1. Mein Text!

Numerische Mathematik für Elektrotechniker

Nichtlineare Gleichungssysteme IV / Nichtlineare Ausgleichsrechnung I

Benjamin Berkels

Karl-Heinz Brakhage, Thomas Jankuhn, Christian Löbbert

Institut für Geometrie und Praktische Mathematik
RWTH Aachen

Wintersemester 2019/2020

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?
- Weshalb verwendet man Dämpfung?
- Wie ist die Problemstellung bei einem nichtlinearen Ausgleichsproblem?
- Wie funktioniert das Gauß-Newton-Verfahren?

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 in 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

$$f_i(x_1, x_2, \ldots, x_{60}) = x_i - 2 + \frac{1}{30} \sum_{j=1}^{60} \cos\left(\frac{\left(i - \frac{1}{2}\right)\left(j - \frac{1}{2}\right)}{3600} \right) x_j^3.$$

Für die Jacobi-Matrix erhält man

$$\left(f'(x)\right)_{i,j} = \frac{\partial f_i(x)}{\partial x_j} = \begin{cases} 1 + \frac{1}{10} \cos\left(\frac{\left(i - \frac{1}{2}\right)^2}{3600} \right) x_i^2 & \text{für } i = j \\ \frac{1}{10} \cos\left(\frac{\left(i - \frac{1}{2}\right)\left(j - \frac{1}{2}\right)}{3600} \right) x_j^2 & \text{für } i \neq j. \end{cases}$$

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 $x^0 = (1, 1, \ldots, 1)^T$:

k	$\|f(x_k)\|$	$\|x_k - x_{k-1}\|$
0	6.95e+00	-
1	9.82e-01	1.17e+00
2	2.20e-02	1.86e-01
3	1.23e-05	4.49e-03
4	3.65e-12	2.43e-06
5	2.77e-15	7.29e-13
6	2.22e-15	2.76e-15

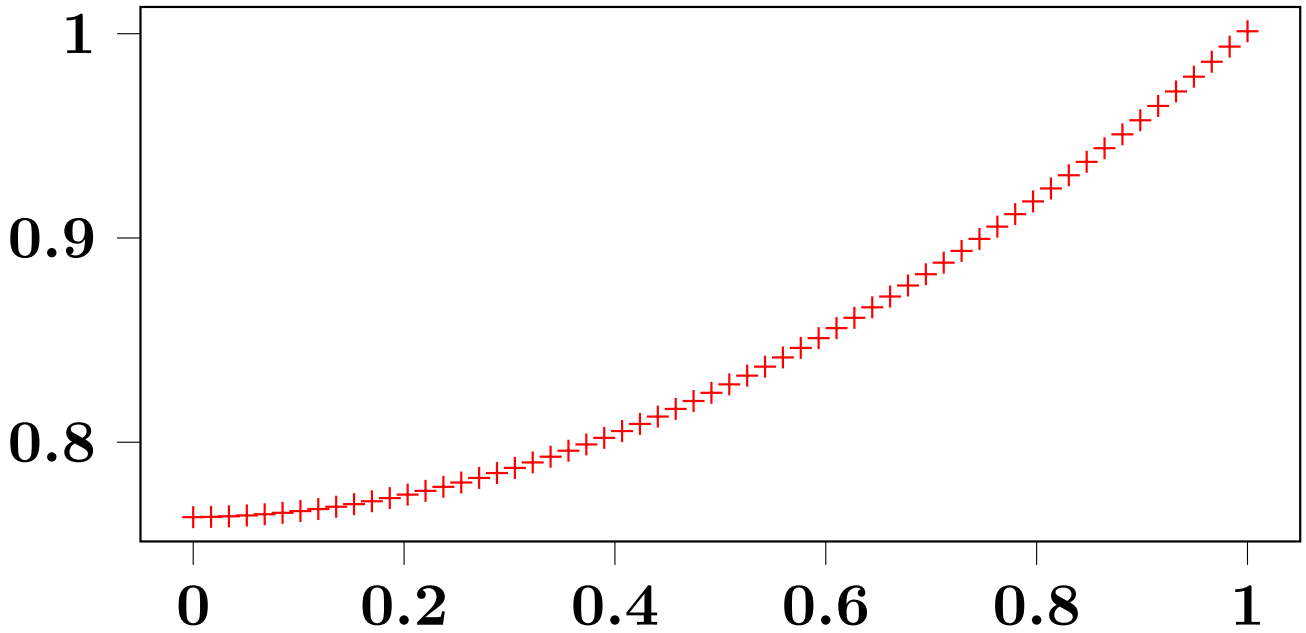
Die dritte Spalte zeigt die Fehlerschätzung

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.

  $$f'(x^0)\, s^k = -f(x^k) \to x^{k+1} = x^k + s^k,\ k = 0, 1, 2, \ldots$$

- 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 (⟶ 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.

