

Kapitel 5

Buch Dahmen-Reusken

RWTH Aachen University

2022

Motivation

1. Die meisten Probleme in der Praxis führen auf **nichtlineare** Gleichungssysteme
2. Je genauer das (mathematische) Modell ist, desto eher ist es nichtlinear:

- ▶ Pendelschwingung: Auslenkungswinkel φ beschrieben durch

$$\ddot{\varphi}(t) + \frac{g}{\ell}\varphi(t) = 0 \quad \text{vs.} \quad \ddot{\varphi}(t) + \frac{g}{\ell}\sin(\varphi(t)) = 0$$

für kleine vs. große Auslenkungen.

Motivation

1. Die meisten Probleme in der Praxis führen auf **nichtlineare** Gleichungssysteme
2. Je genauer das (mathematische) Modell ist, desto eher ist es nichtlinear:

- ▶ Pendelschwingung: Auslenkungswinkel φ beschrieben durch

$$\ddot{\varphi}(t) + \frac{g}{\ell}\varphi(t) = 0 \quad \text{vs.} \quad \ddot{\varphi}(t) + \frac{g}{\ell}\sin(\varphi(t)) = 0$$

für kleine vs. große Auslenkungen.

- ▶ Lineare vs. nichtlineare Diffusion: Temperatur u beschrieben durch

$$u_t = \operatorname{div}(\mathbf{k} \nabla u) \quad \text{vs.} \quad u_t = \operatorname{div}(\mathbf{k}(u) \nabla u)$$

mit Wärmeleitfähigkeit $\mathbf{k}(u) = c_1 + c_2 u + c_3 u^3$.

Motivation

1. Die meisten Probleme in der Praxis führen auf **nichtlineare** Gleichungssysteme
2. Je genauer das (mathematische) Modell ist, desto eher ist es nichtlinear:

- ▶ Pendelschwingung: Auslenkungswinkel φ beschrieben durch

$$\ddot{\varphi}(t) + \frac{g}{\ell}\varphi(t) = 0 \quad \text{vs.} \quad \ddot{\varphi}(t) + \frac{g}{\ell}\sin(\varphi(t)) = 0$$

für kleine vs. große Auslenkungen.

- ▶ Lineare vs. nichtlineare Diffusion: Temperatur u beschrieben durch

$$u_t = \operatorname{div}(\mathbf{k} \nabla u) \quad \text{vs.} \quad u_t = \operatorname{div}(\mathbf{k}(u) \nabla u)$$

mit Wärmeleitfähigkeit $\mathbf{k}(u) = c_1 + c_2 u + c_3 u^3$.

- ▶ Strömungsprobleme, Netzwerkanalyse, ...

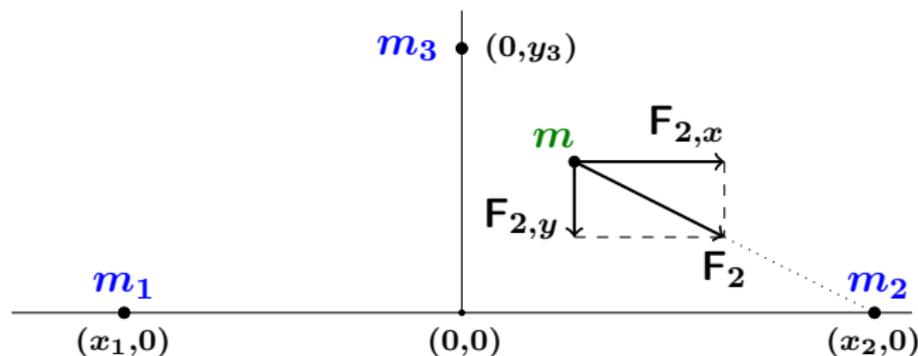
Beispiel 5.1

Für die Gravitationskraft zwischen zwei Punktmassen M_1 und M_2 mit gegenseitigem Abstand r gilt (Newtons Gravitationsgesetz):

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

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

Wir betrachten das folgende Gravitationsfeld:



Beispiel 5.1

Gesucht: Punkt (x, y) , so dass für eine Punktmasse m an der Stelle (x, y) die Gravitationskräfte F_i zwischen m und m_i , $i = 1, 2, 3$, im Gleichgewicht sind.

Hilfsgrößen mit $i = 1, 2, 3$, sind

$$r_i := \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

$$F_i := G \frac{m_i m}{r_i^2}$$

$$F_{i,x} := \frac{F_i(x_i - x)}{r_i}$$

$$F_{i,y} := \frac{F_i(y_i - y)}{r_i}$$

Beispiel 5.1

Die Gleichgewichtsbedingungen sind wie folgt:

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

Beispiel 5.1

Die Gleichgewichtsbedingungen sind wie folgt:

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

Hieraus ergibt sich das System

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

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

Beispiel 5.2

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**

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

erfüllt.

Beispiel 5.2

Das Problem wird, wie in Beispiel 3.7, auf dem Gitter

$$t_j = \left(j - \frac{1}{2}\right) h, \quad j = 1, \dots, n, \quad h = \frac{1}{n},$$

diskretisiert.

Man erhält dann die Gleichungen

$$u_i + h \sum_{j=1}^n \cos(t_i t_j) u_j^3 = 2, \quad i = 1, \dots, n,$$

für die Unbekannten $u_i \approx u(t_i)$, $i = 1, \dots, n$.

Problemstellung

Aufgabe

Zu gegebenem $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ bestimme $x^* = \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}$,

so dass

$$\begin{array}{rcl} f_1(x_1^*, \dots, x_n^*) & = & 0 \\ \vdots & & \vdots \\ f_n(x_1^*, \dots, x_n^*) & = & 0 \end{array}$$

erfüllt ist.

Kompakte Darstellung:

$$f(x^*) = 0$$

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$;

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$;

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

- ▶ **Lineare** Gleichungssysteme: Sonderfall dieser Problemstellung

$$A x^* = b \quad \Leftrightarrow \quad f(x^*) = A x^* - b = 0.$$

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$;

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

- ▶ **Lineare** Gleichungssysteme: Sonderfall dieser Problemstellung

$$A x^* = b \quad \Leftrightarrow \quad f(x^*) = A x^* - b = 0.$$

- ▶ Der Spezialfall $n = 1$ wird oft als **skalare** Gleichung in **einer** Unbekannten bezeichnet.

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$;

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

- ▶ **Lineare** Gleichungssysteme: Sonderfall dieser Problemstellung

$$A x^* = b \quad \Leftrightarrow \quad f(x^*) = A x^* - b = 0.$$

- ▶ Der Spezialfall $n = 1$ wird oft als **skalare** Gleichung in **einer** Unbekannten bezeichnet.
- ▶ Hat man mehr (nichtlineare) Gleichungen als Unbekannte, d.h.

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m \text{ mit } m > n$$

erhält man ein **nichtlineares Ausgleichsproblem**

↪ siehe nächstes Kapitel.

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

Vorgehen: **iterative** Lösungsverfahren, d.h. schrittweise Annäherung an Lösung, bis gewünschte Genauigkeit erreicht ist.

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

Vorgehen: **iterative** Lösungsverfahren, d.h. schrittweise Annäherung an Lösung, bis gewünschte Genauigkeit erreicht ist.

Fragen/Probleme:

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

Vorgehen: **iterative** Lösungsverfahren, d.h. schrittweise Annäherung an Lösung, bis gewünschte Genauigkeit erreicht ist.

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

Vorgehen: **iterative** Lösungsverfahren, d.h. schrittweise Annäherung an Lösung, bis gewünschte Genauigkeit erreicht ist.

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?
- ▶ Unter welchen Bedingungen konvergiert das Verfahren?

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

Vorgehen: **iterative** Lösungsverfahren, d.h. schrittweise Annäherung an Lösung, bis gewünschte Genauigkeit erreicht ist.

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?
- ▶ Unter welchen Bedingungen konvergiert das Verfahren?
- ▶ Wie schnell konvergiert das Verfahren?

Problemstellung

Aufgabe

Gegeben: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

Gesucht: $x^* \in \mathbb{R}^n$, so dass $f(x^*) = 0$.

Problem: analytische Lösung i.A. nicht möglich, d.h. **exakte Lösung in einer endlichen Anzahl von Schritten nicht möglich.**

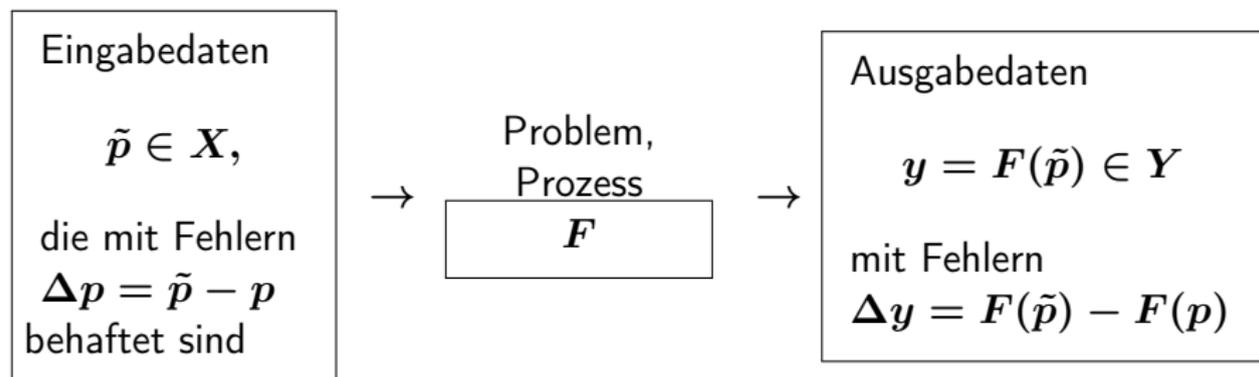
Vorgehen: **iterative** Lösungsverfahren, d.h. schrittweise Annäherung an Lösung, bis gewünschte Genauigkeit erreicht ist.

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?
- ▶ Unter welchen Bedingungen konvergiert das Verfahren?
- ▶ Wie schnell konvergiert das Verfahren?
- ▶ Ist die gewünschte Genauigkeit erreicht?

Kondition eines Nullstellenproblems

Absolute Kondition:



Struktur: $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ durch einen Satz $p = (p_1, \dots, p_m)^T \in \mathbb{R}^m$ von **Parametern** vollständig beschreibbar: $f(x) = f(x; p)$.

Schwierigkeit: eine **explizite Vorschrift** $F : p \rightarrow x^*$ nicht vorhanden.

Kondition eines Nullstellenproblems

Beispiel 5.6: Nullstelle eines Polynoms ($x \in \mathbb{R}$, $p \in \mathbb{R}^{n+1}$)

Die Eingabeparameter sind $p = (p_0, \dots, p_n) \in \mathbb{R}^{n+1}$ und

$$f : \mathbb{R} \times \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad f(x; p) := \sum_{i=0}^n p_i x^i =: P(x).$$

Es gilt: $f(x; p) = 0 \Leftrightarrow P(x) = 0$.

Kondition eines Nullstellenproblems

Beispiel 5.6: Nullstelle eines Polynoms ($x \in \mathbb{R}$, $p \in \mathbb{R}^{n+1}$)

Die Eingabeparameter sind $p = (p_0, \dots, p_n) \in \mathbb{R}^{n+1}$ und

$$f : \mathbb{R} \times \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad f(x; p) := \sum_{i=0}^n p_i x^i =: P(x).$$

Es gilt: $f(x; p) = 0 \Leftrightarrow P(x) = 0$.

Frage: wie hängt $\|x^* - \tilde{x}^*\|$ von $\|p - \tilde{p}\|$ ab?

Annahme: $\det(D_x f(x^*; p)) \neq 0$ (d.h., x^* ist eine einfache Nullstelle).

Mit Taylorentwicklung:

$$\|x^* - \tilde{x}^*\|_{\mathbb{R}^n} \leq \kappa_{\text{abs}}(x^*, p) \|p - \tilde{p}\|_{\mathbb{R}^m}$$

$$\text{mit } \kappa_{\text{abs}}(x^*, p) := \|(D_x f(x^*; p))^{-1} D_p f(x^*; p)\|_{\mathbb{R}^m \rightarrow \mathbb{R}^n}.$$

Kondition eines Nullstellenproblems: Beispiel 5.10

Für $f(x; p) = \sum_{i=0}^n p_i x^i = P(x)$ erhält man

$$D_x f(x^*; p) = P'(x^*),$$

$$D_p f(x^*; p) = (0 \quad x^* \quad (x^*)^2 \quad \dots \quad (x^*)^n).$$

Mit der Norm $\|\cdot\|_1$ in \mathbb{R}^{n+1} , ergibt sich

$$|x^* - \tilde{x}^*| \leq |P'(x^*)|^{-1} \max_{1 \leq j \leq n} |x^*|^j \sum_{j=0}^n |p_j - \tilde{p}_j|.$$

Kondition eines Nullstellenproblems: Beispiel 5.10

Für $f(x; p) = \sum_{i=0}^n p_i x^i = P(x)$ erhält man

$$D_x f(x^*; p) = P'(x^*),$$

$$D_p f(x^*; p) = (0 \quad x^* \quad (x^*)^2 \quad \dots \quad (x^*)^n).$$

Mit der Norm $\|\cdot\|_1$ in \mathbb{R}^{n+1} , ergibt sich

$$|x^* - \tilde{x}^*| \leq |P'(x^*)|^{-1} \max_{1 \leq j \leq n} |x^*|^j \sum_{j=0}^n |p_j - \tilde{p}_j|.$$

Annahme: $P'(x^*) \neq 0$, also x^* ist eine einfache Nullstelle.

Man erwartet eine **schlechte absolute Kondition** in Fällen, in denen

$\kappa_{\text{abs}}(x^*, p) = |P'(x^*)|^{-1} \max_{1 \leq j \leq n} |x^*|^j$ groß ist.

Kondition bei mehrfachen Nullstellen

Annahme: $\det(D_x f(x^*; p)) = 0$ (mehrfache Nullstelle).

Wir betrachten $n = 1$, also $f : \mathbb{R} \rightarrow \mathbb{R}$.

Sei $m \geq 1$, die Vielfachheit der Nullstelle x^* :

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

Kondition bei mehrfachen Nullstellen

Annahme: $\det(D_x f(x^*; p)) = 0$ (mehrfache Nullstelle).

Wir betrachten $n = 1$, also $f : \mathbb{R} \rightarrow \mathbb{R}$.

Sei $m \geq 1$, die Vielfachheit der Nullstelle x^* :

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

Mit Taylorentwicklung:

$$|\tilde{x}^* - x^*| \leq \left(m! \frac{\|D_p f(x^*; p)\|_\infty}{|f^{(m)}(x^*)|} \right)^{\frac{1}{m}} \|\tilde{p} - p\|_1^{\frac{1}{m}}.$$

Kondition bei mehrfachen Nullstellen

Annahme: $\det(D_x f(x^*; p)) = 0$ (mehrfache Nullstelle).

Wir betrachten $n = 1$, also $f : \mathbb{R} \rightarrow \mathbb{R}$.

Sei $m \geq 1$, die Vielfachheit der Nullstelle x^* :

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

Mit Taylorentwicklung:

$$|\tilde{x}^* - x^*| \leq \left(m! \frac{\|D_p f(x^*; p)\|_\infty}{|f^{(m)}(x^*)|} \right)^{\frac{1}{m}} \|\tilde{p} - p\|_1^{\frac{1}{m}}.$$

- ▶ Ein Datenfehler $\|\tilde{p} - p\|_1 = \epsilon \ll 1$ kann wegen des Faktors $\epsilon^{\frac{1}{m}}$ enorm verstärkt werden.
- ▶ Probleme mit mehrfachen Nullstellen sind im Allgemeinen hinsichtlich Störungen in den Eingabedaten **sehr schlecht konditioniert**

Beispiel 5.12

Das Polynom

$$f(x; p) = \sum_{i=0}^3 p_i x^i = x^3 - 3x^2 + 3x - 1$$

hat eine **dreifache Nullstelle** $x^* = 1$.

Wir betrachten eine Störung (nur) des Eingabeparameters p_0 :

$$\tilde{p}_0 = p_0 - \epsilon = -1 - \epsilon, \quad 0 < \epsilon \ll 1.$$

Beispiel 5.12

Das Polynom

$$f(x; p) = \sum_{i=0}^3 p_i x^i = x^3 - 3x^2 + 3x - 1$$

hat eine dreifache Nullstelle $x^* = 1$.

