Numerische Mathematik für Maschinenbauer

Nichtlineare Gleichungssysteme III

A. Reusken

K.-H. Brakhage, Saskia Dietze, Thomas Jankuhn

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Sommersemester 2018

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap 5.6

- Das Newton-Verfahren für Systeme
  - Methode
  - Konvergenzeigenschaften
  - Varianten der Methode

Was Sie mitnehmen sollten:

- Wie funktioniert das allgemeine Newton-Verfahren
- Eigenschaften des Newton-Verfahrens
- Wie sieht das vereinfachte Newton-Verfahren aus
- Weshalb verwendet man Dämpfung

Verständnisfragen zur letzten Vorlesung

Sei x^* eine Nullstelle der Funktion $f(x) = |x|^{2.5} - 3$.

w f hat eine eindeutige Nullstelle x^* in $[0, \infty)$.

f Sei x_0 ein Startwert aus einer hinreichend kleinen Umgebung von x^*, und x_k, $k \geq 1$, die mit dem Newton-Verfahren berechnete Folge.
Es gilt $|x_k - x^*| \approx (x_k - x_{k+1})^2$ für k hinreichend groß.

w Das auf f angewandte Newton-Verfahren konvergiert für jeden Startwert $x_0 > 0$ gegen eine Nullstelle.

Verständnisfragen zur letzten Vorlesung

Es seien $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar und x^* so, dass $\Phi(x^*) = x^*$ gilt.
Für $x_0 \in \mathbb{R}^n$ wird die Fixpunktiteration $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \dots$ definiert.
Weiter sei $\Phi'(x)$ die Ableitung von Φ an der Stelle x.

Es seien $n = 1$, $f : \mathbb{R} \to \mathbb{R}$ zweimal stetig differenzierbar, und $f(x^*) = 0$, $f'(x^*) \neq 0$. Weiter sei $\Phi(x) := x - \frac{f(x)}{f'(x)}$.
Bestimmen Sie $\Phi'(x^*)$. $\boxed{0}$

$\boxed{\text{W}}$ Es seien $f : \mathbb{R}^n \to \mathbb{R}^n$, $M \in \mathbb{R}^{n\times n}$ mit $\det M \neq 0$, und $\Phi(x) := x - Mf(x)$. Das Nullstellenproblem $f(x) = 0$ hat dieselbe Lösungen wie das das Fixpunktproblem $\Phi(x) = x$.

Newton-Verfahren für Systeme

Aufgabe

Sei $f = (f_1, \ldots, f_n)^T : \mathbb{R}^n \to \mathbb{R}^n$ (für $n > 1$) eine zweimal stetig differenzierbare vektorwertige Funktion.
Bestimme

$$x^* = (x_1^*, \ldots, x_n^*)^T \in \mathbb{R}^n, \text{ so dass } f(x^*) = 0$$

erfüllt ist.

- Notation: Wir bezeichnen die Lösung am Iterationsschritt k mit
$$x^k = (x_1^k, \ldots, x_n^k)^T \in \mathbb{R}^n.$$

- Zur Erinnerung: Taylor-Entwicklung (für $i = 1, 2, \ldots, n$)

$$f_i(x) = f_i(x^k) + \sum_{j=1}^{n} \frac{\partial f_i(x^k)}{\partial x_j}(x_j - x_j^k) + \mathcal{O}\left(\|x - x^k\|_2^2\right)$$

Das Newton-Verfahren für Systeme

- Taylor-Entwicklung kompakt

$$f(x) = f(x^k) + f'(x^k)\,(x - x^k) + \mathcal{O}\left(\|x - x^k\|_2^2\right),$$

wobei die Jacobi-Matrix gegeben ist durch

$$f'(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \cdots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

- Für die Nullstelle x^{k+1} der linearen Näherung von f in x^k folgt (vgl. Tangente)

$$0 = f(x^k) + f'(x^k)(x^{k+1} - x^k),$$

und hieraus erhält man

$$x^{k+1} = x^k - \left(f'(x^k)\right)^{-1} f(x^k).$$

Das Newton-Verfahren für Systeme

Algorithmus 5.28. (Newton-Iteration)

`Gegeben: Startwert` x^0.

`Für` $k = 0, 1, 2, \ldots$

1. `Berechne` $f(x^k)$ `und` $f'(x^k)$
2. `Löse das lineare Gleichungssystem in` s^k
$$f'(x^k)\, s^k = -f(x^k).$$
3. `Setze (Newton-Korrektur)`
$$x^{k+1} = x^k + s^k$$

Beachte

- Schritt 2 erfordert die Lösung eines $n \times n$ linearen Gleichungssystems $\Rightarrow$ LR-Zerlegung.
- Die Inverse von $f'(x^k)$ wird nicht explizit berechnet.

