16. VL Einführung in die Num. Mathematik, 8.6.09

Die folgende Tabelle zeigt Intervalle, Gewichtsfunktionen, die zugehörigen Orthogonalpolynome und deren Name

Intervalle	$\varrho(x)$	$p_0, p_1, \dots$	Bezeichnung
$[-1,1]$	1	$1, x, x^2-\frac{1}{3}, \dots$	Legendre
$[-1,1]$	$\frac{1}{\sqrt{1-x^2}}$	$1, x, x^2-\frac{1}{2}$	Tschebyscheff
$[-1,1]$	$(1-x)^\alpha(1+x)^\beta$	$1, \frac{1}{2}(4+\alpha+\beta)x+(\alpha-\beta), \dots$	Jacobi ($\alpha, \beta > -1$)
$]-\infty,\infty[$	e^{-x^2}	$1, x, x^2-\frac{1}{2}, x^3-\frac{3}{2}x, \dots$	Hermite
$[0,\infty[$	$e^{-x}x^\alpha$	$1, x-\alpha-1, \dots$	Laguerre ($\alpha > -1$)

Mit den in der Tabelle angegebenen Polynomen und deren Nullstellen kann man Quadraturformeln für endliche Intervalle und unendliche Intervalle konstruieren.
Die Tschebyscheffpolynome sind trotz der Gewichtsfunktion gegenüber den Legendrepolynomen attraktiv, weil man die Nullstellen des n-ten Tschebyscheffschen Orthogonalpolynoms explizit angeben kann (durch eine Berechnungsformel) ohne die Polynome auszurechnen. Das ist bei den anderen Polynomen aus der Tabelle nicht möglich.

5. Iterative Lösung von Gleichungssystemen

Im Folgenden geht es darum, Gleichungen oder Gleichungssysteme zu lösen. Ist $\vec{G}$ eine Abbildung aus dem $\mathbb{R}^n$ in den $\mathbb{R}^n$, bedeutet

$$\vec{G}(\vec{x}) = \vec{0} \qquad (5.1)$$

gerade ein Gleichungssystem zur Bestimmung einer Nullstelle $\vec{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ der Abb. $\vec{G}$. Definiert man ausgehend von $\vec{G}$ die Abb.

$$\vec{F}(\vec{x}) := \vec{G}(\vec{x}) + \vec{x},$$

dann ist die Lösung von (5.1) gleichbedeutend mit der Bestimmung eines Fixpunktes $\vec{x}$ von $\vec{F}$, also

$$\vec{F}(\vec{x}) = \vec{x} \ \left(\Longleftrightarrow \vec{G}(\vec{x}) = \vec{0}\right). \qquad (5.2)$$

Wenn man keinerlei Vorstellung von der Lösung der Gleichung (5.2) hat, findet man mit der Folge

$$(\vec{x}_k), \quad \vec{x}_0 \in D \subset \mathbb{R}^n, \quad \vec{x}_{k+1} = \vec{F}(\vec{x}_k), \ k = 0,1,2,\dots$$

eine Folge, die, wenn sie konvergiert, im Falle der Stetigkeit der Abb. gegen einen Fixpunkt von $\vec{F}$ konvergiert.

Die Grundlagen der iterativen Lösung von Gleichungen/Gleichungssystemen sollen nur für den Fall $n=1$ dargestellt werden.

Definition 5.1

Sei $f: I \to I$ eine Funktion, die das reelle Intervall I in sich abbildet. Jede Lösung der Gleichung

$$x = f(x) \qquad (5.3)$$

heißt *Fixpunkt* von f. Die Gleichung (5.3) wird *Fixpunktgleichung* genannt.