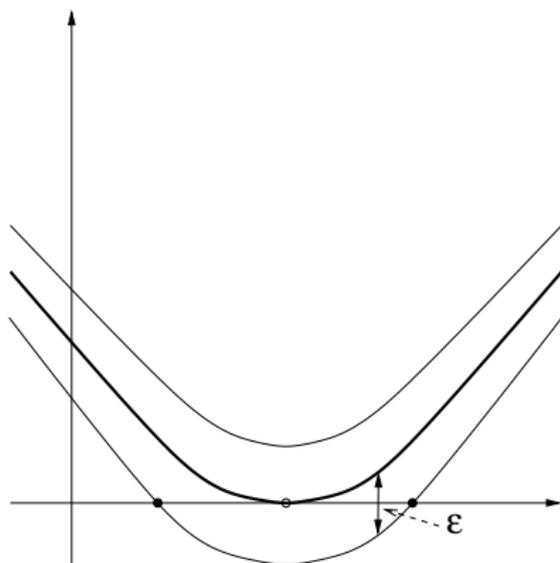
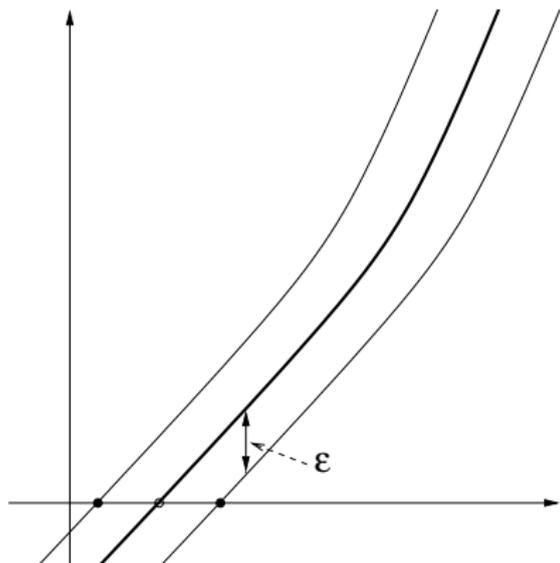
Wir betrachten eine Störung (nur) des Eingabeparameters p_0 :

$$\tilde{p}_0 = p_0 - \epsilon = -1 - \epsilon, \quad 0 < \epsilon \ll 1.$$

Es gilt

$$f(\tilde{x}^*; \tilde{p}) = 0 \Leftrightarrow f(\tilde{x}^*; p) - \epsilon = 0 \Leftrightarrow (\tilde{x}^* - 1)^3 - \epsilon = 0 \Leftrightarrow \tilde{x}^* = 1 + \epsilon^{\frac{1}{3}}.$$

Zum Beispiel für $\epsilon = 10^{-12}$ ergibt sich $|x^* - \tilde{x}^*| = 10^{-4}$.

Kondition bei mehrfachen Nullstellen: $f : \mathbb{R} \rightarrow \mathbb{R}$ 

Unvermeidbarer Fehler aufgrund der Auswertung

Sei $\tilde{f}(x)$ die mit (Rundungs)fehlern behaftete Auswertung der Funktion f an der Stelle x .

Ein auf den \tilde{f} -Werten basiertes Verfahren kann **bestenfalls** eine Annäherung \tilde{x}^* bestimmen, die $\tilde{f}(\tilde{x}^*) = \mathbf{0}$ erfüllt.

Wir verwenden Rückwärtsfehleranalyse:

Die Auswertung von f ist **rückwärtsstabil** wenn Folgendes gilt:

$$\tilde{f}(x) \doteq f(x; \tilde{p}) \quad \text{und} \quad \|\tilde{p} - p\|_1 \leq c_{aus} \text{eps} \|p\|_1$$

mit einer "akzeptablen" Konstante c_{aus} .

Beispiel 5.13

Polynomauswertung

$$f(x; p) = \sum_{i=0}^n p_i x^i$$

Annahme (der Einfachheit halber): Koeffizienten p_i sind Maschinenzahlen.

Auswertung: die Potenzen x^i , $i = 0, \dots, n$, werden gebildet und anschliessend die Summe $\sum_{i=0}^n$ rückwärts (Anfang bei $i = n$) berechnet.

Fehleranalyse:

$$\tilde{f}(x) \doteq \sum_{i=0}^n p_i (1 + \delta_i) x^i, \quad |\delta_i| \leq (n + 1) \text{eps.}$$

Beispiel 5.13

Polynomauswertung

$$f(x; p) = \sum_{i=0}^n p_i x^i$$

Annahme (der Einfachheit halber): Koeffizienten p_i sind Maschinenzahlen.

Auswertung: die Potenzen x^i , $i = 0, \dots, n$, werden gebildet und anschliessend die Summe $\sum_{i=0}^n$ rückwärts (Anfang bei $i = n$) berechnet.

Fehleranalyse:

$$\tilde{f}(x) \doteq \sum_{i=0}^n p_i (1 + \delta_i) x^i, \quad |\delta_i| \leq (n + 1) \text{eps.}$$

Folgerung:

Wegen $\|\tilde{p} - p\|_1 \leq (n + 1) \text{eps} \|p\|_1$ ist für nicht allzugroße Werte von n ist diese **Polynomauswertung rückwärtsstabil**.

Auswertung auf einem Rechner (Bemerkung 5.14)

Sei f eine skalare stetige Funktion mit einer lokal eindeutigen Nullstelle:

$$f(x) = 0 \text{ für } x \in (a, b) \iff x = x^*.$$

Auf einem Rechner (Maschinenzahlen \mathbb{M}):

$$\tilde{f} : (a, b) \rightarrow \mathbb{M}.$$

Auswertung auf einem Rechner (Bemerkung 5.14)

Sei f eine skalare stetige Funktion mit einer lokal eindeutigen Nullstelle:

$$f(x) = 0 \text{ für } x \in (a, b) \iff x = x^*.$$

Auf einem Rechner (Maschinenzahlen \mathbb{M}):

$$\tilde{f} : (a, b) \rightarrow \mathbb{M}.$$

Die Auswertung \tilde{f} ist stückweise konstant:

Es sei $\hat{x} \in \mathbb{M}$, dann gilt $\tilde{f}(x) = \tilde{f}(\hat{x})$ für alle $x \in (a, b)$ für die $\text{fl}(x) = \hat{x}$ gilt.

Auswertung auf einem Rechner (Bemerkung 5.14)

Sei f eine skalare stetige Funktion mit einer lokal eindeutigen Nullstelle:

$$f(x) = 0 \text{ für } x \in (a, b) \iff x = x^*.$$

Auf einem Rechner (Maschinenzahlen \mathbb{M}):

$$\tilde{f} : (a, b) \rightarrow \mathbb{M}.$$

Die Auswertung \tilde{f} ist stückweise konstant:

Es sei $\hat{x} \in \mathbb{M}$, dann gilt $\tilde{f}(x) = \tilde{f}(\hat{x})$ für alle $x \in (a, b)$ für die $\text{fl}(x) = \hat{x}$ gilt.

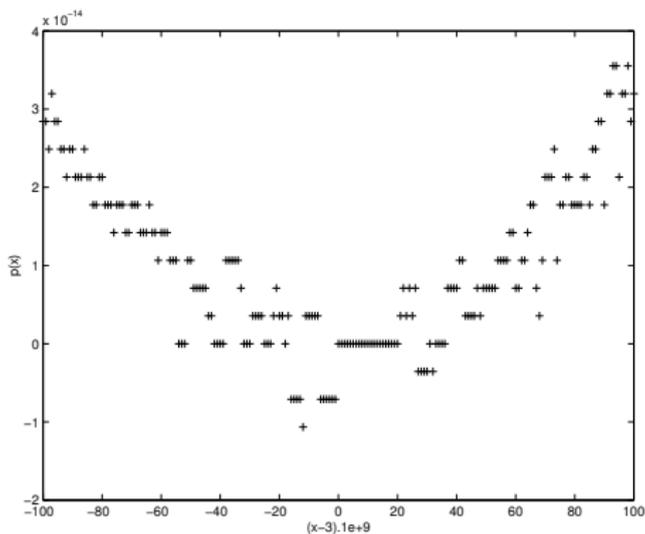
$$\tilde{f}(x) = 0 \text{ für } x \in \mathbb{M}:$$

möglicherweise keine Lösung, oder (sehr) viele Lösungen.

Beispiel 5.15

Sei $P(x) = x^3 - 6x^2 + 9x$ (doppelte Nullstelle $x^* = 3$).

Auswertungen: $P(3 + i * 10^{-9})$, $i = -100, -99, \dots, 99, 100$.

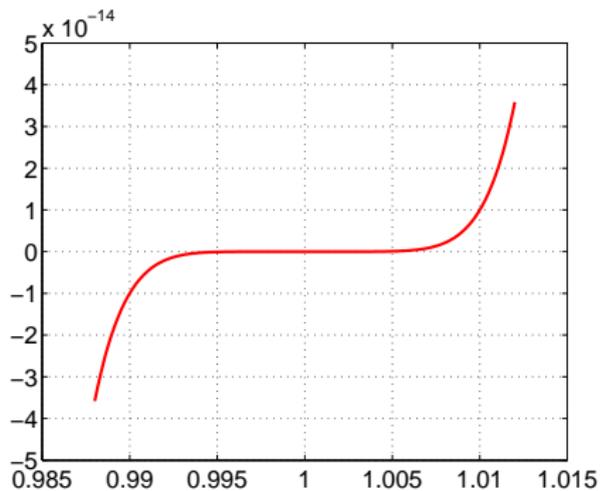


\tilde{P} hat **viele Nullstellen** im Intervall $[3 - 10^{-7}, 3 + 10^{-7}]$

Beispiel: Polynom 7. Grades

Matlab Plot

```
x = 0.988:0.0001:1.012;  
y = (x-1).^7;  
plot(x,y)
```



Eine mehrfache Nullstelle

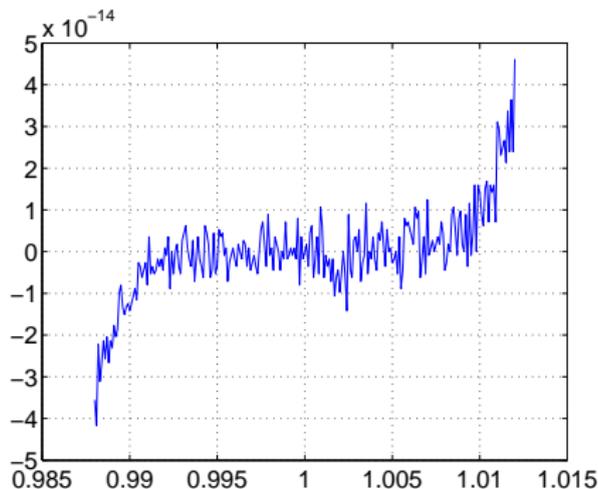
Beispiel: Polynom 7. Grades

Matlab Plot

```
x = 0.988:0.0001:1.012;
```

```
y = x.^7-7*x.^6+21*x.^5-35*x.^4+35*x.^3-21*x.^2+7*x-1;
```

```
plot(x,y)
```



Viele Nullstellen

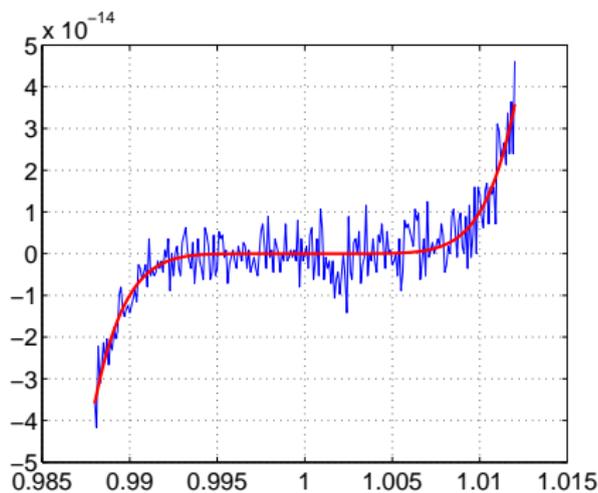
Beispiel: Polynom 7. Grades

Matlab Plot

```
x = 0.988:0.0001:1.012;
```

```
y = x.^7-7*x.^6+21*x.^5-35*x.^4+35*x.^3-21*x.^2+7*x-1;
```

```
plot(x,y)
```



Viele Nullstellen

Fixpunktiteration

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?

Allgemeiner Ansatz für Fixpunktiteration:

Fixpunktiteration

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?

Allgemeiner Ansatz für Fixpunktiteration:

- ▶ Sei $M_x \in \mathbb{R}^{n \times n}$ eine von x abhängige Matrix, die in einer Umgebung der Nullstelle x^* invertierbar ist. Dann folgt

$$f(x^*) = 0 \quad \iff \quad M_{x^*} f(x^*) = 0$$

Fixpunktiteration

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?

Allgemeiner Ansatz für Fixpunktiteration:

- ▶ Sei $M_x \in \mathbb{R}^{n \times n}$ eine von x abhängige Matrix, die in einer Umgebung der Nullstelle x^* invertierbar ist. Dann folgt

$$f(x^*) = 0 \iff M_{x^*} f(x^*) = 0$$

- ▶ Erweitere die Gleichung mit x^* , d.h.

$$M_{x^*} f(x^*) = 0 \iff x^* = x^* - M_{x^*} f(x^*)$$

Fixpunktiteration

Fragen/Probleme:

- ▶ Wie finden wir ein geeignetes iteratives Verfahren?

Allgemeiner Ansatz für Fixpunktiteration:

- ▶ Sei $M_x \in \mathbb{R}^{n \times n}$ eine von x abhängige Matrix, die in einer Umgebung der Nullstelle x^* invertierbar ist. Dann folgt

$$f(x^*) = 0 \iff M_{x^*} f(x^*) = 0$$

- ▶ Erweitere die Gleichung mit x^* , d.h.

$$M_{x^*} f(x^*) = 0 \iff x^* = x^* - M_{x^*} f(x^*)$$

Daraus folgt: Das **Nullstellenproblem**

$$f(x^*) = 0$$

ist **äquivalent** zum **Fixpunktproblem**

$$x^* = \Phi(x^*), \quad \text{mit} \quad \Phi(x) := x - M_x f(x).$$

Fixpunktiteration

Fixpunktiteration

- ▶ Wähle Startwert x_0 in einer Umgebung von x^*
- ▶ Bilde

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, 2, \dots$$

Fixpunktiteration

Fixpunktiteration

- ▶ Wähle Startwert x_0 in einer Umgebung von x^*
- ▶ Bilde

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, 2, \dots$$

Bemerkungen:

1. $\Phi : \mathbb{R} \rightarrow \mathbb{R}$. Die **Steigung von Φ an x^*** entscheidet darüber, ob die Fixpunktiteration gegen x^* konvergiert/divergiert:
 - ▶ $|\Phi'(x^*)| < 1$: x^* anziehend
 - ▶ $|\Phi'(x^*)| > 1$: x^* abstoßend

Fixpunktiteration

Fixpunktiteration

- ▶ Wähle Startwert x_0 in einer Umgebung von x^*
- ▶ Bilde

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, 2, \dots$$

Bemerkungen:

1. $\Phi : \mathbb{R} \rightarrow \mathbb{R}$. Die **Steigung von Φ an x^*** entscheidet darüber, ob die Fixpunktiteration gegen x^* konvergiert/divergiert:
 - ▶ $|\Phi'(x^*)| < 1$: x^* anziehend
 - ▶ $|\Phi'(x^*)| > 1$: x^* abstoßend
2. Durch eine geeignete Wahl von M_x (bzw. Φ) lässt sich die Konvergenz der Fixpunktiteration positiv beeinflussen.

Ein paar Definitionen

Lipschitz-Stetigkeit

Sei $E \subseteq \mathbb{R}^n$. Eine Funktion

$$f : E \rightarrow \mathbb{R}^n$$

heißt **Lipschitz-stetig** auf E , wenn eine Konstante L existiert, so dass

$$\|f(x) - f(y)\| \leq L \|x - y\|$$

für alle $x, y \in E$.

Ein paar Definitionen

Kontraktion

Sei $E \subseteq \mathbb{R}^n$. Eine Abbildung $\Phi : E \rightarrow \mathbb{R}^n$ heißt **Kontraktion** auf E , wenn

$$\|\Phi(x) - \Phi(y)\| \leq L \|x - y\|$$

für alle $x, y \in E$ mit $L < 1$.

Ein paar Definitionen

Kontraktion

Sei $E \subseteq \mathbb{R}^n$. Eine Abbildung $\Phi : E \rightarrow \mathbb{R}^n$ heißt **Kontraktion** auf E , wenn

$$\|\Phi(x) - \Phi(y)\| \leq L \|x - y\|$$

für alle $x, y \in E$ mit $L < 1$.