Beispiel 5.29.

Gegeben sei das nichtlineare Gleichungssystem

$$\begin{aligned} f_1(x_1, x_2) &= 6x_1 - \cos x_1 - 2\,x_2 = 0 \\ f_2(x_1, x_2) &= 8x_2 - x_1 x_2^2 - \sin x_1 = 0 \end{aligned}$$

Führen Sie einen Schritt des Newton-Verfahrens ausgehend vom Startwert $x^0 = (0,0)^T$ durch.

- Berechnung der Jacobi-Matrix

$$f'(x) = \begin{pmatrix} 6 + \sin x_1 & -2 \\ -x_2^2 - \cos x_1 & 8 - 2\,x_1\,x_2 \end{pmatrix}.$$

- Berechnung von $f(x^0)$ und $f'(x^0)$

$$f(x^0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{und} \quad f'(x^0) = \begin{pmatrix} 6 & -2 \\ -1 & 8 \end{pmatrix}.$$

Beispiel 5.29.

Gegeben sei das nichtlineare Gleichungssystem

$$\begin{aligned} f_1(x_1, x_2) &= 6x_1 - \cos x_1 - 2\,x_2 = 0 \\ f_2(x_1, x_2) &= 8x_2 - x_1 x_2^2 - \sin x_1 = 0 \end{aligned}$$

Führen Sie einen Schritt des Newton-Verfahrens ausgehend vom Startwert $x^0 = (0, 0)^T$ durch.

- Berechnung der Newton-Korrektor s^0 aus

$$\begin{pmatrix} 6 & -2 \\ -1 & 8 \end{pmatrix} \begin{pmatrix} s_1^0 \\ s_2^0 \end{pmatrix} = - \begin{pmatrix} -1 \\ 0 \end{pmatrix} \Rightarrow s^0 = \frac{1}{46} \begin{pmatrix} 8 \\ 1 \end{pmatrix}$$

- Berechnung von x^1 ergibt schließlich

$$x^1 = x^0 + s^0 = \frac{1}{46} \begin{pmatrix} 8 \\ 1 \end{pmatrix}$$

Satz 5.31.

Annahmen:

- Gebiet $\boldsymbol{\Omega} \subset \mathbb{R}^{\boldsymbol{n}}$ offen und konvex
- Funktion $\boldsymbol{f} : \boldsymbol{\Omega} \to \mathbb{R}^{\boldsymbol{n}}$ stetig differenzierbar
- Jacobi-Matrix $\boldsymbol{f}'(\boldsymbol{x}) \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}$
  - invertierbar
    $$\det(\boldsymbol{f}'(\boldsymbol{x})) \neq \boldsymbol{0}, \quad \text{für alle } \boldsymbol{x} \in \boldsymbol{\Omega}$$
  - die Inverse beschränkt durch eine Konstante $\boldsymbol{\beta}$
    $$\|(\boldsymbol{f}'(\boldsymbol{x}))^{-1}\| \leq \boldsymbol{\beta} \text{ für alle } \boldsymbol{x} \in \boldsymbol{\Omega}.$$
  - Lipschitz-stetig auf $\boldsymbol{\Omega}$ mit einer Konstanten γ
    $$\|\boldsymbol{f}'(\boldsymbol{x}) - \boldsymbol{f}'(\boldsymbol{y})\| \leq \gamma\|\boldsymbol{x} - \boldsymbol{y}\|, \quad \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\Omega}.$$
- Es existiere eine Lösung $\boldsymbol{x}^*$ von $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ in $\boldsymbol{\Omega}$.

Satz 5.31.

Der Startwert x^0 erfülle

$$x^0 \in K_\omega(x^*) := \{x \in \mathbb{R}^n \quad | \quad \|x^* - x\| < \omega\}$$

mit ω hinreichend klein, so dass $K_\omega(x^*) \subset \Omega$ und

$$\omega \leq \frac{2}{\beta\,\gamma}.$$

Dann gilt für die durch das Newton-Verfahren definierte Folge

$$\left\{x^k\right\}_{k=0}^{\infty} \subset K_\omega(x^*)$$

und sie **konvergiert quadratisch** gegen x^* :

$$\|x^{k+1} - x^*\| \leq \frac{\beta\,\gamma}{2}\,\|x^k - x^*\|^2, \quad k = 0, 1, 2, \ldots$$

Beispiel 5.2. (Erinnerung)

Statt der linearen Integralgleichung im Beispiel 3.3.

$$u(x) + 2\int_0^1 \cos(xt)\, u(t)\, dt = 2, \quad x \in [0,1]$$

ist nun eine nichtlineare Integralgleichung zu lösen:

Gesucht ist eine Funktion $u(x) \geq 0$, die die Integralgleichung

$$u(x) + 2\int_0^1 \cos(xt)\, u(t)^3\, dt = 2, \quad x \in [0,1]$$

erfüllt.

