Numerische Mathematik für Elektrotechniker

Nichtlineare Gleichungssysteme IV / Nichtlineare Ausgleichsrechnung I

A. Schlichting

K.-H. Brakhage, A. El Amouri, T. Tscherpel

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Wintersemester 2018/2019

Heute in der Vorlesung

Themen: Dahmen & Reusken Kap 5.6/6.1-6.2

- Das Newton-Verfahren für Systeme
  - Varianten der Methode und Durchführungshinweise

Kapitel 6: Nichtlineares Ausgleichsproblem

- Das nichtlineare Ausgleichsproblem
- Gauß-Newton-Verfahren

Was Sie mitnehmen sollten:

- Wie funktioniert das allgemeine Newton-Verfahren
- Wie sieht das vereinfachte Newton-Verfahren aus
- Weshalb verwendet man Dämpfung
- Wie ist die Problemstellung bei einem nichtlinearen Ausgleichsproblem
- Wie funktioniert das Gauß-Newton-Verfahren

Beispiel 5.24.: Globale Konvergenz

Man berechne $\sqrt{a}$ für ein $a > 0$ mit Hilfe des Newton-Verfahrens.

Die Wurzel $\sqrt{a}$ von a löst $f(x) := x^2 - a = 0.$

Das Newton-Verfahren ist also die Iteration

$$x_{k+1} = x_k - \frac{x_k^2 - a}{2\,x_k} = \frac{x_k}{2} + \frac{a}{2\,x_k} = \Phi(x_k)\ .$$

Kontraktion

- $\Phi : (\sqrt{a/2}, \infty) \to (\sqrt{a/2}, \infty)$ ist Lipschitz mit $L = \frac{1}{2}$, da

$$|\Phi'(x)| = \left|\frac{1}{2} - \frac{a}{2\,x^2}\right| \leq \frac{1}{2}.$$

Beispiel 5.24.: Globale Konvergenz

Selbstabbildung

- $\Phi : (\sqrt{a/2}, \infty) \to (\sqrt{a/2}, \infty)$ ist eine Selbstabbildung

Banach $\Rightarrow$ Es gibt genau ein $x^* \in (\sqrt{a/2}, \infty)$ mit

$$x^* = \Phi(x^*) = \frac{x^*}{2} + \frac{a}{2\,x^*} \quad \text{also} \quad x^* = \sqrt{a}.$$

- Für $x_0 \in (0, \sqrt{a/2}]$ folgt sofort $x_1 = \Phi(x_0) > \sqrt{a/2}$.

Damit ist das Newtonverfahren global konvergent auf $(0, \infty)$.

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

Ergebnisse für den Startwert $x^0 = (2, 2, \ldots, 2)^T$:

k	$\|f(x^k)\|_2$	$\|x^{k+1} - x^k\|_2$
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	–

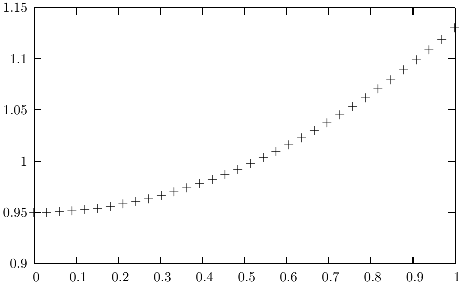
Die dritte Spalte zeigt die Fehlerschätzung

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
- oft neue Berechnung von f' nötig, falls $x_k - x^*$ zu groß (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 damit auch schlechtere Konvergenz
  - zu klein: birgt Gefahr von Auslöschung
- **Siehe auch:** Quasi-Newton-Verfahren (BFGS)

Hinweise zur praktischen Durchführung

3. Homotopieverfahren (stetige Deformation)

Problem: Bestimmung eines "guten" Startwerts.

Ansatz:

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

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

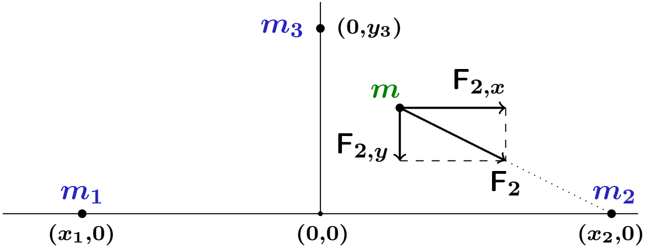
  so dass F für ein μ einfach lösbar ist.