Definition 5.2

Eine auf einer Teilmenge $D \subset \mathbb{R}$ definierte Funktion $f: D \to \mathbb{R}$ heißt *Kontraktion*, wenn eine Konstante $L \in [0,1[$ existiert, so dass f.a. $x_1, x_2 \in D$

$$|f(x_1) - f(x_2)| \le L|x_1 - x_2|$$

gilt (f ist *kontraktiv*)

Satz 5.3 (Banachscher Fixpunktsatz)

Sei $f: I \to I$ eine reellwertige Funktion, die ein abgeschlossenes Intervall I in sich abbildet, und es gelte f.a. $x_1, x_2 \in I$ die Ungleichung

$$|f(x_1) - f(x_2)| \le L|x_1 - x_2| \qquad (5.4)$$

mit einer von x_1, x_2 unabh. Konstanten $L < 1$, d.h. f ist Kontraktion.
Dann hat f in I genau einen Fixpunkt $\hat{x} \in I$ und die durch die Fixpunktiteration $x_{k+1} = f(x_k)$ definierte Iterationsfolge konvergiert für jeden Anfangspunkt $x_0 \in I$ gegen diesen Fixpunkt.

Beweis [illegible]

Bemerkung 5.4 (Banachscher Fixpunktsatz im $\mathbb{R}^n$)

Ist $\vec{F}: A \to A$, $A \subset \mathbb{R}^n$, abgeschlossen, und gilt

$$\|\vec{F}(\vec{x}_1) - \vec{F}(\vec{x}_2)\| \leq L \|\vec{x}_1 - \vec{x}_2\|$$

mit $L < 1$ f.a. $\vec{x}_1, \vec{x}_2 \in A$, dann hat $\vec{F}$ genau einen Fixpunkt $\hat{\vec{x}}$ in A mit

$$\vec{F}(\hat{\vec{x}}) = \hat{\vec{x}}$$

und die durch $\vec{x}_{k+1} = \vec{F}(\vec{x}_k)$ definierte Iterationsfolge konvergiert für jeden Anfangspunkt $\vec{x}_0 \in A$ gegen diesen Fixpunkt. ($\mathbb{R}^n$ ist mit der Metrik $\rho(\vec{x}_1, \vec{x}_2) = \|\vec{x}_1 - \vec{x}_2\|$ ein Banach-Raum).

Bemerkung 5.5

Aus dem Banachschen Fixpunktsatz ergeben sich die Fehlerabschätzungen

$$\|\vec{x}_k - \hat{\vec{x}}\| \leq \frac{L^k}{1-L} \|\vec{x}_1 - \vec{x}_0\| \quad \text{(Apriori-Abschätzung)} \tag{5.5}$$

$$\|\vec{x}_k - \hat{\vec{x}}\| \leq \frac{1}{1-L} \|\vec{x}_{k+1} - \vec{x}_k\| \quad \text{(A-posteriori-Abschätzung)}. \tag{5.6}$$

Die Kontraktivität ist essentiell für die Berechnung von Fixpunkten. Unter bestimmten Voraussetzungen kann man die Existenz einer Kontraktion zeigen.

Satz 5.6

Es sei G eine offene Teilmenge von $\mathbb{R}$ und es sei $f: G \to \mathbb{R}$ stetig diff'bar mit einem Fixpunkt $\hat{x} \in G$.
Wenn $|f'(\hat{x})| < 1$ gilt, dann existiert ein abgeschlossenes Intervall $D \subset G$ mit $\hat{x} \in D$ und $f(D) \subset D$, auf dem f eine Kontraktion ist.

Beweis

Da f' stetig auf der offenen Menge G ist, existiert eine offene Umgebung $K_{\hat{x},\varepsilon} = \{x \mid |x-\hat{x}| < \varepsilon\}$ in G, auf der die Beträge der Ableitung von f immer noch kleiner als 1 sind. Setzt man $D = [\hat{x} - \frac{\varepsilon}{2}, \hat{x} + \frac{\varepsilon}{2}]$, so gilt f.a. $x_1, x_2 \in D$ aufgrund des Mittelwertsatzes der Diff.-Rechnung

$$|f(x_1) - f(x_2| \le k\,|x_1 - x_2|$$

mit $k = \max_{\xi \in D} |f'(\xi)| < 1$. $\square$

Bemerkung 5.7

Ist die Voraussetzung $|f'(\hat{x})| < 1$ des Satzes 5.6 nicht erfüllt, findet man keine Kontraktion. Ist $|f'(\hat{x})| > 1$ dann gilt in der Nähe von $\hat{x}$

$$|f(x) - f(\hat{x})| > |x - \hat{x}|,$$

das rechtfertigt die

Definition 5.8

Ein Fixpunkt $\hat{x}$ heißt anziehender Fixpunkt, wenn $|f'(\hat{x})| < 1$ gilt, und $\hat{x}$ heißt abstoßender Fixpunkt, wenn $|f'(\hat{x})| > 1$ gilt.

