz 5.1 aus der Vorlesung ENM folgt weiter

$$\Phi(x) - \Phi(x^*) = F'(x)^{-1} \left(-F(x^*) + \int_0^1 (F'(x + t(x^* - x)) - F'(x))(x^* - x) dt \right) \quad (3)$$

für alle $x \in B(x^*, r)$. Die Stetigkeit von F' auf der kompakten Menge $B(x^*, r)$ impliziert, dass F' dort auch gleichmäßig stetig ist. Also gibt es zu jedem $\varepsilon > 0$ ein $\delta(\varepsilon) > 0$, so dass auch

$$\|x + t(x^* - x) - x\| \leq \delta(\varepsilon) \quad \text{die Beziehung} \quad \|F'(x + t(x^* - x)) - F'(x)\|_* \leq \varepsilon$$

für beliebige $x \in B(x^*, r)$ und $t \in [0, 1]$ folgt. Damit hat man

$$\lim_{x \rightarrow x^*} \max_{t \in [0, 1]} \|F'(x + t(x^* - x)) - F'(x)\|_* = 0$$

und

$$\lim_{x \rightarrow x^*} \frac{\left\| \int_0^1 (F'(x + t(x^* - x)) - F'(x))(x^* - x) dt \right\|_*}{\|x - x^*\|} = 0$$

Somit erhält man aus Gleichung (3) unter Beachtung von $F(x^*) = 0$ und der Regularität von $F'(x)$

$$\lim_{x \rightarrow x^*} \frac{\|\Phi(x) - \Phi(x^*)\|}{\|x - x^*\| O(x - x^*)} = 0,$$

d.h. Gleichung (2) ist für $G = 0$ erfüllt. □

► **Bemerkung 4.2**

Falls F in einer Umgebung von x^* sogar zweimal stetig differenzierbar und damit Φ dort stetig differenzierbar ist, zeigt Lemma 3.1, dass $\|\Phi'(x)\|_* \leq L$ für alle $x \in D \cap B(x^*, r(L))$ gilt. D.h. die Kontraktionskonstante der Fixpunktabbildung Φ in Gleichung (1) in einer Kugel $B(x^*, r)$ konvergiert gegen 0, wenn man den Radius r gegen 0 gehen lässt. Ferner gibt es Sätze, bei denen unter geeigneten Voraussetzungen eine bestimmte lokale Konvergenzgeschwindigkeit (Q-Ordnung) gezeigt wird (etwa die Q Ordnung 2, wenn insbesondere Φ' stetig ist und $\Phi'(x^*) = 0$ gilt).

Kapitel II

Iterative Verfahren für lineare Gleichungssysteme

Seien eine reguläre Matrix $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$ gegeben. In diesem Kapitel werden iterative Verfahren zur Lösung des linearen Gleichungssystems

$$Ax = b \tag{1}$$

betrachtet.

1. Fixpunktiteration

Grundidee dieser Verfahren ist die geeignete Umformulierung des System $Ax = b$ als Fixpunktaufgabe und die Anwendung des gewöhnlichen Iterationsverfahrens. Die hier betrachtete (zu Gleichung (1) äquivalente) Fixpunktaufgabe lautet

$$x = x - B^{-1}(Ax - b),$$

wobei $B \in \mathbb{R}^{n \times n}$ eine noch zu wählende reguläre Matrix ist. Bei Wahl eines Startpunktes $x^0 \in \mathbb{R}^n$ ergibt sich das gewöhnliche Iterationsverfahren damit zu

$$x^{k+1} := x^k - B^{-1}(Ax^k - b) = (I - B^{-1}A)x^k + B^{-1}b, \quad k = 0, 1, 2, \dots \tag{1}$$

Mit den Bezeichnung $M := I - B^{-1}A$ und $c := B^{-1}b$ untersuchen wir deshalb die Iterationsvorschrift

$$x^{k+1} := Mx^k + c. \tag{2}$$

Die zugehörige Fixpunktabbildung $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist damit offenbar gegeben durch

$$\Phi(x) := Mx + c.$$

Satz 1.1

Es sei $B \in \mathbb{R}^{n \times n}$ regulär und mit $M := I - B^{-1}A$ gelte

$$\lambda := \|M\|_* < 1 \quad (3)$$

wobei $\|\cdot\|_*$ die einer Vektornorm $\|\cdot\|$ zugeordnete Matrixnorm bezeichnet. Dann gilt:

1. Die für eine beliebiges $x^0 \in \mathbb{R}^n$ durch Gleichung (2) erzeugte Folge $\{x^k\}$ konvergiert gegen die eindeutige Lösung x^* des linearen Gleichungssystems Gleichung (1).
2. Die Abschätzungen Gleichung (2) - Gleichung (4) sind für alle $k \in \mathbb{N}$ erfüllt.

Beweis. Direkte Folgerung aus dem Banachschen Fixpunktsatz (Satz 1.2.1) □

► Bemerkung 1.2

In Satz 1.1a) kann die Folgerung Gleichung (3) durch die Bedingung

$$\rho(M) < 1 \quad (4)$$

ersetzt werden. Da

$$\rho(C) \leq \|C\|_* \quad \text{für alle } C \in \mathbb{R}^{n \times n}$$

für jede beliebige zugeordnete Matrixnorm $\|\cdot\|_*$ gilt (vgl. Übungsaufgabe), ist Gleichung (4) eine schwächere Forderung als Gleichung (3). Andererseits gibt es zu jedem Paar $(C, \varepsilon) \in \mathbb{R}^{n \times n} \times (0, \infty)$ eine zugeordnete Matrixnorm $\|\cdot\|_{(C, \varepsilon)}$, so dass

$$\|C\|_{(C, \varepsilon)} \leq \rho(C) + \varepsilon.$$

Dabei ist $\rho(C)$ der Spektralradius der Matrix $C \in \mathbb{R}^{n \times n}$, d.h.

$$\rho(C) := \max_{i=1, \dots, n} |\lambda_i|,$$

wobei $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die Eigenwerte der Matrix $C \in \mathbb{R}^{n \times n}$ bezeichnen. Man kann weiter zeigen, dass Gleichung (4) auch notwendig dafür ist, dass die durch Gleichung (1) erzeugte Folge $\{x^k\}$ für jedes x^0 gegen x^* konvergiert.

Um eine Matrix B zu finden, so dass einerseits der Aufwand pro Iteration Gleichung (1) niedrig und andererseits die Bedingung Gleichung (3) bzw. Gleichung (4) erfüllt ist, betrachten wir die folgende Zerlegung

$$A = L + D + R$$

der Matrix A , wobei $D := \text{diag}(a_{11}, \dots, a_{nn})$ die aus den Diagonalelementen von A bestehende Diagonalmatrix bezeichnet und L bzw. R eine untere bzw. obere Dreiecksmatrix ist mit

$$L = \begin{pmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & & \ddots & \ddots & \\ a_{n1} & \cdots & \cdots & a_{n,n-1} & 0 \end{pmatrix} \quad \text{bzw.} \quad R = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & a_{n-1,n} \\ & & & a_{n,n-1} & 0 \end{pmatrix}.$$

1.1. Das Jacobi-Verfahren

Wir setzen hier voraus, dass D regulär ist und wählen

$$B := D \tag{5}$$

Damit ergibt sich die Iterationsvorschrift

$$x^{k+1} = x^k - D^{-1}(Ax^k - b) = -D^{-1}(L + R)x^k + D^{-1}b. \tag{6}$$

In Gleichung (2) ist entsprechend

$$M := M_J := -D^{-1}(L + R) \text{ und } c := c_J := D^{-1}b$$

zu wählen. Dieses Verfahren heißt Gesamtschrittverfahren oder Jacobi-Verfahren. Der Aufwand pro Schritt (Berechnung von x^{k+1} aus x^k) beträgt $O(n^2)$ bei voll besetzter Matrix A und mindestens $O(n)$, falls A schwach besetzt ist.