- 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

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

- 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

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

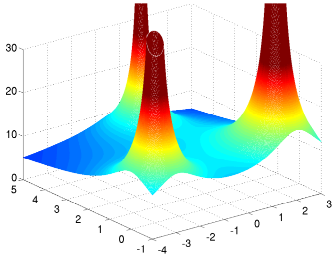
wobei das Potential U gegeben ist durch

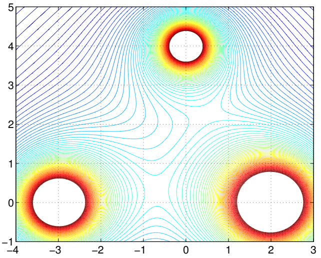
$$U(x,y) := \sum_{i=1}^{3} \frac{m_i}{\left((x_i-x)^2 + (y_i-y)^2\right)^{1/2}}$$

Beispiel 5.1. & 5.34.

- Wahl des Startwerts: $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ ist genau dann Lösung des Systems, wenn $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ ein lokales Minimum, lokales Maximum oder ein Sattelpunkt des Potentials $\boldsymbol{U}$ ist.
- Plot: Das Potential $\boldsymbol{U}$ 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	0.480354952514884 5	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

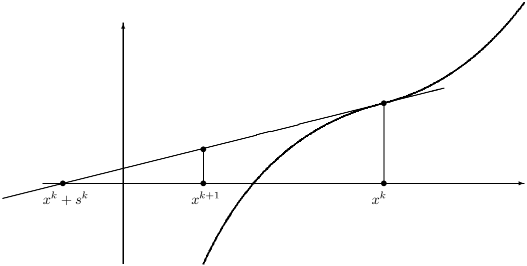
$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^{k} + \boldsymbol{\lambda}\, \boldsymbol{s}^{k}$$

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

Idee: Suche $\boldsymbol{\lambda}$, so dass

$$\left\| \boldsymbol{f}(\boldsymbol{x}^{k+1} \right\| < \left\| \boldsymbol{f}(\boldsymbol{x}^{k}) \right\| \ ,$$

Residuum wird in jedem Schritt verringert.



Matlab-Demo

Nichtlineares Ausgleichsproblem

Kapitel 6:

Nichtlineares Ausgleichsproblem

Satz 5.31.

Annahmen:

$$\vdots$$

- Lipschitz-stetig auf Ω mit einer Konstanten γ

$$\|f'(x) - f'(y)\| \leq \gamma \|x - y\|, \quad x, y \in \Omega.$$

Ein solches γ existiert, wenn f zweimal stetig differenzierbar ist.

$$\vdots$$

Zur Erinnerung: Lineares Ausgleichsproblem (Beispiel 4.2.)

In der Fourieranalyse wird eine T-periodische Funktion f durch eine Linearkombination der T-periodischen trigonometrischen Polynome

$$1, \cos(c\,t), \sin(c\,t), \cos(2\,c\,t), \sin(2\,c\,t), \ldots, \cos(N c\,t), \sin(N c\,t)$$

mit $c := \dfrac{2\,\pi}{T}$ in der Form

$$g_N(t) = \frac{1}{2}\,\alpha_0 + \sum_{k=1}^{N} \Big(\alpha_k\, \cos(k\,c\,t) + \beta_k\, \sin(k\,c\,t) \Big)$$

approximiert.

Zur Erinnerung: Lineares Ausgleichsproblem (Beispiel 4.2.)

Annahme:

Nicht f, sondern nur eine Reihe vom Meßdaten

$$b_i \approx f(t_i), \quad 0 \leq t_1 < t_2 < \cdots < t_m \leq T,$$

ist bekannt, wobei $m > 2N + 1$.

Daraus ergibt sich der

Ansatz zur Bestimmung der Koeffizienten

$$x = (\alpha_0, \alpha_1, \beta_1, \alpha_2, \beta_2 \ldots, \alpha_N, \beta_N)^T:$$

$$x^* = \arg \min_{x \in \mathbb{R}^{2N+1}} \sum_{i=1}^{m} \Big(g_N(t_i) - b_i\Big)^2.$$

Lineares vs. nichtlineares Ausgleichsproblem

Definition

Zu gegebenen $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ und $\boldsymbol{b} \in \mathbb{R}^{m}$, bestimme $\boldsymbol{x}^* \in \mathbb{R}^n$, für das

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{A}\,\boldsymbol{x}^* - \boldsymbol{b}\|_2$$

gilt. Diese Problemstellung heißt das lineare Ausgleichsproblem.