Das Newton-Verfahren zur Lösung nichtlinearer Gleichungen $f(x) = 0$

Die Grundidee der Lösung der nichtlinearen Gleichung besteht in der Näherung einer existierenden Nullstelle durch die sukzessive Lösung linearer Aufgaben. Man approximiert die diff'bare Fkt. f in der Nähe eines geeigneten Startwertes x_0 durch die Tangentenfunktion

$$g(x) = f(x_0) + f'(x_0)(x - x_0) \approx f(x)$$

und bestimmt die Nullstelle von g, also x_1 mit $g(x_1) = 0$, d. h.

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

und allgemein erhält man mit

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \ldots \tag{5.7}$$

eine Newton-Folge, die im Falle der Konvergenz gegen eine Nullstelle von f geht.

Die Folge (5.7) kann man auch anders erhalten. Man definiert die Hilfsfunktion

$$g(x) = x - \frac{f(x)}{f'(x)}, \qquad (5.8)$$

wobei man $f'(x) \neq 0$ voraussetzt. Ist g eine Kontraktion mit $g(I) \subset I$, dann folgt aus dem Banachschen Fixpunktsatz, dass die Folge $x_{k+1} = g(x_k)$ für einen beliebigen Startwert $x_0 \in I$ (abgeschlossenes Intervall) gegen den in I existierenden Fixpunkt $\hat{x}$ von g konvergiert. Und die Fixpunkt-Folge

$$x_{k+1} = g(x_k) = x_k - \frac{f(x_k)}{f'(x_k)}$$

ist eine Newton-Folge zur Berechnung einer Nullstelle von f. Es gilt der folgende

Satz 5.8

Sei $f: I \to \mathbb{R}$ eine auf einem Intervall $I \supset [x_0 - r, x_0 + r]$, $r > 0$, definierte, zweimal stetig diff'bare Funktion mit $f'(x) \neq 0$ f.a. $x \in I$. Weiterhin existiere eine reelle Zahl K, $0 < K < 1$, mit

$$\left| \frac{f(x) f''(x)}{f'(x)^2} \right| \leq K \quad \text{f.a. } x \in I$$

und

$$\left|\frac{f(x_0)}{f'(x_0)}\right| \leq (1-K)r.$$

Dann hat f genau eine Nullstelle $\hat{x}$ in I und die Newton-Folge (5.7) konvergiert quadratisch gegen $\hat{x}$, d.h. es gilt

$$|x_{k+1}-\hat{x}| \leq C(x_k-\hat{x})^2 \quad \text{f.a. } k=0,1,2,\dots$$

mit einer Konstanten C. Außerdem gilt die Fehlerabschätzung

$$|x_k-\hat{x}| \leq \frac{|f(x_k)|}{M} \quad \text{mit } 0<M<\min_{x\in I}|f'(x)|.$$

Beweis

Folgt aus dem Banachschen Fixpunktsatz 5.3 und dem Satz 5.6 über die Existenz einer Kontraktion.

Die quadratische Konvergenz folgt aus:

$$x_{k+1}-\hat{x} = x_k - \frac{f(x_k)}{f'(x_k)} - \hat{x}$$

$$= x_k-\hat{x} - \frac{f(x_k)-f(\overset{\text{Null}}{\hat{x}})}{f'(x_k)}$$

$$= \frac{1}{f'(x_k)}\Big[\underbrace{f'(x_k)(x_k-\hat{x}) - f(x_k)+f(\hat{x})}_{\text{Fehler der Ordnung } O(|x_k-\hat{x}|^2)}\Big]$$

$$\Longrightarrow$$

$$|x_{k+1}-\hat{x}| \leq C(x_k-\hat{x})^2.$$

$\square$

Bemerkung 5.9

Die Voraussetzungen des Satzes 5.8 garantieren die Kontraktivität der Hilfsfunktion $g(x) = x - \frac{f(x)}{f'(x)}$ in eine Umgebung von $\hat{x}$.
D.h. man ist mit dem Newton-Verfahren (Newton-Folge) immer dann erfolgreich, wenn man nur nah genug an der Nullstelle die Iteration beginnt (x_0 nah bei $\hat{x}$).
In diesem Fall ist das Newton-Verfahren auch noch sehr schnell aufgrund der quadratischen Konvergenz.