Satz 1.3

Die Matrix A sei streng diagonaldominant (vgl. Definition 3.1 der Vorlesung ENM). Dann ist die Matrix B aus Gleichung (5) regulär und es gilt

$$\|M_J\|_\infty \leq \lambda_{SD} := \max_{i=1,\dots,n} \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < 1.$$

Beweis. Die Regularität von B ergibt sich sofort aus der strengen Diagonaldominanz von A . Nutzt man die Definition der Zeilensummennorm $\|\cdot\|_\infty$ erhält man sofort

$$\|M_J\|_\infty = \|D^{-1}(L + R)\|_\infty = \max_{i=1,\dots,n} \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| = \lambda_{SD}.$$

Die vorausgesetzte strenge Diagonaldominanz von A sichert $\lambda_{SD} < 1$. □

1.2. Das Gauss-Seidel-Verfahren

Wir setzen hier voraus, dass $L + D$ regulär ist und wählen

$$B := L + D \quad (7)$$

Damit ergibt sich die Iterationsvorschrift

$$x^{k+1} = x^k - (L + D)^{-1}(Ax^k - b) = -(L + D)^{-1}Rx^k + (L + D)^{-1}b. \quad (8)$$

In Gleichung (2) ist entsprechend

$$M := MGS := -(L + D)^{-1}R \text{ und } c := c_{GS} := (L + D)^{-1}b$$

zu wählen. Dieses Verfahren heißt Einzelschrittverfahren oder Gauß-Seidel-Verfahren. Der Aufwand pro Schritt beträgt im ungünstigsten Fall $O(n^2)$. Verbesserungen sind möglich, wenn eine Sparse-Struktur in A ausgenutzt werden kann.

Satz 1.4

Die Matrix A sei streng diagonaldominant (\nearrow Definition 3.1 der Vorlesung ENM). Dann ist die Matrix B aus Gleichung (7) regulär und es gilt

$$\|M_{GS}\|_{\infty} \leq \lambda_{SD} < 1.$$

Beweis. Die Regularität von B folgt sofort aus der strengen Diagonaldominanz von A . Weiter ergibt sich

$$\|M_{GS}\|_{\infty} = \|(L + D)^{-1}R\|_{\infty} = \sup_{\|y\|_{\infty}=1} \|(L + D)^{-1}Ry\|_{\infty}.$$

Um für einen festen Vektor y mit $\|y\|_{\infty} = 1$ eine Abschätzung für die rechte Seite zu erhalten, setzen wir $z := (L + D)^{-1}Ry$. Damit gilt

$$(D + L)z = Ry \quad (9)$$

und

$$z_1 = \frac{1}{a_{11}} \sum_{j=1}^n a_{1j}y_j.$$

Daraus folgt (da $\lambda_{SD} < 1$ wegen der strengen Diagonaldominanz von A)

$$|z_1| \leq \frac{1}{|a_{11}|} \sum_{j=2}^n |a_{1j}| |y_j| \leq \sum_{j=2}^n |a_{1j}| \leq \lambda_{SD} < 1.$$

Nehmen wir nun an, dass

$$|z_1| \leq \text{für } i = 1, \dots, k-1,$$

für ein $k \in \{2, \dots, n\}$ gilt. Dann folgt wegen Gleichung (9) und $\|y\|_\infty = 1$

$$|z_k| = \frac{1}{|a_{kk}|} \left| - \sum_{i=1}^{k-1} a_{ki} z_i + \sum_{i=k+1}^n a_{ki} y_i \right| \leq \frac{1}{|a_{kk}|} \left(\sum_{i=1}^{k-1} |a_{ki}| + \sum_{i=k+1}^n |a_{ki}| \right) \leq \lambda_{SD}.$$