Wesentliche Eigenschaft

Die unbekannten Koeffizienten/Parameter tauchen linear auf
bzw.
es lassen sich entsprechende Parameter definieren/identifizieren.

Beim nichtlinearen Ausgleichsproblem ist dies nicht mehr möglich. . .

Beispiel 6.1.

Differentialgleichung einer gedämpften Schwingung:

$$m\,u'' + b\,u' + D\,u = 0,$$

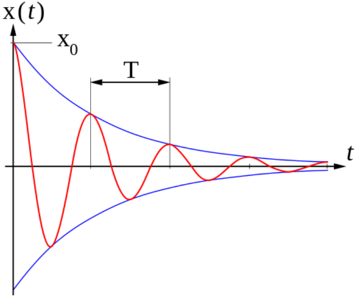
mit Masse m, Dämpfungskonstante b und Federkonstante D.

Lösungen haben die nichtlineare Form:

$$u(t) = u_0\,e^{-\delta\,t}\sin(\omega_d\,t + \varphi_0),$$

wobei:

u_0	$\longrightarrow$	Anfangswert
φ_0	$\longrightarrow$	Nullphasenwinkel
δ	$\longrightarrow$	Abklingkonstante
ω_d	$\longrightarrow$	Eigenkreisfrequenz



Quelle: wikipedia

Beispiel 6.1.

Gegeben:

- 10 Messungen an den Punkten $t_1, t_2, \ldots, t_{10}$ mit zugehörigen Daten $b_1, b_2, \ldots, b_{10}$.
- Modell einer gedämpften Schwingung

$$y(t; x_1, x_2, x_3, x_4) = x_1\, e^{-x_2\, t} \sin(x_3\, t + x_4)$$

mit Parametern $x_1, \ldots, x_4$.

Gesucht:

- Parameter $x_1, \ldots, x_4$, so dass die Summe der Fehlerquadrate

$$\sum_{i=1}^{10} \left(x_1\, e^{-x_2\, t_i}\, \sin(x_3\, t_i + x_4) - b_i\right)^2 = \|F(x)\|_2^2$$

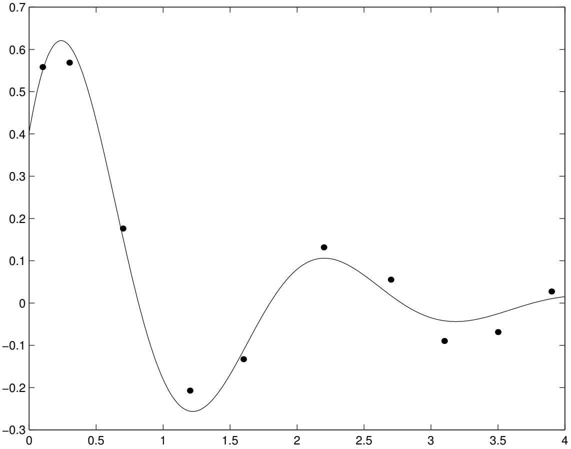
minimal wird.

Hierbei ist $F : \mathbb{R}^4 \to \mathbb{R}^{10}$ nicht linear in x_2, x_3, x_4 definiert

$$\begin{aligned} F_i(x) &= F_i(x_1, x_2, x_3, x_4) \\ &= x_1\, e^{-x_2\, t_i} \sin(x_3\, t_i + x_4) - b_i, \quad i = 1, \ldots, 10. \end{aligned}$$

Beispiel 6.1.

Berechnete Lösung des nichtlinearen Ausgleichsproblems:



Definition

Definiert man allgemein die Abbildung ($m > n$)

$$F : \mathbb{R}^n \to \mathbb{R}^m, \quad F_i(x) := y(t_i; x) - b_i,\ i = 1, \ldots, m,$$

kann das nichtlineare Ausgleichsproblem wie folgt formuliert werden:

Nichtlineares Ausgleichsproblem

Bestimme $x^* \in U \subseteq \mathbb{R}^n$, so dass

$$x^* = \arg \min_{x \in U \subseteq \mathbb{R}^n} \|F(x)\|_2,$$

oder, äquivalent,

$$x^* = \arg \min_{x \in U \subseteq \mathbb{R}^n} \phi(x),$$

wobei $\phi : \mathbb{R}^n \to \mathbb{R},\ \phi(x) := \frac{1}{2}\|F(x)\|_2^2 = \frac{1}{2} F(x)^T F(x).$