- ▶ Φ ist genau dann eine Kontraktion, wenn sie Lipschitz-stetig mit der Konstanten $L \in [0, 1)$ ist.

Selbstabbildung

Eine Abbildung Φ ist eine **Selbstabbildung** auf $E \subset \mathbb{R}^n$, wenn

$$\Phi : E \rightarrow E.$$

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

- ▶ Die Funktion f hat eine eindeutige positive Nullstelle x^* und es gilt $x^* \in [1, 2]$.

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

- ▶ Die Funktion f hat eine eindeutige positive Nullstelle x^* und es gilt $x^* \in [1, 2]$.
- ▶ Mögliche Fixpunktfunktionen sind

$$\Phi_1(x) := x^6 - 1 \quad \text{oder} \quad \Phi_2(x) := (x + 1)^{\frac{1}{6}}.$$

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

- ▶ Die Funktion f hat eine eindeutige positive Nullstelle x^* und es gilt $x^* \in [1, 2]$.
- ▶ Mögliche Fixpunktfunktionen sind

$$\Phi_1(x) := x^6 - 1 \quad \text{oder} \quad \Phi_2(x) := (x + 1)^{\frac{1}{6}}.$$

- ▶ Betrachte $\Phi_1(x)$: wir erhalten

$$|\Phi_1'(x)| = |6x^5| > 1 \quad \text{für } x \in [1, 2],$$

d.h. $\Phi_1(x)$ ist **nicht als Fixpunktfunktion geeignet**.

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

- ▶ Die Funktion f hat eine eindeutige positive Nullstelle x^* und es gilt $x^* \in [1, 2]$.
- ▶ Mögliche Fixpunktfunktionen sind

$$\Phi_1(x) := x^6 - 1 \quad \text{oder} \quad \Phi_2(x) := (x + 1)^{\frac{1}{6}}.$$

- ▶ Betrachte $\Phi_1(x)$: wir erhalten

$$|\Phi_1'(x)| = |6x^5| > 1 \quad \text{für } x \in [1, 2],$$

d.h. $\Phi_1(x)$ ist **nicht als Fixpunktfunktion geeignet**.

- ▶ Betrachte $\Phi_2(x)$: wir erhalten

$$|\Phi_2'(x)| = \left| \frac{1}{6}(x + 1)^{-\frac{5}{6}} \right| \leq \frac{1}{6} \quad \text{für } x \in [1, 2]$$

Beispiel 5.20

und damit (Mittelwertsatz, $\xi \in (1, 2)$)

$$\begin{aligned} |\Phi_2(x) - \Phi_2(y)| &= |\Phi_2'(\xi)(x - y)| \\ &\leq \frac{1}{6} |x - y| \quad \text{für } x, y \in [1, 2]. \end{aligned}$$

Beispiel 5.20

und damit (Mittelwertsatz, $\xi \in (1, 2)$)

$$\begin{aligned} |\Phi_2(x) - \Phi_2(y)| &= |\Phi_2'(\xi)(x - y)| \\ &\leq \frac{1}{6} |x - y| \quad \text{für } x, y \in [1, 2]. \end{aligned}$$

- ▶ Die Funktion $\Phi_2(x)$ ist eine **Selbstabbildung** auf $[1, 2]$, d.h. $\Phi_2 : [1, 2] \rightarrow [1, 2]$.

Beispiel 5.20

und damit (Mittelwertsatz, $\xi \in (1, 2)$)

$$\begin{aligned} |\Phi_2(x) - \Phi_2(y)| &= |\Phi_2'(\xi)(x - y)| \\ &\leq \frac{1}{6} |x - y| \quad \text{für } x, y \in [1, 2]. \end{aligned}$$

- ▶ Die Funktion $\Phi_2(x)$ ist eine **Selbstabbildung** auf $[1, 2]$, d.h. $\Phi_2 : [1, 2] \rightarrow [1, 2]$.
- ▶ Ergebnisse

k	$x_0 = 1.2$ $x_{k+1} = \Phi_2(x_k)$	$x_0 = 1.135$ $x_{k+1} = \Phi_1(x_k)$
0	1.20000000	1.14e+000
1	1.14043476	1.14e+000
2	1.13522949	1.17e+000
3	1.13476890	1.57e+000
4	1.13472810	1.38e+001
5	1.13472448	6.91e+006
6	1.13472416	1.09e+041
7	1.13472414	1.66e+246

Banachscher Fixpunktsatz

Sei X ein linear normierter Raum und $E \subseteq X$ eine vollständige Teilmenge von X . Sei Φ eine **Selbstabbildung** auf E , d.h.

$$\Phi : E \rightarrow E,$$

und ferner eine **Kontraktion auf E** , d.h.

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \quad \text{für alle } x, y \in E,$$

mit $L < 1$.

Banachscher Fixpunktsatz

Sei X ein linear normierter Raum und $E \subseteq X$ eine vollständige Teilmenge von X . Sei Φ eine **Selbstabbildung** auf E , d.h.

$$\Phi : E \rightarrow E,$$

und ferner eine **Kontraktion auf E** , d.h.

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \quad \text{für alle } x, y \in E,$$

mit $L < 1$.

Dann gilt:

1. Es **existiert genau ein Fixpunkt x^*** von Φ in E .

Banachscher Fixpunktsatz

Sei X ein linear normierter Raum und $E \subseteq X$ eine vollständige Teilmenge von X . Sei Φ eine **Selbstabbildung** auf E , d.h.

$$\Phi : E \rightarrow E,$$

und ferner eine **Kontraktion auf E** , d.h.

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \quad \text{für alle } x, y \in E,$$

mit $L < 1$.

Dann gilt:

1. Es **existiert genau ein Fixpunkt x^*** von Φ in E .
2. Für beliebiges $x_0 \in E$ **konvergiert** die Fixpunktiteration

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, 2, \dots$$

gegen den Fixpunkt x^* .

Banachscher Fixpunktsatz

Sei X ein linear normierter Raum und $E \subseteq X$ eine vollständige Teilmenge von X . Sei Φ eine **Selbstabbildung** auf E , d.h.

$$\Phi : E \rightarrow E,$$

und ferner eine **Kontraktion auf E** , d.h.

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \quad \text{für alle } x, y \in E,$$

mit $L < 1$.

3. **A-priori-Fehlerabschätzung:**

$$\|x_k - x^*\| \leq \frac{L^k}{1 - L} \|x_1 - x_0\|.$$

Banachscher Fixpunktsatz

Sei X ein linear normierter Raum und $E \subseteq X$ eine vollständige Teilmenge von X . Sei Φ eine **Selbstabbildung** auf E , d.h.

$$\Phi : E \rightarrow E,$$

und ferner eine **Kontraktion auf E** , d.h.

$$\|\Phi(x) - \Phi(y)\| \leq L\|x - y\| \quad \text{für alle } x, y \in E,$$

mit $L < 1$.

3. **A-priori-Fehlerabschätzung:**

$$\|x_k - x^*\| \leq \frac{L^k}{1 - L} \|x_1 - x_0\|.$$

4. **A-posteriori-Fehlerabschätzung:**

$$\|x_k - x^*\| \leq \frac{L}{1 - L} \|x_k - x_{k-1}\|.$$

Bemerkungen zum Banachschen Fixpunktsatz

Fragen/Probleme:

- ▶ Unter welchen Bedingungen konvergiert iteratives Verfahren?

Bemerkungen zum Banachschen Fixpunktsatz

Fragen/Probleme:

- ▶ Unter welchen Bedingungen konvergiert iteratives Verfahren?

⇒ Banachscher Fixpunktsatz liefert **hinreichende Bedingungen**, damit

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, \dots$$

gegen einen Fixpunkt x^* konvergiert.

Bemerkungen zum Banachschen Fixpunktsatz

Fragen/Probleme:

- ▶ Unter welchen Bedingungen konvergiert iteratives Verfahren?

⇒ Banachscher Fixpunktsatz liefert **hinreichende Bedingungen**, damit

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, \dots$$

gegen einen Fixpunkt x^* konvergiert.

Fragen/Probleme:

- ▶ Ist die gewünschte Genauigkeit erreicht?

Bemerkungen zum Banachschen Fixpunktsatz

Fragen/Probleme:

- ▶ Unter welchen Bedingungen konvergiert iteratives Verfahren?

⇒ Banachscher Fixpunktsatz liefert **hinreichende Bedingungen**, damit

$$x_{k+1} = \Phi(x_k), \quad k = 0, 1, \dots$$

gegen einen Fixpunkt x^* konvergiert.

Fragen/Probleme:

- ▶ Ist die gewünschte Genauigkeit erreicht?

⇒ Wir möchten eine gewünschte Genauigkeit ϵ erreichen, so dass

$$\|x_k - x^*\| \leq \epsilon.$$

Bemerkungen zum Banachschen Fixpunktsatz

Frage

- ▶ Wie viele Iterationen müssen wir durchführen?

Bemerkungen zum Banachschen Fixpunktsatz

Frage

- ▶ Wie viele Iterationen müssen wir durchführen?

⇒ Mit Hilfe der a-priori-Fehlerabschätzung erhalten wir

$$\|x_k - x^*\| \leq \frac{L^k}{1 - L} \|x_1 - x_0\| \stackrel{!}{\leq} \epsilon.$$

Bemerkungen zum Banachschen Fixpunktsatz

Frage

- ▶ Wie viele Iterationen müssen wir durchführen?

⇒ Mit Hilfe der a-priori-Fehlerabschätzung erhalten wir

$$\|x_k - x^*\| \leq \frac{L^k}{1 - L} \|x_1 - x_0\| \stackrel{!}{\leq} \epsilon.$$

und damit ist die maximal benötigte Anzahl an Iterationen

$$k \geq \log(\epsilon(1 - L)/\|x_1 - x_0\|)/\log(L)$$

Bemerkungen zum Banachschen Fixpunktsatz

Frage

- ▶ Wie viele Iterationen müssen wir durchführen?

⇒ Mit Hilfe der a-priori-Fehlerabschätzung erhalten wir

$$\|x_k - x^*\| \leq \frac{L^k}{1 - L} \|x_1 - x_0\| \stackrel{!}{\leq} \epsilon.$$

und damit ist die maximal benötigte Anzahl an Iterationen

$$k \geq \log(\epsilon(1 - L)/\|x_1 - x_0\|)/\log(L)$$

Beachte

Wegen

$$\|x_k - x_{k-1}\| \leq L^{k-1} \|x_1 - x_0\|$$

ist die Schranke in der a-posteriori-Fehlerabschätzung immer besser (d.h. kleiner) als die in der a-priori-Fehlerabschätzung.

Folgerungen aus Banachscher Fixpunktsatz

Folgerung 5.24

Sei $X = \mathbb{R}$, $E = [a, b]$ und Φ auf E stetig differenzierbar.

Es gelte

$$\Phi : [a, b] \rightarrow [a, b] \quad (\text{Selbstabbildung}),$$

und

$$\max_{x \in [a, b]} |\Phi'(x)| =: L < 1.$$

Dann sind alle Voraussetzungen aus BF-Satz erfüllt für $\|\cdot\| = |\cdot|$

Beachte

Nach Mittelwertsatz gilt

$$|\Phi(x) - \Phi(y)| = |\Phi'(\xi)(x - y)| \leq \max_{\xi \in [a, b]} |\Phi'(\xi)| |x - y|,$$

d.h. Φ ist eine Kontraktion.

Folgerungen aus Banachscher Fixpunktsatz

Folgerung 5.25

Sei $E \subseteq X = \mathbb{R}^n$ eine abgeschlossene konvexe Menge, und $\Phi : E \rightarrow \mathbb{R}^n$ sei stetig differenzierbar. Es gelte

$$\Phi : E \rightarrow E \quad (\text{Selbstabbildung}),$$

und bzgl. einer Vektornorm $\|\cdot\|$ auf \mathbb{R}^n gelte für die zugehörige Matrixnorm

$$\max_{x \in E} \|\Phi'(x)\| = L < 1.$$

Dann sind alle Voraussetzungen aus BF-Satz erfüllt.

Hierbei ist

$$\Phi'(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \Phi_1(x) & \cdots & \frac{\partial}{\partial x_n} \Phi_1(x) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} \Phi_n(x) & \cdots & \frac{\partial}{\partial x_n} \Phi_n(x) \end{pmatrix}$$

die Jacobi-Matrix von Φ an der Stelle x .

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

Mögliche Fixpunktfunktion

$$\Phi_2(x) := (x + 1)^{\frac{1}{6}}.$$

Es gilt

$$\Phi_2 : [1, 2] \rightarrow [1, 2] \quad (\text{Selbstabbildung})$$

$$|\Phi_2'(x)| = \left| \frac{1}{6}(x + 1)^{-\frac{5}{6}} \right| \leq \frac{1}{6} \quad \text{für } x \in [1, 2].$$

Beispiel 5.20

Man berechne die positive Nullstelle der Funktion

$$f(x) := x^6 - x - 1.$$

Mögliche Fixpunktfunktion

$$\Phi_2(x) := (x + 1)^{\frac{1}{6}}.$$

Es gilt

$$\Phi_2 : [1, 2] \rightarrow [1, 2] \quad (\text{Selbstabbildung})$$

$$|\Phi_2'(x)| = \left| \frac{1}{6}(x + 1)^{-\frac{5}{6}} \right| \leq \frac{1}{6} \quad \text{für } x \in [1, 2].$$

Daraus folgt:

f hat eine **eindeutige Nullstelle** $x^* \in [1, 2]$.

Fixpunktiteration $x_{k+1} = \Phi_2(x_k)$, $x_0 \in [1, 2]$, **konvergiert gegen** x^* .

Beispiel 5.27

Zeigen Sie, dass das System

$$\begin{aligned}6x &= \cos x + 2y \\ 8y &= xy^2 + \sin x\end{aligned}$$

auf $E = [0, 1] \times [0, 1]$ eine eindeutige Lösung besitzt. Bestimmen Sie diese Lösung bis auf eine Genauigkeit 10^{-3} in der ∞ -Norm.

Beispiel 5.27

Zeigen Sie, dass das System

$$\begin{aligned}6x &= \cos x + 2y \\ 8y &= xy^2 + \sin x\end{aligned}$$

auf $E = [0, 1] \times [0, 1]$ eine eindeutige Lösung besitzt. Bestimmen Sie diese Lösung bis auf eine Genauigkeit 10^{-3} in der ∞ -Norm.

► Fixpunktfunktion:

Beispiel 5.27

Zeigen Sie, dass das System

$$\begin{aligned}6x &= \cos x + 2y \\8y &= xy^2 + \sin x\end{aligned}$$

auf $E = [0, 1] \times [0, 1]$ eine eindeutige Lösung besitzt. Bestimmen Sie diese Lösung bis auf eine Genauigkeit 10^{-3} in der ∞ -Norm.

► Fixpunktfunktion:

$$\Phi(x, y) = \begin{pmatrix} \frac{1}{6} \cos x + \frac{1}{3} y \\ \frac{1}{8} xy^2 + \frac{1}{8} \sin x \end{pmatrix}$$

Beispiel 5.27

Zeigen Sie, dass das System

$$\begin{aligned}6x &= \cos x + 2y \\8y &= xy^2 + \sin x\end{aligned}$$

auf $E = [0, 1] \times [0, 1]$ eine eindeutige Lösung besitzt. Bestimmen Sie diese Lösung bis auf eine Genauigkeit 10^{-3} in der ∞ -Norm.

► Fixpunktfunktion:

$$\Phi(x, y) = \begin{pmatrix} \frac{1}{6} \cos x + \frac{1}{3} y \\ \frac{1}{8} xy^2 + \frac{1}{8} \sin x \end{pmatrix}$$

► Selbstabbildung:

Beispiel 5.27

Zeigen Sie, dass das System

$$\begin{aligned}6x &= \cos x + 2y \\ 8y &= xy^2 + \sin x\end{aligned}$$

auf $E = [0, 1] \times [0, 1]$ eine eindeutige Lösung besitzt. Bestimmen Sie diese Lösung bis auf eine Genauigkeit 10^{-3} in der ∞ -Norm.