Beispiel 5.33.

Aus Beispiel 5.2. ergibt sich für $n = 60$ das Gleichungssystem

$$f_i(x_1, x_2, \ldots, x_{60}) = 0, \qquad i = 1, 2, \ldots, 60,$$

wobei

$$f_i(x_1, x_2, \ldots, x_{60}) = x_i - 2 + \frac{1}{60} \sum_{j=1}^{60} \cos\left(\frac{\left(i - \frac{1}{2}\right)\left(j - \frac{1}{2}\right)}{3600} \right) x_j^3.$$

Für die Jacobi-Matrix erhält man

$$\left(f'(x)\right)_{i,j} = \frac{\partial f_i(x)}{\partial x_j} = \begin{cases} 1 + \frac{1}{20} \cos\left(\frac{\left(i - \frac{1}{2}\right)^2}{3600} \right) x_i^2 & \text{für } i = j \\ \frac{1}{20} \cos\left(\frac{\left(i - \frac{1}{2}\right)\left(j - \frac{1}{2}\right)}{3600} \right) x_j^2 & \text{für } i \neq j. \end{cases}$$

Beispiel 5.33.

In jedem Iterationsschritt des Newton-Verfahrens werden

- die Jacobi-Matrix $\boldsymbol{f}'(\boldsymbol{x}^k)$ und der Funktionswert $\boldsymbol{f}(\boldsymbol{x}^k)$ berechnet,
- das lineare Gleichungssystem $\boldsymbol{f}'(\boldsymbol{x}^k)\,\boldsymbol{s}^k = -\boldsymbol{f}(\boldsymbol{x}^k)$ gelöst,
- $\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{s}^k$ berechnet.

Ergebnisse für den Startwert $\boldsymbol{x}^0 = (2, 2, \dots, 2)^T$:

k	$\lVert f(x^k)\rVert_2$	$\lVert x^{k+1} - x^k\rVert_2$
0	5.87e+01	4.75e+00
1	1.50e+01	2.31e+00
2	2.52e+00	5.78e−01
3	1.31e−01	3.32e−02
4	4.10e−04	1.05e−04
5	4.09e−09	1.05e−09
6	2.51e−15	–

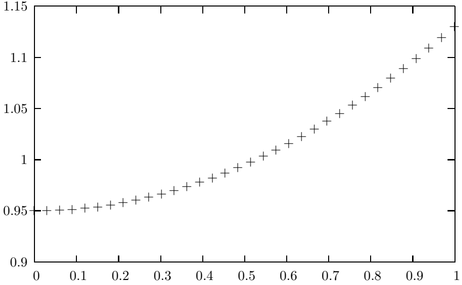
Die dritte Spalte zeigt die Fehlerschätzung

Beispiel 5.33.

Es gilt

$$x_i^6 \approx u(t_i) = u\left(\frac{i-\frac{1}{2}}{60}\right), \qquad i = 1, 2, \ldots, 60.$$

Diese Näherung der Funktion $u(x)$, $x \in [0, 1]$, ist in folgender Abbildung dargestellt:



Hinweise zur praktischen Durchführung

1. Das vereinfachte Newton-Verfahren

Problem: Jeder Schritt erfordert Aufstellen und Lösung des $n \times n$ linearen Gleichungssystems

$$f'(x^k)\, s^k = -f(x^k).$$

Ansatz:

- Aufstellen der Jacobi-Matrix im ersten Schritt $f'(x^0)$.
- Statt $f'(x^k)$ verwende $f'(x^0)$ zur Bestimmung der Newton-Korrektur, d.h.

  $$f'(x^0)\, s^k = -f(x^k) \rightarrow x^{k+1} = x^k + s^k, \; k = 0, 1, 2, \ldots$$
- LR-Berechnung effizient ("mehrere rechte Seiten")

Beachte

- quadratische Konvergenz geht verloren
- evtl. neue Berechnung von f' nach ca. 3-5 Schritten

Hinweise zur praktischen Durchführung

2. Auswertung der Jacobi-Matrix

Problem: Einträge der Jacobi-Matrix, $\partial f_i(x^k)/\partial x_j$, nicht oder nur schwer in geschlossener Form berechenbar.