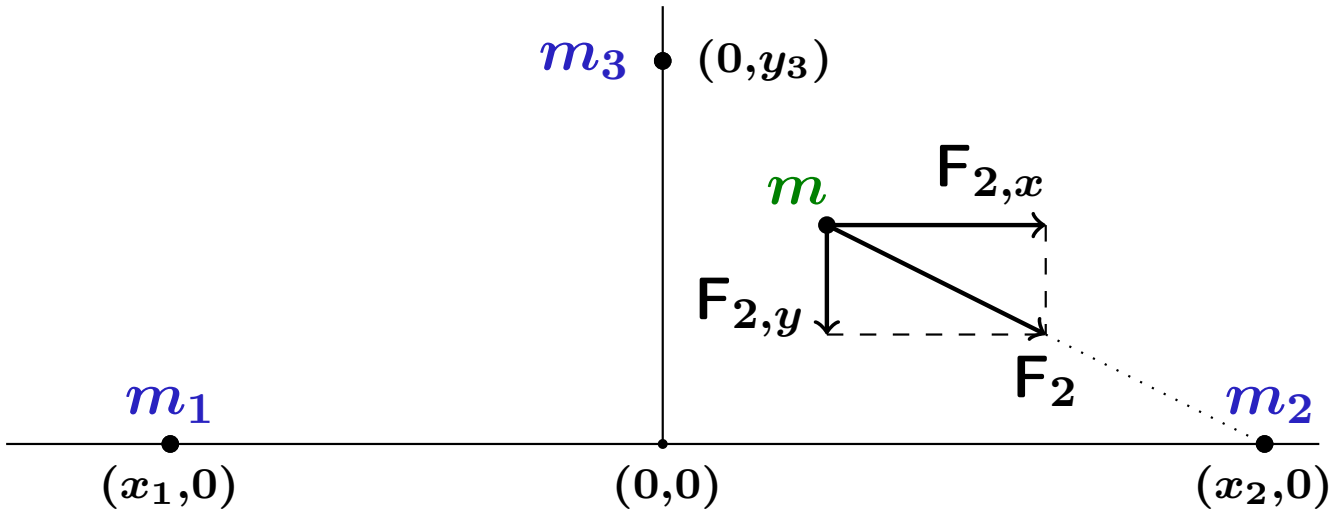
- Beispiel: Nichtlineare Diffusion $u_t = \operatorname{div}(k(u)\,\nabla u)$ mit Wärmeleitfähigkeit $k(u) = c_1 + c_2 u$.
  Wähle als Parameter $\mu = c_2$.

  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

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

- 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

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

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

- Für f_1, f_2 gilt

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

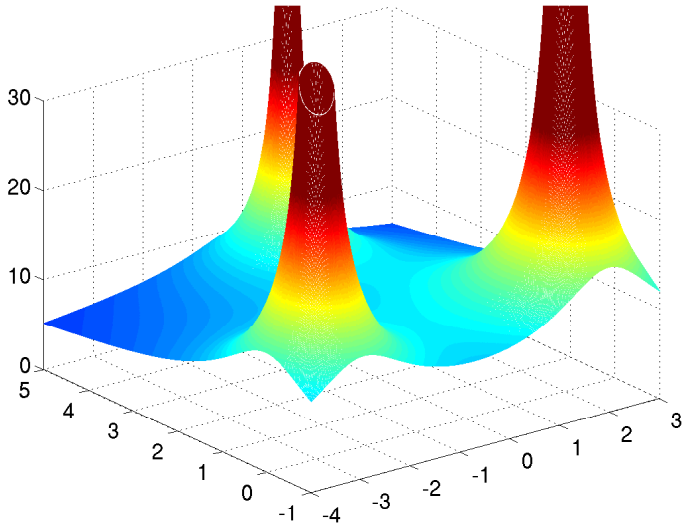
wobei das Potential U gegeben ist durch

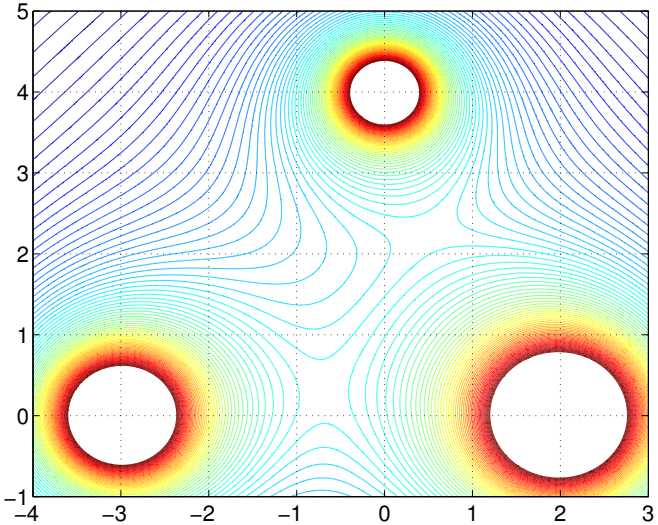
$$U(x,y) := \sum_{i=1}^{3} \frac{m_i}{((x_i - x)^2 + (y_i - y)^2)^{1/2}}$$

