Systeme nichtlinearer Gleichungen

3. Vorlesung
170 004 Numerische Methoden I

Clemens Brand

5. März 2009

Gliederung

Wiederholung

Wichtige Themen zur Fixpunkt-Iteration

- Wann konvergiert Fixpunktiteration
  - anschaulich erklärt
  - mathematisch exakte Konvergenzbedingung
- Was bedeutet „lokale Konvergenz“
- Konvergenzordnung
- Anschauliche Bedeutung von $|\Phi'| < 1$
- Was ist eine kontrahierende Abbildung
- Gibt's zu den vorigen Prüfungsbeispielen Fragen

Konvergenzordnung

Nochmals, auch weils im Skriptum nicht so klar formuliert ist

Neue Fehlerschranke mindestens um Faktor C kleiner als...

- alte Fehlerschranke: *lineare* Konvergenz (wenn $C < 1$)
- das Quadrat des alten Fehlers: *quadratische* Konvergenz; typisch für Newton-Verfahren.
- allgemein: die p-te Potenz des alten Fehlers: *Konvergenz p-ter Ordnung*. Bei Sekanten-Verfahren ist $p \approx 1.61$.

Faustregeln

- Lineare Konvergenz braucht eine fixe Anzahl von Schritten pro gültiger Stelle.
- Ein Newton-Schritt verdoppelt die Zahl der gültigen Stellen.
- Ein Sekanten-Schritt erhöht die gültigen Stellenanzahl um etwa 60%.

Eine Prüfungsfrage

1. Das Newton-Verfahren zur Bestimmung einer Nullstelle der Funktion $f(x)$ lässt sich auch als Fixpunkt-Aufgabe $x = \phi(x)$ auffassen. Wie lautet die Funktion ϕ?
2. Zeigen Sie: ist x eine einfache Nullstelle von f, dann konvergiert das Newton-Verfahren quadratisch.
3. Zeigen Sie am Beispiel der Funktion $f(x) = x^2$: zu einer doppelten Nullstelle konvergiert das Newton-Verfahren nur linear mit Konvergenzfaktor $C = \frac{1}{2}$.

Wiederholung (und Vorschau)

Das haben wir letzte Woche nur kurz angerissen

- Mehrere Gleichungen und Unbekannte
- Unbekannte zu einem Vektor zusammenfassen
- Gleichungssystem als vektorwertige Funktion schreiben
- Begriffe Nullstelle und Fixpunkt lassen sich direkt verallgemeinern
- Auch Fixpunkt-Iteration geht analog

Bei geeigneter Schreibweise ändert sich fast nichts gegenüber dem Fall einer Gleichung und Unbekannten

Skalare und vektorwertige Funktionen

Reellwertige Funktionen, Skalare:	$f : \mathbb{R} \to \mathbb{R}$	,	$y = f(x)$
Vektorwertige Funktionen, Vektoren:	$\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$	,	$\mathbf{y} = \mathbf{f}(\mathbf{x})$

Komponenten eines Vektors $\in \mathbb{R}^n$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{oder } \mathbf{x}^T = [x_1, x_2, \ldots, x_n]$$

Normalerweise ist mit $\mathbf{x}$ ein Spalten-, mit $\mathbf{x}^T$ ein Zeilenvektor gemeint.

Iterationsindizes sind (um sie von Vektorkomponenten zu unterscheiden) in der Regel hochgestellt, in Klammern: $\mathbf{x}^{(k)}, k = 0, 1, 2 \ldots$

Mehrere Unbekannte: Aufgabentypen

Nichtlineares Gleichungssystem in zwei Unbekannten

$$\begin{aligned} 4x - y + xy &= 1 \\ -x + 6y &= 2 - \log(xy) \end{aligned}$$

Nullstelle einer vektorwertigen Funktion $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$

$$\begin{aligned} 4x - y + xy - 1 &= 0 \\ -x + 6y + \log(xy) - 2 &= 0 \end{aligned} \qquad \begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned} \qquad \mathbf{f}(\mathbf{x}) = \mathbf{0}$$

Fixpunkt einer vektorwertigen Funktion $\boldsymbol{\Phi} : \mathbb{R}^2 \to \mathbb{R}^2$

$$\begin{aligned} x_1 &= \frac{1}{4}(x_2 - x_1 x_2 + 1) \\ x_2 &= \frac{1}{6}(x_1 - \log(x_1 x_2) + 2) \end{aligned} \qquad \mathbf{x} = \boldsymbol{\Phi}(\mathbf{x})$$

Beispiel im Skriptum, ab S.21, durchgerechnet!

