

Numerik I

Kurzskriptum nach einer Vorlesung von
Professor Dr. H. M. Möller

Universität Dortmund – Wintersemester 1998/99

Letzte Änderung: 1. Dezember 2001

Dieses Kurzschrift ist aus meiner persönlichen Mitschrift der Vorlesung

Numerik I

bei *Professor Dr. H. M. Möller* im Wintersemester 1997/98 entstanden. Ich habe versucht alles richtig wiederzugeben, kann allerdings keine Garantie darauf geben. Es ist deshalb wahrscheinlich, daß dieses Skriptum Fehler enthält.

Dieses Skriptum darf nur umsonst oder zum Selbstkostenpreis weitergegeben werden. Ich untersage jede kommerzielle Nutzung durch Dritte. Dieses Skriptum ist weder eine offizielle noch eine von *Professor Möller* autorisierte Version. Deshalb ist bei Fehlern zuerst davon auszugehen, daß diese von mir stammen.

Ingo Manfraß

Inhaltsverzeichnis zu Numerik I

<u>INHALTSVERZEICHNIS ZU NUMERIK I.....I</u>	<u>I</u>
<u>§ 1 FEHLERANALYSE</u>	<u>1</u>
1.1. FEHLER	1
1.2. FEHLERFORTPFLANZUNG UND STABILITÄT	1
1.3. RUNDUNGSFEHLER BEI GLEITKOMMAARITHMETIK	2
<u>§ 2 ITERATIVE LÖSUNGEN NICHTLINEARER GLEICHUNGEN.....</u>	<u>4</u>
2.1. FIXPUNKTITERATION:	4
2.2. METRISCHE RÄUME.....	4
2.3. DER BANACHSCHE FIXPUNKTSATZ (BFS)	5
2.4. KONVERGENZORDNUNG.....	6
2.5. DAS VERFAHREN VON NEWTON UND DAS SEKANTENVERFAHREN	6
<u>§ 3 POLYNOME</u>	<u>10</u>
3.1. POLYNOMIALE APPROXIMATION.....	10
<i>ZUNÄCHST TAYLOR</i>	10
<i>KETTENBRUCHENTWICKLUNGEN</i>	10
<i>GLEICHMÄßIGE APPROXIMATION</i>	11
<i>APPROXIMATIONSSATZ VON WEIERSTRAß</i>	11
3.2. AUSWERTUNG VON POLYNOMEN	11
3.3. TSCHEBYSCHJEFF – POLYNOME UND TSCHEBYSCHJEFF-ENTWICKLUNGEN	12
3.4. EINSCHLIEßUNGSSÄTZE FÜR POLYNOMNULLSTELLEN.....	14
3.5. STURMSCHJE KETTE UND DAS BISEKTIONSVERFAHREN	15
3.6. ANWENDUNGEN DES NEWTON – VERFAHRENS	17
<u>§ 4 DIREKTE LÖSUNGEN VON LINEAREN GLEICHUNGSSYSTEMEN</u>	<u>19</u>
4.1. DAS GAUßSCHE ELIMINATIONSVERFAHREN	19
4.2. DIE LR – ZERLEGUNG	20
4.3. DIE CHOLESKY – ZERLEGUNG	21
<i>VERFAHREN VON CHOLESKY CROUT</i>	22
4.4. DAS GAUß – JORDAN – VERFAHREN.....	22
4.5. MATRIXNORMEN	24
4.6. FEHLERABSCHÄTZUNGEN.....	25
4.7. DIE QR –ZERLEGUNG.....	25
<i>STABILE METHODE VON HOUSEHOLDER</i>	26
4.8. LINEARE AUSGLEICHSPROBLEME	27
<u>§ 5 ITERATIVE LÖSUNGEN LINEARER GLEICHUNGEN.....</u>	<u>30</u>
5.1. DAS GESAMT- UND EINZELSCHRITTVERFAHREN.....	30
<i>GESAMTSCHRITTVERFAHREN (GSV, JACOBI – VERFAHREN)</i>	30
<i>EINZELSCHRITTVERFAHREN (ESV, GAUß – SEIDEL – VERFAHREN)</i>	31
5.2. KONVERGENZ VON ITERATIONSVERFAHREN FÜR LINEARE GLEICHUNGSSYSTEME	31
5.3. RELAXATION UND NACHITERATION.....	32
<i>NACHITERATION</i>	33
<u>§ 6 EIGENWERTPROBLEME</u>	<u>34</u>
6.1. GRUNDBEGRIFFE AUS DER ALGEBRA.....	34

Numerik I

6.2. REDUKTION AUF TRIDIAGONAL- BZW. HESSENBERGGESTALT.....	35
<i>ALGORITHMUS VON HYMAN</i>	36
6.3. DIE POTENZMETHODE	36
6.4. DAS QR - VERFAHREN	38
<i>DAS QR – VERFAHREN MIT SHIFT</i>	39
6.5. DAS JACOBI – VERFAHREN	40
<u>INDEX ZU NUMERIK I.....</u>	42

§ 1 Fehleranalyse

1.1. Fehler

- Fehlertypen: - Modellfehler
 (Datenfehler, Idealisierungsfehler)
 - num. Fehler
 (Diskretisierungsfehler, Abbruchfehler, Zuordnungsfehler, Rundungsfehler)

Definition 1.1: x sei eine exakte Größe in einem normierten Raum. x' sei eine Näherung an x . Dann heißt

$$\varepsilon := x - x' \quad \text{absoluter Fehler}$$

$$\delta := \frac{|x - x'|}{|x|} \quad \text{relativer Fehler (falls } x \neq 0)$$

1.2. Fehlerfortpflanzung und Stabilität

Satz 1.2. (Fehlerfortpflanzung bei arithmetischen Operationen)

Die Operanden x_1, x_2 seien näherungsweise bekannt,

$$x_i = x_i' - \varepsilon_i \quad i = 1, 2$$

δ_i sei der relative Fehler. η bzw. ξ sei der absolute bzw. relative Fehler des Resultates bei einer arithm. Grundoperation. Dann gilt, falls der Nenner $\neq 0$:

Add.: $\eta = \varepsilon_1 + \varepsilon_2$

$$\xi = \frac{x_1}{x_1 + x_2} \cdot \delta_1 + \frac{x_2}{x_1 + x_2} \cdot \delta_2$$

Multipl. $\eta = \varepsilon_2 x_1 + \varepsilon_1 x_2$
 $\xi = \delta_1 + \delta_2$

Division: $\eta = \frac{1}{x_2} \cdot \varepsilon_1 - \frac{x_1}{x_2^2} \cdot \varepsilon_2$

$$\xi = \delta_1 - \delta_2$$

wobei die Ergebnisse für die Division und Multiplikation nur genähert sind.

Ferner gilt dies bei arithmetischen Grundoperationen in \mathbf{R} (etc.) und

$$\delta = \frac{|x - x'|}{|x|} \quad \text{falls } x \neq 0.$$

Definition 1.2.1: Ein mathematischer Prozeß heißt gut konditioniert, wenn kleine Änderungen der Daten x_1, \dots, x_n nur kleine Änderungen der exakten Lösung bewirken. Sonst heißt der Prozeß schlecht konditioniert.

$f: \Omega \rightarrow \mathbf{R}$ sei auf $\Omega \subset \mathbf{R}^n$ definiert, Ω offen, konvex. f sei auf Ω stetig differenzierbar.

Datenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \Omega$

$$\text{Fehlervektor } \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \quad \tilde{x} := \mathbf{x} + \boldsymbol{\varepsilon} \in \Omega.$$

Dann folgt aus Taylor:

$$\boldsymbol{\eta} = f(\tilde{x}) - f(x) \doteq \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \cdot \varepsilon_j \quad (\text{absoluter Fehler})$$

$$\boldsymbol{\xi} = \frac{\boldsymbol{\eta}}{f(x)} \doteq \sum_{j=1}^n \frac{x_j}{f(x)} \cdot \frac{\partial f}{\partial x_j}(x) \cdot \boldsymbol{\delta}_j \quad (\text{relativer Fehler})$$

Definition 1.2.2: Sei $f: \Omega \rightarrow \mathbb{R}$ wie oben, dann heißen

$$\boldsymbol{\delta}_i = \frac{\partial f}{\partial x_i}(x) \quad \text{bzw.} \quad \boldsymbol{\tau}_i = \frac{x_i}{f(x)} \cdot \frac{\partial f}{\partial x_i}(x)$$

Konditionszahlen (in Bezug auf die i -te Komponente) bzgl. des absoluten bzw. relativen Fehlers.

Ein Prozeß heißt gut konditioniert, wenn die Beträge der Konditionszahlen klein gegen Eins sind. Andernfalls schlecht konditioniert. \rightarrow sog. „natürliche“ Stabilität oder Instabilität.

Gegensatz dazu ist „numerische“ Stabilität bzw. Instabilität bedingt durch Rechnungsverlauf.

Beispiel 1.2.: $f(x) = \ln(x - \sqrt{x^2 - 1})$ Stichwort: Verlust führender Stellen

Ein Fehler, der bei der Berechnung von φ_i auftritt, geht in die Restabbildung $\varphi_n \circ \dots \circ \varphi_{i+1}$ wie ein Eingabefehler (Datenfehler) ein und wird durch diese Abbildung an das Endresultat weitergegeben.

Definition 1.2.3: Sei $f = \varphi_k \circ \varphi_{k-1} \circ \dots \circ \varphi_1$ die Zerlegung der Abbildung f in Elementaralgorithmen.

Ein Algorithmus $\varphi_k \circ \dots \circ \varphi_1$ zur Berechnung von f heißt numerisch stabil, wenn die Teilalgorithmen

$$h_i := \varphi_k \circ \dots \circ \varphi_i, \quad i = 1, \dots, k-1$$

natürlich stabil sind.

1.3. Rundungsfehler bei Gleitkommaarithmetik

Definition 1.3.1: Sei $g \in \mathbb{N}$ gerade, $t \in \mathbb{N}$, $x \in \mathbb{R} \setminus \{0\}$ mit

$$x = \sigma \cdot g^e \sum_{i=1}^{\infty} \alpha_i g^{-i} \quad \sigma \in \{-1, 1\}, e \in \mathbb{Z}, \alpha_i \in \{0, \dots, g-1\}, \alpha_i \neq 0$$

(g -adische Darstellung, ganz \mathbb{R} darstellbar, aber nicht eindeutig)

Dann

$$rd_t(x) := \begin{cases} \sigma \cdot g^e \sum_{i=1}^t \alpha_i g^{-i} & , \text{ falls } a_{t+1} < \frac{g}{2} \\ \sigma \cdot g^e \left(\sum_{i=1}^t \alpha_i g^{-i} + \frac{1}{g^t} \right) & , \text{ falls } a_{t+1} \geq \frac{g}{2} \end{cases}$$

$rd_t(x)$ heißt der auf t Stellen gerundete Wert von x .

Satz 1.3: Sei $y \in \mathbb{N}$ gerade, $t \in \mathbb{N}$, $x \neq 0$ mit g -adischer Darstellung wie oben.

Dann gilt

$$(i) \quad rd_t(x) = \sigma \cdot g^{e'} \sum_{i=1}^t \alpha'_i g^{-i}$$

$$(ii) \quad |rd_t(x) - x| \leq \frac{1}{2} g^{e-t}$$

$$(iii) \quad \left| \frac{rd_t(x) - x}{x} \right| \leq \frac{1}{2} g^{-t+1}$$

$$(iv) \quad \left| \frac{rd_t(x) - x}{rd_t(x)} \right| \leq \frac{1}{2} g^{-t+1}$$

Bemerkung: In (i) können sich alle Ziffern ändern.

Korollar 1.3: Es gelten die Darstellungen

$$rd_t(x) = x(1 + \varepsilon) = \frac{x}{1 + \xi} \quad \text{mit } \varepsilon, \xi \in \mathbf{R}:$$

$$|\varepsilon| \leq \frac{1}{2} g^{-t+1}, \quad |\xi| \leq \frac{1}{2} g^{-t+1}$$

Definition 1.3.2. Eine g-adische und t-stellige Gleitkommazahl hat die Form

$$x = 0 \text{ oder } x = \sigma \cdot g^e \sum_{i=1}^{\infty} \alpha_i g^{-i}$$

$$\sigma \in \{-1, 1\}, e \in \mathbf{Z}, \alpha_i \in \{0, \dots, g-1\}, \alpha_1 \neq 0$$

σ heißt Vorzeichen

e heißt Exponent

g heißt Basis

$\sum_{i=1}^k \alpha_i g^{-i}$ heißt Mantisse von x .

Maschinenzahlen sind g-adische t-stellige Gleitkommazahlen mit beschränktem Exponenten, etwa $-99 \leq e \leq 99$.

Im folgenden vernachlässigen wir Diskussionen zu Überlauf und Unterlauf.

Arithmetische Grundoperationen zwischen Maschinenzahlen (+), (-), (*), (/) führt im allgemeinen aus Bereich der Maschinenzahl hinaus.

Auf Rechnern wird (+), (-), (*), (/) i. allgemeinen so realisiert, daß

$$x (+) y = rd_t(x + y)$$

$$x (*) y = rd_t(x * y)$$

$$x (/) y = rd_t(x / y).$$

Es folgt:

$$x (+) y = (x + y) (1 + \varepsilon) = x(1 + \varepsilon) + y(1 + \varepsilon)$$

$$x (*) y = x(1 + \varepsilon) y$$

$$x (/) y = x(1 + \varepsilon) / y$$

Wilkinsons backward analysis (Rückwärtsanalyse)

Idee: Arbeitet man mit Daten, die durch Eingangsfehler verfälscht sind, dann sind Rundungsfehler irrelevant, wenn man das numerisch erhaltende Resultat als exaktes Resultat von verfälschten Daten ansehen kann, deren Störungsgrößenordnungsmäßig unterhalb der Eingangsfehler liegen. (man spricht von gutartigem Algorithmus)

Beispiel 1.3. Resultat: Fehler hängt von Summationsreihenfolge ab. Am besten Reihenfolge so wählen, daß die S_i nicht so groß werden. (in numerischen Operationen gilt das Assoziativgesetz nicht)

§ 2 Iterative Lösungen nichtlinearer Gleichungen

2.1. Fixpunktiteration:

Eine Grundaufgabe der Numerik:

Finde ein (alle) $x \in \Omega$ mit $f(x) = 0$, wo $f: \Omega \rightarrow \mathbf{R}$

x ist Nullstelle von f , x ist Lösung von $f(x) = 0$.

Näherungsweise Lösung: zu gegebenem $\varepsilon > 0$ finde $x' \in \Omega$, so daß

$$\exists x \in \Omega: f(x) = 0, \quad |x - x'| < \varepsilon.$$

Damit ist die Aufgabe numerisch gelöst.

Beispiel

2.2. Metrische Räume

2.2.1. Definition Sei E ein Vektorraum über Grundkörper $\mathbf{K} = \{\mathbf{C}, \mathbf{R}\}$. Eine Abbildung $\|\cdot\|: E \rightarrow \mathbf{R}$ heißt Norm, wenn sie erfüllt:

- 1) $\|x\| > 0$ für $0 \neq x \in E$
- 2) $\|\alpha x\| = |\alpha| \|x\|$ für $x \in E, \alpha \in \mathbf{K}$
(absolute [Betrag] Homogenität)
- 3) $\|x + y\| \leq \|x\| + \|y\|$ für beliebige $x, y \in E$

2.2.1. Beispiel $E = \mathbf{R}^n$

$$\|x\|_1 := \sum_{i=1}^n |x_i| \quad L_1 \text{ Norm}$$

$$\|x\|_2 := \sqrt{\sum_{i=1}^n |x_i|^2} \quad L_2 \text{ Norm / Euklidische Norm}$$

$$\|x\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p} \quad L_p \text{ Norm}$$

$$\|x\|_\infty := \max_{i=1}^n \{|x_i|\} \quad L_\infty \text{ Norm, Tschebyscheff-Norm, Maximumsnorm}$$

2.2.1. Satz: Auf dem Vektorraum \mathbf{K}^n sei eine Norm $\|\cdot\|$ gegeben. Dann sind $\|\cdot\|$ und $\|\cdot\|_\infty$ äquivalent, d.h. es gibt Konstanten $\alpha, \beta \in \mathbf{R} : 0 < \alpha < \beta < \infty$, so daß für alle $x \in \mathbf{K}^n$ gilt:

$$\alpha \|x\|_2 \leq \|x\| \leq \beta \|x\|_2$$

Beispiele für Normen

Korollar: Alle Normen im \mathbf{K}^n sind untereinander äquivalent.

2.2.2. Definition Sei E nicht leer. E heißt (zusammen mit einer Abb.

$d: E \times E \rightarrow \mathbf{R}$) metrischer Raum und d Metrik oder Distanzfunktion, wenn gilt:

- 1) $d(x,y) = 0 \Leftrightarrow x = y$
- 2) $d(x,y) \leq d(x,z) + d(y,z)$

Bemerkung: man kann aus Definition 2.2.2. auch die „üblichen“ Axiome herleiten!