Definition

Zur Erinnerung:

Die Funktion $\phi(x) = \frac{1}{2}\|F(x)\|_2^2 = \frac{1}{2}F(x)^T F(x)$ hat in einem Punkt x^* ein lokales Minimum, wenn die folgenden zwei Bedingungen erfüllt sind:

1. $\nabla\phi(x^*) = \mathbf{0}$ (d.h. x^* ist kritischer Punkt von ϕ),
2. $\phi''(x^*) \in \mathbb{R}^{n\times n}$ ist symmetrisch positiv definit.

Es läßt sich durch Nachrechnen bestätigen, dass

$$\begin{aligned} \nabla\phi(x) &= F'(x)^T F(x), \\ \phi''(x) &= F'(x)^T F'(x) + \sum_{i=1}^{m} F_i(x)\, F_i''(x), \end{aligned}$$

mit Jacobi-Matrix $F'(x) \in \mathbb{R}^{m\times n}$

und Hesse-Matrizen $F_i''(x) := \left(\frac{\partial^2 F_i(x)}{\partial x_j \partial x_k}\right)_{1\le j,k\le n} \in \mathbb{R}^{n\times n}$

Das Gauß-Newton-Verfahren

Nichtlineares Ausgleichsproblem (lokal)

Gegeben $F : \mathbb{R}^n \to \mathbb{R}^m$, bestimme $x^* \in U \subset \mathbb{R}^n$, so dass

$$x^* = \arg\min_{x \in U} \|F(x)\|_2.$$

Ansatz:

1. Ersetze $F(x)$ durch lineare Approximation (Taylor-Entwicklung)
2. Schrittweise Annäherung an x^* durch Lösung linearer Probleme in jedem Schritt

Zur Erinnerung:

- Wir bezeichnen die Lösung am Iterationsschritt k mit
$$x^k = (x_1^k, \ldots, x_n^k)^T \in \mathbb{R}^n.$$
- Taylor-Entwicklung
$$F(x) = F(x^k) + F'(x^k)\,(x - x^k) + \mathcal{O}(\|x - x^k\|_2^2).$$

Das Gauß-Newton-Verfahren

Ansatz:

Ersetze $F(x)$ in $\min_{x \in U} \|F(x)\|_2$, durch lineare Approximation

$$\min_{x \in U} \|\underbrace{F(x^k)}_{\in \mathbb{R}^m} + \underbrace{F'(x^k)}_{\in \mathbb{R}^{m \times n}} \underbrace{(x - x^k)}_{\in \mathbb{R}^n}\|_2,$$

Wir setzen $s = x - x^k$ (bzw. $s^k = x^{k+1} - x^k$) und erhalten das lineare Ausgleichsproblem:

Finde $s^k \in \mathbb{R}^n$, so dass

$$s^k = \arg\min_{s \in \mathbb{R}^n} \|F'(x^k)\, s + F(x^k)\|_2$$

Anschließend berechnen wir $\quad x^{k+1} = x^k + s^k$.

Achtung: Es muss überprüft werden ob $x^{k+1} \in U$!

Das Gauß-Newton-Verfahren

Algorithmus 6.3. (Gauß-Newton)

Wähle Startwert x^0.

Für $k = 0, 1, 2, \ldots$:

1. Berechne $F(x^k), F'(x^k)$.
2. Finde s^k, so dass
$$s^k = \arg\min_{s\in\mathbb{R}^n} \|F'(x^k)\, s + F(x^k)\|_2$$
3. Setze $x^{k+1} = x^k + s^k$. Falls $x^{k+1} \notin U$ dann STOPP.

Beachte

- Schritt 2 erfordert die Lösung eines linearen Ausgleichsproblems (Normalgleichung, QR-Zerlegung)
- Falls $F'(x)$ nicht vollen Rang hat, hat das Ausgleichsproblem keine eindeutige Lösung.