► Fixpunktfunktion:

$$\Phi(x, y) = \begin{pmatrix} \frac{1}{6} \cos x + \frac{1}{3} y \\ \frac{1}{8} xy^2 + \frac{1}{8} \sin x \end{pmatrix}$$

► Selbstabbildung:

Für $x \in [0, 1]$ gilt $0 \leq \cos x \leq 1$ und $0 \leq \sin x \leq 1$. Daher gilt

$$\Phi : E \rightarrow E.$$

Beispiel 5.27

► Kontraktion:

Beispiel 5.27

► Kontraktion: Die Jacobi-Matrix ist

$$\Phi'(x, y) = \begin{pmatrix} -\frac{1}{6} \sin x & \frac{1}{3} \\ \frac{1}{8} y^2 + \frac{1}{8} \cos x & \frac{1}{4} xy \end{pmatrix}.$$

Beispiel 5.27

► Kontraktion: Die Jacobi-Matrix ist

$$\Phi'(x, y) = \begin{pmatrix} -\frac{1}{6} \sin x & \frac{1}{3} \\ \frac{1}{8} y^2 + \frac{1}{8} \cos x & \frac{1}{4} xy \end{pmatrix}.$$

Damit erhält man für die ∞ -Norm auf \mathbb{R}^2

$$\begin{aligned} \|\Phi'(x, y)\|_{\infty} &= \max \left\{ \frac{1}{6} |\sin x| + \frac{1}{3}, \frac{1}{8} (|y^2 + \cos x| + 2|xy|) \right\} \\ &\leq \frac{1}{2}. \end{aligned}$$

Beispiel 5.27

► Kontraktion: Die Jacobi-Matrix ist

$$\Phi'(x, y) = \begin{pmatrix} -\frac{1}{6} \sin x & \frac{1}{3} \\ \frac{1}{8}y^2 + \frac{1}{8} \cos x & \frac{1}{4}xy \end{pmatrix}.$$

Damit erhält man für die ∞ -Norm auf \mathbb{R}^2

$$\begin{aligned} \|\Phi'(x, y)\|_\infty &= \max \left\{ \frac{1}{6} |\sin x| + \frac{1}{3}, \frac{1}{8} (|y^2 + \cos x| + 2|xy|) \right\} \\ &\leq \frac{1}{2}. \end{aligned}$$

Wegen Folgerung 5.25 existiert genau eine Lösung in E .

Beispiel 5.27

- **Kontraktion:** Die Jacobi-Matrix ist

$$\Phi'(x, y) = \begin{pmatrix} -\frac{1}{6} \sin x & \frac{1}{3} \\ \frac{1}{8} y^2 + \frac{1}{8} \cos x & \frac{1}{4} xy \end{pmatrix}.$$

Damit erhält man für die ∞ -Norm auf \mathbb{R}^2

$$\begin{aligned} \|\Phi'(x, y)\|_{\infty} &= \max \left\{ \frac{1}{6} |\sin x| + \frac{1}{3}, \frac{1}{8} (|y^2 + \cos x| + 2|xy|) \right\} \\ &\leq \frac{1}{2}. \end{aligned}$$

Wegen Folgerung 5.25 existiert genau eine Lösung in E .

- **Fehlerschätzung:**

Beispiel 5.27

- **Kontraktion:** Die Jacobi-Matrix ist

$$\Phi'(x, y) = \begin{pmatrix} -\frac{1}{6} \sin x & \frac{1}{3} \\ \frac{1}{8} y^2 + \frac{1}{8} \cos x & \frac{1}{4} xy \end{pmatrix}.$$

Damit erhält man für die ∞ -Norm auf \mathbb{R}^2

$$\begin{aligned} \|\Phi'(x, y)\|_\infty &= \max \left\{ \frac{1}{6} |\sin x| + \frac{1}{3}, \frac{1}{8} (|y^2 + \cos x| + 2|xy|) \right\} \\ &\leq \frac{1}{2}. \end{aligned}$$

Wegen Folgerung 5.25 existiert genau eine Lösung in E .

- **Fehlerschätzung:** Mit $\epsilon = 10^{-3}$ und $L = \frac{1}{2}$ benötigt man maximal

$$k \geq \log \left(\frac{0.5 \cdot 10^{-3}}{\|x_1 - x_0\|} \right) / \log \left(\frac{1}{2} \right)$$

Schritte.

Beispiel 5.27

Für den Startwert

$$(x_0, y_0) = (0, 0)$$

erhält man als 1. Iterierte

$$(x_1, y_1) = \Phi(x_0, y_0) = \begin{pmatrix} \frac{1}{6} \cos x_0 + \frac{1}{3} y_0 \\ \frac{1}{8} x_0 y_0^2 + \frac{1}{8} \sin x_0 \end{pmatrix} = \left(\frac{1}{6}, 0 \right)$$

und damit

$$k \geq \log \left(\frac{0.5 \times 10^{-3}}{1/6} \right) / \log \left(\frac{1}{2} \right) = 8.38,$$

d.h. es werden maximal 9 Iterationen benötigt.

Beispiel 5.27

Für den Startwert

$$(x_0, y_0) = (0, 0)$$

erhält man als 1. Iterierte

$$(x_1, y_1) = \Phi(x_0, y_0) = \begin{pmatrix} \frac{1}{6} \cos x_0 + \frac{1}{3} y_0 \\ \frac{1}{8} x_0 y_0^2 + \frac{1}{8} \sin x_0 \end{pmatrix} = \left(\frac{1}{6}, 0 \right)$$

und damit

$$k \geq \log \left(\frac{0.5 \times 10^{-3}}{1/6} \right) / \log \left(\frac{1}{2} \right) = 8.38,$$

d.h. es werden maximal 9 Iterationen benötigt.

Ergebnisse: Siehe folgende Tabelle.

Beispiel 5.27

k	$(x_0, y_0) = (0, 0)$ $(x_k, y_k) = \phi(x_{k-1}, y_{k-1})$	$\frac{0.5}{1-0.5}^*$ $\ (x_k, y_k) - (x_{k-1}, y_{k-1})^T\ _\infty$
0	(0.00000000, 0.00000000)	-
1	(0.16666667, 0.00000000)	1.67e-01
2	(0.16435721, 0.02073702)	2.07e-02
3	(0.17133296, 0.02046111)	6.98e-03
4	(0.17104677, 0.02132096)	8.60e-04
5	(0.17134151, 0.02128646)	2.95e-04
6	(0.17132164, 0.02132275)	3.63e-05
7	(0.17133430, 0.02132034)	1.27e-05
8	(0.17133314, 0.02132189)	1.56e-06
9	(0.17133369, 0.02132175)	5.52e-07

Aus der a-posteriori-Fehlerabschätzung ergibt sich, dass schon für $k = 4$ (statt $k = 9$) die gewünschte Genauigkeit erreicht ist.

Fixpunktiteration für lineare Probleme

Es gilt

$$Ax = b \Leftrightarrow x = x + C(b - Ax) =: \Phi(x)$$

mit einer beliebigen regulären Matrix $C \in \mathbb{R}^{n \times n}$.

Fixpunktiteration:

$$x^{k+1} = x^k + C(b - Ax^k), \quad k = 0, 1, 2, \dots$$

Fixpunktiteration für lineare Probleme

Es gilt

$$Ax = b \Leftrightarrow x = x + C(b - Ax) =: \Phi(x)$$

mit einer beliebigen regulären Matrix $C \in \mathbb{R}^{n \times n}$.

Fixpunktiteration:

$$x^{k+1} = x^k + C(b - Ax^k), \quad k = 0, 1, 2, \dots$$

Für den Fehler gilt

$$\begin{aligned} x^{k+1} - x^* &= (I - CA)^{k+1}(x^0 - x^*) \\ \|x^k - x^*\| &\leq \|I - CA\|^k \|x^0 - x^*\|. \end{aligned}$$

Hinreichend für Konvergenz der Fixpunktiteration:

$$\|I - CA\| < 1$$

Fixpunktiteration für lineare Probleme

Wird zur Lösung **großer dünnbesetzter Gleichungssysteme** verwendet.
Effizienz des Verfahrens hängt ab von:

- ▶ Rechenaufwand pro Iteration
- ▶ Konvergenzgeschwindigkeit der Fixpunktiteration

Fixpunktiteration für lineare Probleme

Wird zur Lösung **großer dünnbesetzter Gleichungssysteme** verwendet.
Effizienz des Verfahrens hängt ab von:

- ▶ Rechenaufwand pro Iteration
- ▶ Konvergenzgeschwindigkeit der Fixpunktiteration

Beispiel: das Gauß-Seidel Verfahren

Sei L der untere Dreiecksanteil der Matrix A . Annahme: $\det(L) \neq 0$.
Wähle $C := L^{-1}$. Beachte:

$$z_k = C(b - Ax^k) \Leftrightarrow Lz_k = b - Ax^k$$

Also: **geringer Rechenaufwand pro Iteration**. Für bestimmte Problemklassen:

$$\|I - CA\| < 1$$

Also: **Konvergenz**.

Stabilität der Fixpunktiteration

Es sei $\mathbf{p} = (p_1, \dots, p_m)^T \in \mathbb{R}^m$ ein Satz von Parametern, der die Iterationsfunktion vollständig beschreibt: $\mathbf{x} = \Phi(\mathbf{x}) = \Phi(\mathbf{x}; \mathbf{p})$

Sei $\tilde{\Phi}(\cdot; \mathbf{p})$ die mit Rundungsfehlern behaftete Auswertung.

Annahme: die Auswertung ist Rückwärtsstabil:

$$\tilde{\Phi}(\mathbf{x}; \mathbf{p}) = \Phi(\mathbf{x}; \tilde{\mathbf{p}}), \quad \text{mit } \|\tilde{\mathbf{p}} - \mathbf{p}\|_{\mathbb{R}^m} \leq c_{aus} \text{eps} \|\mathbf{p}\|_{\mathbb{R}^m}$$

Auf einem Rechner: $\hat{\mathbf{x}}^{k+1} := \tilde{\Phi}(\hat{\mathbf{x}}^k; \mathbf{p}), k = 0, 1, \dots$

Stabilität der Fixpunktiteration

Es sei $\mathbf{p} = (p_1, \dots, p_m)^T \in \mathbb{R}^m$ ein Satz von Parametern, der die Iterationsfunktion vollständig beschreibt: $\mathbf{x} = \Phi(\mathbf{x}) = \Phi(\mathbf{x}; \mathbf{p})$

Sei $\tilde{\Phi}(\cdot; \mathbf{p})$ die mit Rundungsfehlern behaftete Auswertung.

Annahme: die Auswertung ist Rückwärtsstabil:

$$\tilde{\Phi}(\mathbf{x}; \mathbf{p}) = \Phi(\mathbf{x}; \tilde{\mathbf{p}}), \quad \text{mit } \|\tilde{\mathbf{p}} - \mathbf{p}\|_{\mathbb{R}^m} \leq c_{\text{aus}} \text{eps} \|\mathbf{p}\|_{\mathbb{R}^m}$$

Auf einem Rechner: $\hat{\mathbf{x}}^{k+1} := \tilde{\Phi}(\hat{\mathbf{x}}^k; \mathbf{p}), k = 0, 1, \dots$

Man kann zeigen:

Für k hinreichend groß ist der Fehler $\|\hat{\mathbf{x}}^k - \mathbf{x}^*\|$ (maximal) etwa von derselben Größenordnung wie der aufgrund der Kondition des Fixpunktproblems unvermeidbaren Fehler. D.h., falls die Φ -Auswertung rückwärtsstabil ist, ist die **Fixpunktiteration ein stabiles Verfahren**.

Konvergenzordnung und Fehlerschätzung

Ein Maß für die Konvergenzgeschwindigkeit einer Folge ist der Begriff der **Konvergenzordnung**.

Definition 5.32

Es sei $\{x_k\}_{k \in \mathbb{N}}$ eine konvergente Folge mit Grenzwert x^* und mit folgender Eigenschaft: Es existieren $c \in (0, \infty)$, $p \in [1, \infty)$ und $k_0 \in \mathbb{N}$, so dass

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^p$$

für alle $k \geq k_0$ gilt, wobei

$$0 < c < 1 \text{ ist, falls } p = 1.$$

Der maximale p -Wert, für den diese Eigenschaft gilt, wird als **Konvergenzordnung** der Folge bezeichnet.

Konvergenzordnung und Fehlerschätzung

Die Konvergenzordnung eines **iterativen** Verfahrens kann man entsprechend festlegen

Definition

Ein **iteratives Verfahren** zur Bestimmung von $x^* \in \mathbb{R}^n$ (z.B. die Nullstelle einer Funktion) hat die

Konvergenzordnung p ,

wenn es eine Umgebung U von x^* gibt, so dass für alle Startwerte

$$x_0 \in U \setminus \{x^*\}$$

die **von dem Verfahren erzeugte Folge** $\{x_k\}_{k \in \mathbb{N}}$ gegen x^* konvergiert und sie die **Konvergenzordnung p** hat.

Beispiel 5.34

Vergleich der Konvergenzgeschwindigkeit zwischen

1. Verfahren der Ordnung $p = 1$ (lineare Konvergenz), und
2. Verfahren der Ordnung $p = 2$ (quadratische Konvergenz).

Sei $\|x_0 - x^*\| = 0.2$ und $e_k := \|x_k - x^*\|$, dann ergibt sich

Beispiel 5.34

Vergleich der Konvergenzgeschwindigkeit zwischen

1. Verfahren der Ordnung $p = 1$ (**lineare** Konvergenz), und
2. Verfahren der Ordnung $p = 2$ (**quadratische** Konvergenz).

Sei $\|x_0 - x^*\| = 0.2$ und $e_k := \|x_k - x^*\|$, dann ergibt sich

1. Linear: $p = 1$ und $c = \frac{1}{2}$

k	1	2	3	4	5	6
$e_k \leq$	0.1	0.05	0.025	0.0125	0.00625	0.003125

Beispiel 5.34

Vergleich der Konvergenzgeschwindigkeit zwischen

1. Verfahren der Ordnung $p = 1$ (lineare Konvergenz), und
2. Verfahren der Ordnung $p = 2$ (quadratische Konvergenz).

Sei $\|x_0 - x^*\| = 0.2$ und $e_k := \|x_k - x^*\|$, dann ergibt sich

1. Linear: $p = 1$ und $c = \frac{1}{2}$

k	1	2	3	4	5	6
$e_k \leq$	0.1	0.05	0.025	0.0125	0.00625	0.003125

2. Quadratisch: $p = 2$ und $c = 3$

k	1	2	3	4	5	6
$e_k \leq$	0.12	0.0432	0.0056	0.000094	$3 \cdot 10^{-8}$	$2 \cdot 10^{-15}$

Konvergenzordnung einer Fixpunktiteration

Sei
$$\mathbf{x}_{k+1} = \Phi(\mathbf{x}_k), \quad k = 0, 1, \dots,$$

eine **konvergente Fixpunktiteration** mit Fixpunkt \mathbf{x}^* . Mit Hilfe der Taylorreihenentwicklung erhält man

$$\begin{aligned} \mathbf{x}_{k+1} - \mathbf{x}^* &= \Phi(\mathbf{x}_k) - \Phi(\mathbf{x}^*) \\ &= \Phi'(\mathbf{x}^*)(\mathbf{x}_k - \mathbf{x}^*) + \mathcal{O}(\|\mathbf{x}_k - \mathbf{x}^*\|^2). \end{aligned}$$

Konvergenzordnung einer Fixpunktiteration

Sei
$$\mathbf{x}_{k+1} = \Phi(\mathbf{x}_k), \quad k = 0, 1, \dots,$$

eine **konvergente Fixpunktiteration** mit Fixpunkt \mathbf{x}^* . Mit Hilfe der Taylorreihenentwicklung erhält man

$$\begin{aligned}\mathbf{x}_{k+1} - \mathbf{x}^* &= \Phi(\mathbf{x}_k) - \Phi(\mathbf{x}^*) \\ &= \Phi'(\mathbf{x}^*)(\mathbf{x}_k - \mathbf{x}^*) + \mathcal{O}(\|\mathbf{x}_k - \mathbf{x}^*\|^2).\end{aligned}$$

Daraus folgt für die **Konvergenzordnung**:

- ▶ wenn $0 \neq \|\Phi'(\mathbf{x}^*)\| < 1$: **Lineare** Konvergenz ($p = 1$).
- ▶ wenn $\Phi'(\mathbf{x}^*) = 0$, $\Phi''(\mathbf{x}^*) \neq 0$: **Quadratische** Konvergenz ($p = 2$).

Für die meisten in der Praxis benutzten Methoden zur Nullstellenbestimmung gilt $p = 1$ (lineare Konvergenz) oder $p = 2$ (quadratisch).

Fehlerschätzung für skalare Folgen

Definiere $e_k := x^* - x_k$ und $A_k := \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}$

Lemma 5.36

Sei $\{x_k\}_{k=0}^{\infty}$ eine konvergente Folge mit Grenzwert x^* .

Für $p > 1$: $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = 0$ und $\lim_{k \rightarrow \infty} \frac{x_{k+1} - x_k}{e_k} = 1$.