2.2.3. Definition Eine Folge $\{x_n\} \subset E$ heißt eine Cauchy-Folge, falls zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon)$ existiert, so daß $d(x_\mu, x_\nu) < \varepsilon \quad \forall \mu, \nu \geq N$.

$\{x_n\} \subset E$ heißt konvergent gegen $x^* \in E$, wenn
 $d(x_n, x^*) < \varepsilon \quad \forall n \geq N$.

2.2.4. Definition Ein metrischer Raum heißt vollständig, wenn jede Cauchy-Folge in E gegen ein $x^* \in E$ konvergiert.

2.2.5. Definition Ist ein normierter Raum E vollständig bzgl. der Metrik $d(x,y) = \|x - y\|$, dann ist E mit seiner Norm ein Banach Raum.

2.3. Der Banachsche Fixpunktsatz (BFS)

2.3.1. Definition Eine Abbildung φ eines metrischen Raumes E in sich $\varphi: E \rightarrow E$ heißt Lipschitz-beschränkt mit der Lip.-Konstanten $L \geq 0$ falls für alle $x, y \in E$
 $d(\varphi(x), \varphi(y)) \leq L \cdot d(x,y)$.
 φ heißt kontrahierend oder Kontraktion, falls φ Lip.-beschränkt mit Lip.-konstanten $L < 1$ ist.

Bemerkung: 1) Es reicht nicht $L = 1$.

2) Kontrahierende Abb. sind stetig

2.3.1. Banachscher Fixpunktsatz

Ist $\varphi: E \rightarrow E$ eine kontrahierende Abbildung eines vollständigen metrischen Raumes in sich, dann besitzt φ genau einen Fixpunkt $x^* = \varphi(x^*)$. Die Iteration $x_{k+1} = \varphi(x_k)$, $k \geq 0$ konvergiert für jeden Startwert $x_0 \in E$ gegen den Fixpunkt. Bezeichnet L die Kontraktionszahl von φ , dann gelten für $k \geq 1$ die Fehlerabschätzungen:

i) $d(x_k, x^*) \leq L \cdot d(x_{k-1}, x^*) \quad (, L \geq |\varphi'| \text{“})$

ii) $d(x_k, x^*) \leq \frac{L^k}{1-L} \cdot d(x_0, x_1)$ a priori – Abschätzung

iii) $d(x_k, x^*) \leq \frac{L}{1-L} \cdot d(x_k, x_{k-1})$ a posteriori – Abschätzung

Bemerkung: Die Fixpunktfolge kann man auch schreiben als

$$\{x_0, \varphi(x_0), \varphi^2(x_0), \dots\}$$

(oder eine Teilfolge mit Anfangswert x_1 betrachten)

2.3.2. Satz: Sei $I \subseteq \mathbf{R}$ ein Intervall. $\varphi: I \rightarrow \mathbf{R}$ sei stetig diffbar. φ ist kontrahierend, wenn für alle $x \in I$ gilt: $|\varphi'(x)| \leq L < 1$.

Beispiel

In der Praxis ist φ manchmal nicht auf ganz E definiert oder nur in Teilbereichen von E kontrahierend. Daher lokale Variante des BFS:

$D \subset E$, $D \neq \emptyset$ und Kugel in D :

$$K_r(\xi) = \{x \in E \mid d(x, \xi) < r\} \subseteq D.$$

Mit E ist auch $K_r(\xi)$ ein vollständiger metrischer Raum. Daher hat ein φ , das Selbstabbildung auf $K_r(\xi)$ ist und in D kontrahierend ist, einen Fixpunkt, der mit dem Banachschen Fixpunktsatz berechnet werden kann.

Es gilt: $\varphi: K_r(\xi) \rightarrow K_r(\xi)$, wenn $d(\varphi(\xi), \xi) \leq r - Lr$ ist.

2.3.3. Satz: (lokale Variante des BFS):

φ bildet die Teilmenge D des vollständigen metrischen Raum E in E ab. Es mögen $x \in D$, $r > 0$ und $0 \leq L < 1$ existieren mit:

1) $K_r(\xi) = \{x \in E \mid d(x, \xi) \leq r\} \subseteq D$

2) Für $u, v \in K_r(\xi)$: $d(\varphi(u), \varphi(v)) \leq L(d(u, v))$

3) Es gilt die Kugelbedingung. Dann hat φ Fixpunkt in $K_r(\xi)$.

Jede Folge $x_{n+1} = \varphi(x_n)$ mit $x_0 \in K_r(\xi)$ konvergiert gegen den Fixpunkt und es gelten die Fehlerabschätzungen.

(Kugelbedingung : $d(\varphi(u), \varphi(v)) \leq (1-L)r$)

2.3.2. Definition: Es sei $\varphi: E \rightarrow E$ eine Iterationsfunktion in einem metrischen Raum E . x^* sei ein Fixpunkt von φ . Es gebe eine Umgebung $U(x^*)$ von x^* , so daß für alle Startwerte $x_0 \in U(x^*)$ die Iterationsfolge $x_{n+1} = \varphi(x_n)$ $n = 0, 1, \dots$ gegen x^* konvergiert. Dann heißt das durch φ erzeugte Iterationsverfahren lokal konvergent.
Ist $U(x^*) = E$ möglich, dann global konvergent.

2.4. Konvergenzordnung

2.4.1. Definition: Ein durch die Iterationfunktion $\varphi: E \rightarrow E$ mit Fixpunkt $x^* \in E$ gegebenes Iterationsverfahren $x_{n+1} = \varphi(x_n)$ hat mindestens die Konvergenzordnung $s \in \mathbf{R}$, $s \geq 1$, falls für den absoluten Fehler $e_k = d(x_k, x^*)$ gilt:

$$\limsup_{k \rightarrow \infty} \left(\frac{e_{k+1}}{e_k^s} \right) = c \quad \text{mit} \quad \begin{cases} |c| < 1, \text{ für } s = 1 \\ |c| < \infty, \text{ für } s > 1 \end{cases}$$

Der Wert c heißt der asymptotische Fehlerkoeffizient (kurz: Konvergenzfaktor). Ein Iterationsverfahren hat genau die Ordnung s , wenn $c \neq 0$ gilt.

Superlinear, wenn $s = 1$, $c = 0$

linear, wenn $s = 1$, $0 < |c| \leq 1$

quadratisch, wenn $s = 2$, $c \neq 0$.

2.4.1. Satz: $I \subseteq \mathbf{R}$ Intervall, $\varphi: I \rightarrow I$ sei s -mal stetig diffb. Es gelte $\varphi(x^*) = x^*$ und $\varphi'(x^*) = \varphi''(x^*) = \dots = \varphi^{(s-1)}(x^*) = 0$, falls $s > 1$ bzw. $|\varphi'(x^*)| < 1$. Dann hat das Iterationsverfahren $x_{n+1} = \varphi(x_n)$, x_0 geeignet lokal mindestens die Konvergenzordnung s . Gilt nämlich $\varphi^{(s)}(x^*) \neq 0$, dann ist s die genaue Konvergenzordnung.

Bemerkung: Ist $\varphi: I \rightarrow I$ stetig diffbar und $|\varphi'(x)| < 1$, dann gibt der MWS an, daß die Folge $x_{n+1} = \varphi(x_n)$ monoton konvergent, falls $\varphi'(x) > 0$ für alle $x \in I$ und alternierend konvergent, falls $\varphi'(x) < 0$ für alle $x \in I$.

2.5. Das Verfahren von Newton und das Sekantenverfahren

2.5.1. Algorithmus (Newton – Verfahren)

Sei x^* eine Nullstelle von $f: \mathbf{R} \rightarrow \mathbf{R}$. In Umgebung $I_\varepsilon :=]x^* - \varepsilon, x^* + \varepsilon[$ sei f zweimal stetig diffbar.

gesucht: $x_k \in I$: $|x_k - x^*| < \varepsilon$, $f(x_k) = 0$

$$\text{Iteration: } x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad k = 0, 1, \dots$$

2.5.1. Satz: Sei x^* eine NS von f . In einer Umgebung $I_\varepsilon =]x^* - \varepsilon, x^* + \varepsilon[$, $\varepsilon > 0$, von x^* sei f zweimal stetig diffbar und es gelte $f'(x) \neq 0$. Dann kon-

vergiert das Newtonverfahren lokal bei x^* , d.h. $\exists I_\delta \subseteq I_\epsilon$, so daß für jedes $x_0 \in I_\delta$ die Folge $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ gegen x^* konvergiert.

2.5.2. Satz: Sei x^* NS von f . In I_ϵ sei f zweimal stetig diffbar und es gelte $f'(x^*) \neq 0$. Dann ex. $I_\delta =]x^* - \delta, x^* + \delta[\subseteq I_\epsilon$, so daß für jedes $x_0 \in I_\delta$ die Folge $\{x_n\}$ des Newtonverfahrens mindestens quadratisch gegen x^* konvergiert.

Feststellung im Beweis:
$$\limsup_{n \rightarrow \infty} \left(\frac{x^* - x_{n+1}}{(x^* - x_n)^2} \right) = -\frac{1}{2} \frac{f''(x^*)}{f'(x^*)}$$

Bemerkung: Wenn $f(x^*) = f'(x^*) = 0$ dann Sätze 1 und 2 nicht mehr anwendbar!

Sei $f \in C^{m+1} [a,b]$, $m > 1$. Zu $x^* \in [a,b]$ sei

$$f(x^*) = f'(x^*) = \dots = f^{(m-1)}(x^*) = 0 \neq f^{(m)}(x^*).$$

Dann gilt nach Taylor: Bei m -facher Nullstelle gilt:

$$\left(\frac{f(x_n)}{f'(x_n)} \right)' \rightarrow \frac{1}{m}$$

2.5.3. Satz: Sei $f \in C^{(m+1)} [a,b]$ $m > 1$ mit m -facher NS (d.h. Ableitungen wie in (2.5.2.)) $x^* \in [a,b]$. Sei $1 \leq k \leq m$. Dann existiert ein $I_\delta =]x^* - \delta, x^* + \delta[$, so daß für jeden Startwert $x_0 \in I_\delta$ die Folge

$$x_{n+1} = x_n - k \frac{f(x_n)}{f'(x_n)}$$

für $k < m$ linear mit Konvergenzfaktor $1 - \frac{k}{m}$ gegen x^* konvergiert.

Für $k = m$ ist die Konvergenz mindestens quadratisch.

Schätzen der Vielfachheit: man kann nun das obige m auch schätzen. Dazu hilft

die folgende Gleichheit:
$$m \approx -k \frac{x_n - x_{n-1}}{(x_{n+1} - x_n) - (x_n - x_{n-1})}$$

Beispiel

Nachteile des Newton – Verfahren:

- Berechnung von $f'(x_n)$ in jedem Schritt neu
- manchmal f' nicht explizit bekannt
- oder f nicht diffbar
- oder f' nur mühsam zu berechnen

Idee: Sekantenverfahren

2.5.2. Algorithmus: (Sekantenverfahren)

Gegeben: $x_0, x_1 \in I$, $f \in C(I)$, $\epsilon > 0$

Gesucht: $x_k \in I$ mit $|x_k - x^*| < \epsilon$, $f(x^*) = 0$.

Iteration: $x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \cdot f(x_n)$, solange bis

Abbruchkriterium erfüllt.

2.5.3. Algorithmus (Regula falsi)

Gegeben: $x_0, x_1 \in I$, $f \in C(I)$, $\epsilon > 0$, $f(x_0) f(x_1) < 0$

Gesucht: $x_k \in I$ mit $|x_k - x^*| < \epsilon$, $f(x^*) = 0$.

Iteration: $y = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \cdot f(x_n)$, solange bis Abbruchkriterium erfüllt.

$x_{n+1} := y$ falls $f(y) f(x_n) < 0$
 $x_{n+1} := y$ und $x_n := x_{n-1}$ falls $f(y) f(x_n) > 0$.
 Abbruch sobald $|x_{n+1} - x_n| < \epsilon$.

2.5.4. Satz: Es sei $f \in C^2[a,b]$ mit $f(x^*) = 0$ für $x^* = \frac{a+b}{2}$. Außerdem gelte:

- (i) $f'(x) \neq 0 \quad \forall x \in [a,b]$
- (ii) $\min_{a \leq x \leq b} |f'(x)| > \frac{b-a}{2} \max_{a \leq x \leq b} |f''(x)|$

Dann konvergiert das Sekantenverfahren global auf $[a,b]$ gegen x^* und eine Konvergenzordnung ist $\frac{1+\sqrt{5}}{2} \approx 1,618$ wenn $f''(x^*) \neq 0$.

Bemerkung: Ist f in einer Umgebung von x^* streng monoton und strikt konvex, dann hat das Sekantenverfahren die genaue Ordnung $\frac{1+\sqrt{5}}{2} \approx 1,618$.

2.5.5. Satz: Für jede stetige Funktion $f: [a,b] \rightarrow \mathbf{R}$ konvergiert das Iterationsverfahren der Regula falsi für $x_0 = a, x_1 = b$ und $f(a) f(b) < 0$ gegen eine NS x^* von f .

Bemerkung: Ist $f \in C^2[a,b]$ und $f''(x^*) \neq 0$, dann ist die Konvergenzordnung 1.

Betrachte nun mehrdimensionalen Fall:

Für Funktionen $F: \Omega \rightarrow \mathbf{R}^n$ gilt:

$$F(x) = \begin{pmatrix} F_1(x) \\ \vdots \\ F_k(x) \end{pmatrix}, \text{ wo } F_i(x): \Omega \rightarrow \mathbf{R}^n.$$

x^* NS von $F \Leftrightarrow x^*$ ist NS aller F_i .

Beispiel:

Nach Taylor gilt hier entsprechend zu $n = 1$:

$F(x) = F(x^{(v)}) + J_F(x^{(v)}) (x - x^{(v)})$, wo

$$J_F(x^{(v)}) := \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x^{(v)}) & \dots & \frac{\partial F_1}{\partial x_n}(x^{(v)}) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1}(x^{(v)}) & \dots & \frac{\partial F_n}{\partial x_n}(x^{(v)}) \end{pmatrix}, \text{ wenn } x^* \text{ NS, dann}$$

$0 = F(x^{(v)}) + J_F(x^{(v)}) (x^* - x^{(v)})$.

Durch Umformung:

$$x^{(v+1)} = x^{(v)} - J_F(x^{(v)})^{-1} F(x^{(v)}).$$

Numerisch günstiger: löse das LGS: $J_F(x^{(v)}) \Delta^{(v)} = F(x^{(v)})$, dann $x^{(v+1)} = x^{(v)} - \Delta^{(v)}$.

Für multivariablen Newton – Verfahren gilt ebenfalls lokale Konvergenzaussage und die quadratische Konvergenz, falls x^* einfache NS von F (vgl. 2.5.1., 2.5.2)

§ 3 Polynome

3.1. Polynomiale Approximation

Zunächst Taylor

Bei beliebigen Funktionen $f \in C^\infty [a,b]$ treten Probleme auf.

- Taylor Koeffizienten konvergieren oft nur langsam (unter Umständen auch gar nicht) gegen 0.
- Bei zunehmender Entfernung von x zu x_0 wird Approximation durch Taylorpolynom immer schlechter.
- Bei Rechnung mit endlichen Stellenzahl sind die Polynomkoeffizienten auch sehr groß, obwohl Funktionswerte sehr klein. (Auslöschungseffekt)

3.1.1. Beispiel:

Kettenbruchentwicklungen

Mit Kettenbrüchen kann man Funktionen mit hoher Genauigkeit darstellen und relativ schnell auswerten.

3.1.1. Definition: Sei $A := \{a_k\}_{k=1}^\infty$ und $B := \{b_k\}_{k=0}^\infty$ Folgen reeller oder komplexer Zahlen. Dann nennt man

$$k_n := b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{\ddots}{b_n}}}$$

den von A und B erzeugten Kettenbruch der Länge n . Kürzer:

$$k_n = b_0 + \frac{a_1|}{|b_1|} + \dots + \frac{a_n|}{|b_n|}$$

Auswertung: $z^{(0)} := b_n$,

$$z^{(i)} := b_{n-i} + \frac{a_{n-(i-1)}}{z^{(i-1)}} \quad i = 1, \dots, n$$

$$k_n := z^{(n)}.$$

Es kann passieren, daß ein $z^{(i-1)} = 0 \Rightarrow$ Abbruch.

