Kapitel 6

Nichtlineare Gleichungen

6.1 Einführung

Problem:

Gesucht sind Lösungen nichtlinearer Gleichungen bzw. Systeme, das heißt es geht beispielsweise um die Bestimmung der Nullstellen eines Polynoms oder einer stetigen Funktion.

Idee:

Man wende ein iteratives Verfahren an, das heißt, man starte mit einem günstigen "Startwert" x_0 und berechne eine bessere Näherungslösung x_1 nach einer gewissen Vorschrift, zum Beispiel durch

$$x_{k+1} = F(x_k), \qquad k = 0, 1, 2, ...$$

(die Funktion F ist in diesem speziellen Fall nur von x_k abhängig, es ist auch möglich, das sie außerdem noch von $x_{k-1}, x_{k-2}, ...$ abhängt, nötig wird dabei allerdings eine Vorgabe weiterer Startwerte). Wir betrachten das folgende Beispiel:

Beispiel:

Gesucht ist eine Lösung der Gleichung

$$e^{-x} - x = 0.$$

Setzen wir nun

$$F(x) =: e^{-x},$$

so ist also ein x^* gesucht, so dass

$$F(x^*) = x^*.$$

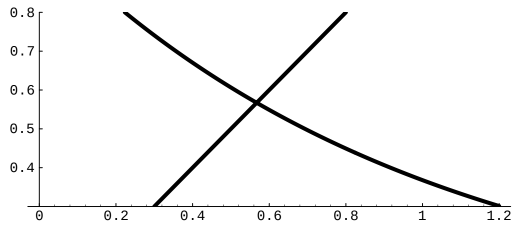
Die Lösung dieses Problems basiert auf dem Banach'schen Fixpunktsatz (Satz 3.2), der besagt: Ist die Funktion $F : D \longrightarrow D$, $D \subseteq \mathbb{R}^n$, eine Kontraktion, das heißt, existiert eine ***Lipschitz-Konstante*** $0 < L < 1$ so, dass

$$\|F(\underline{\mathbf{x}}) - F(\underline{\mathbf{y}})\| \leq L \cdot \|\underline{\mathbf{x}} - \underline{\mathbf{y}}\|$$

für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in D$ gilt, so besitzt F genau einen Fixpunkt $\underline{\mathbf{x}}^* \in D$ und die Folge $\{\underline{\mathbf{x}}_k\}_{k\geq 0}$ mit $\underline{\mathbf{x}}_k = F(\underline{\mathbf{x}}_{k-1})$ mit beliebigem Startwert $\underline{\mathbf{x}}_0 \in D$ konvergiert gegen $\underline{\mathbf{x}}^*$.

Es ist zunächst sicherzustellen, dass der Banach'sche Fixpunktsatz überhaupt in unserem Beispiel anwendbar ist. Wir müssen also ein Intervall $D \subseteq \mathbb{R}$ finden, für das F eine Kontraktion ist. Wir betrachten dazu zunächst das Intervall $[0, 1]$. Da

$$F(0) = 1, \qquad F(1) = \frac{1}{e} \approx 0.367879$$



Darstellung von $f(x) = e^{-x}$ und $f(x) = x$ (siehe Beispiel)

gilt, ist F eine Abbildung von $[0, 1]$ in $[0, 1]$. Nötig für die Kontraktionseigenschaft ist die Existenz einer Konstanten L so, dass

$$\frac{|F(x) - F(y)|}{|x - y|} \leq L < 1.$$

Nach dem Mittelwertsatz der Differentialrechnung existiert ein Wert $\xi \in [0, 1]$ so, dass

$$\frac{|F(x) - F(y)|}{|x - y|} = |F'(\xi)|$$

gilt, das heißt man kann

$$L = \max_{\xi \in [0,1]} |F'(\xi)|$$

wählen. Wegen

$$|F'(\xi)| = |-e^{-\xi}| = e^{-\xi}$$

folgt

$$L = e^{-0} = 1.$$

Damit ist das betrachtete Intervall zu groß, F ist in $[0, 1]$ keine Kontraktion. Durch Verkleinerung des Intervalls (insbesondere an der linken Intervallgrenze, da die Funktion F streng monoton fallend ist) auf das Intervall $[0.2, 1]$ und anschliessende analoge Rechnung erhält man

$$L = e^{-0.2} = 0.81873 < 1,$$

also muss im Intervall $[0.2, 1]$ ein Fixpunkt existieren. Anwendung des Banach'schen Fixpunktsatzes auf die Funktion

$$F : [0.2, 1] \longrightarrow [0.2, 1]$$

mit Startwert 0.5 liefert dann eine Berechnungsvorschrift

$$x_k = F(x_{k-1}), \qquad k = 1, 2, 3, ...$$

und explizit

$$x_k = e^{-x_{k-1}}.$$

Nach Aufruf der folgenden MAPLE-Prozedur erhält man die im Anschluss aufgeführten Ergebnisse:

```
> printlevel:=0;
> x := 0.5;
> for j from 1 to 30 do
      x:=evalf(exp(-x));
      print(j, x)
  end do;
```

x_1	0.6065306597
x_2	0.5452392119
x_3	0.5797030949
$\vdots$	$\vdots$
x_{10}	0.5669072129
$\vdots$	$\vdots$
x_{20}	0.5671424776
$\vdots$	$\vdots$
x_{29}	0.5671432953
x_{30}	0.5671432876

Fazit:

Der Banach'sche Fixpunktsatz selbst kann schon für die Entwicklung eines Iterationsverfahrens benutzt werden.

6.2 Konvergenzordnung und Fehlerabschätzung

Wie gut konvergieren die (mittels des Banachschen Fixpunktsatzes) gewonnenen Iterationsverfahren?