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.
- Plot: 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 Grafik 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	-0.8000000000000000	0.2000000000000000	3.25e-01	1.31e-01
1	-0.6976014350743877	0.2816668886302817	1.03e-02	4.45e-03
2	-0.6941385456976449	0.2844680765354431	1.09e-05	4.09e-06
3	-0.6941346760586007	0.2844693967923931	9.67e-12	4.57e-12
4	-0.6941346760552555	0.2844693967892851	2.56e-16	-

k	x^k	y^k	$\|f(x^k, y^k)\|_2$	$\|(x^k, y^k) - (x^{k+1}, y^{k+1})\|_2$
0	0.5000000000000000	2.2000000000000002	1.87e-01	6.32e-02
1	0.4803549525148845	2.2600665983599457	4.51e-03	2.27e-03
2	0.4825811382211886	2.2596180403489625	4.01e-06	1.75e-06
3	0.4825819025667199	2.2596196187994093	3.13e-12	1.59e-12
4	0.4825819025657873	2.2596196187981272	3.33e-16	-

Hinweise zur praktischen Durchführung

5. Gedämpftes Newton-Verfahren

Problem: Divergenz bei schlechtem Startwert

Ansatz: Man setzt

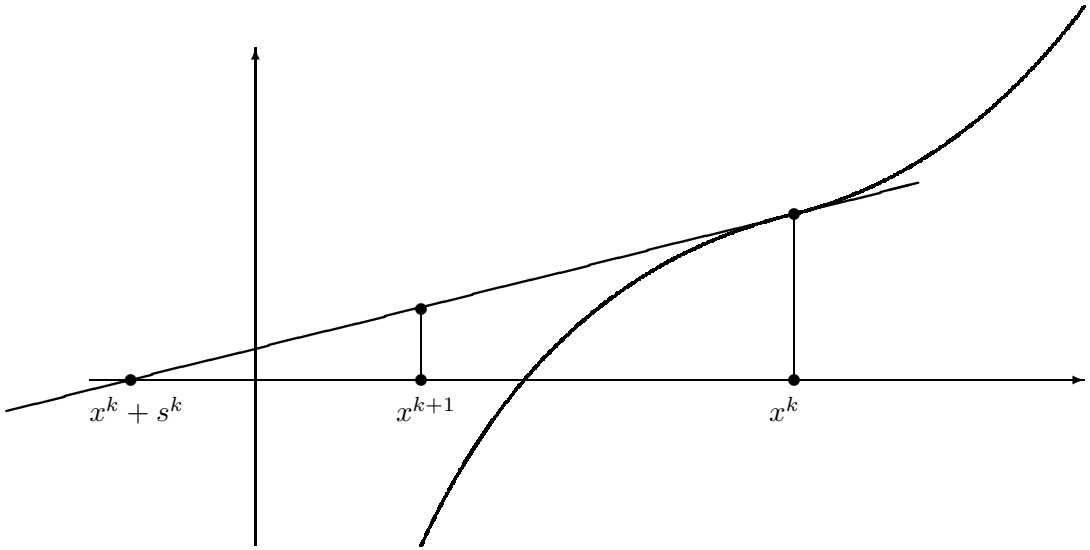
$$x^{k+1} = x^k + \lambda\, s^k$$

für ein passendes $\lambda = \lambda_k,\ 0 < \lambda \leq 1.$

Idee: Suche λ, so dass

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

Residuum wird in jedem Schritt verringert.



Matlab-Demo

`http://www2.imng.uni-stuttgart.de/LstNumGeoMod/TCM/#newton`

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 Konvergenzbereich vergrößern

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^*:

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

w Die Newton-Methode ist lokal quadratisch konvergent.

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

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

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

Nichtlineares Ausgleichsproblem

Kapitel 6:

Nichtlineares Ausgleichsproblem

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

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

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

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

(Gewöhnliche) Differentialgleichung einer gedämpften Schwingung:

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

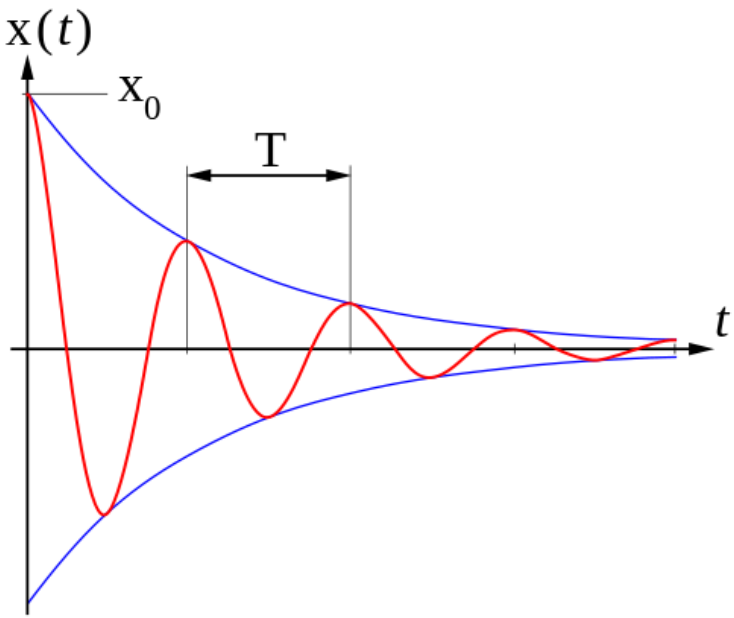
mit Masse m, Dämpfungskonstante b und Federkonstante D.

Lösungen haben die nichtlineare Form:

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

wobei:

$u_0 \longrightarrow$ Anfangswert
$\varphi_0 \longrightarrow$ Nullphasenwinkel
$\delta \longrightarrow$ Abklingkonstante
$\omega_d \longrightarrow$ 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

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

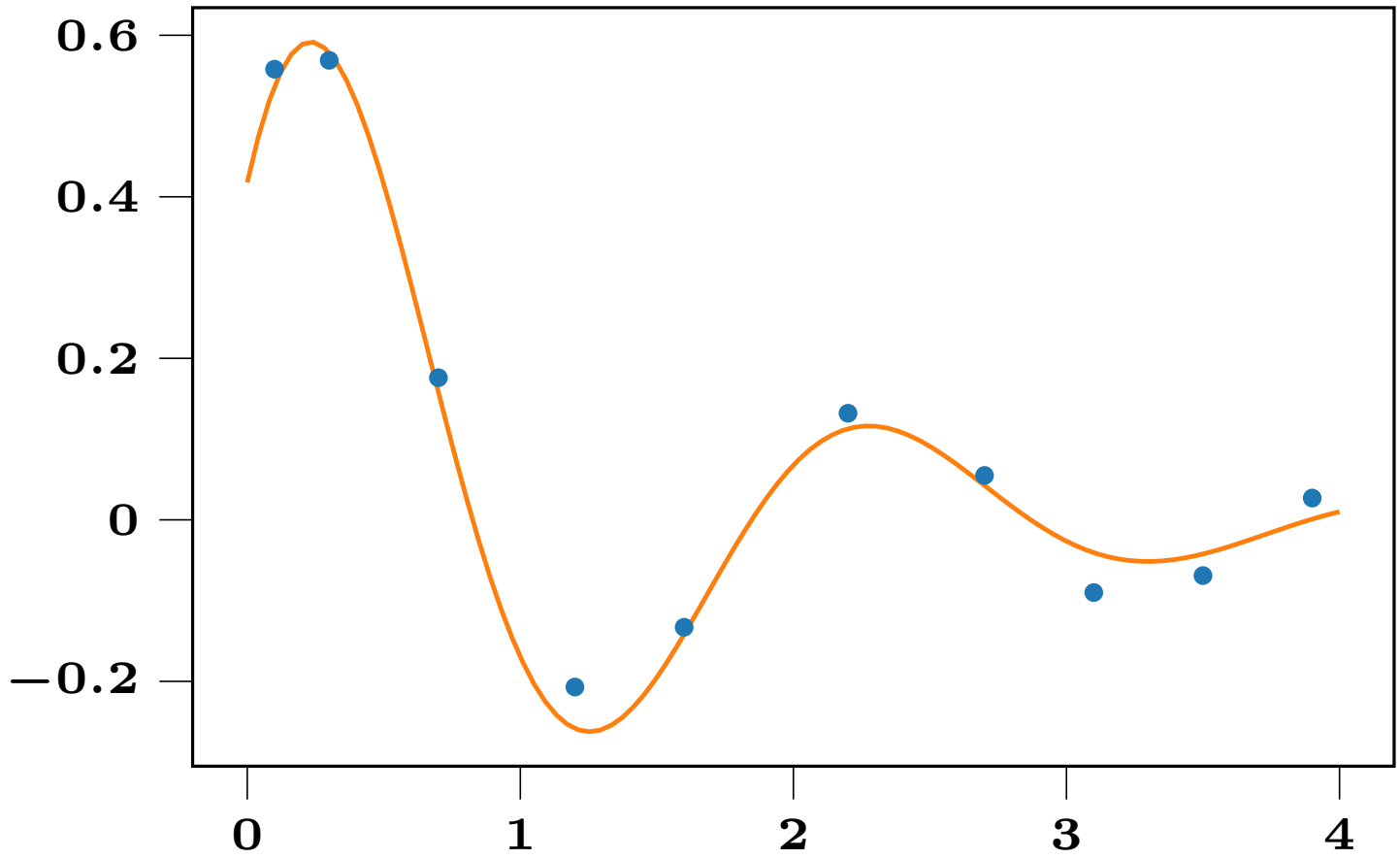
minimal wird.