(Euler und Wallis) Für die Berechnung von k_n betrachtet man die Abschnittskettenbrüche (j -te Näherungsbrüche):

$$k_j := b_0 + \frac{a_1|}{|b_1|} + \dots + \frac{a_j|}{|b_j|} =: \frac{P_j}{Q_j} \quad j = 0, \dots, n$$

3.1.1. Algorithmus (Kettenbruchauswertung):

Gegeben: $A := \{a_k\}_{k=1}^\infty$ und $B := \{b_k\}_{k=0}^\infty$ Folgen in \mathbf{K}

Gesucht: $k_n = b_0 + \frac{a_1|}{|b_1|} + \dots + \frac{a_n|}{|b_n|} \in \mathbf{K}$

Start: $P_{-1} := 1, Q_{-1} := 0, P_0 := b_0, Q_0 := 1$.

Iteration: Für $j = 1, \dots, n$:

$$P_j := b_j P_{j-1} + a_j P_{j-2},$$

$$Q_j := b_j Q_{j-1} + a_j Q_{j-2}.$$

Output: $P_n/Q_n = k_n$.

3.1.2. Beispiel

Gleichmäßige Approximation

zu $f \in C[a,b]$ wird ein Polynom p gesucht mit: $\max_{x \in [a,b]} |f(x) - p(x)| < \varepsilon$

Approximationssatz von Weierstraß

$f \in C[a,b], \varepsilon > 0 \Rightarrow \exists p$ Polynom $\max_{x \in [a,b]} |f(x) - p(x)| < \varepsilon$.

Frage: Wie findet man unter allen Polynomen p von Grad $\leq n$ ein Polynom p^* mit

$$E_n^{[a,b]}(f) := \|f - p^*\|_{\infty}^{[a,b]} \leq \|f - p\|_{\infty}^{[a,b]} \quad ?$$

3.1.3. Beispiel

3.2. Auswertung von Polynomen

$$p_n(x) = (\dots ((a_n x + a_{n-1}) x + a_{n-2}) x \dots) x + a_0$$

3.2.1. Horner – Algorithmus (entspricht dieser Klammerung):

Gegeben: Koeffizientenvektor $(a_0, a_1, \dots, a_n) \in \mathbf{K}^{n+1}, x_0 \in \mathbf{K}$

Gesucht: $p_n(x_0)$ mit $p_n(x) := \sum_{k=0}^n a_k x^k$

Berechnung: $a_n^{(1)} := a_n$, für $k = 1, \dots, n$: $a_{n-k}^{(1)} := x_0 a_{n-k+1}^{(1)} + a_{n-k}$

Output: $a_0^{(1)}$

Aufwand: n Multiplikationen

3.2.1. Horner – Schema

Bemerkung:
$$p_n(x) = \sum_{k=0}^n a_k x^k = (x - x_0) \sum_{k=0}^{n-1} b_k x^k + p_n(x_0)$$

Koeffizientenvergleich liefert: $b_{n-1} = a_n^{(1)} \dots b_0 = a_1^{(1)}, p(x_0) = a_0^{(1)}$

Damit ergibt sich:
$$\frac{p_n(x) - p_n(x_0)}{x - x_0} = \sum_{k=0}^{n-1} b_k x^k \xrightarrow{x \rightarrow x_0} \sum_{k=0}^{n-1} b_k x_0^k = p_n'(x_0)$$

3.2.2. Doppelzeiliges Hornerschema:

Gegeben: (a_n, \dots, a_0) reeller Koeffizientenvektor von $p_n, s, p \in \mathbf{R}$

Gesucht: Zahlen, $a_0^{(1)}, a_1^{(0)} \in \mathbf{R}$ und q^{n-2} , so daß

$$p_n(x) = (x^k - sx - p) q_{n-2}(x) + a_1^{(1)} x + a_0^{(1)}$$

3.2.3. Rechenschema

3.2.4. Satz: Sind α_1, α_2 NS des Polynoms p_n , dann lassen sich diese Nullstellen unter

Verwendung des doppelzeiligen Horner-Schemas mit $s = \alpha_1 + \alpha_2$,

$p = -\alpha_1 \alpha_2$ simultan abspalten. Genau in diesem Fall ist $a_0^{(1)} = a_1^{(1)} = 0$.

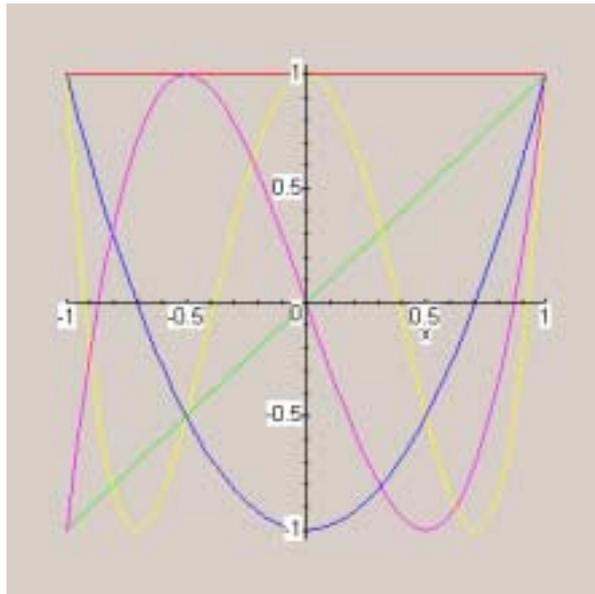
3.3. Tschebyscheff – Polynome und Tschebyscheff-Entwicklungen

3.3.1. Definition: Durch die Rekursion $T_0(x) = 1$, $T_1(x) = x$, $T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x)$, $n \geq 1$ werden Polynome definiert. Sie heißen Tschebyscheff-Polynome erster Art.

Durch die Rekursion $U_0(x) = 1$, $U_1(x) = 2x$,

$U_{n+1}(x) = 2x U_n(x) - U_{n-1}(x)$, $n \geq 1$ werden Polynome definiert. Sie heißen Tschebyscheff-Polynome zweiter Art.

Bemerkung:



Leichte Rechnung zeigt $T_n(x) = 2^{n-1} x^n + \text{Terme niedriger Ordnung}$

$U_n(x) = 2^n x^n + \text{Terme niedriger Ordnung}$

3.3.1. Satz: Die Tschebyscheff – Polynome erster Art haben die Darstellung:

$T_n(x) = \cos(n \arccos(x))$ für $x \in [-1, 1]$, $n \in \mathbf{N}_0$ und die Eigenschaften:

1) $|T_n(x)| \leq 1$ für $x \in [-1, 1]$

2) T_n hat in $[-1, 1]$ die Extremalpunkte: $x_k^{(n)} = \cos\left(\frac{k\pi}{n}\right)$, $k = 0, \dots, n$

und $T_n(x_k^{(n)}) = (-1)^k$

3) T_n hat n einfache Nullstelle, alle liegen in $[-1, 1]$, nämlich

$\tilde{x}_k^{(n)} = \cos\left(\frac{2k-1}{2n} \cdot \pi\right)$, $k = 1, \dots, n$

4) Zwischen je zwei NS von T_{n+1} liegt eine NS von T_n .

Bemerkung: T_n ist gerade für n gerade

T_n ist ungerade für n ungerade

3.3.2. Satz: Unter allen Polynomen $p \in \mathbf{P}_n$, deren Koeffizient bei x^n gleich 1 ist, hat das

Polynom $\frac{1}{2^{n-1}} T_n$ die kleinste Maximumsnorm im Intervall $[-1, 1]$:

$$\min_{p \in \mathbf{P}} \|p\|_{\infty}^{[-1,1]} = \left\| \frac{1}{2^{n-1}} T_n \right\|_{\infty}^{[-1,1]} = \frac{1}{2^{n-1}}.$$

3.4. Einschließungssätze für Polynomnullstellen

3.4.1 Satz: Sei $p_n \in \mathbf{P}_n$. Für jedes $z \in \mathbf{C}$ liegt mindestens eine Nullstelle von p_n in dem Kreis um z mit dem Radius:

$$r := n \cdot \frac{|P_n(z)|}{|P_n'(z)|}.$$

3.4.1. Lemma (Frobenius – Begleitmatrix)

Sei $p_n(z) = \sum_{k=0}^n a_k z^k$ mit $a_n = 1$.

Die Nullstellen von p_n sind Eigenwerte der Matrix

$$A = \begin{pmatrix} 0 & 1 & & & 0 \\ 0 & 0 & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix} \quad \text{Frobenius Begleitmatrix.}$$

Zum Eigenwert λ von A gehört der Eigenvektor $(1, \lambda, \lambda^2, \dots, \lambda^{n-1})^T$.

$$z \cdot \begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}}_{=:A} \cdot \begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ p_n(z) \end{pmatrix} \quad (*)$$

Bemerkung: Durch gliedweises Differenzieren der Identität (*) kann man zeigen, daß

$$V_k(\lambda) = \frac{d^{k-1}}{dz^{k-1}} \frac{1}{(k-1)!} \begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-1} \end{pmatrix}_{z=\lambda} = \frac{1}{(k-1)!} \begin{pmatrix} 0 \\ \vdots \\ (k-1)! \\ \vdots \\ (n-1)! \lambda^{n-k} \\ (n-k)! \end{pmatrix}.$$

Insbesondere $V_1(\lambda) = (1, \lambda, \lambda^2, \dots, \lambda^{n-1})^T$
 $V_2(\lambda) = (0, 1, 2\lambda, \dots, (n-1)\lambda^{n-2})^T$

....

Aus (*) folgt, falls λ m -fache Nullstelle ist

(also $p_n(\lambda) = p_n'(\lambda) = \dots = p_n^{(m-1)}(\lambda) = 0$):

$V_{k-1}(\lambda) + \lambda V_k(\lambda) = A V_k(\lambda).$

Ist λ eine m_i -fache Nullstelle von p_n , $i = 1, \dots, s$ mit $\sum_{i=1}^s m_i = n$

Dann gilt:

$$A[V_1(\lambda_i), V_2(\lambda_i), \dots, V_m(\lambda_i)] = [V_1(\lambda_i), V_2(\lambda_i), \dots, V_m(\lambda_i)] \cdot \underbrace{\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}}_{\text{Jordankästchen } J_i}$$

3.4.2. Satz: (Gerschgorinscher Kreissatz)

Sei $A = (a_{ij})_{i,j=1..n}$ Matrix mit $a_{ij} \in \mathbf{K}$. Sei K_i die abgeschlossene Kreisscheibe um a_{ii} mit Radius $r_i := \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$, $i = 1, \dots, n$.

Dann liegt jeder Eigenwert von A in mindestens einem K_i , d.h.

$$\{\lambda \in \mathbf{C} \mid \lambda \text{ EW}\} \subset \bigcup_{i=1}^n K_i.$$

Bemerkung: A und A^T haben dieselben Eigenwerte.

3.4.3. Satz: Die Vereinigung der Kreisscheiben $K_i^T := \{z \in \mathbf{C} \mid |z - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ki}|\}$

enthält ebenfalls alle Eigenwerte der Matrix A .

Korollar: Alle Eigenwerte von A liegen in $\left(\bigcup_{i=1}^n K_i\right) \cap \left(\bigcup_{i=1}^n K_i^T\right)$.

3.4.4. Satz: Die Nullstellen z_1, \dots, z_n des Polynoms $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ genügen den Abschätzungen:

- (i) $|z_k| \leq \max \left\{ 1, \sum_{j=0}^{n-1} |a_j| \right\}$
- (ii) $|z_k| \leq \max \{ |a_0|, 1+|a_1|, \dots, 1+|a_{n-1}| \}$.

3.5. Sturmsche Kette und das Bisektionsverfahren

Hier alles reell, d.h. Polynome mit reellen Koeffizienten; nur reelle NS werden gezählt.

Sprechweise: $\sigma := (f_0, f_1, \dots, f_n)$ geordnete Liste von Objekten heißt Sequenz.

3.5.1. Definition: Sei $\sigma := (f_0, f_1, \dots, f_n)$ Sequenz reeller Polynome. σ heißt Sturm-sche Kette auf $[a,b]$, wenn gilt:

- (i) $f_0(a) \neq 0, f_0(b) \neq 0$.
- (ii) $f_n(x) \neq 0 \quad \forall x \in [a,b]$
- (iii) $1 \leq k < n, f_k(\xi) = 0 \quad \xi \in [a,b] \Rightarrow f_{k-1}(\xi) \cdot f_{k+1}(\xi) < 0$
- (iv) $\xi \in]a,b[, f_0(\xi) = 0 \Rightarrow f'_0(\xi) f_1(\xi) > 0$.

3.5.2. Definition: Sei $\sigma = (a_0, a_1, \dots, a_n)$ mit $a_i \in \mathbf{R}$. Dann ist $w(\sigma)$ die Anzahl der Zeichenwechsel in der Sequenz ohne Berücksichtigung der Nullstellen.

$$i) w(a_0, a_1, \dots, a_n) = \begin{cases} w(a_1, \dots, a_n) + 1 & \text{falls } a_0 a_1 < 0 \\ w(a_1, \dots, a_n) & \text{falls } a_0 a_1 > 0 \text{ oder } a_0 = 0 \\ w(a_0, a_2, \dots, a_n) & \text{falls } a_0 \neq 0, a_1 = 0 \end{cases}$$

ii) $w(a_n) = 0$ (induktiv)

Abkürzung: $w(\sigma, x) := w(f_0(x), \dots, f_n(x))$, wenn $\sigma = (f_0, \dots, f_n)$

3.5.1. Satz (Sturm):

Sei σ eine Sturmsche Kette auf $[a, b]$. Dann ist $w(\sigma, a) - w(\sigma, b)$ die Anzahl der Nullstellen von f_0 in $[a, b]$.

Konstruktion einer Sturmschen Kette:

f_0 Polynom des Grades n
 $f_1 := f_0'$ Polynom des Grades $n - 1$;

Euklidischen Divisionsalgorithmus:

$$\begin{aligned} f_0 &= q_1 f_1 - f_2 && \text{grad}(f_2) < \text{grad}(f_1) \\ f_1 &= q_2 f_2 - f_3 && \text{grad}(f_3) < \text{grad}(f_2) \\ &\dots \\ f_{s-1} &= q_s f_s - f_{s+1} \\ f_s &= q_{s+1} f_{s+1}. \end{aligned}$$

$$f_{s+1} = \text{ggT}(f_0, f_1). \quad (f_{s+1} \neq 0)$$

Fallunterscheidung:

1. Fall: $f_{s+1} = \text{const} \neq 0$ (dann ist σ eine Sturmsche Kette)
2. Fall: f_{s+1} nicht konstant: Dann wähle $g_k := \frac{f_k}{f_{s+1}}$. Bildet eine

Sturmsche Kette.

Bemerkung: Im 2. Fall ist $\left(\frac{f_0}{f_{s+1}}, \dots, \frac{f_{s+1}}{f_{s+1}} \right) = \sigma'$, $\sigma = (f_0, \dots, f_{s+1})$ eine Sturmsche Kette

und es gilt: $w(\sigma, x) = w(\sigma', x)$ für alle x mit $f_{s+1}(x) \neq 0$.

D. h. zur Berechnung der Wechselzahlen im Satz von Sturm reicht es in jedem Fall, die Wechselzahl $w(\sigma, a) - w(\sigma, b)$ zu berechnen. Sturmsche Ketten auch gut, wenn man nur eine bestimmte („die zweite“) Nullstelle von f_0 ausrechnen will.

3.5.1. Algorithmus (Bisektionsverfahren)

Gegeben: Sturmsche Kette $\sigma := (f_0, f_1, \dots, f_m)$, Intervall $[a, b]$, das alle reellen Nullstelle von f_0 enthält.

$m := w(\sigma, a) - w(\sigma, b)$, $k \in \{1, \dots, m\}$, $\varepsilon > 0$.

Gesucht: Die reelle Nullstelle ξ_k von f_0 mit $\xi_1 < \xi_2 \dots < \xi_m$. Nullstellen von f_0 , genauer: ein Intervall $[a', b']$ mit $\xi_k \in [a', b']$ und $b' - a' < \varepsilon$.

Iteration: Berechne für $c := \frac{a+b}{2}$ die Zahl $w(\sigma, c)$.

Wenn $w(\sigma, c) - w(\sigma, a) < k$, dann ersetze a durch c , d.h. $a := c$,
 $k := k - w(\sigma, c)$.

sonst $b := c$.

Abbruch sobald $b - a < \epsilon$.

3.6. Anwendungen des Newton – Verfahrens

Bisher: reell (Algorithmus 2.5.1)

mehrdim. (Algorithmus 2.5.12)

Newton – Verfahren für komplexe Funktionen:

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)} \quad \text{für } f \in C^1(G), \quad G \subseteq \mathbf{C}.$$

Startwerte abschätzen durch „Höhenlinien“

$$|f|: z \rightarrow |f(z)|.$$

Fehlerabschätzungen mit Satz 3.4.1., aber nur für Polynome $p_n \in \mathbf{P}_n$:

$$|x_k - x^*| \leq n \cdot \frac{|p_n(x)|}{|p_n'(x)|}.$$

Problem: alle Nullstellen eines Polynoms!