Ansatz:

- Annäherung durch numerische Differentiation

$$\frac{\partial f_i(x)}{\partial x_j} \approx \frac{f_i(x + h\, e^j) - f_i(x)}{h},$$

  wobei e^j der j-te Einheitsvektor ist.
- Wahl von h ($\longrightarrow$ Kapitel "Interpolation")
  - zu groß: verringert Genauigkeit der Approximation und damit Konvergenz
  - zu klein: birgt Gefahr der Auslöschung
- **Siehe auch:** Quasi-Newton-Verfahren (BFGS)

Hinweise zur praktischen Durchführung

3. Homotopieverfahren

Problem: Bestimmung eines "guten" Startwerts.

Ansatz:

- Benütze Problemparameter oder künstlich eingeführten Parameter μ zur Definition einer Familie von Problemen

$$F(x, \mu) = 0,$$

  so dass F für ein μ einfach lösbar ist.

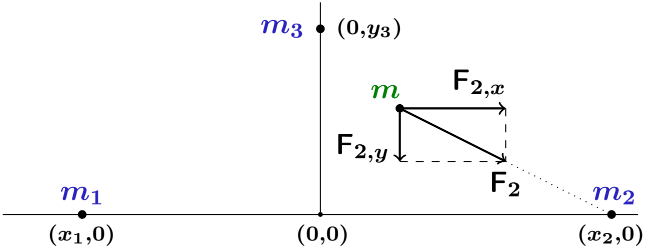
- Beispiel: Nichtlineare Diffusion $u_t = \operatorname{div}(k(u)\,\nabla u)$ mit Wärmeleitfähigkeit $k(u) = c_1 + c_2 u$.
  Wähle als Parameter $\mu = c_2$.

  1. Setze $\mu = 0$. Löse das lineare Problem $F(u, 0) = 0$.
  2. Setze $\mu = \mu + \Delta\mu$ (klein) und nehme alte Lösung als Startwert für das Problem $F(u, \mu) = 0$.
  3. Iteriere bis Newton konvergiert.
     Falls $\mu = c_2$: STOP, sonst gehe zu 2.

Beispiel 5.1. & 5.34.

4. Wahl des Startwerts

Bestimme den Punkt (x, y), so dass für eine Punktmasse m an der Stelle (x, y) die Gravitationskräfte F_i zwischen m und m_i im Gleichgewicht sind.



Beispiel 5.1. & 5.34.

- Gravitationskraft zwischen zwei Punktmassen M_1 und M_2 mit gegenseitigem Abstand r:

$$F = G\frac{M_1 M_2}{r^2},$$

wobei $G = 6.67 \cdot 10^{-11} Nm^2/kg$.

- Hilfsgrößen

$$\begin{aligned} r_i &:= \sqrt{(x - x_i)^2 + (y - y_i)^2}, \\ F_i &:= G\frac{m_i\, m}{r_i^2}, \\ F_{i,x} &:= \frac{F_i(x_i - x)}{r_i},\ F_{i,y} := \frac{F_i(y_i - y)}{r_i},\ i = 1, 2, 3. \end{aligned}$$

- Gleichgewichtsbedingungen

$$F_{1,x} + F_{2,x} + F_{3,x} = 0, \quad F_{1,y} + F_{2,y} + F_{3,y} = 0.$$

Beispiel 5.1. & 5.34.

- Hieraus ergibt sich das nichtlineare Gleichungssystem

$$f_1(x,y) = \sum_{i=1}^{3} \frac{m_i\,(x_i - x)}{\left((x-x_i)^2 + (y-y_i)^2\right)^{3/2}} = 0$$

$$f_2(x,y) = \sum_{i=1}^{3} \frac{m_i\,(y_i - y)}{\left((x-x_i)^2 + (y-y_i)^2\right)^{3/2}} = 0$$

- Für f_1, f_2 gilt

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \nabla U(x,y),$$

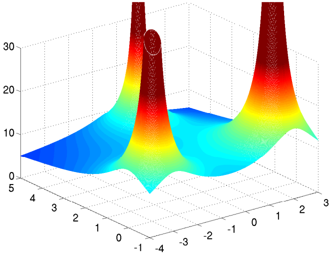
wobei das Potential U gegeben ist durch

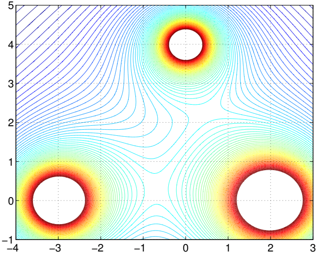
$$U(x,y) := \sum_{i=1}^{3} \frac{m_i}{\left((x_i-x)^2 + (y_i-y)^2\right)^{1/2}}$$

Beispiel 5.1. & 5.34.