Satz 6.1: (**a-prori Fehlerabschätzung**)

Sei $F : D \longrightarrow D$ eine Kontraktion, $D \subseteq \mathbb{R}^n$, und $D \ni x^* = F(x^*)$ (das heißt x^* ist Fixpunkt der Abbildung F) sowie die Iterationsvorschrift $x_k = F(x_{k-1})$, $k = 1, 2, ...$ mit Startwert $x_0 \in D$ gegeben. Dann gilt die sogenannte ***a-priori-Fehlerabschätzung***

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

Beweis:

Es gilt für $m \in \mathbb{N}$

$$\begin{aligned}
\|x_{k+m} - x_k\| &= \|x_{k+m} - x_{k+m-1} + x_{k+m-1} - x_{k+m-2} + x_{k+m-2} - + \cdots - x_k\| \\
&\leq \|x_{k+m} - x_{k+m-1}\| + \|x_{k+m-1} - x_{k+m-2}\| + \cdots + \|x_{k+1} - x_k\| \\
&= \sum_{\mu=k}^{k+m-1} \|x_{\mu+1} - x_\mu\| \\
&= \sum_{\mu=k}^{m+k-1} \|F(x_\mu) - F(x_{\mu-1})\| \\
&\leq \sum_{\mu=k}^{k+m-1} L^\mu \cdot \|x_1 - x_0\| \\
&= \|x_1 - x_0\| \cdot \sum_{\mu=k}^{k+m-1} L^\mu \\
&= L^k \cdot \|x_1 - x_0\| \cdot \sum_{\nu=0}^{m-1} L^\nu \overset{L\leq 1}{=} L^k \cdot \|x_1 - x_0\| \cdot \frac{1-L^m}{1-L}.
\end{aligned}$$

Dann folgt durch Grenzübergang:

$$\|x^* - x_k\| = \lim_{m\to\infty} \|x_{k+m} - x_k\| \le \frac{L^k}{1-L} \cdot \|x_1 - x_0\|.$$

Die a-priori-Fehlerabschätzung gestattet also eine ungefähre Angabe der benötigten Iterationsschritte zur Unterschreitung einer vorgegebenen Toleranz. □

<u>Satz 6.2:</u> (a-posteriori-Fehlerabschätzung)

Mit den Voraussetzungen wie in Satz 6.1 gilt die a-posteriori-Fehlerabschätzung

$$\|x^* - x_k\| \le \frac{L}{1-L} \cdot \|x_k - x_{k-1}\|, \qquad k = 1, 2, \dots .$$

<u>Beweis:</u>

Die Gültigkeit der Abschätzung folgt aus der Tatsache, dass (Idee analog zu oben)

$$\|x_{k+m} - x_k\| \le \sum_{\mu=k}^{k+m-1} \|x_{\mu+1} - x_\mu\|$$

und außerdem

$$\|x_{\mu+1} - x_\mu\| \le L^{\mu-k+1} \cdot \|x_k - x_{k-1}\|$$

für $\mu = k, k+1, k+2, ..., k+m-1$ gilt. Das liefert

$$\begin{aligned} \|x_{k+m} - x_k\| &\le \sum_{\mu=k}^{k+m-1} L^{\mu-k+1} \cdot \|x_k - x_{k-1}\| = \|x_k - x_{k-1}\| \cdot \sum_{\mu=k}^{k+m-1} L^{\mu-k+1} \\ &= \|x_k - x_{k-1}\| \cdot \sum_{\mu=0}^{m-1} L^{\mu+1} \stackrel{L\le 1}{=} \|x_k - x_{k-1}\| \cdot \underbrace{L \cdot \frac{1-L^m}{1-L}}_{\xrightarrow{m\to\infty} \frac{L}{1-L}} . \end{aligned}$$

□

<u>Beispiel:</u>

Wir betrachten erneut das obige Beispiel mit $F(x) = e^{-x}$. Mit dem oben bereits gefundenen $L = 0.81873$ liefert die a-priori-Fehlerabschätzung

$$\begin{aligned} |x^* - x_{30}| &\le \frac{0.81873^{30}}{1-0.81873} \cdot |\underbrace{0.6065306597}_{x_1} - 0.5| \\ &= 0.001456745198. \end{aligned}$$

Für die a-posteriori-Fehlerabschätzung hingegen erhält man

$$\begin{aligned} |x^* - x_{30}| &\le \frac{0.81872}{1-0.81873} \cdot |x_{30} - x_{29}| \\ &= 0.000000036133. \end{aligned}$$

Die a-posteriori-Fehlerabschätzung ist im allgemeinen mit Abstand besser!

Definition 6.3: (Konvergenzordnung)

Ein Iterationsverfahren besitzt die ***Konvergenzordnung p***, falls die von ihm erzeugte Folge $\{x_k\}$ so gegen den Grenzwert x^* konvergiert, dass

$$\lim_{k\to\infty} \frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|^p} < K, \qquad 0 < K < \infty, \quad p \geq 1$$

gilt, und $K < 1$ für $p = 1$. Für $p = 1$ nennt man die Konvergenz ***linear,*** für $p = 2$ heißt die Konvergenz ***quadratisch.***

Satz 6.4:

Sei $D = [a, b] \subset \mathbb{R}$ und $F : D \longrightarrow D$. Weiter sei F auf D eine Kontraktion und einmal stetig differenzierbar mit $F'(x) \neq 0$ für alle $x \in D$. Dann gilt

$$\lim_{k\to\infty} \frac{x_{k+1}-x^*}{x_k-x^*} = F'(x^*).$$

Dabei war $x_{k+1} = F(x_k)$, $x_0 \in D$ der Startwert und $x^* = F(x^*)$ der Fixpunkt der Kontraktion.

Bemerkung:

Da $F'(x)$ stetig ist, ist diese Ableitung auf einem abgeschlossenen Intervall beschränkt und man erhält in diesem Fall sofort lineare Konvergenz für $|F'(x^*)| < 1$.