Fehlerschätzung für skalare Folgen

Definiere $e_k := x^* - x_k$ und $A_k := \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}$

Lemma 5.36

Sei $\{x_k\}_{k=0}^{\infty}$ eine konvergente Folge mit Grenzwert x^* .

Für $p > 1$: $\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = 0$ und $\lim_{k \rightarrow \infty} \frac{x_{k+1} - x_k}{e_k} = 1$.

Aus

$$\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k} = A \in (-1, 1), \quad A \neq 0,$$

folgt, dass die **Konvergenzordnung** der Folge **genau 1** ist und

$$\lim_{k \rightarrow \infty} A_k = A, \quad \lim_{k \rightarrow \infty} \frac{A_k}{1 - A_k} (x_k - x_{k-1}) = 1.$$

Fehlerschätzung für skalare Folgen

Es ergeben sich einfache Fehlerschätzungen (für k hinreichend groß) aus den Resultaten in Lemma 5.36:

$$p = 1 : \quad x^* - x_k \approx \frac{A_k}{1 - A_k} (x_k - x_{k-1}),$$

wobei $A_k = \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}$ etwa konstant sein sollte.

$$p > 1 : \quad x^* - x_k \approx x_{k+1} - x_k.$$

Beachte:

Für $p = 1$ (lineare Konvergenz) ist

$$|x_k - x_{k-1}| \quad \text{oder} \quad |x_{k+1} - x_k|$$

meist **keine** sinnvolle Schätzung der Größe des Fehlers $|x^* - x_k|$.

Beispiel 5.38

Für die Fixpunktiteration $x_{k+1} = \Phi_2(x_k)$ aus Beispiel 5.20 sind einige Resultate in folgender Tabelle zusammengestellt:

k	$x_0 = 0.5, x_{k+1} = \Phi_2(x_k)$	$A_k = \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}$	$\frac{A_k}{1 - A_k}(x_k - x_{k-1})$	$x^* - x_k$
0	0.500000000000	—	—	6.35e-01
1	1.069913193934	—	—	6.48e-02
2	1.128908359044	0.1035161	6.81e-03	5.82e-03
3	1.134208317737	0.0898372	5.23e-04	5.16e-04
4	1.134678435924	0.0887022	4.58e-05	4.57e-05
5	1.134720089466	0.0886023	4.05e-06	4.05e-06
6	1.134723779696	0.0885934	3.59e-07	3.59e-07
7	1.134724106623	0.0885926	3.18e-08	3.18e-08
8	1.134724135586	0.0885926	2.82e-09	2.82e-09
9	1.134724138152	0.0885926	2.49e-10	2.49e-10
10	1.134724138379	0.0885925	2.21e-11	2.21e-11

Fehlerschätzung für Vektorfolgen

Lemma 5.40

Sei $\{x_k\}_{k=0}^{\infty}$ eine konvergente Folge in \mathbb{R}^n mit Grenzwert x^* und Konvergenzordnung $p > 1$. Dann gilt

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|e_k\|} = 1.$$

Aus diesem Resultat ergibt sich folgende Fehlerschätzung:

$$p > 1 : \|x_k - x^*\| \approx \|x_{k+1} - x_k\|, \quad \text{für } k \text{ genügend groß.}$$

Fehlerschätzung für Vektorfolgen

Lemma 5.40

Sei $\{x_k\}_{k=0}^{\infty}$ eine konvergente Folge in \mathbb{R}^n mit Grenzwert x^* und Konvergenzordnung $p > 1$. Dann gilt

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|e_k\|} = 1.$$

Aus diesem Resultat ergibt sich folgende Fehlerschätzung:

$$p > 1 : \|x_k - x^*\| \approx \|x_{k+1} - x_k\|, \quad \text{für } k \text{ genügend groß.}$$

Es sei bemerkt, dass im skalaren Fall **der Fehler** e_k und im vektoriellen Fall die **Größe des Fehlers**, $\|e_k\|$, geschätzt wird.

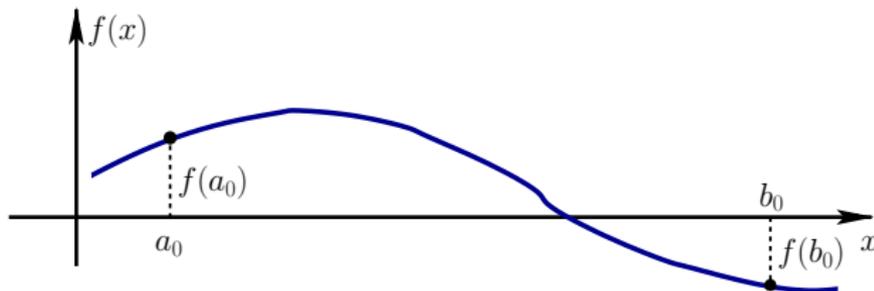
Methode für skalare Gleichungen: Bisektion

Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

- ▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$
- ▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



Methode für skalare Gleichungen: Bisektion

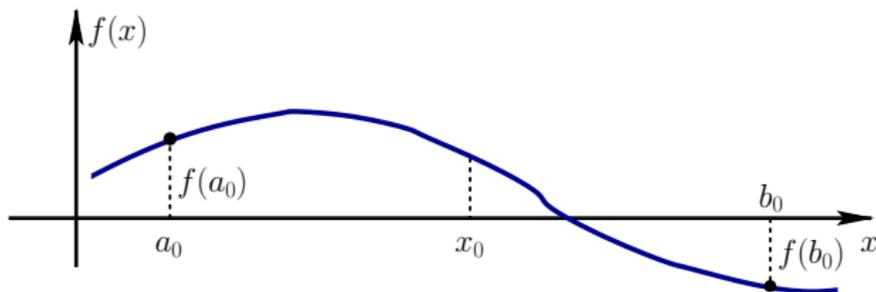
Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$

▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



Methode für skalare Gleichungen: Bisektion

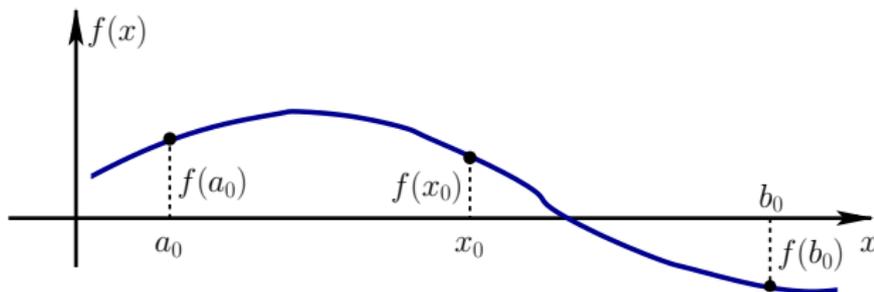
Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$

▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



Methode für skalare Gleichungen: Bisektion

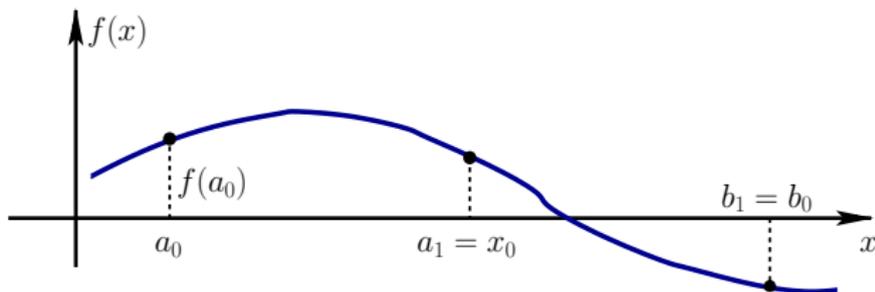
Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$

▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



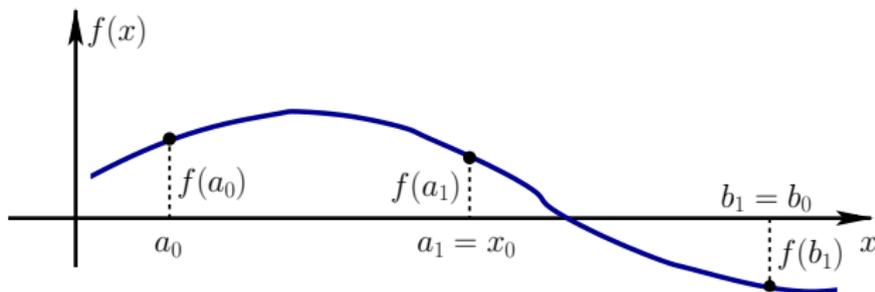
Methode für skalare Gleichungen: Bisektion

Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

- ▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$
- ▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



Methode für skalare Gleichungen: Bisektion

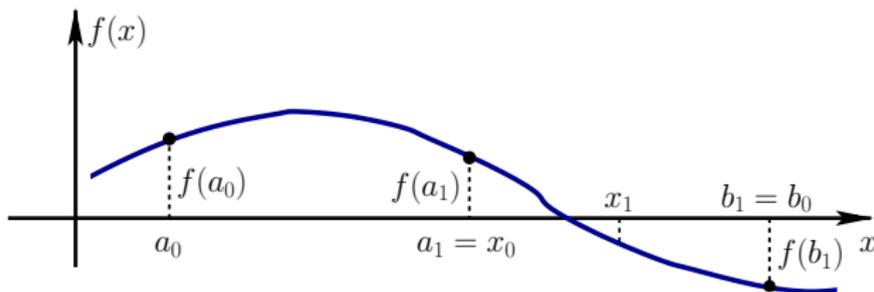
Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$

▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



Methode für skalare Gleichungen: Bisektion

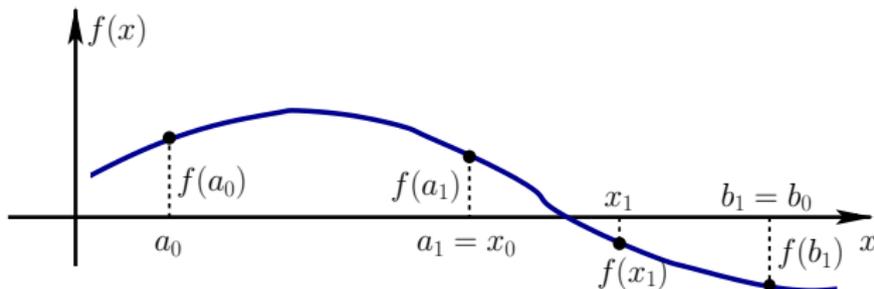
Algorithmus 5.42

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

▶ $x_k = \frac{1}{2}(a_k + b_k)$ und $f(x_k)$

▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.



Beispiel 5.43

Bestimmen Sie die Nullstelle $x^* \in [1, 2]$ der Funktion

$$f(x) = x^6 - x - 1$$

mittels Bisektion (vgl. Beispiel 5.20).

Die Bisektion mit den Startwerten $a_0 = 1$ und $b_0 = 2$ liefert:

k	a_k	b_k	x_k	$b_k - a_k$	$f(x_k)$
0	1.00000	2.00000	1.50000	1.00000	8.89062
1	1.00000	1.50000	1.25000	0.50000	1.56470
2	1.00000	1.25000	1.12500	0.25000	-0.09771
3	1.12500	1.25000	1.18750	0.12500	0.61665
4	1.12500	1.18750	1.15625	0.06250	0.23327
5	1.12500	1.15625	1.14062	0.03125	0.06158
6	1.12500	1.14062	1.13281	0.01562	-0.01958
7	1.13281	1.14062	1.13672	0.00781	0.02062
8	1.13281	1.13672	1.13477	0.00391	0.00043
9	1.13281	1.13477	1.13379	0.00195	-0.00960

Methode für skalare Gleichungen: Newton-Verfahren

Ziel: Konstruiere Φ so, dass die Fixpunktiteration $x_{k+1} = \Phi(x_k)$ möglichst schnell konvergiert.

Methode für skalare Gleichungen: Newton-Verfahren

Ziel: Konstruiere Φ so, dass die Fixpunktiteration $x_{k+1} = \Phi(x_k)$ möglichst schnell konvergiert.

Ansatz:

► Setze $\Phi(x) = x - M_x f(x)$, wobei hier $M_x = g(x)$ (skalar).

Methode für skalare Gleichungen: Newton-Verfahren

Ziel: Konstruiere Φ so, dass die Fixpunktiteration $x_{k+1} = \Phi(x_k)$ möglichst schnell konvergiert.

Ansatz:

- ▶ Setze $\Phi(x) = x - M_x f(x)$, wobei hier $M_x = g(x)$ (skalar).
- ▶ Wähle $g(x)$ so, dass $\Phi'(x^*) = 0$.

Methode für skalare Gleichungen: Newton-Verfahren

Ziel: Konstruiere Φ so, dass die Fixpunktiteration $x_{k+1} = \Phi(x_k)$ möglichst schnell konvergiert.

Ansatz:

- ▶ Setze $\Phi(x) = x - M_x f(x)$, wobei hier $M_x = g(x)$ (skalar).
- ▶ Wähle $g(x)$ so, dass $\Phi'(x^*) = 0$.

Es gilt: $\Phi'(x^*) = 0 \iff g(x^*) = \frac{1}{f'(x^*)}$,

und daraus folgt $\Phi(x) := x - \frac{f(x)}{f'(x)}$.

Methode für skalare Gleichungen: Newton-Verfahren

Ziel: Konstruiere Φ so, dass die Fixpunktiteration $x_{k+1} = \Phi(x_k)$ möglichst schnell konvergiert.

Ansatz:

- ▶ Setze $\Phi(x) = x - M_x f(x)$, wobei hier $M_x = g(x)$ (skalar).
- ▶ Wähle $g(x)$ so, dass $\Phi'(x^*) = 0$.

Es gilt:
$$\Phi'(x^*) = 0 \iff g(x^*) = \frac{1}{f'(x^*)},$$

und daraus folgt
$$\Phi(x) := x - \frac{f(x)}{f'(x)}.$$

Newton-Verfahren

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \dots$$

Geometrische Herleitung des Newton-Verfahrens

$$\text{Es gilt } f(x) = \underbrace{f(x_k) + (x - x_k)f'(x_k)}_{=: T(x)} + \frac{1}{2}(x - x_k)^2 f''(\xi_k)$$

$T(x)$ entspricht der **Tangente** von f bei x_k . So gilt dann

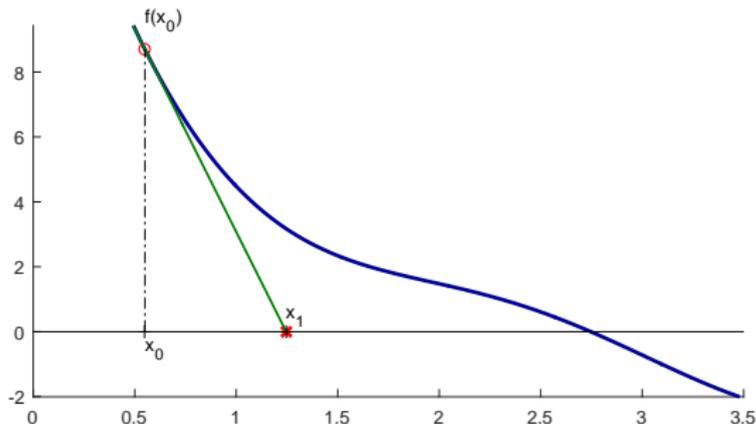
$$T(x_{k+1}) = 0 \quad \Leftrightarrow \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Geometrische Herleitung des Newton-Verfahrens

$$\text{Es gilt } f(x) = \underbrace{f(x_k) + (x - x_k)f'(x_k)}_{=: T(x)} + \frac{1}{2}(x - x_k)^2 f''(\xi_k)$$

$T(x)$ entspricht der **Tangente** von f bei x_k . So gilt dann

$$T(x_{k+1}) = 0 \quad \Leftrightarrow \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

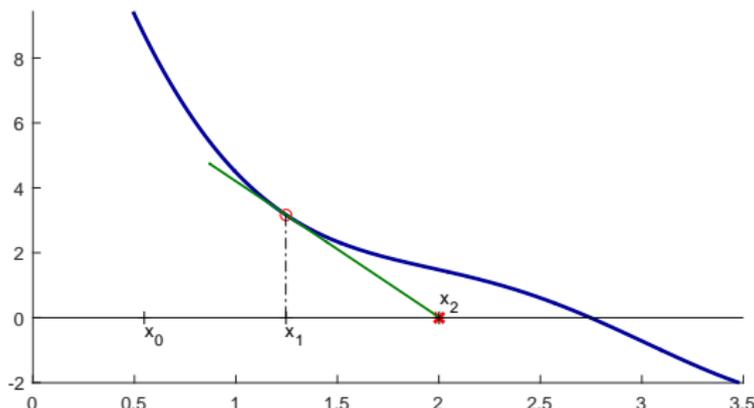


Geometrische Herleitung des Newton-Verfahrens

$$\text{Es gilt } f(x) = \underbrace{f(x_k) + (x - x_k)f'(x_k)}_{=: T(x)} + \frac{1}{2}(x - x_k)^2 f''(\xi_k)$$