Nullstellen und Fixpunkte

Definition

Eine *Nullstelle* einer Funktion $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ ist ein Wert $\mathbf{x}$, für den gilt:

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

Definition

Ein *Fixpunkt* einer Abbildung $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$ ist einen Wert $\boldsymbol{\xi}$, für den gilt:

$$\boldsymbol{\xi} = \mathbf{\Phi}(\boldsymbol{\xi}).$$

(„Funktion“ oder „Abbildung“ meint in diesem Kontext dasselbe.)

Fixpunkt-Iteration

Schreibweise für vektorwertige Funktionen $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$

Gegeben: eine Funktion $\mathbf{\Phi}(\mathbf{x})$ und ein Startwert $\mathbf{x}^{(0)}$.

Ergebnis: falls konvergent, liefert die Fixpunkt-Iteration einen Fixpunkt $\boldsymbol{\xi}$ von $\mathbf{\Phi}$.

Iterationsvorschrift:
für $k = 0, 1, 2 \ldots$

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

Sekantenmethode

Sekantenmethode ist zweidimensionales Fixpunkt-Verfahren

Die Sekantenmethode berechnet aus zwei Näherungen $x^{(0)}, x^{(1)}$ eine verbesserte Näherung, rechnet dann mit zwei neuen Näherungen weiter. Fasse die beiden Näherungen als Komponenten eines Vektors auf. Die Schreibweise

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} x_2 \\ x_2 - f(x_2)\frac{x_1 - x_2}{f(x_1) - f(x_2)} \end{bmatrix}$$

formuliert die Sekantenmethode als zweidimensionale Fixpunkt-Iteration

$$\mathbf{x}^{(k+1)} = \mathbf{\Phi}(\mathbf{x}^{(k)}) \text{ für } k = 0, 1, 2 \ldots$$

Vektornormen: wozu?

Jedes iterative Verfahren braucht eine *Abbruchbedingung*.
Zum Beispiel: Hör auf,

- wenn der Fehler klein genug ist; oder
- wenn der Funktionswert (fast) 0 ist; oder
- wenn sich Werte (fast) nicht mehr ändern.

$$|f(x^{(k)})| < \epsilon, \qquad |x^{(k+1)} - x^{(k)}| < \epsilon$$

In solchen Bedingungen tritt die *Betrags-Funktion* auf.

Wie entscheidet man, ob ein *Vektor* von Werten klein genug ist; oder ob zwei Vektoren sich fast nicht mehr unterscheiden?

Vektornormen

Eine *Norm* ist eine Maßzahl für die „Größe" eines Vektors. Für einen Vektor $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ ist

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| \qquad \textit{Einsnorm}$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} (x_i)^2} \qquad \textit{euklidische Norm, Zweinorm}$$

$$\|\mathbf{x}\|_\infty = \max_i |x_i| \qquad \textit{Unendlich-Norm, Maximums-Norm}$$

In MATLAB:
$\|\mathbf{x}\|_1 =$ `norm(x,1)`, $\|\mathbf{x}\|_2 =$ `norm(x)` oder `norm(x,2)`,
$\|\mathbf{x}\|_\infty =$ `norm(x,inf)`.

Eine *Norm* kann auch den *Abstand* zwischen zwei Punkten $\mathbf{x}$ und $\mathbf{y}$ messen:

$$\mathrm{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

- Taxis in Manhattan messen Strecken in der 1-Norm.
- Abstand in der Luftlinie entspricht der 2-Norm.

Matrixnormen

Achtung—Ergänzung zum Skriptum!

- Ein Hauptberuf von Matrizen ist, Vektoren zu multiplizieren.
- Das Ergebnis einer Matrix-Vektor-Multiplikation ist wieder ein Vektor; der ist i.a. länger, kürzer und verdreht gegenüber Ausgangsvektor
- Eine gegebene Matrix kann aber Vektoren nicht beliebig stark verlängern. Es gibt für jede Matrix einen „Maximal-Verlängerungs-Faktor“
- Eine *Matrixnorm* misst die „Größe“ einer Matrix, das heißt, wie „stark“ sie auf Vektoren wirkt.