Seien ξ_1, \dots, ξ_s bekannte Nullstellen.

$$\text{Es gilt: } x_{k+1} = x_k - \frac{1}{\underbrace{\frac{f'(x_k)}{f(x_k)} - \sum_{i=1}^s \frac{1}{x_k - \xi_i}}_{= \sum_{i=s+1}^n \frac{1}{x_k - \xi_i} \cdot \frac{g'(x_k)}{g(x_k)}}}$$

$$\text{mit } g(x) = a_n (x - \xi_{s+1}) \dots (x - \xi_n) = \frac{f(x)}{(x - \xi_1) \cdot \dots \cdot (x - \xi_s)}$$

Dieses Verfahren nennt man Deflationsverfahren.

(mindestens 2x so langsam, wie das reelle Verfahren)

Problem: Finde die komplexen Nullstellen eines Polynoms $p_n(x) = \sum_{k=0}^n a_k x^k$, mit $a_k \in \mathbf{R}$

unter Vermeidung der komplexen Arithmetik.

Lösung mit Bairstow – Verfahren.

Algorithmus von Bairstow:

Gegeben: Koeffizientenvektor $(a_0, \dots, a_n)^T \in \mathbf{R}^{n+1}$ von $p_n(x) = \sum_{k=0}^n a_k x^k$,

Näherung ξ an eine komplexe Nullstelle x^* von p_n , $\epsilon > 0$.

Gesucht: Näherung $\xi_k, \bar{\xi}_k$ an x^* , \bar{x}^* mit $|\xi_k - x^*| < \epsilon$.

Start: $s_0 := 2 \operatorname{Re} \xi$, $p_0 := -|\xi|^2$

$$\text{Iteration: } k = 0, 1, \dots \quad J(s_k, p_k) := \begin{pmatrix} p_k c_{3k} & c_{2k} \\ c_{2k} + s_k c_{3k} & c_{3k} \end{pmatrix}.$$

Sei (ϵ_k, δ_k) Lösung von $J(s_k, p_k) \begin{pmatrix} \epsilon_k \\ \delta_k \end{pmatrix} = - \begin{pmatrix} b_0(s_k, p_k) \\ b_1(s_k, p_k) \end{pmatrix}$

und $s_{k+1} = s_k + \epsilon_k$.

$p_{k+1} = p_k + \delta_k$.

$\Rightarrow \xi_{k+1}, \bar{\xi}_{k+1} = \frac{s_{k+1}}{2} \pm \sqrt{p_{k+1} + \frac{s_{k+1}^2}{4}}$, Wenn $|\xi_{k+1} - \xi_k| < \epsilon$, dann Abbruch.

$$\text{Bemerkung: } L_i^{-1} = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & -l_{i+1,i} & \ddots & \\ 0 & -l_{n,i} & 0 & 0 & 1 \end{pmatrix}$$

4.1.1. Satz: Jede untere Dreiecksmatrix mit normierter Diagonale, also:

$$L = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ * & & & 1 \end{pmatrix} \text{ lässt sich als Produkt von elementaren unteren Drei-}$$

ecksmatrizen schreiben. $L = L_1 \dots L_{n-1}$.

Gaußelimination

4.2. Die LR – Zerlegung

4.2.1. Definition: Sei $A \in \mathbf{K}^{n \times n}$ und $L, R \in \mathbf{K}^{n \times n}$ mit $L = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ * & & & 1 \end{pmatrix}$ (Linke – untere

– Dreiecksmatrix) und $R = \begin{pmatrix} * & \dots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & * \end{pmatrix}$ (Rechte – obere – Dreiecks-

matrix) gilt $A = L R$, so heißt dieses Darstellung die LR – Zerlegung von A.

4.2.1. Satz: Es sei $A \in \mathbf{K}^{n \times n}$ eine nicht-singuläre Matrix, die nach Durchführung der Gaußelimination auf die obere Dreiecksmatrix R übergeführt sein möge. Ferner sei $P = P_{n-1} \dots P_1$ das Produkt aller benötigter Permutationsmatrizen. Dann gilt

$$P A = L R,$$

mit L untere Dreiecksmatrix mit Koeffizienten aus dem Pivotisierungsverfahren.

Bemerkung: $\det(A) = (-1)^k \det(R) = (-1)^k \prod_{j=1}^n a_{jj}^{(k)}$ wobei k die Anzahl der erforderlichen

Zeilenvertauschungen ist.

4.2.2. Satz: Eine reguläre Matrix $A \in \mathbf{K}^{n \times n}$ besitzt genau dann eine LR – Zerlegung, wenn alle Hauptminoren (d.h. Unterdeterminanten) von Null verschieden sind.

Berechnung der LR – Zerlegung nach Crout:

Für $i = 1, \dots, n$ gilt für die Einträge von R

$$r_{ik} := a_{ik} - \sum_{j=1}^{i-1} l_{ij} r_{jk} \quad k = i, i+1, \dots, n \text{ und } l_{ij} \text{ Faktor aus Gaußelim. (1. Zeile bleibt)}$$

Für die Einträge der Unteren – Dreiecksmatrix L:

$$l_{ki} := \frac{1}{r_{ii}} \left(a_{ki} - \sum_{j=1}^{i-1} l_{kj} r_{jk} \right) \quad k = i+1, \dots, n.$$

4.2.3. Satz: Wenn für $A \in \mathbf{K}^{n \times n}$ mit $\det(A) \neq 0$ eine LR Zerlegung existiert, dann ist sie eindeutig.

Bemerkung: Gleichungssystem $Ax = b$ ist leicht lösbar, wenn $A = LR$ oder $PA = LR$ bekannt ist:

$$\begin{aligned} PAx &= Pb \\ L(Rx) &= Pb \end{aligned}$$

Schritt 1: Löse $Ly = Pb$ (Stichwort: Vorwärtseinsetzen)

Schritt 2: Löse $Rx = y$ (Stichwort: Rückwärtseinsetzen)

Gaußelimination ist LR – Zerlegung und Vorwärts- und Rückwärtseinsetzen mit Vertauschen der Reihenfolge der Rechenoperationen.

Komplexität (Anzahl der Multiplikationen / Divisionen und der Additionen)

LR: Berechnung von

$$l_{jk} := \frac{a_{jk}^{(k)}}{a_{kk}^{(k)}} \quad j = k+1, \dots, n$$

$$a_{jl}^{(k+1)} = a_{jl}^{(k)} - l_{jk} a_{kl}^{(k)} \quad j = k+1, \dots, n; \quad l = k+1, \dots, n$$

Berechnung von $A^{(k+1)}$ erfordert daher $\frac{1}{3}n^3 + O(n^2)$ Multiplikationen / Divisionen und genauso viele Additionen.

Rückwärts- bzw. Vorwärtseinsetzen:

Es ergibt sich insgesamt: n^2 Multiplikationen / Divisionen und $n^2 - n$ Additionen

4.2.4. Definition: Eine Matrix $A \in \mathbf{K}^{n \times n}$ hat obere Bandbreite k und untere Bandbreite m , wenn ihre Elemente a_{ij} für Indizes i, j mit $j - i > k$ und $i - j > m$ alle Null sind.

4.2.4. Satz: $A \in \mathbf{A}^{n \times n}$ habe Zerlegung $A = LR$. Hat A obere Bandbreite k und untere Bandbreite m , dann auch L und R (entsprechend!)

4.3. Die Cholesky – Zerlegung

4.3.1. Definition: Eine symmetrische Matrix $A \in \mathbf{R}^{n \times n}$ mit der Eigenschaft

$$x^T A x > 0 \quad \text{für } x \neq 0$$

heißt positiv definite Matrix.

4.3.1. Satz: Sei $A \in \mathbf{R}^{n \times n}$ symmetrisch. Dann sind äquivalent:

- (i) A ist positiv definit
- (ii) Alle Eigenwerte von A sind größer Null
- Ist A pos. definit, dann gilt außerdem $\det(A) > 0$.

Beispiel

4.3.2. Satz: Sei A positiv definit. Dann

- (i) Alle Hauptuntermatrizen A_k sind positiv definit
- (ii) $a_{ii} a_{jj} > a_{ij}^2$ für $i, j = 1, \dots, n, i \neq j$
- (iii) $a_{ii} > 0$ für $i = 1, \dots, n$

Folge: $\det(A_k) > 0 \quad \forall k \Rightarrow A$ hat LR-Zerlegung.

4.3.3. Satz (Cholesky – Zerlegung)

Jede positiv definite Matrix $A \in \mathbf{R}^{n \times n}$ kann eindeutig zerlegt werden in die

Form $A = G G^T$ mit einer unteren Dreiecksmatrix G mit positiven Diagonalelementen.

Bemerkung: Aus $A = G G^T$ folgt $(G := L \cdot \text{diag}(\sqrt{r_{11}}, \dots, \sqrt{r_{mm}}))$

$$a_{ki} = a_{ik} = \sum_{l=1}^n g_{il} g_{kl} \quad 1 \leq i \leq k \leq n$$

Ferner gilt: $g_{il} = 0$ für $l > i$ und damit

$$a_{ik} = \sum_{l=1}^i g_{il} g_{kl} \quad \text{und} \quad a_{ii} = g_{i1}^2 + \dots + g_{ii}^2.$$

Verfahren von Cholesky Crout

Für $i = 1, \dots, n$ berechne $g_{ii} = \sqrt{a_{ii} - \sum_{l=1}^{i-1} g_{il}^2}$

Für $k = i+1, \dots, n$ berechne $g_{ki} = \frac{1}{g_{ii}} \left(a_{ik} - \sum_{l=1}^{i-1} g_{il} g_{kl} \right)$

Satz 4.3.4: Bei der Berechnung der Cholesky Zerlegung einer positiv definiten $n \times n$ Matrix sind

$$\frac{n(n-1)(n+4)}{6} \quad \text{Multiplikationen / Divisionen}$$

$$\frac{n(n+1)(n-1)}{6} \quad \text{Additionen}$$

und n Wurzeln erforderlich.

4.4. Das Gauß – Jordan – Verfahren

Mit Hilfe dieses Verfahrens kann man nun eine Matrix invertieren bzw. überprüfen, ob sie invertierbar ist.

Algorithmus von Gauß – Jordan:

Gegeben: $A = A^{(1)} \in \mathbf{K}^{n \times n}$, invertierbar

Gesucht: A^{-1}

Für $l = 1, \dots, n$ bilde $A^{(l)}$ nach folgenden Regeln:

1) Pivot – Suche:

Bestimme Zeilindex $p_i \in \{l+1, \dots, n\}$ so daß

$$\left| a_{p_i, l+1}^{(l)} \right| = \max_{l < k \leq n} \left| a_{k, l+1}^{(l)} \right|$$

2) Vertausche $l+1$ -te mit p_i -ter Zeile von $A^{(k)}$,

Das gibt eine Matrix $\tilde{A}^{(l)}$ und Komponenten \tilde{a}_{ij}

3) Austauschschritt:

a) Spaltenregel

$$a_{i, l+1}^{(l+1)} := \frac{\tilde{a}_{i, l+1}^{(l)}}{\tilde{a}_{l+1, l+1}^{(l)}} \quad i = 1, \dots, n; i \neq l+1$$

b) Zeilenregel:

$$a_{l+1,k}^{(l+1)} := \frac{\tilde{a}_{l+1,k}^{(l)}}{\tilde{a}_{l+1,l+1}^{(l)}} \quad k = 1, \dots, n; k \neq l+1$$

c) Pivotregel:

$$a_{l+1,l+1}^{(l+1)} := \frac{1}{\tilde{a}_{l+1,l+1}^{(l)}}$$

d) Rechenregel:

$$a_{ik}^{(l+1)} = a_{ik}^{(l)} - \frac{\tilde{a}_{i,l+1}^{(l)} \cdot \tilde{a}_{l+1,k}^{(l)}}{\tilde{a}_{l+1,l+1}^{(l)}} \quad i, k = 1, \dots, n, i, k \neq l+1$$

$$A^{-1} = A^{(n+1)} \cdot P_{P_{n-1}, n-1} \cdots P_{P_1, 1}$$

Variante: Gauß – Jordan – Tableau:

x_1	x_2	x_3	...	x_n	
a_{11}	a_{12}	a_{13}		a_{1n}	y_1
a_{21}	a_{22}	a_{23}		a_{2n}	y_2
...			
a_{n1}	a_{n2}	a_{n3}	...	a_{nn}	y_n

x_1	x_2	y_2	...	x_n	
a'_{11}	a'_{12}	a'_{13}		a'_{1n}	y_1
a'_{21}	a'_{22}	a'_{23}		a'_{2n}	x_3
...			
a'_{n1}	a'_{n2}	a'_{n3}	...	a'_{nn}	y_n

usw., die a'_{ij} werden dann durch die Pivot-, Spalten-, Zeilen- und Rechteckregel bestimmt.

Gesamtaufwand Gauß – Jordan:

n mal Austauschschritt $\Rightarrow n^3$ Multiplikationen / Divisionen

$n(n-1)^2$ Additionen

Bemerkung: Gauß – Jordan nur anwenden, wenn wirklich A^{-1} erforderlich ist. Im allgemeinen reicht LR Zerlegung! (Grund: hoher Aufwand)

4.5. Matrixnormen

Definition 4.5.1: Eine Abbildung $N: \mathbf{C}^{n \times n} \rightarrow \mathbf{R}_+$? $[0, \infty[$ heißt Matrixnorm, wenn gilt:

- i) $N(A) = 0 \Leftrightarrow A = 0 \in \mathbf{C}^{n \times n}$
- ii) $\lambda \in \mathbf{C}, A \in \mathbf{C}^{n \times n} \Rightarrow N(\lambda A) = |\lambda| N(A)$
- iii) $A, B \in \mathbf{C}^{n \times n} \Rightarrow N(A+B) \leq N(A) + N(B)$
- iv) $A, B \in \mathbf{C}^{n \times n} \Rightarrow N(A \cdot B) \leq N(A) \cdot N(B)$ (Submultiplikativität)

Beispiele:

i) Zeilensummennorm:

$$A \rightarrow \|A\|_{\infty} = \max_{i=1}^n \sum_{k=1}^n |a_{ik}|$$

ii) Spaltensummennorm:

$$A \rightarrow \|A\|_1 = \max_{k=1}^n \sum_{i=1}^n |a_{ki}|$$

iii) Frobeniusnorm:

$$A \rightarrow \|A\|_F = \sqrt{\sum_{i=1}^n \sum_{k=1}^n |a_{ik}|^2}$$

Satz 4.5.1. Ist $\|\cdot\|: \mathbf{C}^n \rightarrow \mathbf{R}$ eine Vektornorm, dann erhält man durch

$$\text{lub}(A) := \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

eine Matrixnorm auf $\mathbf{C}^{n \times n}$. Sie heißt die zu $\|\cdot\|$ gehörende lub – Norm. (least upper bound)

Bemerkung: $\|Ax\| \leq \text{lub}(A) \|x\|$

wenn $\|Ax\| \leq K(A) \|x\|$, dann $\text{lub}(A) \leq K(A)$

Definition 4.5.2: Sei $A \in \mathbf{C}^{n \times n}$ und $\lambda_1, \dots, \lambda_n \in \mathbf{C}$ die Eigenwerte von A. Dann ist

$$\rho(A) := \max_{i=1}^n |\lambda_i|$$

der Spektralradius von A.

Satz 4.5.2: $\text{lub}_2(A) = \sqrt{\rho(A^H A)} = \sqrt{\max_{\|x\|_2=1} \|A^H A x\|_2}$

$$A = (a_{ik}) \Rightarrow A^H = (\bar{a}_{ki}).$$

Bemerkung: $\|A\|_2 = \text{lub}_2(A)$ heißt Spektralnorm

$(A^H A)^H = A^H A$ und damit sind Eigenwerte ≥ 0

Satz 4.5.3. Es gilt:

$$\text{i) } \text{lub}_{\infty}(A) := \max_{\|x\|_{\infty}=1} \|Ax\|_{\infty} = \max_{i=1}^n \sum_{k=1}^n |a_{ik}|$$

$$\text{ii) } \text{lub}_1(A) := \max_{\|x\|_1=1} \|Ax\|_{\infty} = \max_{k=1}^n \sum_{i=1}^n |a_{ki}|$$

Satz 4.5.4. Für die Frobenius Norm gilt:

$$\text{lub}_2(A) \leq \|A\|_F \leq \sqrt{\text{Rg}(A)} \cdot \text{lub}_2(A)$$

Definition 4.5.3. $N: \mathbf{K}^{n \times n} \rightarrow \mathbf{R}_+$ sei Matrixnorm. N heißt passend zur Vektornorm

$\|\cdot\|: \mathbf{K}^n \rightarrow \mathbf{R}$, wenn gilt:

$$\|Ax\| \leq N(A) \|x\| \quad \forall A \in \mathbf{K}^{n \times n}, x \in \mathbf{K}^n$$

Bemerkung: $\text{lub}_{\|\cdot\|}$ ist passend zu $\|\cdot\|$.