$T(x)$ entspricht der **Tangente** von f bei x_k . So gilt dann

$$T(x_{k+1}) = 0 \quad \Leftrightarrow \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

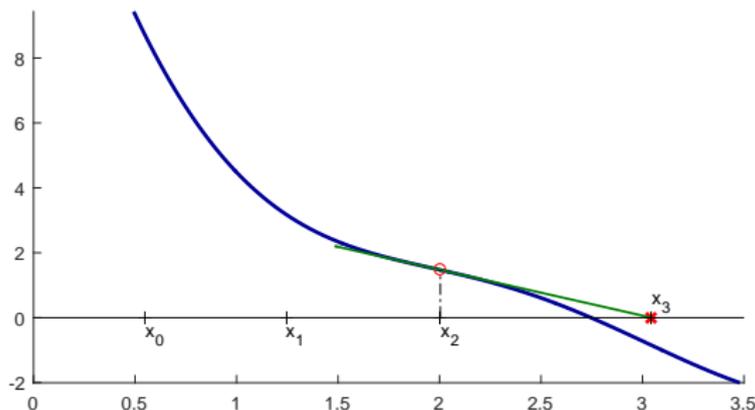


Geometrische Herleitung des Newton-Verfahrens

$$\text{Es gilt } f(x) = \underbrace{f(x_k) + (x - x_k)f'(x_k)}_{=: T(x)} + \frac{1}{2}(x - x_k)^2 f''(\xi_k)$$

$T(x)$ entspricht der **Tangente** von f bei x_k . So gilt dann

$$T(x_{k+1}) = 0 \quad \Leftrightarrow \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

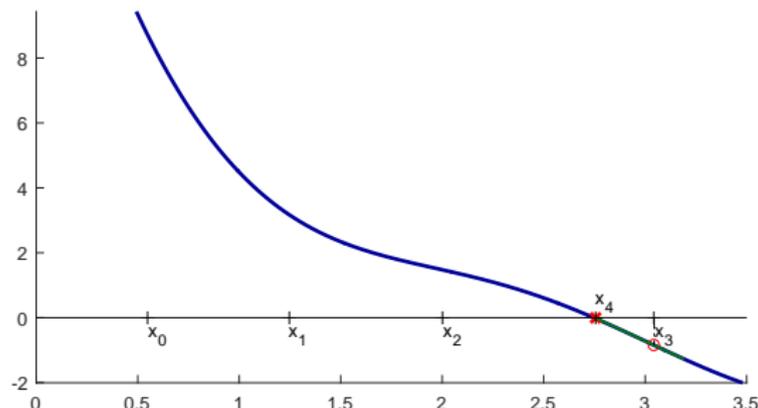


Geometrische Herleitung des Newton-Verfahrens

$$\text{Es gilt } f(x) = \underbrace{f(x_k) + (x - x_k)f'(x_k)}_{=: T(x)} + \frac{1}{2}(x - x_k)^2 f''(\xi_k)$$

$T(x)$ entspricht der **Tangente** von f bei x_k . So gilt dann

$$T(x_{k+1}) = 0 \quad \Leftrightarrow \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

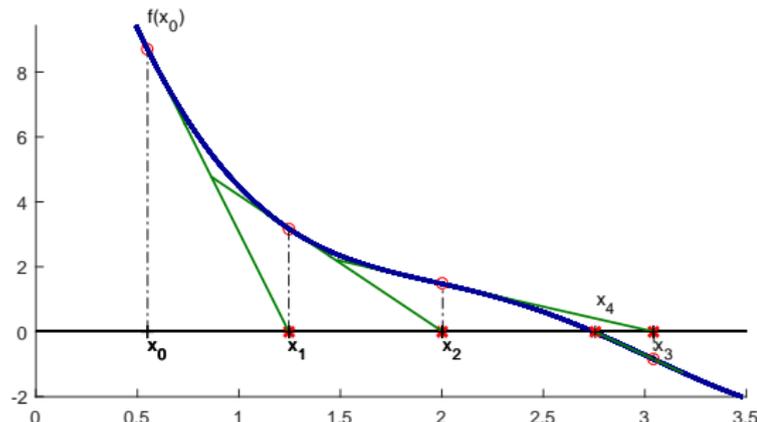


Geometrische Herleitung des Newton-Verfahrens

$$\text{Es gilt } f(x) = \underbrace{f(x_k) + (x - x_k)f'(x_k)}_{=: T(x)} + \frac{1}{2}(x - x_k)^2 f''(\xi_k)$$

$T(x)$ entspricht der **Tangente** von f bei x_k . So gilt dann

$$T(x_{k+1}) = 0 \quad \Leftrightarrow \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$



Konvergenz Newton-Verfahren

Satz 5.44

Sei f zweimal stetig differenzierbar $U = [a, b]$, $x^* \in (a, b)$, und es gelte

$$\begin{aligned}f(x^*) &= 0 \\f'(x^*) &\neq 0.\end{aligned}$$

Für $x_k \in U$ und

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}$$

gilt

$$x_{k+1} - x^* = \frac{1}{2} \frac{f''(\xi_k)}{f'(x_k)} (x_k - x^*)^2, \quad \xi_k \in U.$$

Also ist das [Newton-Verfahren lokal quadratisch konvergent](#).

Beispiel 5.45

Bestimmen Sie die Nullstelle $x^* \in [1, 2]$ der Funktion

$$f(x) = x^6 - x - 1$$

mittels Newton-Verfahrens (vgl. Beispiel 5.20).

Beispiel 5.45

Bestimmen Sie die Nullstelle $x^* \in [1, 2]$ der Funktion

$$f(x) = x^6 - x - 1$$

mittels Newton-Verfahrens (vgl. Beispiel 5.20).

Die Newton-Iteration $x_{k+1} = x_k - \frac{x_k^6 - x_k - 1}{6x_k^5 - 1}$ liefert:

k	x_k mit $x_0 = 0.5$	x_k mit $x_0 = 2$	$x_{k+1} - x_k$ mit $x_0 = 2$
0	0.5000000000000000	2.0000000000000000	-3.19e-01
1	-1.32692307692308	1.68062827225131	-2.50e-01
2	-1.10165080870249	1.43073898823906	-1.76e-01
3	-0.92567640260338	1.25497095610944	-9.34e-02
4	-0.81641531662254	1.16153843277331	-2.52e-02
5	-0.78098515830640	1.13635327417051	-1.62e-03
6	-0.77810656986872	1.13473052834363	-6.39e-06
7	-0.77808959926268	1.13472413850022	-9.87e-11
8	-0.77808959867860	1.13472413840152	0.00e+00
9	-0.77808959867860	1.13472413840152	-

Beispiel 5.46

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

Beispiel 5.46

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

Ansatz: Die Wurzel von a , \sqrt{a} , ist Lösung von

$$f(x) := x^2 - a = 0.$$

Beispiel 5.46

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

Ansatz: Die Wurzel von a , \sqrt{a} , ist Lösung von

$$f(x) := x^2 - a = 0.$$

Das Newton-Verfahren ergibt hier $x_{k+1} = \frac{1}{2}(x_k + a/x_k)$ und liefert für $a = 2$ die Resultate:

	x_k	$x_{k+1} - x_k$	$\sqrt{2} - x_k$
0	100.00000000000000	-5.00e+01	-9.86e+01
1	50.01000000000000	-2.50e+01	-4.86e+01
2	25.02499600079984	-1.25e+01	-2.36e+01
3	12.55245804674590	-6.20e+00	-1.11e+01
4	6.35589469493114	-3.02e+00	-4.94e+00
5	3.33528160928043	-1.37e+00	-1.92e+00
6	1.96746556223115	-4.75e-01	-5.53e-01
7	1.49200088968972	-7.58e-02	-7.78e-02
8	1.41624133203894	-2.03e-03	-2.03e-03
9	1.41421501405005	-1.45e-06	-1.45e-06
10	1.41421356237384	-	-7.45e-13

Beispiel 5.46: Globale Konvergenz

Verfahren konvergiert für **jeden Startwert $x_0 > 0$** .

► Für $x_0 > 0$ gilt

$$x_{k+1} - \sqrt{a} = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right) - \sqrt{a} = \frac{1}{2x_k} (x_k - \sqrt{a})^2 \geq 0$$

also $x_k \geq \sqrt{a}$ für alle $k \geq 1$.

Beispiel 5.46: Globale Konvergenz

Verfahren konvergiert für jeden Startwert $x_0 > 0$.

► Für $x_0 > 0$ gilt

$$x_{k+1} - \sqrt{a} = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right) - \sqrt{a} = \frac{1}{2x_k} (x_k - \sqrt{a})^2 \geq 0$$

also $x_k \geq \sqrt{a}$ für alle $k \geq 1$.

Damit ergibt sich, für $k \geq 1$:

►

$$0 \leq x_{k+1} - \sqrt{a} = \frac{1}{2} \frac{x_k - \sqrt{a}}{x_k} (x_k - \sqrt{a}) \leq \frac{1}{2} (x_k - \sqrt{a}).$$

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

Newton-Verfahren Matlab-Demo 5.48

Konvergenzverhalten des Newton-Verfahrens:

Newton-Verfahren Matlab-Demo 5.48

Konvergenzverhalten des Newton-Verfahrens:

- ▶ Im Allgemeinen nur **lokale** Konvergenz
- ▶ Divergenz kann auftreten
- ▶ Manchmal **globale** Konvergenz
- ▶ **Lokale quadratische** Konvergenz
- ▶ “Endlose” Iteration möglich
- ▶ Überspringen von Nullstellen

Merke:

- ▶ Quadratische Konvergenz nur lokal
- ▶ Guter Startwert ist wichtig

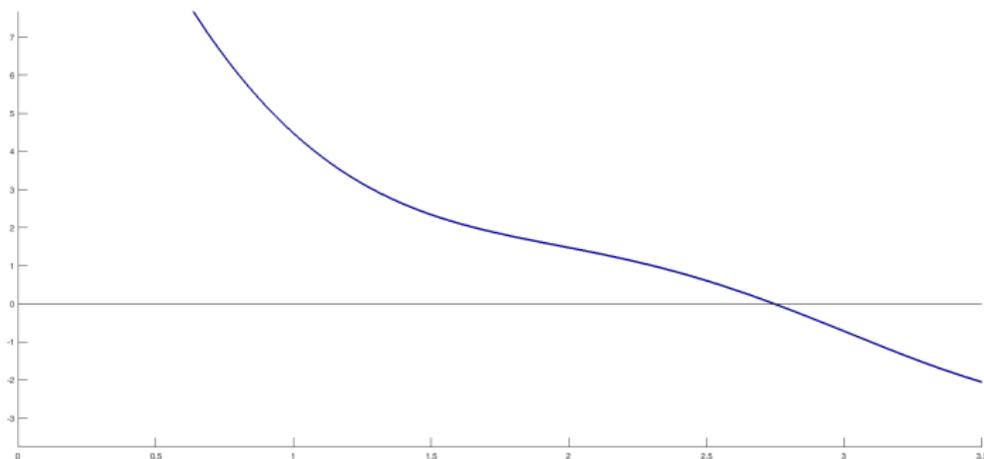
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



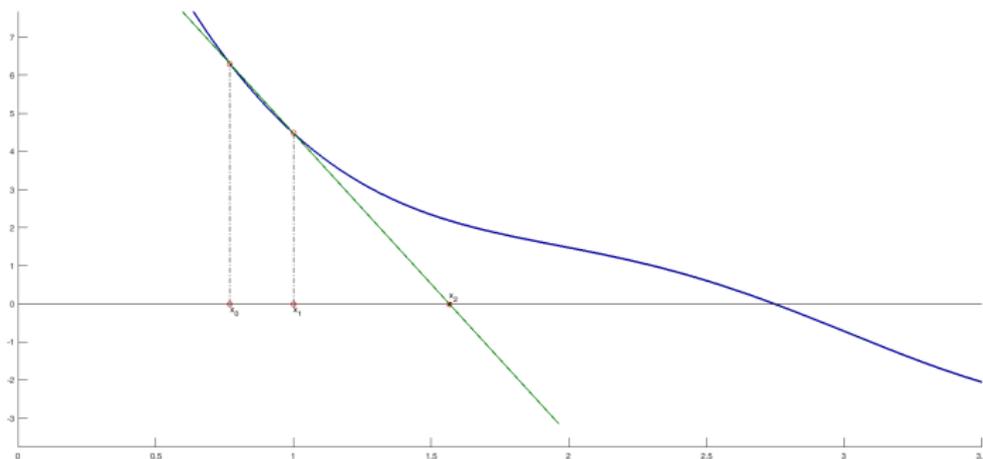
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



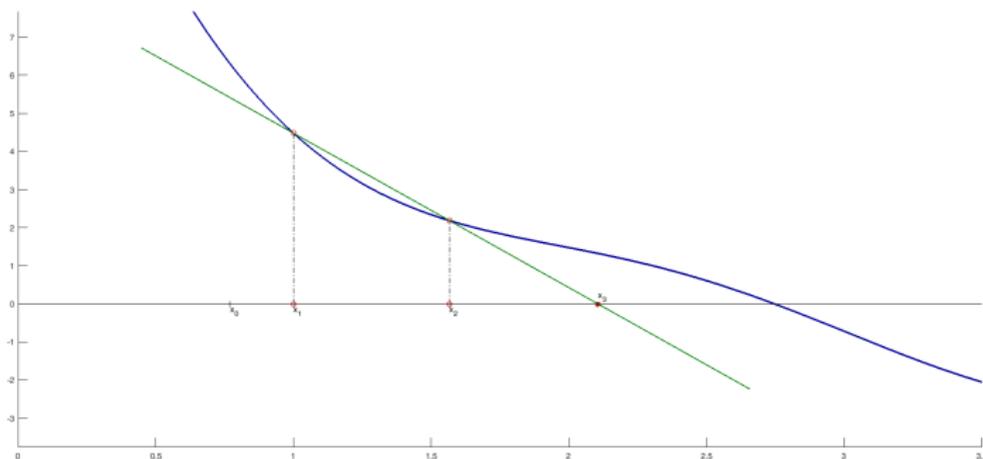
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



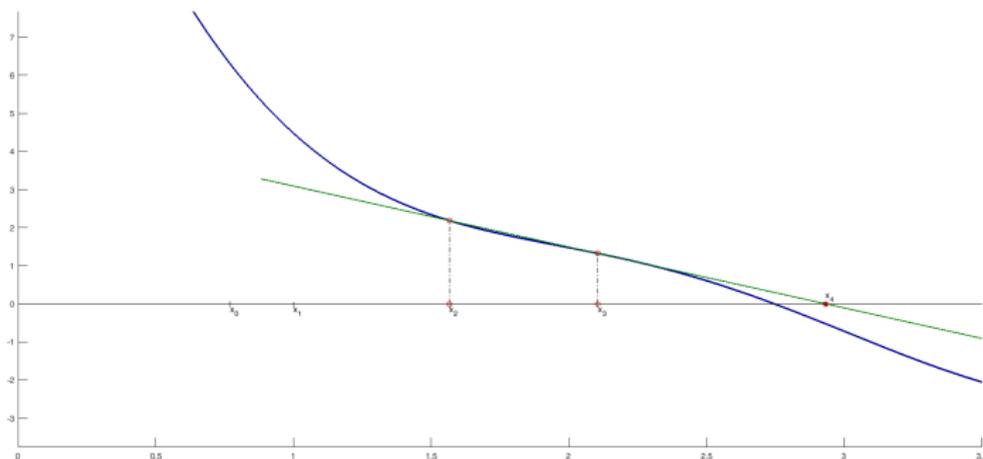
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



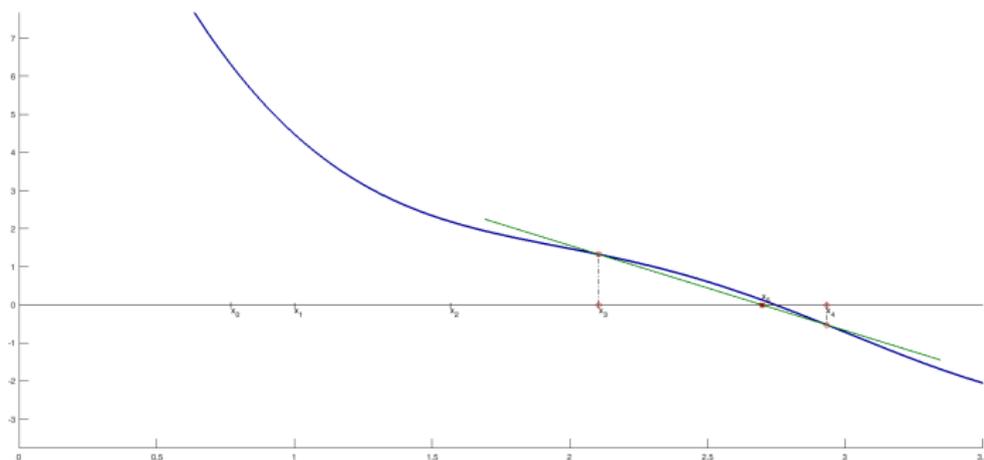
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