Bemerkungen

- "Analogie" nichtlineare Gleichungssysteme.
- In einem kritischen Punkt x^* von ϕ muss die Ableitung

$$\nabla\phi(x) = F'(x)^T F(x)$$

gleich Null $\in \mathbb{R}^{m\times n}$ sein.
Als Abbruchkriterium für das Verfahren wird daher

$$\|F'(x^k)^T F(x^k)\|_2 \leq \varepsilon$$

häufig benutzt, wobei ε eine vorgegeben Toleranz ist.

- Der Erfolg des Gauß-Newton-Verfahrens hängt von der Wahl des Startwerts ab (vgl. Newton-Verfahren).
- Die Bedingung $\mathbf{Rang}(F'(x)) = n$ kann man weglassen, wenn der Zusatz "mit minimaler 2-Norm" aufgenommen wird.

Analyse der Gauß-Newton-Methode

Sei $\boldsymbol{x}^*$ ein kritischer Punkt von ϕ, der in einer Umgebung $\boldsymbol{U}$ eindeutig ist.

Annahme:

$$\mathbf{Rang}(\boldsymbol{F}'(\boldsymbol{x})) = \boldsymbol{n} \quad \text{für alle} \quad \boldsymbol{x} \in \boldsymbol{U} \ .$$

Für $\boldsymbol{x}^k \in \boldsymbol{U}$ hat das lineare Ausgleichsproblem die eindeutige Lösung

$$\boldsymbol{s}^k = -[\boldsymbol{F}'(\boldsymbol{x}^k)^T \, \boldsymbol{F}'(\boldsymbol{x}^k)]^{-1} \, \boldsymbol{F}'(\boldsymbol{x}^k)^T \, \boldsymbol{F}(\boldsymbol{x}^k)$$

Analyse der Gauß-Newton-Methode

Deshalb gilt für die Gauß-Newton-Iteration:

$$\begin{aligned} x^{k+1} &= x^k - [F'(x^k)^T F'(x^k)]^{-1} F'(x^k)^T F(x^k) \\ &= x^k - [F'(x^k)^T F'(x^k)]^{-1} \nabla\phi(x^k) \\ &= \Phi(x^k)\ , \end{aligned}$$

mit

$$\Phi(x) := x - [F'(x)^T F'(x)]^{-1} \nabla\phi(x)\ .$$

Es gilt:

$$x = \Phi(x) \ \Leftrightarrow\ \nabla\phi(x) = 0 \ \Leftrightarrow x = x^*.$$

Die Gauß-Newton-Methode ist also eine Fixpunktiteration.

Beispiel 6.4.

Lösen Sie das nichtlineare Ausgleichsproblem

$$x^* = \arg \min_{x \in [0,2\pi]} \|F(x)\|_2,$$

wobei

$$F(x) := \begin{pmatrix} a + r\,\cos(x) \\ r\,\sin(x) \end{pmatrix}, \text{ mit } a > r > 0.$$

- Für die Jacobi-Matrix erhält man

$$F'(x) = r \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix}, \quad F'(x)^T\, F'(x) = r^2.$$

- Außerdem ergibt sich

$$\phi(x) = \frac{1}{2}\,\|F(x)\|_2^2 = \frac{1}{2}\,(a^2 + 2\,a\,r\,\cos(x) + r^2)$$

und damit

$$\nabla\phi(x) = -a\,r\,\sin(x).$$

Beispiel 6.4.

- Für die Iterationsfunktion zu F erhält man schließlich

$$\begin{aligned}\Phi(x) &= x - \left(F'(x)^T F'(x)\right)^{-1} \nabla\phi(x^k) \\ &= x + \frac{a}{r}\sin(x)\end{aligned}$$

- Es gibt zwei kritische Punkte von ϕ

$$\begin{aligned} x^* &= 0 \quad \text{(lokales Maximum)}, \\ x^* &= \pi \quad \text{(lokales Minimum)}. \end{aligned}$$

- In den kritischen Punkten $x^* = 0$, $x^* = \pi$ gilt

$$|\Phi'(x^*)| = \left|1 + \frac{a}{r}\cos(x^*)\right|.$$

und damit

$$\begin{aligned} |\Phi'(x^*)| &= \frac{a+r}{r} > 1 && \text{für } x^* = 0 \text{ (lokales Max)} \\ |\Phi'(x^*)| &= \frac{a-r}{r} = \frac{a}{r} - 1 && \text{für } x^* = \pi \text{ (lokales Min)} \end{aligned}$$

Beispiel 6.4.