4.6. Fehlerabschätzungen

Gleichungssystem $Ax = b$.

Störung von b : $A(x + \Delta_1 x) = b + \Delta b$ $\|\Delta_1 x\| \leq ?$

Störung von A : $(A + \Delta A)(x + \Delta_2 x) = b$ $\|\Delta_2 x\| \leq ?$

Störung von A und b : $(A + \Delta A)(x + \Delta_3 x) = b + \Delta b$ $\|\Delta_3 x\| \leq ?$

Definition 4.6.1.: Für eine reguläre Matrix $A \in \mathbf{K}^{n \times n}$ und eine Matrixnorm

$N: \mathbf{K}^{n \times n} \rightarrow \mathbf{R}_+$ heißt

$$\text{cond}_N(A) := N(A) N(A^{-1})$$

die Kondition von A bezüglich N .

Bemerkung: $N(E) \geq 1$ und damit $1 \leq \text{cond}(A)$

Satz 4.6.1: Sei $A \in \mathbf{K}^{n \times n}$ und $0 \neq b \in \mathbf{K}^n$, $x + \Delta x$ sei Lösung von $A(x + \Delta x) = b + \Delta b$ und $Ax = b$. N sei eine zu $\|\cdot\|$ passende Matrixnorm. Dann gilt

$$\|\Delta x\| \leq N(A^{-1}) \|\Delta b\| \quad \text{absoluter Fehler}$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \cdot \frac{\|\Delta b\|}{\|b\|}$$

Lemma 4.6.2. Sei $A \in \mathbf{K}^{n \times n}$. Falls Matrixnorm N existiert mit $N(A) < 1$, dann existiert auch $(E + A)^{-1}$ und es gibt

$$N((E + A)^{-1}) \leq \frac{1}{1 - N(A)}$$

Satz 4.6.3. Sei $A, \Delta A \in \mathbf{K}^{n \times n}$, $b \in \mathbf{K}^{n \times n}$, A regulär, $Ax = b$ und

$$(A + \Delta A)(x + \Delta x) = b.$$

Sei $\|\cdot\|$ Norm auf \mathbf{K}^n , N passende Matrixnorm und

$$N(A^{-1}) N(\Delta A) < 1$$

Dann gilt:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \text{cond}(A) \frac{N(\Delta A)}{N(A)}} \frac{N(\Delta A)}{N(A)}$$

Bemerkung: Ganz allgemein (Störung mit ΔA und Δb).

$$(A + \Delta A)(x + \Delta x) = b + \Delta b \quad Ax = b$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \text{cond}(A) \frac{N(\Delta A)}{N(A)}} \left(\frac{N(\Delta A)}{N(A)} + \frac{\|\Delta b\|}{\|b\|} \right)$$

Satz 4.6.4: Ist $A \in \mathbf{K}^{n \times n}$ Zeilenäquivalent

$$\exists c \in \mathbf{R}_+: \sum_{j=1}^n |a_{ij}| = c, \text{ dann gilt } i = 1, \dots, n$$

dann gilt für jede Diagonalmatrix D

$$\text{cond}_\infty(A) \leq \text{cond}_\infty(DA).$$

4.7. Die QR-Zerlegung

Bei LR Zerlegung bzw. Gaußelimination

$$A_{\text{neu}} = L^{-1} A_{\text{alt}} \quad (\text{z.B. } A_{\text{neu}} = A^{(k+1)} \quad A_{\text{alt}} = A^{(k)})$$

$$A_{\text{alt}} = A \quad A_{\text{neu}} = R$$

$$\Rightarrow \text{cond}(A_{\text{neu}}) = N(A_{\text{neu}}) N(A_{\text{neu}}^{-1}) \leq \text{cond}(L) \text{cond}(A_{\text{alt}})$$

Im allgem. $\text{cond}(L) > 1$ und Konditionszahlverstärkung um den Faktor $\text{cond}(L)$ tritt oft auf.

Bei unitärem Q gilt: $\|Q\|_2 = 1$

Satz 4.7.1. $A \in \mathbb{C}^{n \times n}$, Q unitär $\Rightarrow \text{cond}_2(QA) = \text{cond}_2(A)$.

Satz 4.7.2. Jede reguläre Matrix $A \in \mathbb{C}^{n \times n}$ läßt sich schreiben als

$$A = QR \quad (\text{QR Zerlegung})$$

Q unitär, R obere Δ -Matrix mit pos. Diagonale

Satz 4.7.3: Die QR-Zerlegung von einer regulären Matrix A mit $A = QR$, Q unitär, R obere Δ -Matrix ist eindeutig, wenn $r_{kk} > 0$ ($k = 1, \dots, n$)

stabile Methode von Householder

Definition 4.7.1. Sei $w \in \mathbb{C}^n$, $\|w\|_2 = 1$. Dann heißt

$$H_w = E - 2ww^H = \begin{pmatrix} 1 - 2|w_1|^2 & -2w_1\overline{w_2} & & \\ -2w_2\overline{w_1} & 1 - 2|w_2|^2 & & \\ & & \ddots & \\ & & & 1 - 2|w_n|^2 \end{pmatrix}$$

elementare hermitesche Matrix bzw. Householdermatrix.

Die Abbildung

$$\phi_n: \mathbb{C}^n \rightarrow \mathbb{C}^n:$$

$$z \rightarrow H_w z \text{ heißt}$$

Householder-Transformation.

Satz 4.7.4.: Es gelten die Aussagen

- i) H_w ist hermitesch und unitär
- ii) H_w ist involutiv ($H_w^2 H = E$)
- iii) H_w hat Eigenwert -1 und EV w und $(n-1)$ fachen Eigenwert 1 mit Eigenraum $V = \{u \in \mathbb{C}^n \mid w^H u = 0\}$
- iv) Sei $w = (0, w_2, \dots, w_n)^T$, $\tilde{w} = (w_2, \dots, w_n)^T$

$$\text{Dann gilt } Hw = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & H\tilde{w} & & \\ 0 & & & \end{pmatrix}.$$

Bemerkung: Householder-Trafo ist Spiegelung an Hyperebene $V = \{u_1, \dots, u_{n-1}\}$

Frage: Gibt es zu $a, b \in \mathbb{C}^n$ ein $w \in \mathbb{C}^n$, $\|w\|_2 = 1$ mit $H_w a = b$?

Antwort: Householder-Algorithmus

$$a \in \mathbb{C}^n, a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \text{ gesucht ist ein } \alpha \in \mathbb{C}, w \in \mathbb{C}^n: w^H w = 1 \text{ und}$$

$$H_w a = \alpha e_1.$$

$$\Rightarrow \alpha = \begin{cases} -\|a\|_2 \frac{a_1}{|a_1|}, & a_1 \neq 0 \\ -\|a\|_2, & a_1 = 0 \end{cases}$$

$$\beta := \|a - \alpha e_1\|_2 = \sqrt{(|a_1| + \|a\|_2)^2 + |a_2|^2 + \dots + |a_n|^2} \text{ und}$$

$$w = \frac{1}{\beta} (a - \alpha e_1)$$

Satz 4.7.5. Zu einer beliebigen Matrix $A \in \mathbb{C}^{m \times n}$ eine unitäre Matrix $Q \in \mathbb{C}^{m \times m}$ und eine obere Dreiecksmatrix $R \in \mathbb{C}^{m \times n}$ mit $A = QR$.

Verfahren: Sei H_{w_i} eine Householdermatrix, die die i -te Spalte gemäß des obigen Algorithmus umformt. Dann ist $R = H_{w_s} \cdots H_{w_1} A$ und

$$Q = H_{w_1} \cdots H_{w_s}$$

Algorithmus stabiler und allgemeiner als Schmidtsches Orthogonalisierungsverfahren!

4.8. Lineare Ausgleichsprobleme

Methode der kleinsten Quadrate (von Gauß)

Meßpunkte x_i

Meßwerte y_i

Aus physikalisch oder ökonom. Gründen muß $y_i = \alpha + \beta x_i$ gelten:

$$\begin{aligned} r_i &= \alpha + \beta x_i - y_i && \text{Residuum} \\ r &= (r_1, \dots, r_m)^T && \text{Residuenvektor} \end{aligned}$$

Gesucht: α^*, β^* mit $\min_{\alpha, \beta} \sum_{i=1}^m (\alpha + \beta x_i - y_i)^2 = \sum_{i=1}^m (\alpha^* + \beta^* x_i - y_i)^2$

Methode der kleinsten Quadrate (least – square – Problem)

Matrixform:
$$r = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

gesucht: $\min_{(\alpha, \beta) \in \mathbb{R}^2} r^T r = \min_{(\alpha, \beta) \in \mathbb{R}^2} \|r\|_2^2$

Allgemeiner: gesucht ist $\min_{(\alpha, \beta) \in \mathbb{R}^2} \|Ax - b\|_2^2$ lineares Ausgleichsprobl.

Verallgemeinerung der Lösung linearer Gleichungssysteme auf „Lösung“ von überbestimmten und unlösbaren Problemen.

Im folgenden Beschränkung auf \mathbb{R}^n .

Im \mathbb{R}^n gilt: $x, y \in \mathbb{R}^n \Rightarrow \langle x, y \rangle := \sum_{i=1}^n x_i y_i$ und $\|x\|_2 = \sqrt{\langle x, x \rangle}$ d.h.

euklidische Norm stammt ab vom inneren Produkt.

Satz 4.8.1. Sei V ein Vektorraum mit inneren Produkt $U_n \subseteq V$ sei $n - \dim$

Untervektorraum von V . V hat Norm $\|x\| = \sqrt{\langle x, x \rangle}$.

Zu jedem $f \in V$ gibt es genau ein $h^* \in U_n$:

(*) $\langle f - h^*, h \rangle = 0 \quad \forall h \in U_n$

(**) h^* eindeutig bestimmt.

Bemerkung: Wegen (*) heißt h^* orthogonale Projektion von f von U_n . Wegen (**) heißt h^* die beste Approximation von f in U_n .

Satz 4.8.2. Sei $A \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^m$ mit $m > n$ gegeben. Betrachte das lin.

Ausgleichsproblem:

(#) $\|Ax^* - b\|_2^2 = \min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$.

Dann gilt:

- i) x^* löst (#) genau dann, wenn es die Normalgleichung $A^t A x^* = A^t b$ erfüllt.
- ii) Die Menge $L := \{x \in \mathbf{R}^n \mid A^t A x = A^t b\}$ ist ein nichtleerer affiner Raum, d.h. $x_1, x_2 \in L \Rightarrow (1 - \lambda) x_1 + \lambda x_2 \in L \forall \lambda \in \mathbf{R}$.
- iii) (#) hat genau dann eine eindeutige Lösung, wenn $\text{Rg}(A) = n$ (vollster Spaltenrang)
- iv) Unter allen Lösungen # gibt es genau eine Lösung x^+ mit minimaler euklidischer Norm.

Bemerkung: x^+ heißt Pseudonormlösung von # bzw. von $Ax = b$.

Bemerkung: Ist die QR – Zerlegung von A bekannt, $A = QR$, Q orthogonal.

Definition 4.8.1: Sei $A \in \mathbf{R}^{m \times n}$. Die Darstellung

$$A = U \Sigma V^t$$

mit orthogonalen Matrizen $U \in \mathbf{R}^{m \times m}$, $V \in \mathbf{R}^{n \times n}$ und Diagonalmatrix $\Sigma \in \mathbf{R}^{m \times n}$ Singulärwertzerlegung von A.

Lemma 4.8.3. Sei $A \in \mathbf{R}^{m \times n}$. Die positiven Eigenwerte von $A^t A$ und AA^t stimmen überein und haben dieselben Vielfachheit, d.h.

$$\dim \text{Kern}(\lambda E_n - A^t A) = \dim \text{Kern}(\lambda E_m - AA^t).$$

Bemerkung: AA^t und $A^t A$ sind reell symmetrisch \Rightarrow alle Eigenwerte reell.

Lemma 4.8.4. Für $A \in \mathbf{R}^{m \times n}$ ist $\text{Rg}(A) = \text{Rg}(A^t) = \text{Rg}(AA^t) = \text{Rg}(A^t A)$ die Anzahl der positiven Eigenwerte von $A^t A$ bzw. AA^t mit Vielfachheit $\dim \text{Kern}(\lambda E_n - A^t A)$ gezählt.

Satz 4.8.5.: Sei $A \in \mathbf{R}^{m \times n}$ mit $r = \text{Rg}(A)$. Dann existieren orthogonale Matrizen

$$U = (U_1, \dots, U_m) \in \mathbf{R}^{m \times m}, V = (V_1, \dots, V_n) \in \mathbf{R}^{n \times n} \text{ derart, daß}$$

$$\Sigma := U^t A V = \text{diag}(\sigma_1, \dots, \sigma_{\min(m,n)}) \in \mathbf{R}^{m \times n} \text{ mit}$$

$$\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_{\min(m,n)} = 0.$$

Bemerkung: Σ hat die Gestalt:

$$\Sigma = \begin{pmatrix} \sigma_1 & & 0 & \dots & 0 \\ & \ddots & \vdots & & \vdots \\ & & \sigma_r & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbf{R}^{m \times n}.$$

Definition 4.8.2: Die positiven Eigenwerte von AA^t (und $A^t A$) sind eindeutig bestimmt. Damit sind auch $\sigma_1, \dots, \sigma_r$ sind eindeutig durch A bestimmt. Sie heißen Singulärwerte von A. (Nur U und V sind nicht eindeutig)

Man kann auch mit Hilfe der Singulärwertzerlegung das Minimalwertproblem aus Satz 4.8.2 bestimmen:

$$\text{Es ergibt sich: } V^T x = \left(\frac{u_1^T b}{\sigma}, \dots, \frac{u_r^T b}{\sigma}, a_{r+1}, \dots, a_n \right)^T. \text{ (} a_i \text{ beliebig).}$$

Unter allen Lösungen hat x^+ minimale euklidische Norm (Setze $a_i = 0$)

Definition 4.8.3: Zu $A \in \mathbf{R}^{m \times n}$ und $\text{Rg}(A) = r$ sei $U^t A V = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_{\min(m,n)}, 0, \dots, 0)$ eine nach Satz 4.8.5 existierende Singulärwertzerlegung mit

$$\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_{\min(m,n)}.$$

Mit der Matrix $\Sigma^+ = \begin{pmatrix} \frac{1}{\sigma_1} & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \frac{1}{\sigma_r} & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbf{R}^{m \times n}$

heißt $A^+ = V \Sigma^+ U^T \in \mathbf{R}^{n \times m}$ Pseudoinverse (bzw. Moore Penrose – Inverse).

Satz 4.8.6: Sei $A \in \mathbf{R}^{m \times n}$, $r = \text{Rg}(A)$.

Dann gilt:

i) Die eindeutig bestimmte Lösung x^* minimaler Norm des Ausgleichsproblems $\min_{x \in \mathbf{R}^n} \|Ax - b\|_2^2$ ist gegeben durch $x^+ = A^+b$.

ii) A^+ ist eindeutig bestimmt durch A bestimmt

iii) $\text{Rg}(A) = n \Rightarrow A^+ = (A^T A)^{-1} A^T$.

iv) $m = n$, A regulär $\Rightarrow A^+ = A^{-1}$

v) Es gelten die vier Moore – Penrose Bedingungen:

$$AA^+ = (AA^+)^T$$

$$A^+A = (A^+A)^T$$

$$AA^+A = A$$

$$A^+AA^+ = A^+$$

vi) Erfüllt B die vier Moore Penrose Bedingung, dann ist $B = A^+$.

Fazit: Pseudoinverse kann auf 3 Weisen erklärt werden:

i) Matrix zur Abbildung $b \rightarrow x^+$

ii) $A^+ = V \Sigma^+ U^T$

iii) Durch die Moore – Penrose – Bedingungen.

Satz 4.8.7.: Sei $A \in \mathbf{R}^{m \times n}$ mit Pseudoinverser $A^+ \in \mathbf{R}^{n \times m}$. Dann hat A^T die Pseudoinverse $(A^+)^T$: kurz: $(A^T)^+ = (A^+)^T$.

Satz 4.8.8 Sei $A \in \mathbf{R}^{m \times m}$. Dann ist $x \rightarrow AA^T x$ bzw. $x \rightarrow A^+ A x$ die orthogonale Projektion auf $\text{Bild}(A)$ bzw. $\text{Bild}(A^+)$.