Newton-Verfahren für nichtlineare Gleichungssysteme $\vec{F}(\vec{x}) = \vec{0}$

Satz 5.9.
$\vec{F}: D \to \mathbb{R}^n$, $D \subset \mathbb{R}^n$ sei zweimal stetig partiell diffbar und besitze eine Nullstelle $\hat{\vec{x}} \in D$.
Weiterhin sei $\vec{F}'(\vec{x})$ für jedes $\vec{x} \in D$ regulär.
Dann folgt:
Es gibt eine Umgebung U von $\hat{\vec{x}}$, so dass die Newton-Folge

$$\vec{x}_{k+1} = \vec{x}_k - [\vec{F}'(\vec{x}_k)]^{-1} \vec{F}(\vec{x}_k), \quad k = 0, 1, 2, \dots \qquad (5.9)$$

von einem beliebigen Startpunkt $\vec{x}_0 \in U$ ausgehend gegen die Nullstelle $\hat{\vec{x}}$ konvergiert.
Die Konvergenz ist quadratisch, d.h. es gibt eine Konstante $C > 0$ mit

$$\|\vec{x}_k - \hat{\vec{x}}\| \le C \|\vec{x}_{k-1} - \hat{\vec{x}}\|^2 \quad , k = 1,2,3,\dots$$

und es gilt die Fehlerabschätzung

$$\|\vec{x}_k - \hat{\vec{x}}\| \le \|\vec{f}(\vec{x}_k)\| \sup_{\vec{x} \in \mathring{D}} \|[\vec{F}'(\vec{x})]^{-1}\| \; ,$$

wobei auf der rechten Seite die Matrixnorm

$$\|A\| = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2} \quad \text{für eine } (n\times n)\text{-Matrix}$$

$A = (a_{ij})$ verwendet wurde.

Der Beweis dieses Satzes verläuft analog zum Nachweis von Satz 5.8 für den eindimensionalen Fall.

Beispiel

Zur Bestimmung der Stützstellen (Nullstellen des n-ten Tschebyscheff-Polynoms) $x_1,\dots,x_n$ und der Integrationsgewichte $G_1,\dots,G_n$ ist das Gleichungssystem

$$\sum_{k=1}^{n} G_k x_k^j = \int_{-1}^{1} x^j \, dx \quad , j = 0,1,\dots,2n-1, \qquad (5.10)$$

zu lösen. Mit $b_j = \int_{-1}^{1} x^j \, dx = \frac{1}{j+1}(1-(-1)^{j+1})$ bedeutet das die Nullstelle der Abb. $\vec{F} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$

$$\vec{F}(G_1,\dots,G_n,x_1,\dots,x_n) = \begin{pmatrix} G_1 x_1^0 + G_2 x_2^0 + \dots + G_n x_n^0 - b_0 \\ \vdots \\ G_1 x_1^{2n-1} + G_2 x_2^{2n-1} + \dots + G_n x_n^{2n-1} - b_{2n-1} \end{pmatrix}$$

zu bestimmen. Mit der Ableitungsmatrix $\vec{F}' = (J_{ik})$

$$J_{ik} = \begin{cases} x_k^{i-1} & , 1 \le k \le n, \; i = 1,\dots,2n \\ 0 & , i = 1, \; n+1 \le k \le 2n \\ G_{k-n}(i-1) x_{k-n}^{i-2} & , i = 2,\dots,2n, \; n+1 \le k \le 2n \end{cases}$$

kann man mit dem Newton-Verfahren (5.9) verwenden, das Gleichungssystem (5.10) lösen.