Das Gauß-Newton-Verfahren hat in diesem Beispiel folgende Eigenschaften:

1. Das lokale Maximum ist abstoßend
2. Die Methode ist linear konvergent in einer Umgebung des lokalen Minimums (wenn $a < 2\,r$),

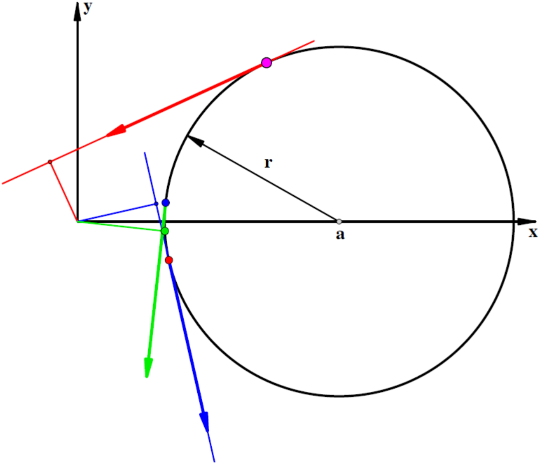
   Falls $a = r$ liegt sogar quadratische Konvergenz vor.

   oder
3. das lokale Minimum ist auch abstoßend (wenn $a > 2\,r$).

Man kann zeigen, dass ähnliche Eigenschaften in einem allgemeinen Rahmen gültig sind.

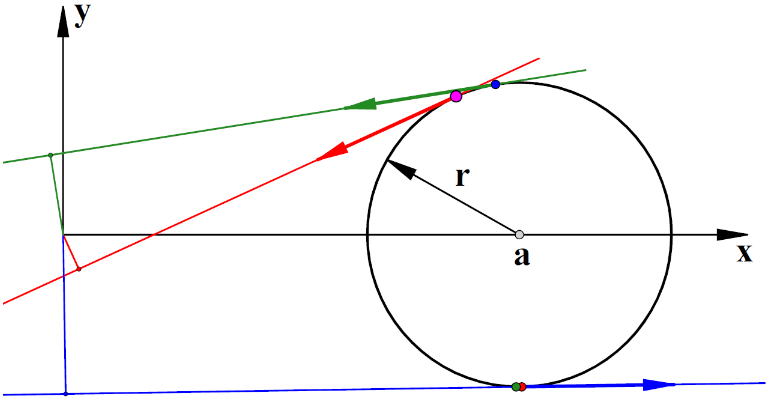
Beispiel 6.4.

Konvergenter Fall: $a < 2r$



Beispiel 6.4.

Divergenter Fall: $a > 2r$



Folgerung 6.6.: Konvergenzanalyse

Untersuchen die Lipschitzkonstante von

$$\Phi(x) = x - [F'(x)^T F'(x)]^{-1} \nabla\phi(x) \quad \text{mit} \quad \nabla\phi(x) = F'(x)^T F(x).$$

Ableitung in x^* unter Ausnutzung $\nabla\phi(x^*) = 0$ ergibt sich zu

$$\Phi'(x^*) = -\left[F'(x)^T F'(x)\right]^{-1} \sum_{i=1}^{m} F_i(x^*)\, F_i''(x^*)\,.$$

Im Normalfall gilt für ein Ausgleichsproblem $F(x^*) \neq 0$.
Also $\Phi'(x^*) \neq 0$ und damit:

Falls die Gauß-Newton-Methode konvergiert, ist die Konvergenz im allgemeinen nicht schneller als linear.

Falls $F(x^*) = 0$ und Gauß-Newton-Methode konvergiert, dann ist die Konvergenz sogar quadratisch (Newton-Fall).

Folgerung 6.6.: Spektralradius

Sei $\rho(\Phi'(x^*))$ der Spektralradius von $\Phi'(x^*)$, das ist betragsgrößte Eigenwert von

$$\Phi'(x^*) = -\left[F'(x)^T F'(x)\right]^{-1} \sum_{i=1}^{m} F_i(x^*)\, F_i''(x^*)\,.$$

Wenn der kritische Punkt x^* lokales Maximum oder Sattelpunkt ist, dann gilt

$$\rho(\Phi'(x^*)) \geq 1 \quad \text{und deshalb} \quad \|\Phi'(x^*)\| \geq \rho(\Phi'(x^*)) \geq 1\,.$$

Beachte

Kritischen Punkte von F, welche kein (lokales) Minimum sind, sind für das Gauß-Newton-Verfahren abstoßend.