§ 5 Iterative Lösungen linearer Gleichungen

5.1. Das Gesamt- und Einzelschrittverfahren

Betrachte spezielles Gleichungssystem:

$$\begin{aligned} \text{Fixpunktgleichung: } & x = Tx + u \quad T \in \mathbf{K}^{n \times n}, u \in \mathbf{K}^n \\ \text{Iterationsvorschrift: } & x^{(k+1)} = T x^{(k)} + u \quad k = 0, 1, 2, \dots \\ & x^{(0)} \in \mathbf{K}^n \\ & \mathbf{K}^n \rightarrow \mathbf{K}^n \\ & \phi: x \rightarrow T x + u \end{aligned}$$

ϕ ist kontrahierend, wenn für eine Matrix – Vektornorm Paar (passendes) $N, \|\cdot\|$ gilt:

$$\|\phi(x) - \phi(y)\| = \|T(x-y)\| \leq N(T) \|x-y\| \quad N(T) < 1$$

Damit gilt: Banachsches – Fixpunktproblem: Konvergenz für beliebige $x^{(0)} \in \mathbf{K}^n$ gegen Fixpunkt.

Problem: Wie kriegt man $Ax = b$ auf Fixpunktform?

1. Methode: $A = D - B$, mit $D := \text{diag}(a_{11}, \dots, a_{nn})$ und $B := D - A$.

Wenn $a_{kk} \neq 0$, für $k = 1, \dots, n$.

$$\begin{aligned} Ax = b & \Leftrightarrow x = D^{-1}Bx + D^{-1}b, \\ \text{d.h. } T &= D^{-1}B \quad \text{und } u = D^{-1}b. \end{aligned}$$

2. Methode: $A = D - L - U$ mit $D := \text{diag}(a_{11}, \dots, a_{nn})$ und

L strikte untere – Dreiecksmatrix

U strikte obere – Dreiecksmatrix.

Wenn $a_{kk} \neq 0$, für $k = 1, \dots, n$.

$$\begin{aligned} Ax = b & \Leftrightarrow x = (D - L)^{-1} Ux + (D - L)^{-1}b, \\ \text{d.h. } T &= (D - L)^{-1} U, \quad u = (D - L)^{-1}b \end{aligned}$$

Verfahren:

Gesamtschrittverfahren (GSV, Jacobi – Verfahren)

$A = D - B$, $D := \text{diag}(a_{11}, \dots, a_{nn})$, $a_{kk} \neq 0, \forall k = 1, \dots, n$

Iterationsvorschrift: (kurz) $x^{(k+1)} = D^{-1}Bx^{(k)} + D^{-1}b$.

ausführlich:

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{a_{11}} (0 - a_{12}x_2^{(k)} - \dots - a_{1n}x_n^{(k)} + b_1) \\ x_2^{(k+1)} &= \frac{1}{a_{22}} (a_{21}x_1^{(k)} - 0 - \dots - a_{2n}x_n^{(k)} + b_2) \\ &\dots \\ x_n^{(k+1)} &= \frac{1}{a_{nn}} (a_{n1}x_1^{(k)} - \dots - 0 + b_n) \end{aligned}$$

Iterationsmatrix:

$$T_{GSV} = D^{-1}B \text{ (Jacobi - Matrix)}$$

Einzelschrittverfahren (ESV, Gauß – Seidel – Verfahren)

$A = D - L - U$, D Diagonalmatrix, L und U strikte untere bzw. obere Dreiecksmatrix.

$x^{(0)} \in \mathbf{K}^n$

Iterationsvorschrift (kurz): $x^{(k+1)} = (D - L)^{-1}U x^{(k)} + (D - L)^{-1}b$

ausführlich:

$$x_1^{(k+1)} = \frac{1}{a_{11}} (0 - a_{12}x_2^{(k)} - \dots - a_{1n}x_n^{(k)} + b_1)$$

$$x_2^{(k+1)} = \frac{1}{a_{22}} (a_{21}x_1^{(k+1)} - 0 - \dots - a_{2n}x_n^{(k)} + b_2)$$

....

$$x_n^{(k+1)} = \frac{1}{a_{nn}} (a_{n1}x_1^{(k+1)} - \dots - a_{n,n-1}x_{n-1}^{(k+1)} - 0 + b_n)$$

(Achtung: man muß nicht $(D - L)$ invertieren !!!!)

Iterationsmatrix:

$$T_{ESV} = (D - L)^{-1} U$$

Wann konvergieren diese Verfahren ?

5.2. Konvergenz von Iterationsverfahren für lineare Gleichungssysteme

Für den Fehler $e^{(k)}$ gilt:

$$x = Tx + u$$

$$x^{(k+1)} = Tx^{(k)} + u \quad \text{wobei, } e^{(k)} := x^{(k)} - x. \Rightarrow e^{(k+1)} = Te^{(k)} = T^k e^{(0)}$$

Konvergenz tritt ein, wenn $\lim_{k \rightarrow \infty} T^k = 0$, d.h. $\lim_{k \rightarrow \infty} T_{ij}^k = 0 \quad \forall i, j$.

Satz 5.2.1: Das Iterationsverfahren $x^{(k+1)} = Tx^{(k)}$ konvergiert für beliebige Startwerte $x^{(0)} \in \mathbf{K}^n$ genau dann gegen die Lösung von $x = Tx + u$, wenn $\rho(T) < 1$.

Lemma 5.2.2.: $A \in \mathbf{K}^{n \times n}$, N Matirxnorm $\Rightarrow \rho(A) \leq N(A)$.

Satz 5.2.3. Gilt: $|a_{ii}| > \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}| \quad i = 1, \dots, n$, dann konvergiert das GSV bei beliebigem

Startvektor $x^{(0)}$.

Bemerkung: $\|T\|$ gibt das Maß für die Geschwindigkeit mit der der Fehler gegen Null geht (falls $\|T\| < 1$)

Satz 5.2.4. (Kriterium von Sassenfeld)

Sei $A = E - L - U$ (E jetzt Einheitsmatrix)

Es gelte: $1 > \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}| \quad i = 1, \dots, n$, dann gilt:

$$\|(E - L)^{-1} U\|_{\infty} \leq \|L + U\|_{\infty} < 1, \text{ d.h. } \|T_{ESV}\| \leq \|T_{GSV}\| < 1$$

Satz 5.2.5 (Reich – Ostrowski)

Ist $A \in \mathbf{C}^{n \times n}$ hermitesch und positiv definit, dann konvergiert das ESV bei jedem Startwert $x^{(0)} \in \mathbf{C}^n$.

Bemerkung: Das GSV konvergiert für positiv definite hermitesche Matrizen nicht unbedingt.

5.3. Relaxation und Nachiteration

5.3.1 Algorithmus (GSV mit Relaxation, JOR – Verfahren)

Gegeben: $A \in \mathbf{K}^{n \times n}$, $D = \text{diag}(a_{11}, \dots, a_{nn})$, $b \in \mathbf{K}^n$, Startwert $x^{(0)} \in \mathbf{K}^n$.

Iteration: $x^{(k+1)} = [(1-\omega)E + \omega D^{-1} B] x^{(k)} + \omega D^{-1} b$

Komponentenweise:

$$x_1^{(k+1)} = (1-\omega) x_1^{(k)} - \frac{\omega}{a_{11}} a_{12} x_2^{(k)} - \dots - \frac{\omega}{a_{1n}} a_{1n} x_n^{(k)} + \frac{\omega}{a_{11}} b_1$$

....

$$x_n^{(k+1)} = - \frac{\omega}{a_{nn}} a_{n1} x_1^{(k)} - \dots - \frac{\omega}{a_{nn}} a_{n(n-1)} x_{n-1}^{(k)} + (1-\omega) x_n^{(k)} + \frac{\omega}{a_{nn}} b_n$$

$T_{\text{GSV}}(\omega) = (1-\omega)E + \omega D^{-1}B$ (JOR – Matrix)

Satz 5.3.1.: $T_{\text{GSV}} = T_{\text{GSV}}(1)$ haben die Eigenwerte $\lambda_1, \dots, \lambda_n$ und zugehörige Eigenvektoren $x^{(1)}, \dots, x^{(n)}$. Dann hat $T_{\text{GSV}}(\omega)$ die Eigenwerte $\mu_i(\omega) = 1 - \omega + \omega \lambda_i$ mit zugehörigen Eigenvektoren $x^{(i)}$, $i = 1, \dots, n$.

Satz 5.3.2.: Die Iterationsmatrix des GSV besitze die reellen Eigenwerte

$$-1 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n < 1.$$

Dann wird der Spektralradius des JOR – Verfahrens minimal für

$$\omega^* = \frac{2}{2 - \lambda_1 - \lambda_n} \quad \text{mit } \rho(T_{\text{GSV}}(\omega)) = \frac{\lambda_n - \lambda_1}{2 - \lambda_1 - \lambda_n}.$$

5.3.2. Algorithmus: (ESV mit Relaxation, SOR – Verfahren)

Gegeben: $A = D - L - U$, $D = \text{diag}(a_{11}, \dots, a_{nn})$, L, U strikte untere bzw. obere Dreiecksmatrix, D regulär, $b \in \mathbf{K}^n$, $x^{(0)} \in \mathbf{K}^n$, Relaxationswert ω .

Iteration: (Komponentenweise)

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}} \left(- \sum_{j < i} (a_{ij} x_j^{(k+1)}) - \sum_{i < j} (a_{ij} x_j^{(k)}) + b_i \right)$$

$$x_i^{(k+1)} = (1-\omega) x_i^{(k)} + \omega \tilde{x}_i^{(k+1)}$$

mit $i = 1, \dots, n$.

Iterationsmatrix: $T_{\text{ESV}}(\omega) = (D - \omega L)^{-1} [(1-\omega)D + \omega U]$

Satz 5.3.3. (Kahau):

$$\rho(T_{\text{ESV}}(\omega)) \geq |1 - \omega|$$

Gleichheit nur dann, wenn alle EW von $T_{\text{ESV}}(\omega)$ betragsmäßig gleich $|1 - \omega|$ sind.

Folge: Konvergenz des SOR – Verfahrens höchstens für $\omega \in]0, 2[$.

Satz 5.3.4. (Reich – Ostrowski)

Ist $A \in \mathbf{R}^{n \times n}$ symm., dann gilt:

Das SOR – Verfahren konvergiert genau dann für alle $\omega \in]0, 2[$, wenn A positiv definit ist.

Nachiteration

Statt $A = LR$ mit numerischer Rechnung nun $A = \tilde{L}\tilde{R} + F$ mit Fehlermatrix F , $\|F\|_\infty$ sehr klein. Lösung \tilde{x} aus $\tilde{L}\tilde{R}\tilde{x} = b$.

Für reguläres A ist \tilde{R} (und \tilde{L}) regulär, wenn $\|F\|_\infty$ genügend klein.
Es ergibt sich die Iteration:

Iterationsvorschrift der Nachiteration:

$$x^{(k+1)} = x^{(k)} - \tilde{R}^{-1}\tilde{L}^{-1}(Ax^{(k)} - b)$$

Iterationsmatrix:

$$T_N = E - \tilde{R}^{-1}\tilde{L}^{-1}A$$

Weil $Ax^{(k)} \approx b$ gilt, tritt Stellenauslöschung bei $Ax^{(k)} - b$ ein. Daher Rechnung mit doppelter Stellenzahl, bei $r^{(k)} := Ax^{(k)} - b$.

Danach berechnet man Korrektur

$$z^{(k)} = \tilde{R}^{-1}\tilde{L}^{-1}r^{(k)} = x^{(k)} - x^{(k+1)}$$

(Stichwort: Vorwärts- und Rückwärtseinsetzen)

§ 6 Eigenwertprobleme

6.1. Grundbegriffe aus der Algebra

$A \in \mathbb{C}^{n \times n}$, $\lambda \in \mathbb{C}$ ist Eigenwert, wenn ein $x \in \mathbb{C}^n$, $x \neq 0$, ex. mit $Ax = \lambda x$. x heißt Eigenvektor von A zum Eigenwert λ .

$U := \{x \mid Ax = \lambda x\}$ heißt Eigenraum zum Eigenwert λ . Vielfachheit von λ ist die Dimensionalension von U .

U ist Untervektorraum von \mathbb{C}^n , der invariant unter A ist, d.h. $x \in U \Rightarrow Ax \in U$.

Die Eigenwerte sind Nullstellen des charakteristischen Polynoms $p(\lambda) = \det(A - \lambda E_n)$.

Gilt $Ax = \lambda x \Rightarrow T A T^{-1}Tx = \lambda Tx$, d.h. Tx ist EV zum EW λ .

Satz 6.1.1. (Komplexe Schur – Zerlegung)

Sei $A \in \mathbb{C}^{n \times n}$. Dann ex. eine unitäre Matrix $U \in \mathbb{C}^{n \times n}$ derart, daß

$$U^H A U = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Dabei sind $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A (nicht notwendig verschieden)

Satz 6.1.2. (reelle Schur – Zerlegung)

Sei $A \in \mathbb{R}^{n \times n}$. Dann ex. eine orthogonale Matrix $Q \in \mathbb{R}^{n \times n}$ mit

$$Q^T A Q = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1s} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & R_{ss} \end{pmatrix},$$

wobei die R_{kk} reelle 1×1 oder 2×2 Matrizen sind

$$R_{kk} = (\lambda_k) \in \mathbb{R}^{1 \times 1}$$

$$R_{kk} \in \mathbb{R}^{2 \times 2} \Rightarrow R_{kk} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \text{ wo } \alpha \pm i\beta \text{ EW von } A \text{ ist.}$$

Satz 6.1.3 (Eigenwertzerlegung hermitescher Matrizen)

Zu jeder hermiteschen Matrix aus $\mathbb{C}^{n \times n}$ gibt es eine unitäre Matrix $U \in \mathbb{C}^{n \times n}$ mit $U^H A U = D = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Bemerkung: Man nennt A diagonalisierbar wegen $U^H A U = D$. u_1, \dots, u_n ist eine ON Basis von Eigenvektoren von A .

$A = U D U^H$ heißt die Eigenwertzerlegung von A .

Definition 6.1.1.: Eine Matrix $\mathbb{C}^{n \times n}$ heißt normal, wenn $A A^H = A^H A$ gilt.

Satz 6.1.4. $A \in \mathbb{C}^{n \times n}$ ist genau dann normal, wenn es eine unitäre Matrix $U \in \mathbb{C}^{n \times n}$ gibt mit $U^H A U = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_k \in \mathbb{C}$, $k = 1, \dots, n$.

Bemerkung: Jede hermitesche Matrix ist normal. Unitäre Matrizen sind normal.

Folge: Normale Matrizen (und nur die) sind unitär diagonalisierbar. Sie besitzen n Eigenvektoren, die eine ON Basis des \mathbb{C}^n bilden.

6.2. Reduktion auf Tridiagonal- bzw. Hessenberggestalt

Definition 6.2.1: Eine Matrix $B = (b_{ij})_{i,j=1}^n \in \mathbb{C}^{n \times n}$ der Form

$$B = \begin{pmatrix} b_{11} & & \cdots & & b_{1n} \\ b_{21} & b_{22} & & & \vdots \\ 0 & \ddots & \ddots & & \\ 0 & 0 & b_{n,n-1} & b_{nn} & \end{pmatrix}, \text{ d.h. } b_{ij} = 0 \text{ f\u00fcr } i > j+1 \text{ (obere) Hessen-}$$

bergmatrix. B hei\u00dft nicht – zerfallen, falls $b_{j+1,j} \neq 0$ f\u00fcr $j = 1, \dots, n-1$.

Bemerkung: Gilt $B^H = B$, dann ist B tridiagonal.

Satz 6.2.1: Zu $A \in \mathbb{C}^{n \times n}$ ex. eine (obere) Hessenbergmatrix B und Householdermatrizen $H_1, \dots, H_{n-2} \in \mathbb{C}^{n \times n}$, so da\u00df $B = H_{n-2} \dots H_1 A H_1 \dots H_{n-2}$

Bemerkung: Beweis ist konstruktiv. Mann nennt das resultierende Verfahren: „Verfahren von Hyman“ oder „Reduktionsverfahren von Householder“
Wenn $A = A^H$, dann auch $B = B^H$.

F\u00fcr hermitesche Matrizen liefert Satz 6.2.1 eine Tridiagonalmatrix.

Lemma 6.2.2.: Sei B eine zerfallene Hessenbergmatrix $B = (b_{jk}) \in \mathbb{C}^{n \times n}$ $b_{l+1,l} = 0$ f\u00fcr ein $l \in \{1, \dots, n-1\}$.

Dann gilt: $\det(B - \lambda E_n) = \det(B_l - \lambda E_{n-l}) \det(C_{n-l} - \lambda E_{n-l})$
(Determinantenregel)

Es gilt (ab sofort B nicht – zerfallen)

$x \neq 0$

$$(B - \lambda E)x = 0$$

Damit kann man herleiten:

$$-q + (b_{11} - \lambda)x_1 + b_{12}x_2 + \dots + b_{1n}x_n = 0$$

$$b_{21}x_1 + (b_{22} - \lambda)x_2 + \dots + b_{2n}x_n = 0$$

....

$$b_{n,n-1}x_{n-1} + (b_{nn} - \lambda)x_n = 0$$

Wenn $q = q(\lambda) = 0$, dann ist λ EW von $(x_1, \dots, x_n)^T$.