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

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

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

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

oder, äquivalent,

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

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

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

$$x^* \in \operatorname*{arg\,min}_{x \in \mathbb{R}^n} \|F(x)\|_2.$$

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 $\boldsymbol{F(x)}$ in $\min\limits_{x\in\mathbb{R}^n} \|\boldsymbol{F(x)}\|_2$, durch lineare Approximation

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

Wir setzen $s = x - x^k$ und erhalten das lineare Ausgleichsproblem:

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

$$s^k \in \operatorname*{arg\,min}_{s\in\mathbb{R}^n} \|F'(x^k)\, s + F(x^k)\|_2$$

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

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
$$s^k \in \arg\min_{s \in \mathbb{R}^n} \|F'(x^k)\, s + F(x^k)\|_2$$
3. Setze $x^{k+1} = x^k + s^k$.

Beachte

- Schritt 2 erfordert die Lösung eines linearen Ausgleichsproblems (Normalgleichung, QR-Zerlegung, SVD)
- Falls $F'(x)$ nicht vollen Rang hat, hat das Ausgleichsproblem keine eindeutige Lösung. $\rightarrow$ Wähle s^k mit minimaler 2-Norm.

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).

Analyse der Gauß-Newton-Methode

Sei x^* ein kritischer Punkt von ϕ, der in einer Umgebung U eindeutig ist.

Annahme:

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

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

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

(Lösung der Normalgleichung)

Analyse der Gauß-Newton-Methode

Deshalb gilt für die Gauß-Newton-Iteration:

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

x^* Fixpunkt von Φ
$|\Phi'(x^*)| < 1 \Rightarrow$ Konv.

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

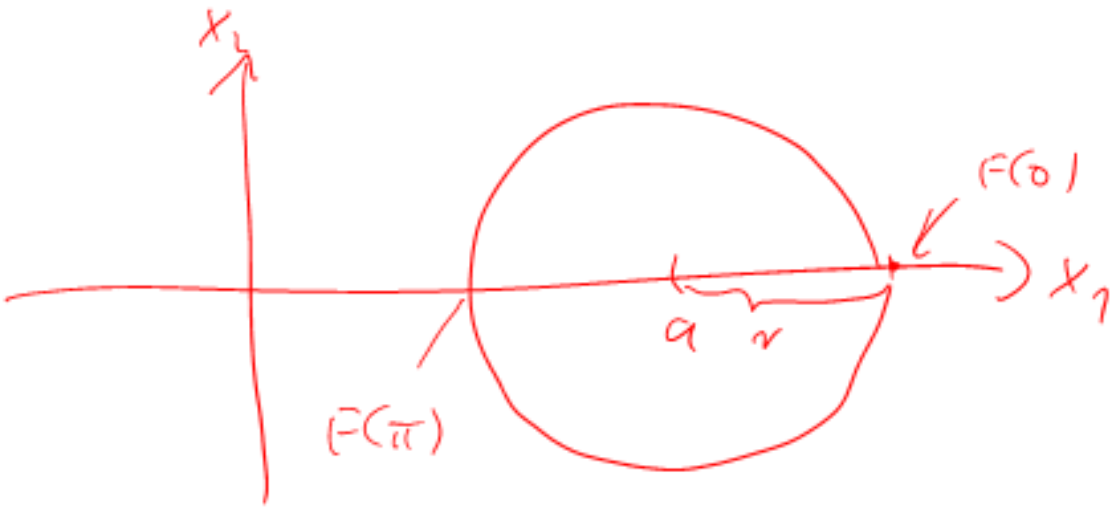
Lösen Sie das nichtlineare Ausgleichsproblem

$F(x) = F(x+2\pi)$!