Sehr günstige Eigenschaft, da (lokales) Minimum gesucht wird und das Verfahren somit falsche kritischen Punkte ignoriert.

Folgerung 6.6.: Konvergenzkriterium

Die Größe $\rho(\Phi'(x^*))$ entscheidet über die lokale Konvergenz des Gauß-Newton-Verfahrens.

Für ein lokales Minimum x^* der Funktion ϕ ist die lokale Konvergenz des Gauß-Newton-Verfahrens gesichert, falls

$$\rho(\Phi'(x^*)) < 1 \quad .$$

Sei x^* ein lokales Minimum von ϕ, wofür gilt:

$$\rho(\Phi'(x^*)) > 1.$$

Dann ist $\|\Phi'(x^*)\| > 1$ auch für jede Operatornorm $\|\cdot\|$.

Ein lokales Minimum von ϕ kann für die Gauß-Newton-Methode also abstoßend sein.

Beispiel 6.7.

Das Gauß-Newton-Verfahren angewandt auf das Problem der gedämpften Schwingung in Beispiel 6.1.

k	$\|F(x^k)\|_2$	$\|\nabla\phi(x^k)\|_2$	$\|\nabla\phi(x^k)\|_2/\|\nabla\phi(x^{k-1})\|_2$
0	0.35035332090089	1.45e−01	-
1	0.34106434131008	1.33e−01	0.91
2	0.22208131421995	4.88e−02	0.37
3	0.16802866234936	1.02e−01	2.08
4	0.09190056278958	1.80e−01	0.18
5	0.08902339976144	1.18e−03	0.07
6	0.08895515308450	3.81e−04	0.32
7	0.08894991006370	1.15e−04	0.30
8	0.08894937563528	4.07e−05	0.35
9	0.08894931422207	1.38e−05	0.34
10	0.08894930687791	4.85e−06	0.35
11	0.08894930599062	1.68e−06	0.35
12	0.08894930588306	5.87e−07	0.35

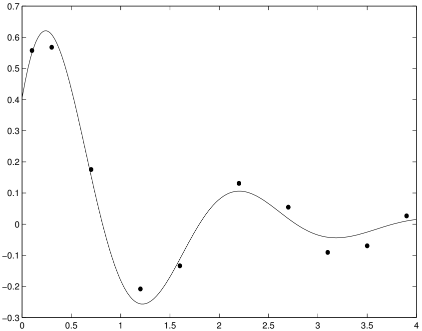
In der letzten Spalte der Tabelle sieht man das lineare Konvergenzverhalten des Gauß-Newton-Verfahrens.

Beispiel 6.7.

Die berechneten Parameterwerte x^* aus dem 12 Iterationsschritt liefern die Lösung

$$y(t;x^*) = x_1^*\, e^{-x_2^*\, t}\, \sin\big(x_3^*\, t + x_4^*\big),$$

die im folgenden Plot dargestellt ist.



Zusammenfassung

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

Zusammenfassung

- Beim Gauß-Newton-Verfahren wird das nichtlineare Ausgleichsproblem über eine Folge linearer Ausgleichsprobleme gelöst.
- Lokale Konvergenz im allgemeinen nur 1. Ordnung, falls $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}$, sogar von 2. Ordnung.
- Es kann lokale Divergenz auftreten.
- Matrix $\boldsymbol{F}'(\boldsymbol{x}^{\boldsymbol{k}})$ kann Rang $< \boldsymbol{n}$ haben.

Verständnisfragen

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

$$x_0 \in U, \quad x_{k+1} = x_k - \left(f'(x_k)\right)^{-1} f(x_k) \quad \text{für } k \geq 0.$$

$\boxed{w}$ Die Newton-Methode ist lokal quadratisch konvergent.

$\boxed{w}$ Wenn das Newton-Verfahren konvergiert, dann gilt für genügend große Werte von k : $||x_k - x^*|| \approx ||x_k - x_{k+1}||$.

$\boxed{f}$ Beim Newton-Verfahren kann eine Dämpfungsstrategie benutzt werden, die dazu dient für jeden Startwert die Konvergenz des Verfahrens zu gewährleisten.

$\boxed{f}$ Beim Newton-Verfahren kann eine Dämpfungsstrategie benutzt werden, die dazu dient Auslöschungseffekte bei der Berechnung der Korrektur zu vermeiden.