Somit hat man induktiv $|z_k| \leq \lambda_{SD}$ für $k = 1, \dots, n$ und damit

$$\|(L + D)^{-1} R y\|_\infty = \|z\|_\infty \leq \lambda_{SD}$$

für beliebige y mit $\|y\|_\infty = 1$. □

1.3. SOR-Verfahren

Um dieses verfahren zu beschreiben, nehmen wir an, dass für ein $\omega \neq 0$ die Matrix

$$B := L + \frac{1}{\omega} D \tag{10}$$

regulär ist. Damit ergibt sich die Iterationsvorschrift

$$x^{k+1} := x^k - \left(L + \frac{1}{\omega} D \right)^{-1} (Ax^k - b) = M(\omega)x^k + c(\omega)$$

$$M(\omega) := I - \left(L + \frac{1}{\omega} D^{-1} A \right) = \left(L + \frac{1}{\omega} D \right)^{-1} \tag{11}$$

und

$$c(\omega) := \left(L + \frac{1}{\omega} D \right)^{-1} b. \tag{12}$$

Für $\omega = 1$ erhält man offenbar als Spezialfall das Gauß-Seidel-Verfahren, so dass der folgende Satz auch dafür Anwendung finden kann. Man beachte dazu Bemerkung 1.2.

Satz 1.5

Die Matrix A sei symmetrisch und positiv definit. Dann ist die Matrix B aus Gleichung (10) regulär (für jedes $\omega \neq 0$). Falls $\omega \in (0, 2)$, dann gilt

$$\rho(M(\omega)) < 1$$

und umgekehrt.

Beweis. Da A positiv definit ist, gilt $e_i^T A e_i = a_{ii} > 0$ für $i = 1, \dots, n$. Also ist D positiv definit und damit B regulär für alle $\omega \neq 0$.

Sei $\lambda \in \mathbb{C}$ ein Eigenwert von $M(\omega)$ und $z \in \mathbb{C}^n$ ein zugehöriger Eigenvektor. Mit

$$A = A - M(\omega)^T AM(\omega) + M(\omega)^T AM(\omega)$$

sowie (unter Berücksichtigung der Definition von M und von $A = A^T$ und $R = L^T$)

$$\begin{aligned} A - M(\omega)^T AM(\omega) &= A - (I - B^{-1}A)^T A (I - B^{-1}A) \\ &= AB^T A + AB^{-1}A - AB^{T-1}AB^{-1}A \\ &= (B^{-1}A)^T (B + B^T - A)(B^{-1}A) \\ &= (B^{-1}A)^T \left(L + \frac{1}{\omega}D + L^T + \frac{1}{\omega}D - L - D - L^T \right) (B^{-1}A) \\ &= (B^{-1}A)^T \left(\frac{2-\omega}{\omega}D \right) (B^{-1}A) \end{aligned}$$

ergibt sich daher

$$z^H Az = (AB^{T-1}z)^H \left(\frac{2-\omega}{\omega}D \right) (B^{-1}Az) + z^H M(\omega)^T AM(\omega)z.$$

Da die Diagonalmatrix D positiv-definit ist, besitzt $\frac{2-\omega}{\omega}D$ dieselbe Eigenschaft für $\omega \in (0, 2)$. Es folgt

$$(AB^{T-1}z)^H \left(\frac{2-\omega}{\omega}D \right) (B^{-1}Az) > 0$$

und damit

$$|\lambda| < 1. \tag{13}$$

Also gilt $\rho(M(\omega)) = \max_{i=1,\dots,n} |\lambda_i| < 1$, sofern $\omega \in (0, 2)$. Die Umkehrung der Aussage ergibt sich aus dem Satz von KAHAN (↗ Übungsaufgabe). \square

Es ist nun naheliegend, dass man $\omega \in (0, 2)$ so wählen möchte, dass $\rho(\omega)$ möglichst klein ist. Dies ist in bestimmten Fällen näherungsweise möglich, ansonsten beschränkt man sich auf geeignete Heuristiken zur Wahl von ω . Auf der Fixpunktiteration Gleichung (2) beruhende Verfahren werden häufig auch Splitting-Methoden genannt. Es gibt noch weitere solche Verfahren, auf die hier nicht eingegangen wird.