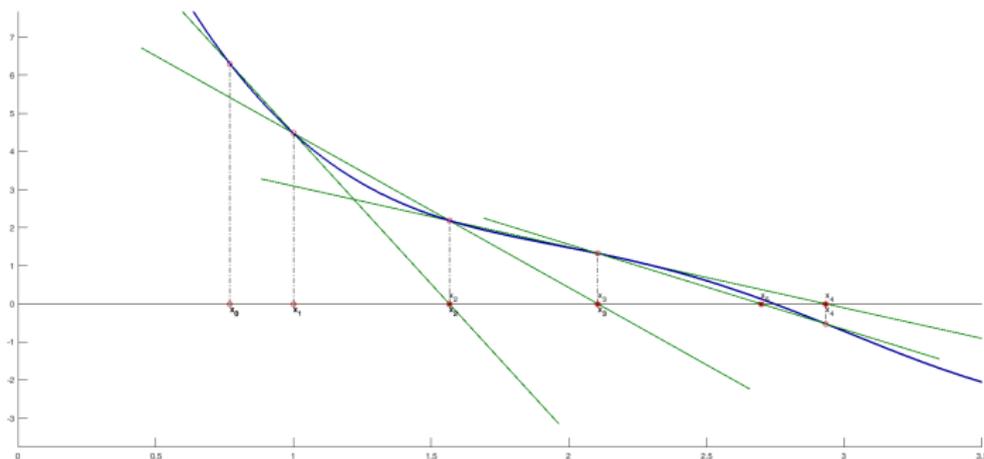
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



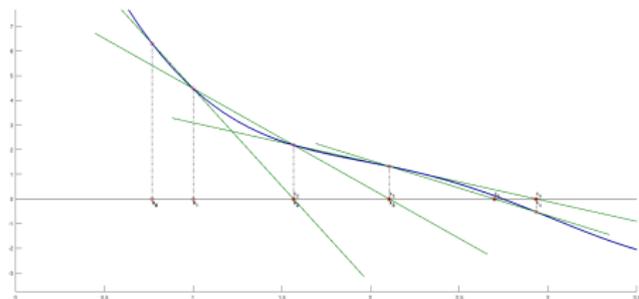
Methode für skalare Gleichungen: Sekanten-Verfahren

Idee:

- ▶ Ersetze **Tangente** $T(x)$ im Newton-Verfahren durch **Sekante**

$$S(x) = f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} .$$

- ▶ **Nullstelle** der Sekante ergibt neue Annäherung x_{k+1}



Der **Sekantenanstieg**

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

entspricht gerade $f'(x_k)$ im Newton-Verfahren.

Sekanten-Verfahren

Sekanten-Verfahren

$$\begin{aligned}x_{k+1} &= f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} \\&= x_k - f(x_k) \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right) \\&= \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}\end{aligned}$$

Sekanten-Verfahren

Sekanten-Verfahren

$$\begin{aligned}x_{k+1} &= f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} \\ &= x_k - f(x_k) \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right) \\ &= \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}\end{aligned}$$

Vorteile gegenüber Newton-Verfahren

- ▶ Berechnung der Ableitung $f'(x)$ wird vermieden.
- ▶ Effizienter, wenn Auswertung von $f'(x)$ und $f(x)$ etwa gleich teuer.

Sekanten-Verfahren

Sekanten-Verfahren

$$\begin{aligned}x_{k+1} &= f(x_k) \frac{x - x_{k-1}}{x_k - x_{k-1}} + f(x_{k-1}) \frac{x_k - x}{x_k - x_{k-1}} \\ &= x_k - f(x_k) \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right) \\ &= \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}\end{aligned}$$

Vorteile gegenüber Newton-Verfahren

- ▶ Berechnung der Ableitung $f'(x)$ wird vermieden.
- ▶ Effizienter, wenn Auswertung von $f'(x)$ und $f(x)$ etwa gleich teuer.

Nachteile gegenüber Newton-Verfahren

- ▶ Konvergenzordnung $p \approx 1.6$.
- ▶ Verfahren benötigt zwei Startwerte.

Beispiel 5.49

Bestimmen Sie die Nullstelle $x^* \in [1, 2]$ der Funktion

$$f(x) = x^6 - x - 1$$

Das Sekanten-Verfahren mit den Startwerten $x_0 = 2$ und $x_1 = 1$ liefert:

k	x_k	$x_{k+1} - x_k$
0	2.0000000000000000	-1.00e+00
1	1.0000000000000000	1.61e-02
2	1.01612903225806	1.74e-01
3	1.19057776867664	-7.29e-02
4	1.11765583094155	1.49e-02
5	1.13253155021613	2.29e-03
6	1.13481680800485	-9.32e-05
7	1.13472364594870	4.92e-07
8	1.13472413829122	1.10e-10
9	1.13472413840152	-

Die Werte in der dritten Spalte ergeben eine Fehlerabschätzung.

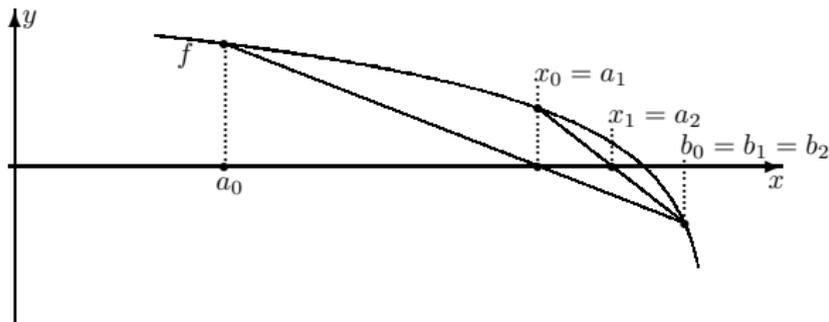
Regula-Falsi

Mischung der Bisektion und des Sekanten-Verfahrens.

Annahme: $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Statt der Mitte des Intervalls wird die **Nullstelle der Sekante** durch die Punkte $(a_0, f(a_0))$, $(b_0, f(b_0))$ bestimmt.

Man wählt nun ein neues Intervall $[a_1, b_1] = [a_0, x_0]$ oder $[a_1, b_1] = [x_0, b_0]$ für das $f(a_1)f(b_1) \leq 0$ gilt, usw.



Regula-Falsi

Algorithmus 5.50

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

- ▶ $x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$, und $f(x_k)$
- ▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.

Regula-Falsi

Algorithmus 5.50

Gegeben $a_0 < b_0$ mit $f(a_0)f(b_0) < 0$.

Für $k = 0, 1, 2, \dots$ berechne:

- ▶ $x_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$, und $f(x_k)$
- ▶ Setze $a_{k+1} = a_k, b_{k+1} = x_k$ falls $f(x_k)f(a_k) \leq 0$
 $a_{k+1} = x_k, b_{k+1} = b_k$ sonst.

$x^* \in (a_k, b_k)$ für alle k und $a_k \rightarrow x^*$ oder $b_k \rightarrow x^*$ für $k \rightarrow \infty$.

- ▶ $x^* \in (a_k, b_k)$: die Methode ist **sehr zuverlässig**.
- ▶ Im Allg. schnellere Konvergenz als bei Bisektion, aber **$p = 1$**

Newton-Verfahren für Systeme

Aufgabe

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

Bestimme

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

Newton-Verfahren für Systeme

Aufgabe

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

Bestimme

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

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

Newton-Verfahren für Systeme

Aufgabe

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

Bestimme

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

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

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

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

Das Newton-Verfahren für Systeme

- ▶ Taylor-Entwicklung kompakt

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

wobei die **Jacobi-Matrix** gegeben ist durch

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

Das Newton-Verfahren für Systeme

- ▶ Taylor-Entwicklung kompakt

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

wobei die **Jacobi-Matrix** gegeben ist durch

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

- ▶ Für die Nullstelle x^{k+1} der **linearen Näherung** von f in x^k folgt

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

und hieraus erhält man

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

Das Newton-Verfahren für Systeme

Algorithmus 5.51 (Newton-Iteration)

Gegeben: Startwert x^0 .

Für $k = 0, 1, 2, \dots$

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)

$$x^{k+1} = x^k + s^k$$

Das Newton-Verfahren für Systeme

Algorithmus 5.51 (Newton-Iteration)

Gegeben: Startwert x^0 .

Für $k = 0, 1, 2, \dots$

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.
- ▶ Die Inverse von $f'(x^k)$ wird nicht explizit berechnet.

Beispiel 5.53

Gegeben sei das nichtlineare Gleichungssystem

$$f_1(x_1, x_2) = 6x_1 - \cos x_1 - 2x_2 = 0$$

$$f_2(x_1, x_2) = 8x_2 - x_1x_2^2 - \sin x_1 = 0$$

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

Beispiel 5.53

Gegeben sei das nichtlineare Gleichungssystem

$$f_1(x_1, x_2) = 6x_1 - \cos x_1 - 2x_2 = 0$$

$$f_2(x_1, x_2) = 8x_2 - x_1x_2^2 - \sin x_1 = 0$$

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

- Berechnung der Jacobi-Matrix

$$f'(x) = \begin{pmatrix} 6 + \sin x_1 & -2 \\ -x_2^2 - \cos x_1 & 8 - 2x_1x_2 \end{pmatrix}.$$

Beispiel 5.53

Gegeben sei das nichtlineare Gleichungssystem

$$f_1(x_1, x_2) = 6x_1 - \cos x_1 - 2x_2 = 0$$

$$f_2(x_1, x_2) = 8x_2 - x_1x_2^2 - \sin x_1 = 0$$

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

- Berechnung der Jacobi-Matrix

$$f'(x) = \begin{pmatrix} 6 + \sin x_1 & -2 \\ -x_2^2 - \cos x_1 & 8 - 2x_1x_2 \end{pmatrix}.$$

- 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 5.53

Newton-Iteration für das nichtlineare Gleichungssystem

$$f_1(x_1, x_2) = 6x_1 - \cos x_1 - 2x_2 = 0$$

$$f_2(x_1, x_2) = 8x_2 - x_1x_2^2 - \sin x_1 = 0$$

- Berechnung der Newton-Korrektur s^0 aus

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

Beispiel 5.53

Newton-Iteration für das nichtlineare Gleichungssystem

$$f_1(x_1, x_2) = 6x_1 - \cos x_1 - 2x_2 = 0$$

$$f_2(x_1, x_2) = 8x_2 - x_1x_2^2 - \sin x_1 = 0$$

- ▶ 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

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

Beispiel 5.53

Newton-Iteration für das nichtlineare Gleichungssystem

$$f_1(x_1, x_2) = 6x_1 - \cos x_1 - 2x_2 = 0$$

$$f_2(x_1, x_2) = 8x_2 - x_1x_2^2 - \sin x_1 = 0$$

Weitere Iterationen:

k	x^k	$f(x^k)$
0	(0, 0)	(-1, 0)
1	(0.173913043478261, 0.021739130434783)	(0.150, 0.008)e-01
2	(0.171334222062832, 0.021321946986676)	(0.328, 0.050)e-05
3	(0.171333648176505, 0.021321814151379)	(0.162, 0.022)e-12
4	(0.171333648176476, 0.021321814151372)	(-0.138, 0.000)e-16

Satz 5.54

Annahmen:

- ▶ Gebiet $\Omega \subset \mathbb{R}^n$ offen und konvex
- ▶ Funktion $f : \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar

Satz 5.54

Annahmen:

- ▶ Gebiet $\Omega \subset \mathbb{R}^n$ offen und konvex
- ▶ Funktion $f : \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar
- ▶ Jacobi-Matrix $f'(x) \in \mathbb{R}^{n \times n}$
 - ▶ invertierbar

$$\det(f'(x)) \neq 0, \quad \text{für alle } x \in \Omega$$

- ▶ die Inverse beschränkt durch eine Konstante β

$$\|(f'(x))^{-1}\| \leq \beta \quad \text{für alle } x \in \Omega.$$

- ▶ Lipschitz-stetig auf Ω mit einer Konstanten γ

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

Satz 5.54

Annahmen:

- ▶ Gebiet $\Omega \subset \mathbb{R}^n$ offen und konvex
- ▶ Funktion $f : \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar
- ▶ Jacobi-Matrix $f'(x) \in \mathbb{R}^{n \times n}$
 - ▶ invertierbar

$$\det(f'(x)) \neq 0, \quad \text{für alle } x \in \Omega$$

- ▶ die Inverse beschränkt durch eine Konstante β

$$\|(f'(x))^{-1}\| \leq \beta \quad \text{für alle } x \in \Omega.$$

- ▶ Lipschitz-stetig auf Ω mit einer Konstanten γ

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

- ▶ Es existiere eine Lösung x^* von $f(x) = 0$ in Ω .

Satz 5.54

Der Startwert x^0 erfülle

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

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

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

Satz 5.54

Der Startwert x^0 erfülle

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

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

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

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

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

Beispiel 5.2 (Erinnerung)

Statt der **linearen** Integralgleichung im Beispiel 3.7 ist nun eine **nichtlineare** Integralgleichung zu lösen:

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

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

erfüllt.

Beispiel 5.56

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

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

wobei

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

Beispiel 5.56

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

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

wobei

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

Für die **Jacobi-Matrix** erhält man

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

Beispiel 5.56

In jedem Iterationsschritt des Newton-Verfahrens werden

- ▶ die Jacobi-Matrix $f'(x^k)$ und der Funktionswert $f(x^k)$ berechnet,

Beispiel 5.56

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,

Beispiel 5.56

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.

Beispiel 5.56

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

k	$\ f(x^k)\ _2$	$\ x^{k+1} - x^k\ _2$
0	5.57e-01	4.59e-01
1	7.53e-02	2.01e-02
2	1.50e-04	3.83e-05
3	5.46e-10	1.40e-10
4	2.49e-15	2.49e-15

Die dritte Spalte zeigt die Fehlerschätzung

Beispiel 5.56

Es gilt

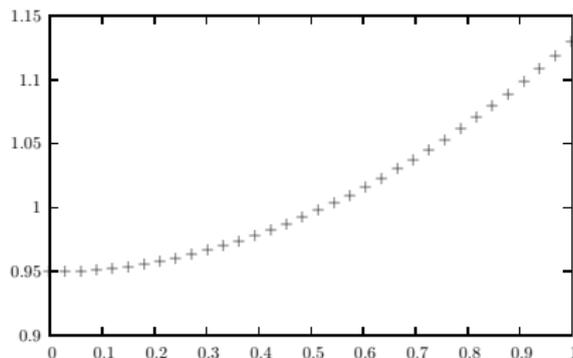
$$x_i^4 \approx u(t_i) = u\left(\frac{i - \frac{1}{2}}{60}\right), \quad i = 1, 2, \dots, 60.$$

Beispiel 5.56

Es gilt

$$x_i^4 \approx u(t_i) = u\left(\frac{i - \frac{1}{2}}{60}\right), \quad i = 1, 2, \dots, 60.$$

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



Hinweise zur praktischen Durchführung

1. 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.

Hinweise zur praktischen Durchführung

1. 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.

Hinweise zur praktischen Durchführung

1. 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
 - ▶ zu groß: **verringert Genauigkeit der Approximation** und damit auch schlechtere Konvergenz
 - ▶ zu klein: birgt Gefahr der **Auslöschung**

Hinweise zur praktischen Durchführung

2. Das vereinfachte Newton-Verfahren

Problem: Jeder Schritt erfordert Aufstellen und Lösung von

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

Hinweise zur praktischen Durchführung

2. Das vereinfachte Newton-Verfahren

Problem: Jeder Schritt erfordert Aufstellen und Lösung von

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

Ansatz:

- ▶ Aufstellen der Jacobi-Matrix im ersten Schritt $f'(x^0)$.