Beweis:

Sei $x_0 \neq x^*$, dann folgt, dass $x_k \neq x^*$ für $k \geq 1$ gilt, denn angenommen es existiert ein k so, dass $x_k = x^*$ und $x_{k-1} \neq x^*$ gilt. Dann folgt

$$F(x_k) - F(x_{k-1}) = x_k - x_k = 0. \tag{1}$$

Andererseits ist aber nach dem Mittelwertsatz der Differentialrechnung

$$F(x_k) - F(x_{k-1}) = (x_k - x_{k-1}) \cdot F'(\xi) \tag{2}$$

für ein $\xi \in [a, b]$. Da weiter nach Voraussetzung

$$\underbrace{x_k}_{=x^*} - \underbrace{x_{k-1}}_{\neq x^*} \neq 0$$

und

$$F'(x) \neq 0 \qquad \forall \quad x \in [a, b]$$

gilt, ergibt sich aus (1) und (2) ein Widerspruch, das heißt die Annahme war falsch. Nutzt man erneut den Mittelwertsatz der Differentialrechnung, so erhält man mit $\varepsilon_k := x_k - x^*$

$$\begin{aligned}
\varepsilon_{k+1} &:= x_{k+1} - x^* \\
&= F(x_k) - F(x^*) \\
&= F(x^* + \varepsilon_k) - F(x^*) \\
&\overset{Taylor}{=} F(x^*) + \varepsilon_k \cdot F'(x^* + \theta \cdot \varepsilon_k) - F(x^*) \qquad (\theta \in (0,1)) \\
&= \varepsilon_k \cdot F'(x^* + \theta \cdot \varepsilon_k).
\end{aligned}$$

Da nach Voraussetzung $\varepsilon_k = x_k - x^* \neq 0$ galt, folgt aus der obigen Gleichungskette

$$\frac{\varepsilon_{k+1}}{\varepsilon_k} = F'(x^* + \theta \cdot \varepsilon_k).$$

Für $k \to \infty$ strebt ε_k gegen 0 (Fixpunkteigenschaft von x^*), und aufgrund der Stetigkeit von F' erhalten wir

$$\lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = F'(x^*).$$

Damit ist der Beweis beendet. □

Was ist aber zu tun, wenn $F'(x^*) = 0$ gilt ?

Satz 6.5:

Sei $F : [a,b] \longrightarrow [a,b]$ gegeben und auf dem Intervall $[a,b]$ eine Kontraktion sowie zweimal stetig differenzierbar. Ferner sei $F'(x^*) = 0$ und $F''(x) \neq 0$ für alle $x \in [a,b]$. Dann gilt

$$\lim_{k\to\infty} \frac{x_{k+1} - x^*}{(x_k - x^*)^2} = \frac{1}{2} \cdot F''(x^*),$$

das heißt, es liegt quadratische Konvergenz vor.

Beweis:

Es gilt

$$\begin{aligned}
& x_{k+1} - x^* \\
= \quad & F(x_k) - F(x^*) \\
= \quad & F(x^* + (x_k - x^*)) - F(x^*) \\
\stackrel{Taylor}{=} \quad & F(x^*) + (x_k - x^*) \cdot \underbrace{F'(x^*)}_{=0} + \frac{1}{2} \cdot (x_k - x^*)^2 \cdot F''(x^* + \theta \cdot (x_k - x^*)) - F(x^*) \\
= \quad & \frac{1}{2} \cdot (x_k - x^*)^2 \cdot F''(x^* + \theta \cdot (x_k - x^*))
\end{aligned}$$

mit einem $\theta \in (0,1)$.Wir zeigen wieder, dass für einen Startwert $x_0 \neq x^*$ auch alle weiteren x_k, $k \geq 1$ ungleich x^* sind: Angenommen, es gäbe ein ein $k \in \mathbb{N}$ mit $x_k \neq x^*$ und $x_{k+1} = x^*$. dann folgt aus der obigen Rechnung, dass

$$\frac{1}{2} \cdot (x_k - x^*)^2 \cdot F''(x^* + \theta \cdot (x_k - x^*)) = 0$$

im Widerspruch zu $(x_k - x^*) \neq 0$ und $F''(x) \neq 0$ für alle $x \in [a,b]$. Also ist eine Division durch den Faktor $(x_k - x^*)^2$ erlaubt und die obige Gleichung ist äquivalent zu

$$\frac{x_{k+1} - x^*}{(x_k - x^*)^2} = \frac{1}{2} \cdot F''(x^* + \theta \cdot (x_k - x^*)).$$

Da F'' nach Voraussetzung stetig in $[a,b]$ ist und die Folge $\{x_k\}$ gegen x^* konvergiert, ergibt sich

$$\lim_{k\to\infty} F''(x^* + \theta \cdot (x_k - x^*)) = \frac{1}{2} \cdot F''(x^*),$$

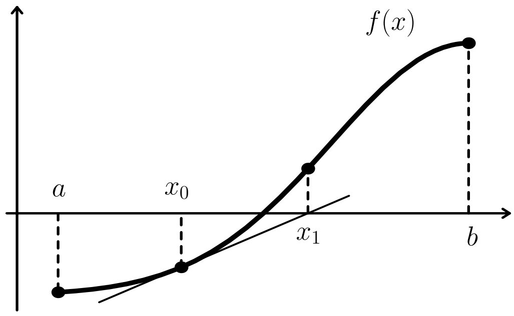
damit ist der Beweis beendet. □

6.3 Das Newton-Verfahren

Es sei $f(x)$ im Intervall $[a, b]$ stetig differenzierbar und es existiere eine Nullstelle von $f(x)$ in $[a, b]$.