Bei gegebenem x_n haben wir das Gleichungssystem f\u00fcr q, x_1, \dots, x_{n-1} mit Koeff. Matrix

$$\tilde{B}(\lambda) = \begin{pmatrix} -1 & b_{11} - \lambda & b_{12} & \cdots & b_{1,n-1} \\ 0 & b_{21} & & & \vdots \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & b_{n,n-1} \end{pmatrix}$$

$\det(\tilde{B}(\lambda)) = -b_{21} \dots b_{n,n-1} \neq 0$ (n. Voraussetzung).

Es gilt damit: $\tilde{B}(\lambda) \begin{pmatrix} q \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = x_n \begin{pmatrix} -b_{1n} \\ \vdots \\ \vdots \\ -(b_{nn} - \lambda) \end{pmatrix}$. Setze x_n (o.B.d.A als 1)

Es ergibt sich:

Algorithmus von Hyman

Gegeben: B obere Hessenbergmatrix, nicht zerfallen, Startwert $\lambda^{(0)}$.

Gesucht: Verbesserte Näherung $\lambda^{(1)}$ aus Eigenwert von A und Näherung an zugehörigen Eigenvektor.

Schritt 1: Löse $(B - \lambda^{(0)}E) x - q_1 e_1 = 0$

Schritt 2: Löse $(B - \lambda^{(0)}E) x' - q' e_1 = x \quad x'_n = 0$

Schritt 3: Setze $\lambda^{(1)} = \lambda^{(0)} - \frac{q}{q'}$

Schneller, falls B Tridiagonalmatrix ist:

$$B_n = \begin{pmatrix} a_1 & b_1 & & 0 \\ \bar{b}_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ 0 & & \bar{b}_{n-1} & a_n \end{pmatrix} \quad a_i \in \mathbf{R}, b_k \in \mathbf{C} \setminus \{0\}$$

Satz 6.2.3. Das char. Polynom $p_n(\lambda) = \det(B_n - \lambda E)$ läßt sich mit der Rekursion

$$p_0(\lambda) = 1, p_1(\lambda) = a_1 - \lambda, p_k(\lambda) = (a_k - \lambda) p_{k-1}(\lambda) - |b_{k-1}|^2 p_{k-2}(\lambda) \quad k = 2, \dots, n.$$

Bemerkung: Aus Rekursionsformel bekommt man durch Ableiten

$$p_0' = 0, p_1'(\lambda) = -\lambda,$$

$$p_n'(\lambda) = (a_n - \lambda) p'_{k-1}(\lambda) - |b_{k-1}|^2 p'_{k-2}(\lambda) - p_{k-1}(\lambda)$$

Satz 6.2.4. $B_n \in \mathbf{C}^{n \times n}$ sei eine nicht – zerfallende hermitesche Tridiagonalmatrix B_k ihre k – reihige Hauptmatrix. Dann besitzt p_k , das charakteristische Polynom von B_k , reelle einfache Nullstellen $\lambda_j^{(k)}, j = 1, \dots, k$ und für $k = 1, \dots, n-1$ trennen die Nullstellen von p_k die von p_{k+1} , d.h.

$$\lambda_1^{(k+1)} < \lambda_1^{(k)} < \lambda_2^{(k+1)} < \dots < \lambda_k^{(k)} < \lambda_{k+1}^{(k+1)}.$$

Satz 6.2.5. Seien p_0, p_1, \dots, p_n die zu einer hermiteschen Tridiagonalmatrix gehörende Rekursionsfolge gemäß Satz 6.2.4. Dann bilden

$$p_0, -p_1, \dots, (-1)^n p_n$$

auf allen Intervallen $[a, b]$ mit $p_n(a) \neq 0 \neq p_n(b)$ eine Sturmsche Kette.

Folge: Man kann mit dem Bisektionsverfahren etwa den k -ten Eigenwert der Matrix B berechnen. Andere Möglichkeit: Newton.

6.3. Die Potenzmethode

6.3.1 Algorithmus nach von Mises / Geiringer

Gegeben: $A \in \mathbf{C}^{n \times n}$ mit dominanten EW $\lambda_1, \varepsilon > 0$

Gesucht: Näherung $\lambda^{(k)}$ an λ_1 und $x^{(k)}$ an zugehörigen EV.

Setze $z^{(0)} := (1, \dots, 1)^T$ (oder ähnlich, aber $\neq 0$)

$$x^{(0)} := \frac{1}{\|z^{(0)}\|} z^{(0)}$$

Für $k = 1, 2, \dots$ setze $z^{(k)} = A x^{(k-1)}$

$$x^{(k)} = \frac{1}{\|z^{(k)}\|} z^{(k)}$$

$\lambda^{(k)}$ ist Mittelwert von $\left\{ \frac{z_i^{(k)}}{x_i^{(k-1)}} \mid x_i^{(k-1)} \neq 0 \right\}$

Abbruch sobald $|\lambda^{(k-1)} - \lambda^{(k)}| < \varepsilon$.

Dann ist $\lambda^{(k)}$ Näherung an λ_1 und $x^{(k)}$ Näherung an EV.

Satz 6.3.1: Sei $A \in \mathbb{C}^{n \times n}$ diagonalähnlich mit Eigenwerten

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

und zugehörige Eigenvektoren v_1, \dots, v_n (lin. unabh.) Ist $\lambda_1 = |\lambda_1| e^{i\varphi}$,

$0 \leq \varphi < 2\pi$ und hat der Startwert $z^{(0)}$ bzgl. v_1 eine nicht – verschwindene

$$\text{Komponente } z^{(0)} = \sum_{j=1}^n c_j v_j \quad c_1 \neq 0,$$

dann gilt

$$e^{-ik\varphi} x^{(k)} = \frac{c_1 v_1}{\|c_1 v_1\|_2} + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \quad \text{für } k \rightarrow \infty.$$

Folge: $z^{(k+1)} = Ax^{(k)} = e^{ik\varphi} \frac{\lambda_1 c_1 v_1}{\|c_1 v_1\|_2} + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$ für $k \rightarrow \infty$.

Satz 6.3.2: Sei $A \in \mathbb{C}^{n \times n}$ und $0 \neq x \in \mathbb{C}^{n \times n}$. Dann hat das lineare Ausgleichsproblem

$$\min_{\lambda \in \mathbb{C}} \|Ax - \lambda x\|_2^2 \text{ die eindeutige Lösung } \lambda = \frac{x^H Ax}{x^H x}.$$

Definition 6.3.1: Für $A \in \mathbb{K}^{n \times n}$ und $0 \neq x \in \mathbb{K}^n$ heißt $R_A(x) = \frac{x^H Ax}{x^H x}$ Rayleigh –

Quotient von x und A .

Satz 6.3.3: Für $A \in \mathbb{C}^{n \times n}$, $A = A^H$ mit Eigenwerten $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_n$ gilt:

$$R_A(x) \in \mathbb{R} \text{ für } x \in \mathbb{C}^n \setminus \{0\}$$

$$\lambda_n = \min_{x \neq 0} R_A(x)$$

$$\lambda_1 = \max_{x \neq 0} R_A(x)$$

Satz 6.3.4: Unter den Voraussetzungen und Bezeichnungen von Satz 6.3.1 und

$$\text{zusätzlich } A = A^T \text{ gilt } R_A(x^{(k)}) = \lambda_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \quad k \rightarrow \infty.$$

Probleme: 2 gleichgroße EW Trick: betrachte $(A - \sigma E)$

Zur Berechnung des betragskleinsten EW:

von Mises – Iteration: $z^{(k)} = A^{-1} x^{(k-1)}$

$$x^{(k)} = \frac{1}{\|z^{(k)}\|} z^{(k)}, \quad k = 1, 2, \dots$$

Verfahren von Wielandt (allgemeine Form):

Gegeben: $A \in \mathbb{C}^{n \times n}$, $x^{(0)} \neq 0$, $\sigma^{(0)}$ Schätzungen an Eigenwert λ_j von A , so daß

$$|\lambda_j - \sigma^{(0)}| < |\lambda_k - \sigma^{(0)}| \quad \text{für alle EW } \lambda_k \neq \lambda_j.$$

Gesucht: Verbesserte Näherung an λ_j und Näherung an zugehörigen EV.

$$\text{Iteration: } z^{(1)} = (A - \sigma^{(0)}E)^{-1} x^{(0)}, \quad x^{(1)} = \frac{1}{\|z^{(1)}\|} z^{(1)}$$

$$\mu_1 \text{ sei der Mittelwert von } \left\{ \frac{z_i^{(1)}}{x_i^{(0)}} \mid x_i^{(0)} \neq 0 \right\}$$

$$\sigma^{(1)} = \sigma^{(0)} + \frac{1}{\mu_1}.$$

$$\text{Allgemein: } z^{(k+1)} = (A - \sigma^{(k)}E)^{-1} x^{(k)}, \quad x^{(k+1)} = \frac{1}{\|z^{(k+1)}\|} z^{(k+1)}$$

$$\mu_{k+1} \text{ sei der Mittelwert von } \left\{ \frac{z_i^{(k+1)}}{x_i^{(k)}} \mid x_i^{(k)} \neq 0 \right\}$$

$$\sigma^{(k+1)} = \sigma^{(k)} + \frac{1}{\mu_{k+1}}.$$

Abbruch, sobald $|\sigma^{(k+1)} - \sigma^{(k)}|$ klein genug.

Verfahren von Wielandt für hermitesche Matrizen:

Näherung an EW $\sigma^{(k+1)} = \sigma^{(k)} + \frac{1}{\mu_{k+1}}$

$$\mu_{k+1} = x^{(k+1)H} A x^{(k+1)} \text{ statt Mittelwert}$$

6.4. Das QR - Verfahren

6.4.1. Algorithmus (Teilraumiteration mittels orthonormaler Basen)

Gegeben: $A \in \mathbb{C}^{n \times n}$, $p \in \{1, \dots, n\}$

Gesucht: Näherung an p EW und zugehörige EV

Start: $Q_0 \in \mathbb{C}^{n \times p}$ mit orthonorm. Spalten ($Q_0^H Q_0 = E_p$)

Iteration: Für $k = 1, 2, \dots$ $Z_k = A Q_{k-1}$ (Potenzierschritt)

QR – Zerlegung (z.B. Satz 4.7.5)

$Z_k =: Q_k R_k$ (Normierungsschritt)

Abbruch mit geeignetem Kriterium

6.4.2. Algorithmus (Grundform des QR – Verfahrens):

gegeben: $A_1 := A \in \mathbb{C}^{n \times n}$

gesucht: Näherungen an $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, wo die λ_i die Eigenwerte von A sind.

Iteration: Für $k = 1, 2, \dots$ bestimme die QR – Zerlegung

$$A_k =: Q_k R_k$$

$$A_{k+1} := R_k Q_k$$

Abbruch sobald $\|R_k - D\|$ klein genug

Satz 6.4.1: Sei $A = QR$ und $B=RQ$. Ist A hermitesch bzw. obere Hessenbergmatrix, dann auch B .

Lemma 6.4.2: Sei $A \in \mathbb{C}^{n \times n}$ regulär, Q_k und R_k seien durch das QR – Verfahren erzeugt ($k= 1, 2, \dots$).

$$\text{Man setze } \hat{Q}_0 = E, \hat{Q}_k = Q_1 \cdots Q_k, \hat{Q}_k = (\hat{q}_1^{(k)}, \dots, \hat{q}_n^{(k)}), \hat{R}_k = R_k \cdots R_1$$

Dann gilt für alle $k \in \mathbb{N}$:

$$1) \hat{Q}_k^H A \hat{Q}_k = A_{k+1} \quad (A_{k+1} \text{ ist unitärähnlich zu } A)$$

$$2) A \hat{Q}_k = \hat{Q}_{k+1} R_{k+1}$$

$$3) \text{ Sei } S_m = \text{spann} \{e_1, \dots, e_m\} \text{ (Einheitsvektoren)}$$

Dann gilt:

$$A^{(k+1)} S_m = \text{spann}\{A^{(k+1)} e_1, \dots, A^{(k+1)} e_m\} = \text{spann}\{\hat{q}_1^{(k+1)}, \dots, \hat{q}_m^{(k+1)}\},$$

d.h. die ersten m – Spalten von \hat{Q}_{k+1} bilden eine Orthonormalbasis der ersten m Spalten von $A^{(k+1)}$.

$$4) A^{k+1} = \hat{Q}_{k+1} \hat{R}_{k+1} \quad (\text{Potenzieren}) \quad (\text{QR – Zerlegung von } A^{k+1})$$

Definition 6.4.1: Seien L, U zwei m – dim. Unterräume des C^n . Dann wird ihr Abstand definiert durch

$$d(L, U) = \|P_L - P_U\|_2, \quad \text{wobei } P_L \text{ bzw. } P_U \text{ die orthogonale Projektion des } C^n \text{ auf } L \text{ bzw. } U \text{ ist.}$$

Bemerkung: Ist q_1, \dots, q_m eine Orthonormalbasis von L , dann ist $x \rightarrow \sum_{i=1}^m (q_i^H x) q_i$ die

orthogonale Projektion von C^n auf L . Die zugehörige Matrix

$$P_L = \sum_{i=1}^m q_i q_i^H$$

Satz 6.4.3: $A \in C^{n \times n}$ diagonalähnlich mit EW $\lambda_1, \dots, \lambda_n$ und zugeh. Eigenvektoren u_1, \dots, u_n . Es sei $|\lambda_1| \geq \dots \geq |\lambda_n| > 0$.

Für ein $m \in \{1, \dots, n-1\}$ sei $|\lambda_m| > |\lambda_{m+1}|$.

$$U_m := \text{spann}\{u_1, \dots, u_m\}$$

$$T_m := \text{spann}\{u_{m+1}, \dots, u_n\}$$

Sei S_m m – dim. Teilraum von C^n mit $S_m \cap T_m = \{0\}$.

Dann gilt:

$$d(A^k S_m, U_m) = O\left(\left|\frac{\lambda_{m+1}}{\lambda_m}\right|^k\right) \quad \text{für } k \rightarrow \infty.$$

Satz 6.4.4: Sei $A \in C^{n \times n}$ diagonalähnlich mit EW $\lambda_1, \dots, \lambda_n$ und zugehörigen EV u_1, \dots, u_n . Es sei $|\lambda_1| > \dots > |\lambda_n| > 0$.

Für $m = 1, \dots, n-1$ sei

$$S_m = \text{spann}\{e_1, \dots, e_m\}, U_m = \text{spann}\{u_1, \dots, u_m\}, T_m = \text{spann}\{u_{m+1}, \dots, u_n\}$$

$$S_m \cap T_m = \{0\} \quad m = 1, \dots, n-1.$$

Dann gilt für die durch das QR – Verfahren erzeugte Folge $\{A_k\}_{k=1}^\infty, A_k =$

$$\left(a_{ij}^{(k)}\right)_{j=1}^n, \quad \lim_{k \rightarrow \infty} a_{ij}^{(k)} = 0 \quad 1 \leq j < i \leq n$$

$$\lim_{k \rightarrow \infty} a_{ii}^{(k)} = \lambda_i \quad i = 1, \dots, n$$

Das QR – Verfahren mit shift

Im allgemeinen ist die Konvergenz des QR – Verfahrens zu langsam.

- Reduzieren des Aufwandes durch Transformation auf obere Hessenbergform
- Verfahren konvergiert schneller mit shift:

$$A_k - \sigma_k E = Q_k R_k \quad \text{QR – Zerlegung}$$

↑ geshiftete Matrix

$$A_{k+1} = R_k Q_k + \sigma_k E$$

A_k habe die EW $\mu \Rightarrow A_k - \sigma_k E$ hat die EW $\mu - \sigma_k$.

Shiftwahl $A: \sigma_k = a_{nn}^{(k)}$ Durch diese Shiftwahl wird erreicht, daß $a_{n,n-1}^{(k)}$ quadratisch gegen 0 konvergiert.

Shiftwahl B: σ_k als den Eigenwert von $\begin{pmatrix} a_{n-1,n-1}^{(k)} & a_{n-1,n}^{(k)} \\ a_{n,n-1}^{(k)} & a_{n,n}^{(k)} \end{pmatrix}$, der am nächsten an $a_{nn}^{(k)}$ liegt.

Bei Tridiagonalmatrizen erreicht man quadratische Konvergenz für $a_{n-1,n}^{(k)} \rightarrow 0$ und $a_{n,n-1}^{(k)} \rightarrow 0$ für $k \rightarrow \infty$.

6.5. Das Jacobi – Verfahren

$A \in \mathbf{R}^{n \times n}$, $A = A^T \Rightarrow$ alle EW reell (A ist insbesondere normal!)