2. Krylov-Raum-basierte Verfahren

2.1. Krylov-Räume

2.2. Basisalgorithmen zur Lösung von $Ax = b$

2.3. Das CG-Verfahren

2.4. Fehlerverhalten des CG-Verfahrens

2.5. Vorkonditionierung

2.6. Ausblick und Anmerkungen

Kapitel III

Numerische Behandlung von Anfangswertaufgaben

1. Aufgabe und Lösbarkeit

Es seien $a, b \in \mathbb{R}$ mit $a < b$, eine stetige Funktion $f: [a, b] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ und $y^0 \in \mathbb{R}^m$ gegeben. Unter Anfangswertaufgabe (AWA) 1. Ordnung versteht man das Problem, eine stetige Funktion $y: [a, b] \rightarrow \mathbb{R}^m$ zu ermitteln, so dass y auf (a, b) stetig differenzierbar ist und

$$y'(x) = f(x, y(x)) \quad \text{mit} \quad y(a) = y^0$$

für alle $x \in [a, b]$ gilt. Eine solche Funktion wollen wir Lösung der AWA nennen. Kürzer schreibt man für die AWA auch

$$y' = f(x, y) \quad \text{mit} \quad y(a) = y^0 \tag{1}$$

Die Existenz und Eindeutigkeit einer Lösung einer AWA hängen von den Eingangsinformationen a, b, f und y^0 ab. Es gilt folgender Satz zur (globalen) Existenz und Eindeutigkeit einer Lösung auf $[a, b]$:

Satz 1.1 (Picard-Lindelöf: globale Version)

Es sein $f: [a, b] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig und es existiere $L > 0$, so dass

$$\|f(x, y) - f(x, z)\| \leq L\|y - z\| \quad \forall (x, y), (x, z) \in [a, b] \times \mathbb{R}^m \tag{2}$$

Dann besitzt Gleichung (1) für jedes $y^0 \in \mathbb{R}^m$ eine eindeutige Lösung.

Die Bedingung Gleichung (2) ist eine globale Lipschitz-Bedingung an f bezüglich der zweiten Veränderlichen. Es ist leicht, AWA anzugeben, in denen diese Bedingung nicht erfüllt ist und keine Lösung in ganz $[a, b]$ existiert, zum Beispiel

$$y' = y^2 \quad \text{mit} \quad y(0) = 1$$

Dafür erhält man für beliebige $x, y, z \in \mathbb{R}$

$$|f(x, y) - f(x, z)| = |y^2 - z^2| = |y + z||y - z|$$

das heißt die Bedingung Gleichung (2) kann in diesem Beispiel (global) nicht gelten. Die Lösung der AWA lautet $y(x) = \frac{-1}{x-1}$ für $x \in [0, 1)$. Für Intervalle $[0, b]$ mit $b \geq 1$ existiert keine Lösung. Eine Abschwächung der Lipschitz-Bedingung Gleichung (2) gestattet folgender

Satz 1.2 (Picard-Lindelöf: lokale Version)

Es sei $f: [a, b] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig und zu jeder kompakten Menge $Y \subset \mathbb{R}^m$ existiere $L_Y > 0$, so dass

$$\|f(x, y) - f(x, z)\| \leq L_y \|y - z\| \quad \forall (x, y), (x, z) \in [a, b] \times Y$$

Dann gibt es für jedes $y^0 \in \mathbb{R}^m$ ein Teilintervall $I \subseteq [a, b]$ mit $a \in I$, so dass die AWA Gleichung (1) auf I eine eindeutige Lösung besitzt.

Seien $g: [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}$ stetig und $\eta \in \mathbb{R}^n$. Jede explizite Differentialgleichung n -ter Ordnung

$$y^{(n)} = g(x, y, y', y'', \dots, y^{(n-1)})$$

mit den Anfangsbedingungen

$$y(a) = \eta_1, \quad y'(a) = \eta_2, \quad y''(a) = \eta_3, \quad \dots \quad y^{(n-1)}(a) = \eta_n$$

kann mittels Substitution

$$y_1 = y, \quad y_2 = y', \quad y_3 = y'', \quad \dots \quad y_n = y^{(n-1)}$$

in eine AWA 1. Ordnung überführt werden:

$$\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} y_2 \\ \vdots \\ y_n \\ g(x, y_1, \dots, y_n) \end{pmatrix} \quad \text{mit} \quad \begin{pmatrix} y_1(a) \\ \vdots \\ y_n(a) \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