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

wobei

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



Beispiel 6.4

Lösen Sie das nichtlineare Ausgleichsproblem

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

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

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

$= r^2(-\sin^2(x)) + r^2(\cos^2(x))$

Beispiel 6.4

Lösen Sie das nichtlineare Ausgleichsproblem

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

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

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

Beispiel 6.4

- Für die Iterationsfunktion zu $\boldsymbol{F}$ erhält man schließlich

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

$= \frac{1}{r^2}$ $\quad = -a\,r\,\sin(x)$

Beispiel 6.4

- Für die Iterationsfunktion zu $\boldsymbol{F}$ erhält man schließlich

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

- Es gibt zwei kritische Punkte von ϕ in $[\mathbf{0}, \mathbf{2\pi})$

$$\boldsymbol{x^* = 0} \quad \text{(lokales Maximum)},$$
$$\boldsymbol{x^* = \pi} \quad \text{(lokales Minimum)}.$$

- Wegen

$$\boldsymbol{\Phi'(x) = 1 + \frac{a}{r}\, \cos(x).}$$

gilt in den kritischen Punkten $\boldsymbol{x^* = 0,\ x^* = \pi}$

$$\boldsymbol{|\Phi'(0)| = \frac{a + r}{r} > 1}$$

$$\boldsymbol{|\Phi'(\pi)| = \frac{a - r}{r} = \frac{a}{r} - 1}$$

Beispiel 6.4

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

1. Das lokale Maximum $x^* = 0$ ist abstoßend
2. Die Methode ist linear konvergent in einer Umgebung des lokalen Minimums $x^* = \pi$, falls $a < 2r$.

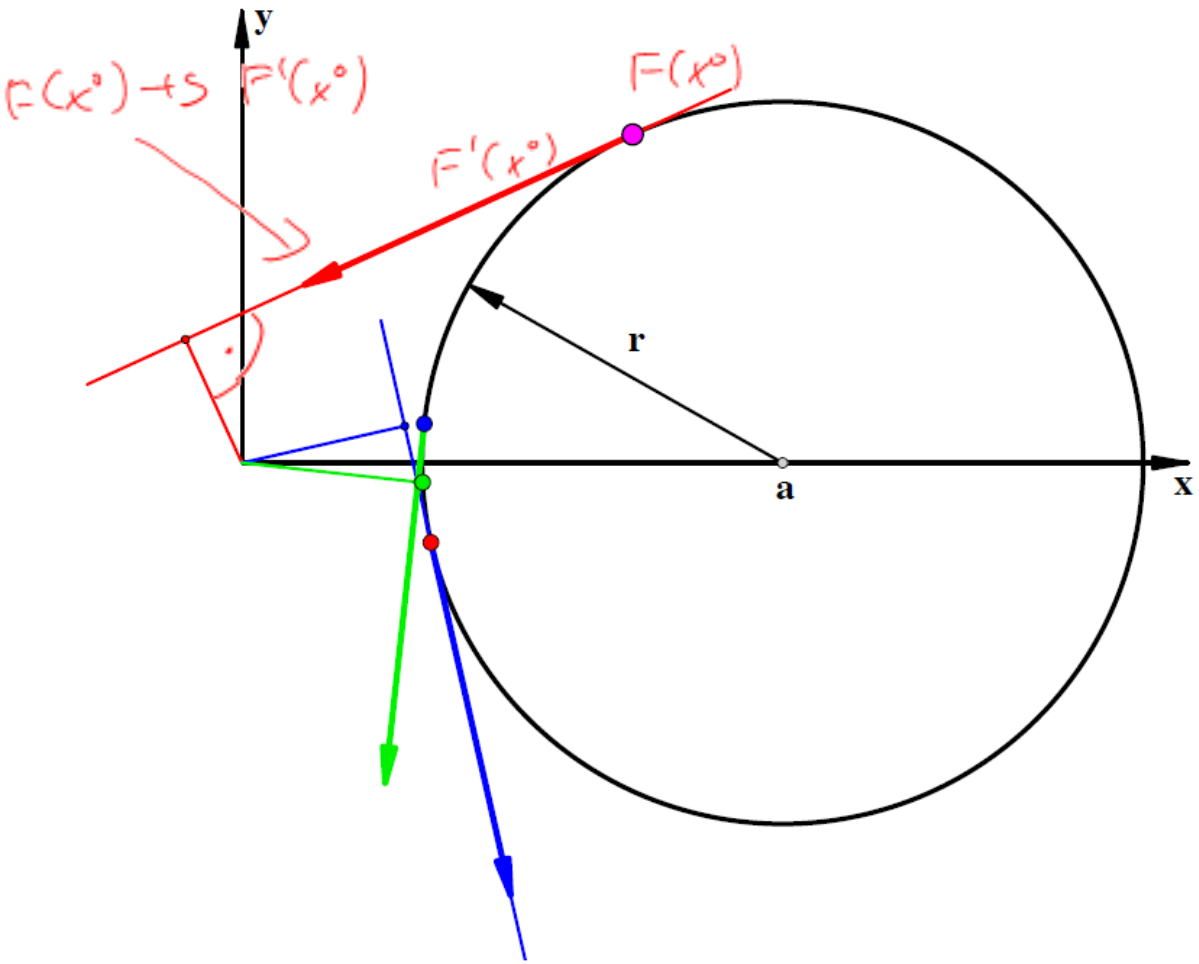
   Falls $a = r$ liegt sogar quadratische Konvergenz vor.

   oder
3. das lokale Minimum ist auch abstoßend, falls $a > 2r$.