Idee:

Man wähle einen Startwert $x_0 \in [a, b]$. Die Werte x_{k+1} $(k = 0, 1, 2, ...)$ werden dann so gewählt, dass sie jeweils die Schnittpunkte der x-Achse mit den Tangenten an die Funktion $f(x)$ im Punkt x_k sind.



Geometrische Darstellung des Newton-Verfahrens

Allgemein lautet die Tangentengleichung für die zu betrachtende Tangente

$$y \; = \; f(x_k) + (x - x_k) \cdot f'(x_k),$$

dabei soll nun der Wert, für den $y = 0$ gilt, der nächste Iterationspunkt sein, das heißt, es gilt $y = 0$ an der Stelle x_{k+1}. Damit lässt sich x_{k+1} durch Umstellen der Gleichung berechnen:

$$\begin{aligned} 0 \; &= \; f(x_k) + (x_{k+1} - x_k) \cdot f'(x_k) \\ \Longrightarrow \quad x_{k+1} \; &= \; x_k - \frac{f(x_k)}{f'(x_k)} \qquad\qquad (k = 0, 1, 2, ...). \end{aligned}$$

Somit ergibt sich hier direkt die Iterationsformel des Newton-Verfahrens. Dieses Verfahren ist eine ***Fixpunktiteration***. Setzt man

$$F(x) \; = \; x - \frac{f(x)}{f'(x)}$$

so ist x^* mit

$$F(x^*) \; = \; x^*$$

gesucht, denn

$$F(x^*) \; = \; x^* \quad \Longleftrightarrow \quad x^* - \frac{f(x^*)}{f'(x^*)} = x^* \quad \Longleftrightarrow \quad \frac{f(x^*)}{f'(x^*)} = 0.$$

Nötige Voraussetzung bei der Anwendung des Newton-Verfahrens ist, dass $f'(x) \neq 0$ gilt für alle $x_k,\; k = 0, 1, \ldots$.

Satz 6.6:

Sei $f \in C^2[A, B]$, $f : (A, B) \longrightarrow \mathbb{R}$, $(A, B) \subset \mathbb{R}$. Für ein Intervall $(a, b) \subset (A, B)$ gelte

(i) $f'(x) \neq 0 \quad \forall\ x \in [a, b]$, d.h., f ist in $[a, b]$ monoton.

(ii) Es ist
$$\frac{|f(x) \cdot f''(x)|}{|f'(x)|^2} \leq L < 1 \qquad \forall\ x \in [a, b].$$
(Diese Bedingung sichert die Kontraktionseigenschaft für $F(x)$.)

(iii) Es sei
$$x - \frac{f(x)}{f'(x)} := F(x) \in [a, b] \qquad \forall\ x \in [a, b],$$
d.h., $F(x)$ ist eine Abbildung von $[a, b]$ in $[a, b]$.

Dann existiert genau eine Nullstelle $x^* \in [a, b]$ und für einen beliebigen Startwert $x_0 \in (a, b)$ konvergiert das Newton-Verfahren gegen x^* mit $F(x^*) = x^*$ und die Folge $\{x_k\}$ ist mindestens quadratisch konvergent.

Bemerkung:

Ist x^* eine Nullstelle von $f \in C^2[A, B]$ mit $f'(x^*) \neq 0$, dann existiert immer ein Intervall (a, b) mit $x^* \in (a, b)$ so, dass (i)-(iii) erfüllt sind. Die Eigenschaft (i) folgt für ein Intervall
$$[x^* - \varepsilon,\ x^* + \varepsilon]$$
aus der Stetigkeit von f'. Aufgrund der Stetigkeit von f'' auf diesem abgeschlossenen Intervall lässt sich eine Abschätzung der Form
$$|f''(x)| \leq C$$
mit positiver Konstante C finden, da außerdem $f'(x^*) \neq 0$ vorausgesetzt war, ist auf diesem Intervall auch eine Abschätzung der Form
$$|f'(x)| \geq c$$
mit positiver Konstante c gültig. Damit gilt
$$\frac{|f(x) \cdot f''(x)|}{|f'(x)|^2} \leq \frac{|f(x)| \cdot C}{c^2}.$$
Da $f(x^*) = 0$ war, kann man ein ε finden, für das
$$|f(x^* + h)| < \frac{c^2}{C}$$
für alle $|h| < \varepsilon$ gilt. Damit gilt auch Eigenschaft (ii).
Wählt man dann speziell $\varepsilon > 0$ so, dass für alle $x \in [x^* - \varepsilon,\ x^* + \varepsilon]$ die Eigenschaft
$$c \leq |f'(x)| \leq 2c \qquad (c > 0)$$
gilt (auch das ist aufgrund der Stetigkeit von f' und der daraus folgenden Beschränktheit auf abgeschlossenen Intervallen immer möglich) und ist o.B.d.A. $f'(x) > 0$ für alle $x \in [x^* - \varepsilon,\ x^* + \varepsilon]$, dann ergibt sich (da x^* ja nach Voraussetzung die Nullstelle von f ist), dass
$$\begin{cases} f(x) < 0 & \text{für} \quad x < x^*, \\ f(x) > 0 & \text{für} \quad x > x^*, \end{cases}$$

(streng monotones Wachstum!). Für $x \in [x^*,\ x^* + \varepsilon]$ gilt dann

$$\begin{aligned} x - \frac{f(x)}{f'(x)} \quad &\in \quad \left[x - \frac{|f(x)|}{c},\ x - \frac{|f(x)|}{2c} \right] \\ &\subset \quad \left[x^* - \frac{|f(x)|}{c} + |x - x^*|,\ x^* - \frac{|f(x)|}{2c} + |x - x^*| \right]. \end{aligned}$$