2. Einschrittverfahren

2.1. Grundlagen

Anstelle der gesuchten Lösungsfunktion $y: [a, b] \rightarrow \mathbb{R}^m$ einer AWA ist man an möglichst guten Näherungen $y^k \in \mathbb{R}^m$ ($k = 0, 1, 2, \dots, N$) für die Funktionswerte $y(x_k) \in \mathbb{R}^m$ der Funktion y an Gitterpunkten $x_k \in [a, b]$ interessiert. Auf Grundlage der Paare (x_k, y^k) ($k = 0, 1, \dots, N - 1$) ist es auch möglich, eine Näherungsfunktion y zu erzeugen (etwa durch Interpolation).

Einschrittverfahren bilden eine Klasse von Verfahren, die Näherungen y^k zu erzeugen. Das Gitter $\{x_0, \dots, x_N\}$ ist so gewählt, dass

$$x_0 = a < x_1 < x_2 < \dots < x_{N-1} < x_N = b$$

Außerdem setzen wir

$$h_k = x_{k+1} - x_k \quad \text{für } k = 0, \dots, N - 1$$

und bezeichnen h_k als Schrittweite. Falls $h = h_0 = \dots = h_{N-1}$, so heißen die Gitterpunkte bzw. das Gitter gleichabständig oder äquidistant.

Ein Verfahren zur Erzeugung einer Folge y^0, \dots, y^N heißt Einschrittverfahren für das AWA Gleichung (1), wenn

$$y^{k+1} = y_k + h_k \Phi(x_k, y_k, y^{k+1}, h_k) \quad \text{für } k = 0, \dots, N - 1 \quad (1)$$

Dabei bezeichnet $\Phi(x, y, z, h)$ den Funktionswert einer Verfahrensfunktion

$$\Phi : [a, b] \times \mathbb{R}^m \times \mathbb{R}^m \times (0, b - a] \rightarrow \mathbb{R}^m$$

die das jeweilige Einschrittverfahren definiert. Man beachte, dass y^0 bereits durch die Anfangsbedingung in Gleichung (1) gegeben ist. Ein Einschrittverfahren heißt implizit, falls Φ tatsächlich von z abhängt. Dann ist zur Bestimmung von y^{k+1} aus Gleichung (1) die Lösung eines im Allgemeinen nichtlinearen Gleichungssystems erforderlich. Falls Φ nicht von z abhängt, heißt das Einschrittverfahren explizit. Das explizite EULER-Verfahren (auch Polygonzugverfahren genannt) ist gegeben durch

$$\Phi(x, y, z, h) = f(x, y) \quad (2)$$

das heißt

$$y^{k+1} = y^k + h_k f(x_k, y^k)$$

Für das implizite EULER-Verfahren gilt die Vorschrift

$$y^{k+1} = y^k + h_k f(x_k + h_k, y^{k+1})$$

Um die Güte der Näherungen y^k zu beurteilen, untersuchen wir zunächst den lokalen Diskretisierungsfehler eines Einschrittverfahrens.

2.2. Lokaler Diskretisierungsfehler und Konsistenz

Definition 2.1 (lokaler Diskretisierungsfehler)

Seien $y: [a, b] \rightarrow \mathbb{R}^m$ Lösung der Differentialgleichung $y' = f(x, y)$ und Φ die Verfahrensfunktion eines Einschrittverfahrens. Für $x \in [a, b)$ und $h > 0$ mit $x + h \leq b$ heißt

$$\Delta(x, h) = y(x + h) - \left(y(x) + h\Phi(x, y(x), y(x + h), h) \right) \quad (3)$$

lokaler Diskretisierungsfehler und

$$\frac{\Delta(x, h)}{h} = \frac{y(x + h) - y(x)}{h} - \Phi(x, y(x), y(x + h), h) \quad (4)$$

relativer lokaler Diskretisierungsfehler des Einschrittverfahrens.