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

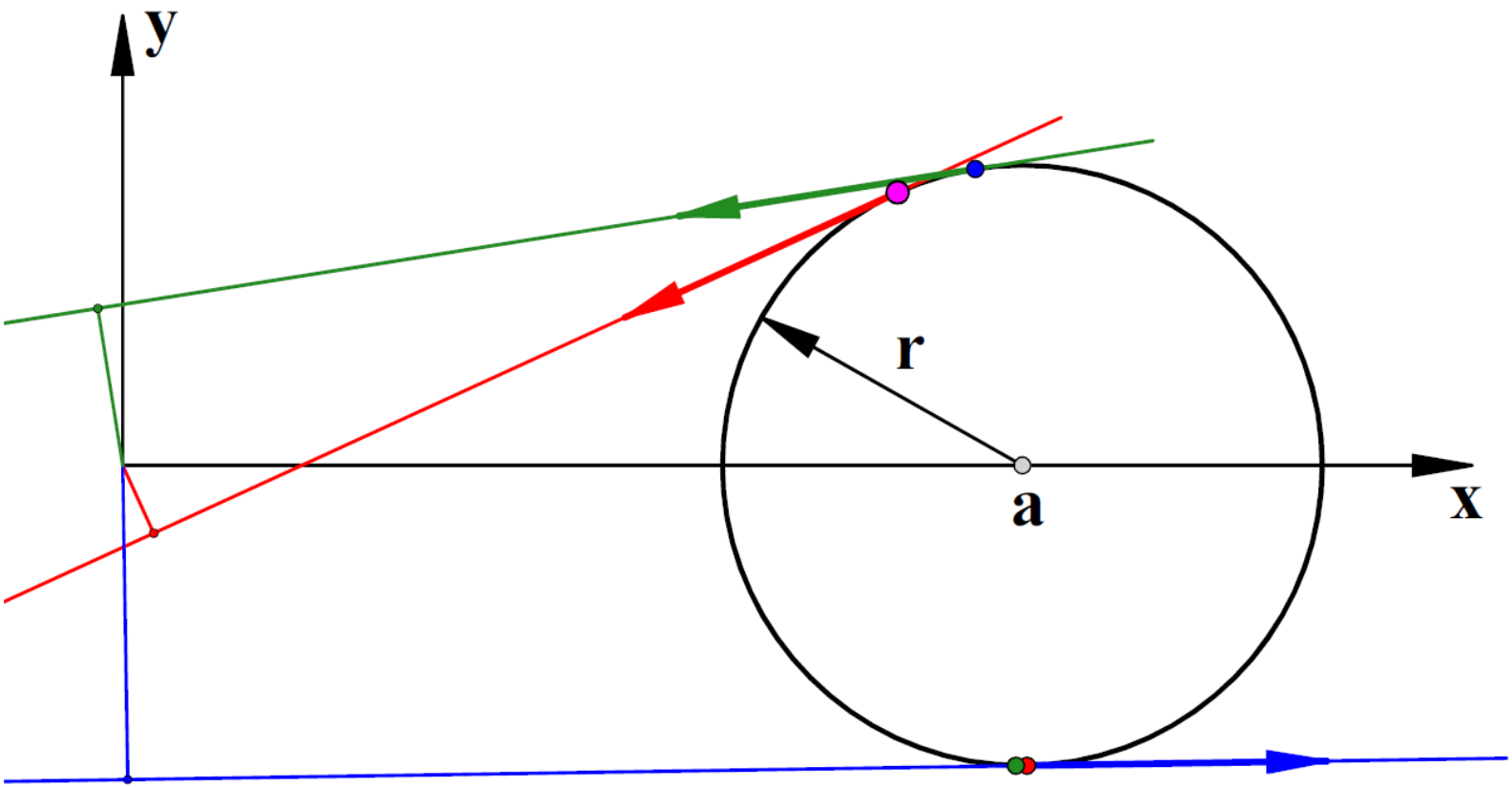
Beispiel 6.4

Konvergenter Fall: $a < 2r$



Beispiel 6.4

Divergenter Fall: $a > 2r$



Folgerung 6.6

Für die allgemeine Gauß-Newton-Iterationsfunktion Φ gilt

$$\|\Phi'(x^*)\|_A = \rho(K)\,\|F(x^*)\|_2\ ,$$

$$\|\Phi'(x^*)\| \geq \rho(K)\,\|F(x^*)\|_2,$$

mit $A^2 = F'(x^*)^T F'(x^*)$, $\|\Phi'(x^*)\|_A = \|A\Phi'(x^*)A^{-1}\|$ und

$$K := A^{-1}\left(\sum_{i=1}^{m} \frac{F_i(x^*)}{\|F(x^*)\|_2}\,F_i''(x^*)\right)\,A^{-1}$$

für jede Operatornorm $\|\cdot\|$.