- Wahl des Startwerts: (x^*, y^*) ist genau dann Lösung des Systems, wenn (x^*, y^*) ein lokales Minimum, lokales Maximum oder ein Sattelpunkt des Potentials U ist.
- Das Potential U hat zwei Sattelpunkte und keine lokalen Maxima oder Minima.

  Das System hat also genau zwei Lösungen.





Beispiel 5.1. & 5.34.

- Anhand der Graphik kann man geeignete Startwerte wählen.
- Ergebnisse des Newton-Verfahrens:

k	x^k	y^k	$\|f(x^k,y^k)\|_2$	$\|(x^k,y^k)-(x^{k+1},y^{k+1})\|_2$
0	−0.800000000000000	0.200000000000000	3.25e−01	1.31e−01
1	−0.697601435074387	0.281666888630281	1.03e−02	4.45e−03
2	−0.694138545697644	0.284468076535443	1.09e−05	4.09e−06
3	−0.694134676058600	0.284469396792393	9.67e−12	4.57e−12
4	−0.694134676055255	0.284469396789285	2.02e−16	-

k	x^k	y^k	$\|f(x^k,y^k)\|_2$	$\|(x^k,y^k)-(x^{k+1},y^{k+1})\|_2$
0	0.500000000000000	2.200000000000000	1.87e−01	6.32e−02
1	0.4803549525148845	2.260066598359946	4.51e−03	2.27e−03
2	0.4825811382211886	2.259618040348963	4.01e−06	1.75e−06
3	0.4825819025667199	2.259619618799409	3.13e−12	1.59e−12
4	0.4825819025657873	2.259619618798127	3.33e−16	-

Hinweise zur praktischen Durchführung

5. Gedämpftes Newton-Verfahren

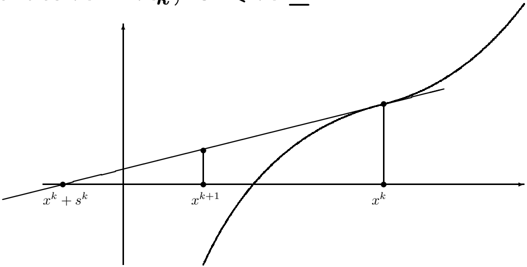
Problem: Divergenz bei schlechtem Startwert

Ansatz:

Man setzt

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{\lambda}\, \boldsymbol{s}^k$$

für ein passendes $\boldsymbol{\lambda} = \boldsymbol{\lambda}_k,\ \boldsymbol{0} < \boldsymbol{\lambda} \leq \boldsymbol{1}.$



Matlab-Demo

Zusammenfassung

- Newton-Verfahren: Konvergenzordnung: $p = 2$, lokale Konvergenz
  - Stoppkriterium: $\|x^* - x^k\| \approx \|x^{k+1} - x^k\|$
- Vereinfachtes Newton-Verfahren: feste Jacobi-Matrix
- Wahl des Startvektors: Homotopieverfahren, problemabhängig
- über Dämpfung kann man den Einzugsbereich vergrößern

Verständnisfragen

Es sei $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ zweimal stetig differenzierbar in einer Umgebung $\boldsymbol{U}$ von $\boldsymbol{x}^*$ und es gelte $\boldsymbol{f}(\boldsymbol{x}^*) = \boldsymbol{0}$.
Wir nehmen an, dass $\det(\boldsymbol{f}'(\boldsymbol{x})) \neq 0$ für alle $\boldsymbol{x} \in \boldsymbol{U}$, und betrachten die Newton-Methode zur Bestimmung von $\boldsymbol{x}^*$:

$$\boldsymbol{x}_0 \in \boldsymbol{U}, \quad \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - (\boldsymbol{f}'(\boldsymbol{x}_k))^{-1} \boldsymbol{f}(\boldsymbol{x}_k) \quad \text{für } k \geq 0.$$

- w Die Newton-Methode ist lokal quadratisch konvergent.
- w Wenn das Newton-Verfahren konvergiert, dann gilt für genügend große Werte von k : $||\boldsymbol{x}_k - \boldsymbol{x}^*|| \approx ||\boldsymbol{x}_k - \boldsymbol{x}_{k+1}||$.
- f Beim Newtonverfahren kann eine Dämpfungsstrategie benutzt werden, die dazu dient für jeden Startwert die Konvergenz des Verfahrens zu gewährleisten.
- f Beim Newtonverfahren kann eine Dämpfungsstrategie benutzt werden, die dazu dient Auslöschungseffekte bei der Berechnung der Korrektur zu vermeiden.