Ziel: Mit „unendlich vielen“ elementaren Orthogonalen Matrizen $J^{(k)}$

$$\lim_{k \rightarrow \infty} (J^{(k)T} \dots J^{(1)T} A J^{(1)} \dots J^{(k)}) = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Definition 6.5.0: Ein elementare Orthogonalmatrix ist eine Matrix der Form

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}$$

Definition 6.5.1: Sei $c = \cos \varphi$, $s = \sin \varphi$, $1 \leq \mu < \nu \leq n$. Dann heißt die Matrix

$$J_{\mu\nu}(\varphi) = \left(\begin{array}{ccc|ccc} 1 & & & & & 0 \\ & c & \dots & s & & \\ & & 1 & & & \\ & -s & \dots & c & & \\ \hline 0 & & & & & 1 \end{array} \right) \text{Jivens – Rotation oder Jacobi –}$$

Transformation.

Lemma 6.5.1: Sei $A \in \mathbf{R}^{n \times n}$, $A = A^T$. $A = (a_{ij})_{i,j=1}^n$. Für $B = (b_{ij})_{i,j=1}^n$ und

$B := J_{\mu\nu}^T(\varphi) A J_{\mu\nu}(\varphi)$ gilt:

$$b_{\mu\mu} = c^2 a_{\mu\mu} - 2 c s a_{\mu\nu} + s^2 a_{\nu\nu}$$

$$b_{\nu\nu} = s^2 a_{\mu\mu} + 2 c s a_{\mu\nu} + c^2 a_{\nu\nu}$$

$$b_{\mu\nu} = b_{\nu\mu} = c s (a_{\mu\mu} - a_{\nu\nu}) + (c^2 - s^2) a_{\mu\nu}$$

$$b_{i\mu} = b_{\mu i} = c a_{i\mu} - s a_{i\nu} \quad i \neq \mu, \nu$$

$$b_{i\nu} = b_{\nu i} = s a_{i\mu} + c a_{i\nu} \quad i \neq \mu, \nu$$

$$b_{ij} = a_{ij} \text{ sonst}$$

„Maß“ für die Abweichung von der Diagonalgestalt: $S(A) = \sum_{\substack{i,j=1 \\ i \neq j}}^n a_{ij}^2$

Lemma 6.5.2: $J_{\mu\nu}(\varphi)$, A, und B wie oben. $a_{\mu\nu} \neq 0$ $\mu < \nu$.

Der Drehwinkel $\varphi \in]-\frac{\pi}{4}, \frac{\pi}{4}[$ sei durch

$$\cot(2\varphi) = \frac{a_{\nu\nu} - a_{\mu\mu}}{2a_{\mu\nu}}, \text{ falls } a_{\nu\nu} \neq a_{\mu\mu}$$

$$\varphi = \frac{\pi}{4}, \text{ falls } a_{\nu\nu} = a_{\mu\mu}$$

definiert.

Dann gilt:

i) $b_{\mu\nu} = b_{\nu\mu} = 0$

ii) $S(B) = S(A) - 2 a_{\mu\nu}^2 < S(A)$

iii) Ist $|a_{\mu\nu}| = \max_{1 \leq i < j \leq n} |a_{ij}|$, dann gilt $S(B) \leq \left(1 - \frac{2}{n^2 - n}\right) \cdot S(A)$

Bemerkung: Bestimmung von φ aus $\cot(2\varphi) = \frac{a_{\nu\nu} - a_{\mu\mu}}{2a_{\mu\nu}} =: \vartheta$ ist für kleine $|\varphi|$

ungünstig.

besseres Vorgehen: nach Rutishauser:

$$\cot(2\varphi) = \dots = \frac{1 - \tan^2 \varphi}{2 \tan \varphi}, \text{ d.h. } t := \tan \varphi \text{ erfüllt } t^2 + 2\vartheta t - 1 = 0.$$

Als $t = \tan \varphi$ wählt man die Betragsmäßig kleinere Lösung $t =$

$$\frac{\text{sign}(\vartheta)}{|\vartheta| + \sqrt{1 + \vartheta^2}} = -\vartheta + \text{sign}(\vartheta)\sqrt{1 + \vartheta^2}$$

3.Bin.

$\vartheta = 0 \Rightarrow t = 1.$

$$c = \cos \varphi = \frac{1}{\sqrt{1 + t^2}}, \quad s = \sin \varphi = c t.$$

Algorithmus: (klassisches Jacobi – Verfahren)

Gegeben: $A \in \mathbf{R}^{n \times n}$, $A = A^T$

$$A^{(0)} := A = (a_{ij}), \quad \varepsilon > 0 \text{ (Abbruchbedingung)}$$

Gesucht: Näherungen $\hat{\lambda}_1, \dots, \hat{\lambda}_n$ an die EW von A mit $|\hat{\lambda}_i - \lambda_i| < \varepsilon.$

Iteration: Für $k = 0, 1, 2, \dots$

i) Bestimme μ, ν so, daß $|a_{\mu\nu}^{(k)}| = \max_{1 \leq i < j \leq n} |a_{ij}^{(k)}|$

ii) Bestimme t, s, c nach Rutishauser

iii) $A^{(k+1)} = J_{\mu\nu}^T(\varphi) A^{(k)} J_{\mu\nu}(\varphi)$

Abbruch sobald $S(A^{(k+1)}) < \varepsilon^2.$

Satz 6.5.3: Beim Jacobi – Verfahren konvergiert die Folge orthogonal ähnlicher Matrizen $\{A^{(k)}\}$ elementweise gegen eine Diagonalmatrix, deren Elemente die EW von A sind.

Index zu Numerik I

#

Σ 26

A

a posteriori – Abschätzung	5
a priori – Abschätzung	5
A^+	27
Abbruchfehler	1
abgeschlossen	14
Abschätzung	14
Abschnittskettenbruch	9
absolute Fehler	23
absoluter Fehler	1, 2
Addition	20, 21
affiner Raum	26
Algebra	31
Algorithmus von Bairstow	16
Algorithmus von Gauß-Jordan ..	20
Algorithmus von Hyman	33
allgemeine Störung	23
Anwendungen des Newton- Verfahrens	16
Anzahl	19
Approximation	25
Approximationssatz von Weierstraß	10

Ä

äquivalent	4
------------------	---

A

Arithmetische Grundoperation	3
arithmetische Operation	1
Assoziativgesetz	3
asymptotische Fehlerkoeffizient ..	6
Aufwand	10
Ausgleichsproblem	25
Auslöschungseffekt	9
Auswertung	12
Auswertung von Polynomen	10

B

backward analysis	3
Bairstow-Verfahren	16
Banach Raum	5
Banachsche Fixpunktsatz	5
Banachsches-Fixpunktproblem ..	28
Bandbreite	19
Basis	3
Begleitmatrix	13

BFS	5
Bisektionsverfahren	14, 15, 33

C

Cauchy-Folge	4
charakteristische Polynom	31
Cholesky	20
Cholesky-Zerlegung	19, 20
Clenshaw –Algorithmus	13
cond(L)	23
cond _N (A)	23
Crout	20

D

Das Gesamt- und Einzelschrittverfahren	27
Datenfehler	1, 2
Datenvektor	1
Deflationsverfahren	16
det	19
Determinantenregel	32
Diagonalelemente	20
Diagonalmatrix	23
Die QR –Zerlegung	23
Direkte Lösungen von linearen Gleichungssystemen	17
Diskretisierung	17
Diskretisierungsfehler	1
Distanzfunktion	4
Division	20, 21
Divisionsalgorithmus	15
Doppelzeiliges Hornerschema ..	10
Dreiecksmatrix	18

E

Eigenraum	24, 31
Eigenvektor	13, 30
Eigenwert	13, 19, 24, 30, 33
Eigenwerte	14
Eigenwertprobleme	31
Eingabefehler	2
Eingangsfehler	3
Einheitsmatrix	29
Einschließungssätze für Polynomnullstellen	13
Einzelschrittverfahren	28
Elementaralgorithmus	2
elementare hermitesche Matrix ..	24
elementare untere Dreiecksmatrix	18
ESV	28, 29
ESV mit Relaxation	30
Euklidische Divisionsalgorithmus	15
Euklidische Norm	4
Euler	9
Exponent	3

Extremalpunkt	11
---------------------	----

F

Fehlerabschätzungen	23
Fehleranalyse	1
Fehlerfortpflanzung bei arithmetischen Operationen	1
Fehlerfortpflanzung und Stabilität	1
Fehlerkoeffizient	6
Fehlermatrix	30
Fehlertypen	1
Fehlervektor	2
Fixpunktgleichung	27
Fixpunktiteration	4
Folge	9
Fourier – Tschebyscheff – Reihe	12
Fourier – Tschebyscheff Koeffizient	12
Frobenius – Begleitmatrix	13
Frobeniusnorm	22

G

g-adisch	3
g-adische Darstellung	2
Gauß	20, 25
Gaußelimination	18, 23
Gaußelimination	18
Gauß-Jordan-Tableau	21
Gauß-Jordan-Verfahren	20
Gaußsche Eliminationsverfahren	17
Gauß-Seidel-Verfahren	28
Genauigkeit	9
geordnete Liste	14
gerade	11
Gerschgorinscher Kreissatz	14
gerundete Wert	2
Gesamtschrittverfahren	28
Geschwindigkeit	29
Gewichtsfunktion	12
ggT	15
Gleichmäßige Approximation ...	10
gleichmäßige Konvergenz	12
Gleichungssystem	19, 32
Gleitkommazahl	3
global	8
global konvergent	6
Grad	10
Grundbegriffe aus der Algebra ..	31
Grundkörper	4
GSV	28, 29
GSV mit Relaxation	29
gut konditioniert	1
gutartiger Algorithmus	3

H

Hauptminor	18
------------------	----

Numerik I

Hauptunterdeterminante	20
hermitesch	24, 29
Hessenberggestalt	32
Hessenbergmatrix	32
Höhenlinien	16
Horner – Algorithmus	10
Horner – Schema	10
Householder	32
Householder-Algorithmus	24
Householdermatrix	24, 32
Householder-Transformation	24
Hyman	32, 33
Hyperebene	24

I

Idealisierungsfehler	1
Identität	13
Instabilität	2
Intervall	12
invariant	31
involutiv	24
Iteration	5, 10
Iterative Lösungen linearer Gleichungen	27
Iterative Lösungen nichtlinearer Gleichungen	4
Iterationsverfahren	8

J

Jacobi-Matrix	28
Jacobi-Verfahren	28
JOR – Verfahren	29
Jordan	20
Jordankästchen	14

K

Kahau	30
Kern	26
Kettenbruch	9
Kettenbruchauswertung	9
Kettenbruchentwicklung	9
Koeffizient	14
Koeffizientenvektor	10, 16
Koeffizientenvergleich	10
komplexe Arithmetik	16
komplexe Nullstellen	16
Komplexe Schur – Zerlegung	31
komplexe Zahlen	9
Komplexität	19
Kondition	23
konditioniert	1
Konditionszahl	2
Konditionszahlverstärkung	23
Konstruktion einer Sturmschen Kette	15
kontrahierend	5, 27
Kontraktion	5
Kontraktionszahl	5
konvergent	4
Konvergenz	30

Seite 43

Konvergenz von Iterationsverfahren für lineare Gleichungssysteme	29
Konvergenzaussage	9
Konvergenzfaktor	6, 7
Konvergenzordnung	6, 8
konvex	1, 8
Kreissatz	14
Kreisscheibe	14
Kriterium von Sassenfeld	29
Kugel	5

L

L_1 Norm	4
L_2 Norm	4
Länge	9
least upper bound	22
least-square-Problem	25
linear	6, 7
Lineare Ausgleichsprobleme	25
Lip.-Konstante	5
Lipschitz-beschränkt	5
Liste	14
lokal konvergent	6
lokale Variante des BFS	5
L_p Norm	4
LR – Zerlegung	18
LR Zerlegung	19
LR-Zerlegung	20
$\text{lub}_2(A)$	22
lub-Norm	22
$\text{lub}_1(A)$	22
$\text{lub}_\infty(A)$	22
L_∞ Norm	4

M

Mantisse	3
Maschinenzahl	3
mathematischer Prozeß	1
Matrix	17
Matrixnorm	22, 23
Maximumsnorm	4
Maximumsnorm	12
mehrdimensionale Fall	8
Meßpunkt	25
Meßwert	25
Methode der kleinsten Quadrate	25
Metrik	4
Metrische Räume	4
metrischer Raum	4
Minimalwertproblem	26
Modellfehler	1
monoton	8
Moore Penrose-Inverse	27
Multiplikation	10
multivariables Newton-Verfahren	9
Multiplikation	20, 21

N

Nachiteration	30
Näherung	16
Näherungsbruch	9
natürliche Instabilität	2
natürliche Stabilität	2
Newton	33
Newton – Verfahren	6
Newton-Verfahren	7
Newton-Verfahren für komplexe Funktionen	16
Norm	4
Normalgleichung	26
normierte Diagonale	18
Nullstelle	6, 11, 14, 31
numerische Instabilität	2
Numerische Lösung mit Diskretisierung	17
numerische Stabilität	2
numerischer Fehler	1

O

obere Bandbreite	19
obere Dreiecksmatrix	25
Objekt	14
offen	1
Ordnung	6, 11
orthogonal	26
orthogonale Projektion	25
Orthogonalsystem	12
Ostrowski	29, 30

P

passend	22
Permutationsmatrix	17, 18
Pivotisierungsverfahren	18
Pivotregel	21
Polynom	9, 10, 11
Polynomiale Approximation	9
Polynomkoeffizient	9
positiv definit	19, 20, 29, 30
Pseudoinverse	27
Pseudonormlösung	26

Q

QR-Zerlegung	24
quadratisch	6, 7
quadratische Konvergenz	9

R

Radius	13, 14
Rechenregel	21
Rechenschema	10
Rechnungsverlauf	2
Reduktion auf Tridiagonal- bzw. Hessenberggestalt	32

Reduktionsverfahren von	
Householder	32
reelle Zahlen	9
Regula falsi	7, 8
regulär	18, 30
Reich-Ostrowski	29, 30
Reihenfolge	3
Rekursion	11, 33
relativer Fehler	1, 2
Relaxation	29, 30
Relaxationswert	30
Resdenvektor	25
Residuum	25
Rückwärtsanalyse	3
Rückwärtseinsetzen	19, 31
Rundungsfehler	3
Rundungsfehler bei	
Gleitkommaarithmetik	2

S

Sassenfeld	29
Schätzen der Vielfachheit	7
Schema	10
schlecht konditioniert	2
Schmidtsches	
Orthogonalisierungsverfahren	25
Schur-Zerlegung	31
Sehnentangentenverfahren	6
Seidel	28
Sekantenverfahren	7, 8
Sequenz	14
singulär	18
Singulärwerte	26
Singulärwertzerlegung	26
SOR – Verfahren	30
Spaltenregel	21
Spaltensummennorm	22
Spektralnorm	22
Spektralradius	22, 30
Spiegelung	24
stabile Methode von Householder	
.....	24

Stabilität	2
Startvektor	29
Stellenzahl	31
stetig differenzierbar	1
Störung	23
streng monoton	8
strikt konvex	8
strikte Dreiecksmatrix	28
Sturmsche Kette	14, 15, 33
Submultiplikativität	22
Summationsreihenfolge	3
Superlinear	6
symmetrisch	26
symmetrische Matrix	19

T

Taylor	2, 7, 8, 9
Taylorpolynom	9
Teilalgorithmus	2
$T_{ESV}(\omega)$	30
$T_{GSV}(\omega)$	30
tridiagonal	32
Tridiagonalgestalt	32
Tridiagonalmatrix	32
Tschebyscheff-Entwicklung	11
Tschebyscheff-Norm	4
Tschebyscheff-Polynom	11
Tschebyscheff-Polynome erster	
Art	11
Tschebyscheff-Polynome zweiter	
Art	11
t-stellige Gleitkommazahl	3

Ü

überbestimmte Probleme	25
Überlauf	3

U

ungerade	11
unitär	23, 24, 31
Unterdeterminante	18
untere Bandbreite	19
untere Dreiecksmatrix	18, 20
Unterlauf	3

V

Vektornorm	22, 27
Vektorraum	4
Vereinigung	14
Verfahren von Cholesky Crout	20
Verfahren von Hyman	32
Verlust führender Stellen	2
Vermeidung der komplexen	
Arithmetik	16
Vielfachheit	7, 31
vollständig	5
Vorwärtseinsetzen	19, 31
Vorzeichen	3

W

Wallis	9
Wechselzahl	15
Weierstraß	10
Wilkinsons backward analysis	3
Wurzel	20

Z

Zeilenäquivalent	23
Zeilenregel	21
Zeilensummennorm	22
Zeilenvertauschung	18
zerfallend	32
Zuordnungsfehler	1