Verschiedene Matrixnormen

Achtung—Ergänzung zum Skriptum!

Eine *Matrixnorm* misst die „Größe" einer Matrix. Die $1-, 2-$ und ∞-Normen lassen sich von den entsprechenden Vektornormen ableiten: geben für die Rechenoperation $\mathbf{y} = A \cdot \mathbf{x}$ an, um wieviel $\mathbf{y}$ gegenüber $\mathbf{x}$ *maximal vergrößert* wird.

$\|A\|_1$	*Einsnorm*: maximale Spaltenbetragssumme
$\|A\|_2$	*Zweinorm*: größter Singulärwert
$\|A\|_\infty$	*Unendlich-Norm*: maximale Zeilenbetragssumme
$\|A\|_F$	*Frobenius-Norm:* $\sqrt{\sum a_{ij}^2}$

MATLAB: $\|A\|_1 =$ `norm(A,1)`, ..., $\|A\|_F =$ `norm(A,'fro')`.

Fixpunkt-Iteration konvergiert für kontrahierende Abbildungen

Die Funktion $\mathbf{\Phi}(x)$ besitze einen Fixpunkt $\boldsymbol{\xi}$: $\mathbf{\Phi}(\boldsymbol{\xi}) = \boldsymbol{\xi}$. Sei ferner B eine *Umgebung* um den Fixpunkt $\boldsymbol{\xi}$ in der Form $B = \{\mathbf{x} : \|\boldsymbol{\xi} - \mathbf{x}\| < r\}$, sodass $\mathbf{\Phi}$ in B eine *kontrahierende Abbildung* ist, d. h.es gilt

$$\|\mathbf{\Phi}(\mathbf{x}) - \mathbf{\Phi}(\mathbf{y})\| \leq C\|\mathbf{x} - \mathbf{y}\| \,, \quad C < 1$$

für alle $\mathbf{x}, \mathbf{y} \in B$ in (irgend-)einer Norm $\| \cdot \|$.
Dann konvergiert die Fixpunkt-Iteration $\mathbf{x}^{(k+1)} = \mathbf{\Phi}(\mathbf{x}^{(k)})$ mindestens linear gegen $\boldsymbol{\xi}$ für alle $\mathbf{x}^{(0)} \in B$.

Fixpunkt-Iteration konvergiert für $\|D_\phi\| < 1$

Achtung—Ergänzung zum Skriptum!

Die Matrix der partiellen Ableitungen

$$D_\phi = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \frac{\partial \phi_1}{\partial x_2} & \dots & \frac{\partial \phi_1}{\partial x_n} \\ \frac{\partial \phi_2}{\partial x_1} & \frac{\partial \phi_2}{\partial x_2} & \dots & \frac{\partial \phi_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \frac{\partial \phi_n}{\partial x_2} & \dots & \frac{\partial \phi_n}{\partial x_n} \end{bmatrix}$$

heißt die *Jacobi-Matrix* der Abbildung $\mathbf{\Phi}(\mathbf{x})$. Ist in einen Fixpunkt von $\mathbf{\Phi}(\mathbf{x})$ (irgend-) eine Norm $\|D_\phi\| < 1$, dann konvergiert die Fixpunkt-Iteration für Startwerte in einer Umgebung des Fixpunktes.

Beispiel von vorhin

Die Funktion $\mathbf{\Phi}$ ist hier ein Vektor aus zwei reellwertigen Funktionen ϕ_1 und ϕ_2, der Vektor $\mathbf{x}$ hat zwei Komponenten x_1 und x_2.

$$\mathbf{\Phi}(\mathbf{x}) = \begin{bmatrix} \phi_1(x_1, x_2) \\ \phi_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} \frac{1}{4}(x_2 - x_1 x_2 + 1) \\ \frac{1}{6}(x_1 - \log(x_1 x_2) + 2) \end{bmatrix}$$

$$D_\phi = \begin{bmatrix} \frac{-x_2}{4} & \frac{1-x_1}{4} \\ \frac{1-\frac{1}{x_1}}{6} & \frac{-1}{6\,x_2} \end{bmatrix}$$

Ausgewertet für $\mathbf{x} = \begin{bmatrix} 0,35 \\ 0,64 \end{bmatrix}$ ($\approx$ Fixpunkt) $D_\phi = \begin{bmatrix} -0,160 & 0,163 \\ -0,310 & -0,260 \end{bmatrix}$

$\|D_\phi\|_1 = 0,4695 \quad \|D_\phi\|_2 = 0,4051 \quad \|D_\phi\|_\infty = 0,5699 \quad \|D_\phi\|_F = 0,4644$

Newton-Verfahren für Systeme

Gegeben eine differenzierbare vektorwertige Funktion $\mathbf{f}(\mathbf{x})$ und ein Startwert $\mathbf{x}^{(0)}$. Gesucht eine Nullstelle von $\mathbf{f}$.