Hinweise zur praktischen Durchführung

2. Das vereinfachte Newton-Verfahren

Problem: Jeder Schritt erfordert Aufstellen und Lösung von

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

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.

$$f'(x^0)s^k = -f(x^k) \rightarrow x^{k+1} = x^k + s^k, \quad k = 0, 1, 2, \dots$$

Hinweise zur praktischen Durchführung

2. Das vereinfachte Newton-Verfahren

Problem: Jeder Schritt erfordert Aufstellen und Lösung von

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

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.

$$f'(x^0)s^k = -f(x^k) \rightarrow x^{k+1} = x^k + s^k, \quad k = 0, 1, 2, \dots$$

- ▶ LR -Berechnung effizient ("mehrere rechte Seiten")

Hinweise zur praktischen Durchführung

2. Das vereinfachte Newton-Verfahren

Problem: Jeder Schritt erfordert Aufstellen und Lösung von

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

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.

$$f'(x^0)s^k = -f(x^k) \rightarrow x^{k+1} = x^k + s^k, \quad k = 0, 1, 2, \dots$$

- ▶ *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

3. Wahl eines “guten” Startwertes: Homotopieverfahren

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

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

so dass F für einen μ -Wert einfach lösbar ist.

Beispiel. Nichtlineare Integralgleichung aus Beispiel 5.2:

$$x_i + h \sum_{j=1}^n \cos(t_i t_j) x_j^\mu - 2 = 0, \quad i = 1, 2, \dots, n,$$

Hinweise zur praktischen Durchführung

3. Wahl eines “guten” Startwertes: Homotopieverfahren

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

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

so dass F für einen μ -Wert einfach lösbar ist.

Beispiel. Nichtlineare Integralgleichung aus Beispiel 5.2:

$$x_i + h \sum_{j=1}^n \cos(t_i t_j) x_j^\mu - 2 = 0, \quad i = 1, 2, \dots, n,$$

1. Setze $\mu = 1$. Löse das **lineare** Problem $F(x, 1) = 0$.
2. Setze $\mu_{\text{neu}} = 1 + \Delta\mu$ (klein) und nehme alte Lösung als Startwert der Newton-Iteration für das Problem $F(x, \mu_{\text{neu}}) = 0$.
3. Wiederhole bis $\mu_{\text{neu}} = 3$.

Hinweise zur praktischen Durchführung

4. Gedämpftes Newton-Verfahren

Problem: Divergenz bei schlechtem Startwert

Hinweise zur praktischen Durchführung

4. Gedämpftes Newton-Verfahren

Problem: Divergenz bei schlechtem Startwert

Ansatz: Man setzt

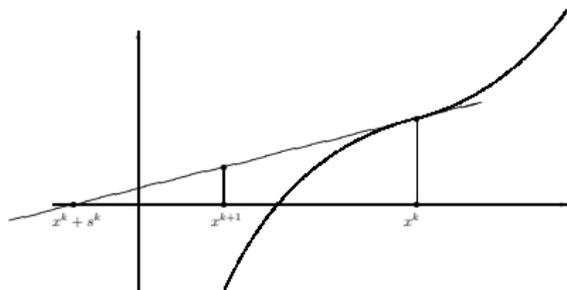
$$x^{k+1} = x^k + \lambda s^k$$

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

Idee: Suche λ , so dass

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

Residuum wird in jedem Schritt verringert.



Hinweise zur praktischen Durchführung

4. Gedämpftes Newton-Verfahren

Problem: Divergenz bei schlechtem Startwert

Ansatz: Man setzt

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

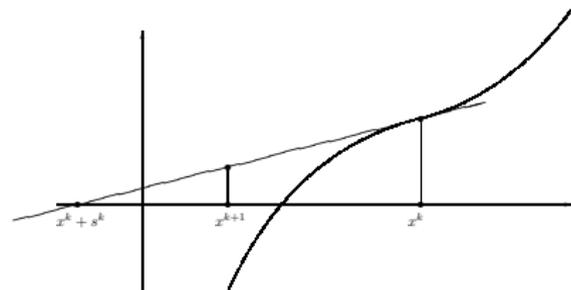
für ein passendes $\lambda = \lambda_k$, $0 < \lambda \leq 1$.

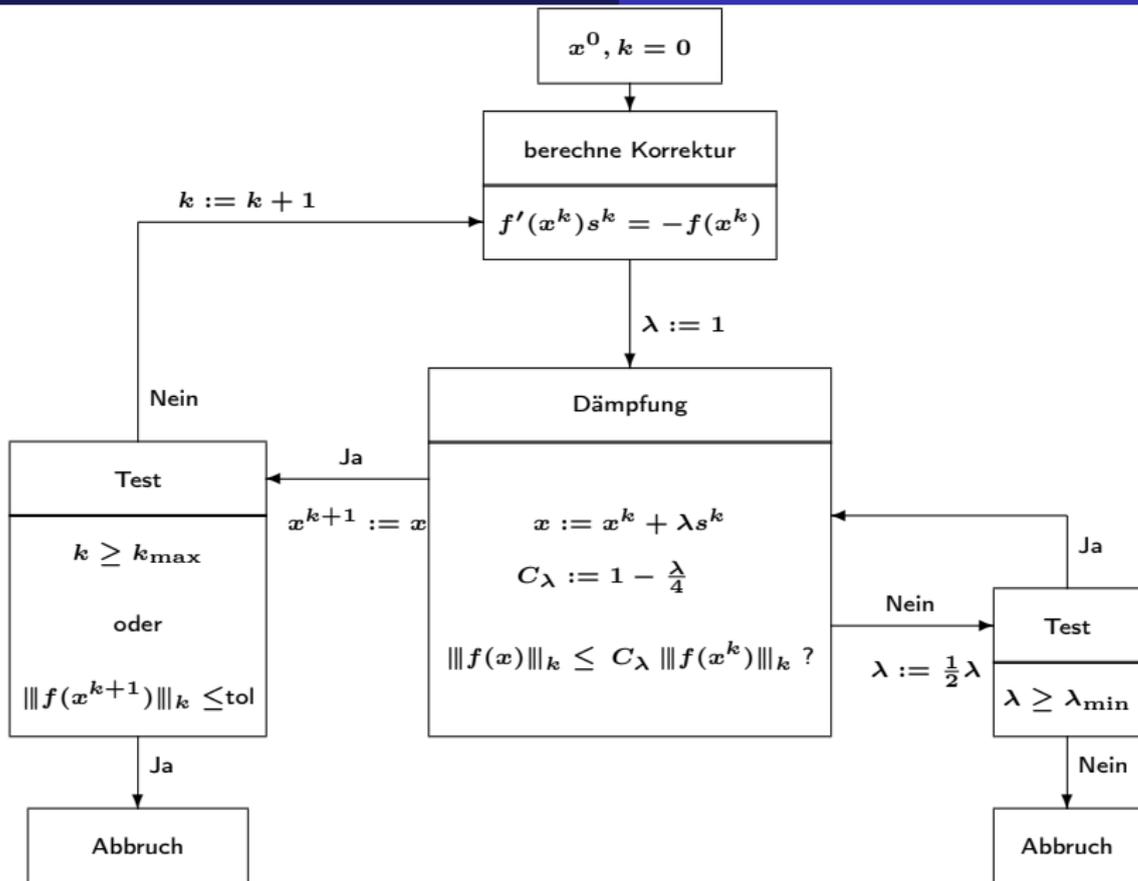
Idee: Suche λ , so dass

$$(1) \quad \|f(\mathbf{x}^{k+1})\| < \|f(\mathbf{x}^k)\|,$$

Residuum wird in jedem Schritt verringert.

In der Praxis: eine geeignete Norm $\|\cdot\|$, so dass (1) **affin invariant** ist.

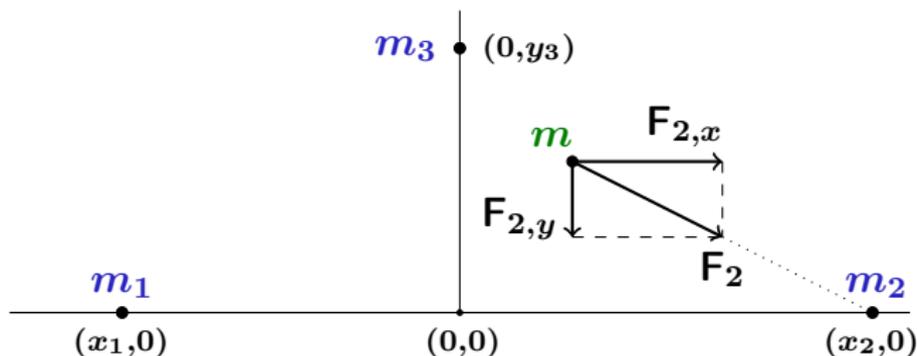




Beispiel 5.1 & 5.59

5. Wahl des Startwertes: basiert auf Hintergrundinformation

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

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

- ▶ Hilfsgrößen

$$r_i := \sqrt{(x - x_i)^2 + (y - y_i)^2},$$

$$F_i := G \frac{m_i m}{r_i^2},$$

$$F_{i,x} := \frac{F_i(x_i - x)}{r_i}, \quad F_{i,y} := \frac{F_i(y_i - y)}{r_i}, \quad i = 1, 2, 3.$$

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

- ▶ Hilfsgrößen

$$r_i := \sqrt{(x - x_i)^2 + (y - y_i)^2},$$

$$F_i := G \frac{m_i m}{r_i^2},$$

$$F_{i,x} := \frac{F_i(x_i - x)}{r_i}, \quad F_{i,y} := \frac{F_i(y_i - y)}{r_i}, \quad i = 1, 2, 3.$$

- ▶ 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 & 5.59

- Hieraus ergibt sich das **nichtlineare Gleichungssystem**

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

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

Beispiel 5.1 & 5.59

- Hieraus ergibt sich das **nichtlineare Gleichungssystem**

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

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

- 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

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

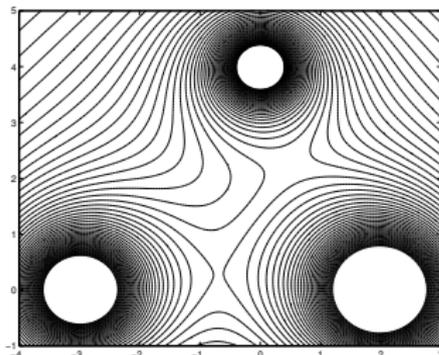
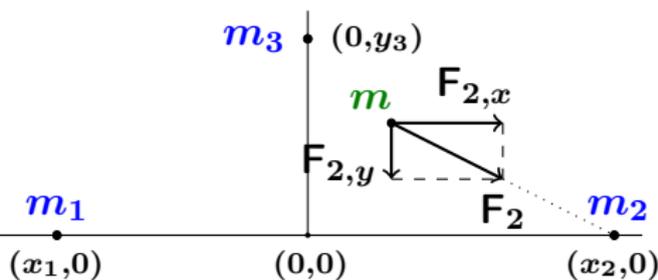
Beispiel 5.1 & 5.59

- ▶ **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.

Beispiel 5.1 & 5.59

- **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 & 5.59

- ▶ 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	-0.8000000000000000	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	-

Deflation

Nullstellen eines reellen Polynoms

$$P_n(x) = \sum_{j=0}^n a_j x^j, \quad a_j \in \mathbb{R}, \quad a_n \neq 0.$$

Deflation: Abspaltung eines linearen Faktors $x - z$.

$$P_n(x) = (x - z)P_{n-1}(x) + R, \quad P_{n-1}(x) = \sum_{j=0}^{n-1} b_j x^j, \quad R \in \mathbb{R}$$

Deflation

Nullstellen eines reellen Polynoms

$$P_n(x) = \sum_{j=0}^n a_j x^j, \quad a_j \in \mathbb{R}, \quad a_n \neq 0.$$

Deflation: Abspaltung eines linearen Faktors $x - z$.

$$P_n(x) = (x - z)P_{n-1}(x) + R, \quad P_{n-1}(x) = \sum_{j=0}^{n-1} b_j x^j, \quad R \in \mathbb{R}$$

Polynomdivision eines linearen Faktors

Eingabe: Koeffizienten $a_0, \dots, a_n \in \mathbb{R}$, $z \in \mathbb{R}$.

$$b_{n-1} = a_n;$$

$$\text{Für } j = n - 2, \dots, 0: \quad b_j = a_{j+1} + z b_{j+1};$$

$$R = a_0 + z b_0$$

Beachte: z Nullstelle von $P_n \iff R = 0$.

Newton-Verfahren

Das Newton-Verfahren ist auch zur Berechnung **komplexer** Nullstellen eines Polynoms verwendbar.

Beispiel 5.65. Das Polynom $P_3(x) = x^3 - x^2 + x - 1$ hat die Nullstellen $1, i, -i$. Das Newton-Verfahren

$$x_{k+1} = x_k - \frac{x_k^3 - x_k^2 + x_k - 1}{3x_k^2 - 2x_k + 1} = \frac{2x_k^3 - x_k^2 + 1}{3x_k^2 - 2x_k + 1}, \quad k = 0, 1, \dots,$$

mit Startwert $x_0 = 0.4 + 0.75i$ liefert die Resultate

k	x_k	$ i - x_k $	$ x_{k+1} - x_k $
0	0.400000000000000 + 0.750000000000000 i	4.72e-01	7.74e-01
1	-0.36104836292270 + 0.61085408548207 i	5.31e-01	4.79e-01
2	0.10267444513356 + 0.72886626636306 i	2.90e-01	4.58e-01
3	-0.01987923527724 + 1.17013991538812 i	1.71e-01	1.46e-01
4	0.00377579358344 + 1.02575250192764 i	2.60e-02	2.54e-02
5	0.00048863011493 + 1.00054628083004 i	7.33e-04	7.33e-04
6	0.00000056371102 + 0.99999979344332 i	6.00e-07	6.00e-07
7	-0.000000000000037 + 0.99999999999984 i	4.00e-13	4.00e-13
8	0.000000000000000 + 1.000000000000000 i	5.55e-17	-

Bairstow-Verfahren

Das Rechnen mit komplexen Zahlen kann vermieden werden.

Skalierung: $a_n = 1$.

Seien $z_1 = u_1 + i v_1$, $\bar{z}_1 = u_1 - i v_1$ Nullstellen von P_n .

Das Produkt

$$(x - z_1)(x - \bar{z}_1) = x^2 - 2u_1x + u_1^2 + v_1^2$$

ist ein **quadratischer** Teiler von P_n mit **reellen** Koeffizienten 1 , $-2u_1$ und $u_1^2 + v_1^2$.

Bairstow-Verfahren

Das Rechnen mit komplexen Zahlen kann vermieden werden.

Skalierung: $a_n = 1$.

Seien $z_1 = u_1 + i v_1$, $\bar{z}_1 = u_1 - i v_1$ Nullstellen von P_n .

Das Produkt

$$(x - z_1)(x - \bar{z}_1) = x^2 - 2u_1x + u_1^2 + v_1^2$$

ist ein **quadratischer** Teiler von P_n mit **reellen** Koeffizienten 1 , $-2u_1$ und $u_1^2 + v_1^2$.

Grundidee: statt der komplexen Nullstellen z_1, \bar{z}_1 wird der quadratische Faktor bestimmt.

Bairstow-Verfahren

Für gegebene $r, s \in \mathbb{R}$ und

$$P_n(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0 \quad (n \geq 2), \quad q_{r,s}(x) = x^2 - rx - s,$$

bestimme $P_{n-2}(x) = x^{n-2} + b_{n-3}x^{n-3} + \dots + b_0$ und $A, B \in \mathbb{R}$ so dass

$$P_n(x) = q_{r,s}(x)P_{n-2}(x) + Ax + B \quad \text{für alle } x \in \mathbb{R}.$$

Durch Koeffizientenvergleich ergibt sich

Polynomdivision eines quadratischen Faktors

Eingabe: Koeffizienten a_0, \dots, a_{n-1} ($a_n = 1$), $r, s \in \mathbb{R}$

$$b_{n-3} = a_{n-1} + r;$$

$$b_{n-4} = a_{n-2} + rb_{n-3} + s;$$

$$\text{Für } j = n-5, \dots, 0: \quad b_j = a_{j+2} + rb_{j+1} + sb_{j+2};$$

$$A = a_1 + rb_0 + sb_1;$$

$$B = a_0 + sb_0;$$

Bairstow-Verfahren

Beachte: P_{n-2} , A und B hängen von r und s ab, und

$$q_{r,s} \text{ teilt genau dann } P_n, \text{ wenn } A = B = 0.$$

Die Aufgabe, solche r und s zu finden, kann man als 2×2 -Nullstellenproblem

$$\begin{cases} A(r, s) = 0 \\ B(r, s) = 0 \end{cases}$$

formulieren. Hierauf lässt sich das Newton-Verfahren anwenden.

Insgesamt ergibt sich das **Bairstow-Verfahren**.