Aus dem Mittelwertsatz der Differentialrechnung und der Stetigkeit von f', folgt außerdem für eine Zwischenstelle $\xi \in (x^*,\, x)$

$$\begin{aligned} \frac{|f(x) - \overbrace{f(x^*)}^{=0}|}{|x - x^*|} \ &= \ |f'(\xi)| \\ \Longrightarrow \quad c \cdot |x - x^*| \ &\le \ |f(x)| \ \le \ 2c \cdot |x - x^*|, \end{aligned}$$

so dass

$$x - \frac{f(x)}{f'(x)} \in \left[x^* - |x - x^*|,\ x^* + \frac{1}{2}|x - x^*| \right] \subset \left[x^* - \varepsilon,\ x^* + \frac{\varepsilon}{2} \right].$$

Analog folgt für $x \in [x^* - \epsilon, x^*]$

$$x - \frac{f(x)}{f'(x)} \in \left[x^* - \frac{\varepsilon}{2},\ x^* + \varepsilon \right].$$

Damit ist auch Eigenschaft (iii) erfüllt.

<u>Beweis von Satz 6.6:</u>

1) Nach Eigenschaft (iii) ist $F(x)$ eine Abbildung von $[a, b]$ in $[a, b]$. Nach dem Mittelwertsatz der Differentialrechnung existiert ein $\xi \in (x, y)$ mit

$$\begin{aligned} \frac{|F(x) - F(y)|}{|x - y|} \ &= \ |F'(\xi)| \\ \Longleftrightarrow \quad |F(x) - F(y)| \ &= \ |F'(\xi)| \cdot |x - y|. \end{aligned}$$

F ist eine Kontraktion, wenn wir zeigen können, dass

$$\max_{\xi \in [a,b]} |F'(\xi)| \ \le \ L \ < \ 1$$

gilt. Aus

$$F(x) \ = \ x - \frac{f(x)}{f'(x)}$$

folgt

$$F'(\xi) = 1 - \frac{f'(\xi)^2 - f(\xi) \cdot f''(\xi)}{f'(\xi)^2} = \frac{f(\xi) \cdot f''(\xi)}{f'(\xi)^2}.$$

Nach Eigenschaft (ii) ist $F(x)$ also eine Kontraktion. Da F eine Abbildung von $[a, b]$ in $[a, b]$ ist, existiert nach dem Banach'schen Fixpunktsatz also ein Fixpunkt $x^* \in [a, b]$ mit $F(x^*) \ = \ x^*$. Damit erhält man unter Berücksichtigung der Eigenschaft (i),

$$f'(x^*) \ \neq \ 0,$$

dass

$$x^* - \frac{f(x^*)}{f'(x^*)} \ = \ x^* \quad \Longleftrightarrow \quad \frac{f(x^*)}{f'(x^*)} \ = \ 0 \quad \Longrightarrow \quad f(x^*) \ = \ 0$$

gilt, das heißt, der Fixpunkt x^* ist gleichzeitig die gesuchte Nullstelle der Funktion f im Intervall $[a, b]$. Die Eindeutigkeit das Fixpunktes folgt aus Satz 3.2.

2) Das Newton-Verfahren hat die Konvergenzordnung 2, denn eine Taylor-Entwicklung der Funktion f (Entwicklungspunkt x_k) ergibt sich zu

$$\begin{aligned} f(x^*) &= f(x_k + (x^* - x_k)) \\ &= f(x_k) + f'(x_k) \cdot (x^* - x_k) + \frac{1}{2} \cdot f''(x_k + \theta \cdot (x^* - x_k)) \cdot (x^* - x_k)^2, \end{aligned}$$

wobei $\theta \in (0,1)$ gilt. Wegen $f(x^*) = 0$ ergibt sich mit der Iterationsformel

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

aus der Taylor-Entwicklung

$$\begin{aligned} & 0 = f(x_k) + f'(x_k) \cdot (x^* - x_k) + \frac{1}{2} \cdot (x^* - x_k)^2 \cdot f''(x_k + \theta \cdot (x^* - x_k)) \\ \iff & \underbrace{\frac{f(x_k)}{f'(x_k)} + (x^* - x_k)}_{x^* - x_{k+1}} = -\frac{1}{2} \cdot (x^* - x_k)^2 \cdot f''(x_k + \theta \cdot (x^* - x_k)) \cdot \frac{1}{f'(x_k)} \\ \iff & x^* - x_{k+1} = -\frac{1}{2} \cdot (x^* - x_k)^2 \cdot f''(x_k + \theta \cdot (x^* - x_k)) \cdot \frac{1}{f'(x_k)} \\ \implies & \frac{|x^* - x_{k+1}|}{|x^* - x_k|^2} = \frac{1}{2} \cdot |f''(x_k + \theta \cdot (x^* - x_k))| \cdot \left| \frac{1}{f'(x_k)} \right| \\ \implies & \lim_{k \to \infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^2} = \frac{1}{2} \cdot \left| \frac{f''(x^*)}{f'(x^*)} \right| < \infty \end{aligned}$$

das heißt, es ergibt sich die Konvergenzordnung 2. □

Beispiel:

Gesucht ist die m-te Wurzel aus $K > 0$, also die Nullstelle der Funktion

$$f(x) = x^m - K, \qquad (m \geq 2).$$

Diese ist im Intervall $(0, \infty)$ eindeutig. Zunächst ist vor der Anwendung des Newton-Verfahrens sicherzustellen, dass alle in Satz 6.6 genannten Bedingungen erfüllt sind. Es ist also zuerst ein geeignetes Intervall zu wählen, so dass dort (i)-(iii) gelten.