Im Normalfall ist $F(x^*) \neq 0,\ K \neq 0$ und deshalb $\Phi'(x^*) \neq 0$.

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

Folgerung 6.6

Wenn der kritische Punkt x^* ein lokales Maximum von ϕ oder ein Sattelpunkt ist, gilt

$$\rho(K)\,\|F(x^*)\|_2 \geq 1$$

und deshalb

$$\|\Phi'(x^*)\| \geq 1$$

für jede Operatornorm $\|\cdot\|$.

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

Die Größe $\rho(K)\,\|F(x^*)\|_2$ 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

$$\rho(K)\,\|F(x^*)\|_2 < 1.$$

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

$$\rho(K)\,\|F(x^*)\|_2 > 1.$$

Dann ist $\|\Phi'(x^*)\| > 1$ 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	0.3504832688821768	1.45e-01	-
1	0.3425956295622140	1.35e-01	0.93
2	0.2226700964271932	4.87e-02	0.36
3	0.1704909542025844	1.04e-01	2.13
4	0.0921417452120145	1.86e-02	0.18
5	0.0891134468759020	1.21e-03	0.07
6	0.0890416386791181	3.94e-04	0.32
7	0.0890360284049988	1.20e-04	0.31
8	0.0890354511735490	4.24e-05	0.35
9	0.0890353846566892	1.44e-05	0.34
10	0.0890353766986606	5.06e-06	0.35
11	0.0890353757375046	1.75e-06	0.35
12	0.0890353756210519	6.11e-07	0.35

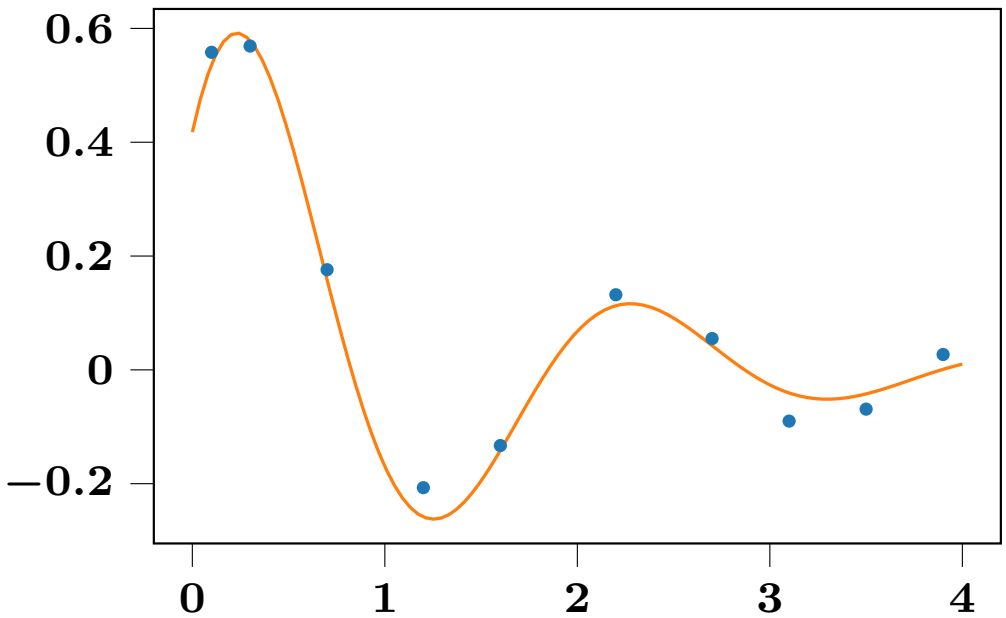
In der letzten Spalte der Tabelle sieht man das lineare Konvergenzverhalten des Gauß-Newton-Verfahrens.

Beispiel 6.7

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

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

die im folgenden Plot dargestellt ist.



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 $\boldsymbol{F}'\left(\boldsymbol{x}^{\boldsymbol{k}}\right)$ kann Rang $< \boldsymbol{n}$ haben.

Verständnisfragen

Es sei $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^m$ mit $\boldsymbol{m} > \boldsymbol{n}$.

Wir betrachten das (nichtlineare) Ausgleichsproblem:

$$\text{Bestimme } x^* \in \mathbb{R}^n \text{ so, dass } x^* = \underset{x \in \mathbb{R}^n}{\arg\min} \|F(x)\|_2.$$

f Die Gauß-Newton-Methode ist immer konvergent in einer hinreichend kleinen Umgebung von x^*.

w Die Gauß-Newton-Methode kann man als Fixpunktiteration darstellen.

f Die Konvergenzordnung der Gauß-Newton-Methode ist in der Regel 2.