Der lokale Diskretisierungsfehler gibt also die Abweichung zwischen exakter Lösung $y(x + h)$ an der Stelle $x + h$ und der Näherung an dieser Stelle an, wobei angenommen wird, dass die Näherung unter Verwendung der exakten Lösung $y(x)$ (und ggf. $y(x + h)$) berechnet wird. Die Bezeichnung relativer Diskretisierungsfehler ist bezüglich der Schrittweite h zu verstehen.

Definition 2.2 (konsistent, Konsistenzordnung)

Ein Einschrittverfahren heißt konsistent zur Differentialgleichung $y' = f(x, y)$, wenn

$$\lim_{h \downarrow 0} \left\| \frac{\Delta(x, h)}{h} \right\| = 0 \quad \forall x \in [a, b)$$

für jede Lösung $y: [a, b] \rightarrow \mathbb{R}^m$ der Differentialgleichung gilt. Gibt es außerdem $p \geq 1$, $M > 0$, $\tilde{h} > 0$, so dass

$$\left\| \frac{\Delta(x, h)}{h} \right\| \leq Mh^p \quad \forall (x, h) \in [a, b) \times (0, \tilde{h}) \text{ mit } x + h \leq b$$

für jede Lösung $y: [a, b] \rightarrow \mathbb{R}^m$ der Differentialgleichung gilt, so hat das Einschrittverfahren (für diese Differentialgleichung) die Konsistenzordnung p .

Satz 2.3

Sei $f: [a, b] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig differenzierbar. Dann hat das explizite EULER-Verfahren die Konsistenzordnung 1.

Beweis. Mit Gleichung (2) folgt

$$\Delta(x, h) = y(x + h) - y(x) - hf(x, y(x))$$

Da y die Differentialgleichung $y' = f(x, y)$ löst und f stetig differenzierbar ist, muss y zweimal stetig differenzierbar sein. Aus der TAYLOR-Formel erhält man für $i \in \{1, \dots, m\}$

$$\begin{aligned} \Delta(x, h)_i &= y'_i(x)h + \frac{1}{2}y''_i(\xi_i(x, h))h^2 - hf_i(x, y(x)) \\ &= \frac{1}{2}y''_i(\xi_i(x, h))h^2 \end{aligned}$$

für ein $\xi_i(x, h) \in (x, x + h)$. Die Stetigkeit von y'' auf $[a, b]$ und Division durch h liefert die Behauptung mit $M = \frac{1}{2} \max_{1 \leq i \leq m} \max_{\xi \in [a, b]} \|y''_i(\xi)\|$ und $\tilde{h} = b - a$. \square

2.3. Konvergenz von Einschrittverfahren

2.4. Stabilität gegenüber Rundungsfehlern

2.5. Runge-Kutta-Verfahren

3. Mehrschrittverfahren

3.1. Grundlagen

3.2. Konsistenz- und Konvergenzordnung für lineare MSV

4. A-Stabilität

5. Einblick: Steife Probleme

6. Ausblick

Anhang

Index

- EULER-Verfahren, [18](#)
- Anfangswertaufgabe, [16](#)
 - Lösung, [16](#)
- Einschrittverfahren, [18](#)
 - explizit, [18](#)
 - implizit, [18](#)
- Einzelschrittverfahren, [13](#)
- Fixpunkt, [3](#)
- Fixpunktabbildung, [3](#)
- Fixpunktaufgabe, [3](#)
- Fixpunktiteration, [5](#)
- Gauß-Seidel-Verfahren, [13](#)
- Gesamtschrittverfahren, [12](#)
- gewöhnliches Iterationsverfahren, [5](#)
- Gitter, [18](#)
 - äquidistant, [18](#)
 - gleichabständig, [18](#)
- Gitterpunkten, [18](#)
- Jacobi-Verfahren, [12](#)
- konsistent, [19](#)
- Konsistenzordnung, [19](#)
- Kontraktionskonstante, [4](#)
- kontraktiv, [4](#)
- lokaler Diskretisierungsfehler, [19](#)
- Nullstellenaufgabe, [3](#)
- Polygonzugverfahren, [18](#)
- Schrittweite, [18](#)
- selbstabbildend, [4](#)
- Spektralradius, [11](#)
- Verfahrensfunktion, [18](#)