(i) Es gilt

$$f'(x) = m \cdot x^{m-1} \neq 0 \qquad \forall \quad x > 0.$$

(ii) Es gilt

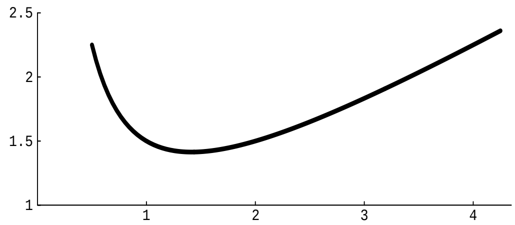
$$\begin{aligned} \frac{f(x) \cdot f''(x)}{f'(x)^2} &= \frac{(x^m - K) \cdot m(m-1) \cdot x^{m-2}}{m^2 \cdot x^{2m-2}} \\ &= \frac{(x^m - K) \cdot (m-1)}{m \cdot x^m} \\ &= \frac{m-1}{m} \cdot \left(1 - \frac{K}{x^m}\right). \end{aligned}$$

Um Eigenschaft (ii) aus Satz 6.6 zu erfüllen, muss dieser Quotient durch eine Konstante $L < 1$ abgeschätzt werden können. Es ist

$$\left| 1 - \frac{K}{x^m} \right| \leq 1$$

für

$$K \leq 2x^m \iff x^m \geq \frac{K}{2}.$$



$F(x) = x/2 + 1/x$ (siehe Beispiel)

Damit ist beispielsweise das Intervall $[a, b]$ mit $a^m \geq \frac{K}{2}$ und $b = \infty$ zulässig.
Wir wählen ein a mit $\frac{K}{2} \leq a^m < K$.
(iii) Zu überprüfen ist, ob für dieses Intervall $[a, b]$ die Bedingung (iii) gilt, d.h., ob

$$\begin{aligned} F(x) &:= x - \frac{f(x)}{f'(x)} \\ &= x - \frac{x^m - K}{m \cdot x^{m-1}} \\ &= x - \frac{x}{m} + \frac{K}{m \cdot x^{m-1}} \\ &= \frac{m-1}{m} \cdot x + \frac{K}{m \cdot x^{m-1}} \end{aligned}$$

eine Abbildung von $[a, b]$ in $[a, b]$ ist. $F(x)$ ist monoton wachsend für $x^m > K$ ($x \geq \sqrt[m]{K}$) und monoton fallend für $a^m \leq x^m \leq K$ ($a \leq x \leq \sqrt[m]{K}$), denn es ist

$$F'(x) = \frac{m-1}{m} + \frac{K}{m} \cdot (-m+1) \cdot x^{-m} = \frac{m-1}{m} \cdot \left(1 - \frac{K}{x^m}\right).$$

Zu Veranschaulichung betrachte die folgende Skizze ($m = 2$, $K = 2$, $f(x) = x^2 - 2$, $F(x) = \frac{1}{2}x + \frac{1}{x}$):

Es gilt (wieder im allgemeinen Fall!) für $\sqrt[m]{K/2} \leq a < \sqrt[m]{K}$:

$$\begin{aligned} F(a) &= \frac{m-1}{m} \cdot a + \frac{K}{m \cdot a^{m-1}} \\ &\overset{a^m < K}{\geq} \frac{m-1}{m} \cdot a + \frac{a^m}{m \cdot a^{m-1}} \\ &= \frac{m-1}{m} \cdot a + \frac{a}{m} \\ &= a \end{aligned}$$

und für $x = \sqrt[m]{K}$

$$\begin{aligned} F\left(\sqrt[m]{K}\right) &= \frac{m-1}{m} \cdot \sqrt[m]{K} + \frac{K}{m \cdot \left(\sqrt[m]{K}\right)^{m-1}} \\ &= \frac{m-1}{m} \cdot \sqrt[m]{K} + \frac{1}{m} \cdot \frac{\left(\sqrt[m]{K}\right)^m}{\left(\sqrt[m]{K}\right)^{m-1}} \\ &= \frac{m-1}{m} \cdot \sqrt[m]{K} + \frac{1}{m} \cdot \sqrt[m]{K} \\ &= \sqrt[m]{K} \\ &> a. \end{aligned}$$

Alle nötigen Bedingungen aus Satz 6.6 sind somit erfüllt auf dem Intervall $[a, \infty)$. Die Iterationsvorschrift lautet in diesem Fall explizit

$$\begin{aligned} x_{k+1} &= x_k - \frac{f(x_k)}{f'(x_k)} \\ &= x_k - \frac{x_k^m - K}{m \cdot x_k^{m-1}} \\ &= \frac{1}{m} \cdot \left(\frac{K}{x_k^{m-1}} + (m-1) \cdot x_k \right). \end{aligned}$$

Im Fall $m = 2$ (Quadratwurzel) ergibt sich daraus speziell

$$x_{k+1} = \frac{1}{2} \cdot \left[\frac{K}{x_k} + x_k \right],$$

wobei z.B. der Startwert $x_0 = K$ gewählt werden kann. Nach dieser Formel berechnet ein Taschenrechner/Computer die Quadratwurzel.
Für die Funktion

$$f(x) = x^2 - 2$$

mit dem Startwert $x_0 = 2$ und der Iterationsvorschrift

$$x_{k+1} = \frac{1}{2} \cdot \left[\frac{2}{x_k} + x_k \right] \qquad k = 0, 1, 2, ...,$$

finden wir

$$x_1 = 1.5,$$

$$x_2 = \frac{1}{2} \cdot \left[\frac{2}{\frac{3}{2}} + \frac{3}{2} \right] = \frac{1}{2} \cdot \frac{17}{6} = \frac{17}{12} = 1.4166\overline{6}$$

und

$$x_3 = \frac{1}{2} \cdot \left[\frac{2}{\frac{17}{12}} + \frac{17}{12} \right] = \frac{1}{2} \cdot \frac{577}{204} = \frac{577}{408} = 1.414215686,$$

dabei ist die Berechnung im dritten Iterationsschritt bereits auf 6 Nachkommastellen genau!