Iterationsvorschrift

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

mit $\Delta\mathbf{x}^{(k)}$ *als Lösung von* $D_f(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} = -\mathbf{f}(\mathbf{x}^{(k)})$

Auch dieses Verfahren ist ein Fixpunktverfahren. Die Iterationsfunktion lässt sich formal schreiben (Vorsicht—rechnerische Ausführung in dieser Form ungünstig!)

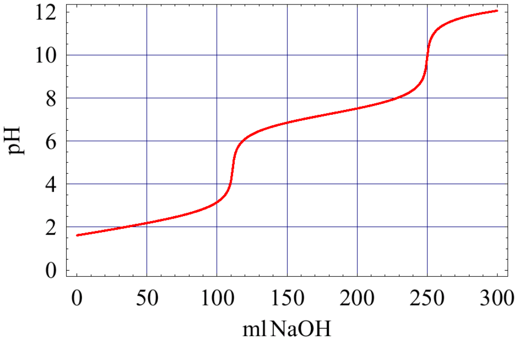
$$\mathbf{\Phi}(\mathbf{x}) = \mathbf{x} - D_f(\mathbf{x})^{-1}\mathbf{f}(\mathbf{x})$$

Beispiel:

Titration von 0,1m-Phosphorsäure mit 1m-NaOH

Ein System sechs nichtlinearer Gleichungen bestimmt die Konzentration der undissoziierten H_3PO_4; der Kationen H^+, Na^+; und Anionen OH^-, $H_2PO_4{}^-$, $HPO_4{}^{--}$, $PO_4{}^{---}$.

$H^+ \cdot OH^-$	$=$	K_w	Ionenprodukt des Wassers
$H^+ \cdot H_2PO_4{}^-$	$=$	$K_{s1} \cdot H_3PO_4$	P-Säure, 1. Dissoziationsstufe
$H^+ \cdot HPO_4{}^{--}$	$=$	$K_{s2} \cdot H_2PO_4{}^-$	2. Dissoziationsstufe
$H^+ \cdot PO_4{}^{---}$	$=$	$K_{s2} \cdot HPO_4{}^{--}$	3. Dissoziationsstufe
$H_3PO_4 + H_2PO_4{}^- + HPO_4{}^{--} + PO_4{}^{---}$	$=$	C_{0PS}	P-Bilanz
$H_2PO_4{}^- + 2HPO_4{}^{--} + 3PO_4{}^{---} + OH^-$	$=$	$Na^+ + H^+$	Elektroneutralität



Die Lösung dieses Gleichungssystems für gegebene Na^+-Konzentration bestimmt den pH-Wert. Im Verlauf der Titration schwanken die einzelnen Konzentrationen über viele Zehnerpotenzen. Vorgefertigte Lösungsverfahren haben damit große Schwierigkeiten.

Quelle: Dr. Josef Draxler, Verfahrenstechnik des industriellen Umweltschutzes, Leoben, 2003

Lösung nichtlinearer Gleichungssysteme: Übersicht der Methoden

- Fixpunkt-Iteration: Allgemeine Formulierung; kein Rezept, um günstiges $\mathbf{\Phi}$ zu finden.
- Newton-Raphson: Standard-Verfahren. Varianten:
  - *gedämpft*: langsamere, aber verlässlichere Konvergenz.
  - *fixe Jacobi-Matrix*: lin. Konvergenz, weniger Rechenaufwand
- MATLAB Symbolic Toolbox: `solve` liefert auch numerische Werte.
- Umformen auf Minimierungsaufgabe: finde Minimum einer skalaren Funktion in mehreren Variablen

$$\|\mathbf{f}(\mathbf{x})\|_2 \to \min$$

Heißt *Nichtlineare unrestringierte Optimierung*, *unconstrained nonlinear optimization*, umfangreiches Gebiet, viele Methoden.
In MATLAB `fminsearch(@fun,x0)` verfügbar.