6.4 Intervallschachtelung, Regula falsi, Sekantenmethode

Das Newton-Verfahren zur Berechnung einer Nullstelle ist nicht immer anwendbar (z.B., wenn f nicht differenzierbar ist. Wir betrachten daher noch einige weitere Methoden zur Berechnung der Nullstelle einer Funktion. Es sei im Folgenden $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine nichtlineare, stetige Funktion.

Intervallschachtelung

Sei das Intervall $I_0 = [a, b]$ bekannt mit $f(a) \cdot f(b) < 0$, das heißt in einem der beiden Endpunkte ist der Funktionswert negativ, im anderen positiv. Für stetiges f existiert damit eine Nullstelle von f in $[a, b]$.

Algorithmus:

O.B.d.A. sei $f(a) < 0$, $f(b) > 0$. Man geht dann in der folgenden Weise vor:

(i) Solange $|b - a| > toleranz$
  1. Berechne $x = a + (b - a) \cdot 0.5$.
  2. Berechne $f(x)$.
     Falls $f(x) < 0$ gilt, so setze $a := x$.
     Falls $f(x) > 0$ gilt, so setze $b := x$.

(ii) Drucke a, b.

Ist $L_0 = b - a$, so ergibt sich nach k Iterationsschritten

$$L_k = \frac{b-a}{2^k},$$

also

$$|x_k - x^*| \le \frac{b-a}{2^{k+1}},$$

dabei bezeichnet x_k jeweils den Mittelpunkt des Intervalls I_k. Die Fehlerschranke nimmt mit dem Faktor $\frac{1}{2}$ ab, die Konvergenzordnung ist 1.

Methode der Regula falsi

Voraussetzung:

Das Startintervall $[a, b]$ mit $f(a) \cdot f(b) < 0$ ist bekannt. Setze dann $x_0 := a$, $x_1 := b$.

Idee:

Man bestimme den Wert x_2 als Nullstelle der Geraden g_1 durch die Punkte $(x_0, f(x_0))$ und $(x_1, f(x_1))$, das heißt, g_1 besitzt die Gestalt

$$g_1(x) = f(x_0) + (x - x_0) \cdot \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Für $g_1(x_2) = 0$ folgt

$$\begin{aligned} & f(x_0) + (x_2 - x_0) \cdot \frac{f(x_1) - f(x_0)}{x_1 - x_0} = 0 \\ \iff \quad & x_2 = x_0 - f(x_0) \cdot \frac{x_1 - x_0}{f(x_1) - f(x_0)} = \frac{x_0 \cdot f(x_1) - x_1 \cdot f(x_0)}{f(x_1) - f(x_0)}. \end{aligned}$$

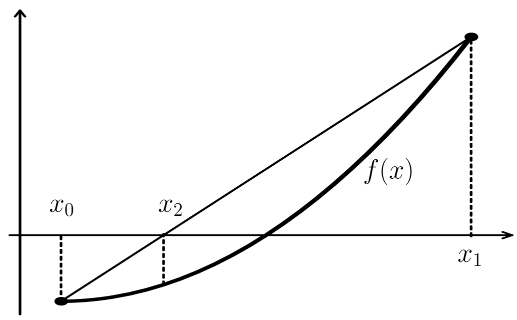
Es sei $f(x_0) < 0$, $f(x_1) > 0$. Falls $f(x_2) > 0$ gilt, so wähle $I_1 := [x_0, x_2]$, sonst wähle $I_1 := [x_2, x_1]$. Man kann so folgenden Algorithmus verwenden:

Algorithmus:

Bei gegebenen $x_0 < x_1$, $y_0 = f(x_0)$, $y_1 = f(x_1)$ mit $y_0 \cdot y_1 < 0$ berechne man für $k = 1, 2, 3, ...$

$$x_{k+1} := \frac{x_{k-1} \cdot y_k - x_k \cdot y_{k-1}}{y_k - y_{k-1}}, \qquad y_{k+1} := f(x_{k+1}).$$

Falls dann $y_{k+1} = 0$ (aufgrund von Rundungsfehlern in der Realität wohl besser $|y_{k+1}| < toleranz$) gilt, dann stoppt der Algorithmus, x_{k+1} ist dann die gesuchte Nullstelle.



Geometrische Darstellung der Regula falsi Methode

Ansonsten

- falls $y_{k+1} \cdot y_{k-1} < 0$ gilt, so setze $x_k := x_{k-1}$, $y_k := y_{k-1}$
- falls $y_{k+1} \cdot y_{k-1} > 0$ gilt, so vertausche die Punkte (x_{k+1}, y_{k+1}) und (x_k, y_k)

Um die Konvergenzordnung dieser Methode zu bestimmen, benötigen wir folgendes Lemma.

Lemma 6.7:

Seien $x_0, x_1, x_2 \in \mathbb{R}$ mit $x_0 < x_1 < x_2$. Ist $f \in C^2[x_0, x_2]$, so existiert eine Zwischenstelle $\xi \in [x_0, x_2]$ mit

$$f[x_0, x_1, x_2] = \frac{1}{2} f''(\xi).$$

Dabei ist

$$f[x_0, x_1, x_2] := \frac{f[x_0, x_1] - f[x_1, x_2]}{x_0 - x_2}$$

und

$$f[x_0, x_1] := \frac{f(x_0) - f(x_1)}{x_0 - x_1}, \qquad f[x_1, x_2] =: \frac{f(x_1) - f(x_2)}{x_1 - x_2},$$

das heißt, wir erhalten explizit die ***dividierte Differenz zweiter Ordnung***

$$\begin{aligned} f[x_0, x_1, x_2] &= \frac{\left(\frac{f(x_0) - f(x_1)}{x_0 - x_1}\right) - \left(\frac{f(x_1) - f(x_2)}{x_1 - x_2}\right)}{x_0 - x_2} \\ &= \frac{f(x_0)}{(x_0 - x_1)(x_0 - x_2)} + \frac{f(x_1)}{(x_1 - x_0)(x_1 - x_2)} + \frac{f(x_2)}{(x_2 - x_0)(x_2 - x_1)}. \end{aligned}$$

Beweis:

Für beliebiges $a \in [x_0, x_2]$ gilt

$$\begin{aligned} & f(x) = f(a) + \int_a^x \underbrace{f'(t)}_{f'(t)\cdot 1} \, dt \\ \Longleftrightarrow \quad & f(x) \stackrel{p.I.}{=} f(a) + f'(t) \cdot (t - x)\Big|_{t=a}^{x} - \int_a^x f''(t) \cdot (t - x) \, dt \qquad (*) \\ \Longleftrightarrow \quad & f(x) = f(a) + f'(a) \cdot (x - a) + \int_a^x f''(t) \cdot (x - t) \, dt. \end{aligned}$$

Setzt man jetzt

$$\omega_0 = \frac{1}{(x_0-x_1)(x_0-x_2)}, \qquad \omega_1 = \frac{1}{(x_1-x_0)(x_1-x_2)}, \qquad \omega_2 = \frac{1}{(x_2-x_0)(x_2-x_1)},$$

so erhält man für die dividierte Differenz zweiter Ordnung

$$\begin{aligned} f[x_0,x_1,x_2] &= f(x_0)\cdot\omega_0 + f(x_1)\cdot\omega_1 + f(x_2)\cdot\omega_2 \\ &\overset{(*)}{=} \sum_{k=0}^{2}\Big[f(a) + f'(a)\cdot(x_k-a) + \int_a^{x_k} f''(t)\cdot(x_k-t)\,dt\Big]\cdot\omega_k. \end{aligned}$$

Unter Verwendung der Tatsache, dass

$$\begin{aligned} \omega_0+\omega_1+\omega_2 &= \frac{1}{(x_0-x_1)(x_0-x_2)} + \frac{1}{(x_1-x_0)(x_1-x_2)} + \frac{1}{(x_2-x_0)(x_2-x_1)} \\ &= \frac{(x_1-x_2)-(x_0-x_2)+(x_0-x_1)}{(x_0-x_1)(x_0-x_2)(x_1-x_2)} \\ &= 0 \end{aligned}$$

sowie

$$\begin{aligned} \omega_0\cdot x_0 + \omega_1\cdot x_1 + \omega_2\cdot x_2 &= \frac{x_0(x_1-x_2)+x_1(x_2-x_0)+x_2(x_0-x_1)}{(x_0-x_1)(x_0-x_2)(x_1-x_2)} \\ &= \frac{x_0x_1 - x_0x_2 + x_1x_2 - x_1x_0 + x_2x_0 - x_2x_1}{(x_0-x_1)(x_0-x_2)(x_1-x_2)} \\ &= 0 \end{aligned}$$

gilt, ergibt sich

$$\sum_{k=0}^{2}\big[f(a) + (x_k-a)\cdot f'(a)\big]\cdot\omega_k = 0,$$

das heißt, es folgt

$$f[x_0,x_1,x_2] = \sum_{k=0}^{2}\omega_k\cdot\int_a^{x_k} f''(t)\cdot(x_k-t)\,dt.$$

Wählt man jetzt $a = x_1$, dann erhält man

$$\int_{x_1}^{x_0} f''(t)\cdot\frac{x_0-t}{(x_0-x_1)(x_0-x_2)}\,dt + \int_{x_1}^{x_2} f''(t)\cdot\frac{x_2-t}{(x_2-x_0)(x_2-x_1)}\,dt = \int_{x_0}^{x_2} f''(t)\cdot K(t)\,dt$$

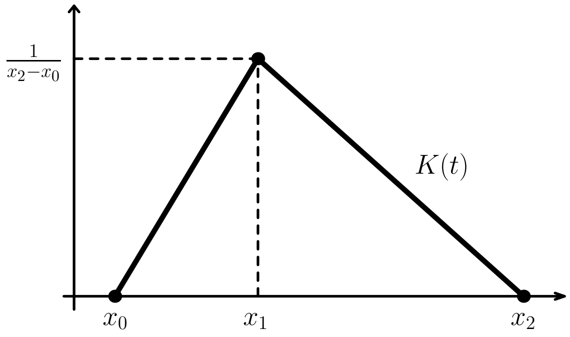
mit

$$K(t) = \begin{cases} \dfrac{t-x_0}{(x_0-x_1)(x_0-x_2)} \geq 0 & \text{für} \quad x_0 \leq t \leq x_1, \\[2ex] \dfrac{x_2-t}{(x_2-x_0)(x_2-x_1)} \geq 0 & \text{für} \quad x_1 < t \leq x_2. \end{cases}$$

Die Funktion $K(t)$ heißt ***Kern,*** in diesem Fall ist der Kern positiv, dies wird in der folgenden Skizze deutlich:

Wegen

$$\Big(\min_{x\in[x_0,x_2]} f''(x)\Big)\int_{x_0}^{x_2} K(t)\,dt \leq \int_{x_0}^{x_2} f''(t)\,K(t)\,dt \leq \Big(\max_{x\in[x_0,x_2]} f''(x)\Big)\int_{x_0}^{x_2} K(t)\,dt$$



Kern $K(t)$.

folgt dann aus der Stetigkeit von f''

$$f[x_0, x_1, x_2] \;=\; f''(\xi) \cdot \underbrace{\int\limits_{x_0}^{x_2} K(t)\, dt}_{=\frac{1}{2},\, vgl. Skizze}$$

mit einer Zwischenstelle $\xi \in [x_0, x_2]$. $\square$

Satz 6.8:

Sei $f \in C^2[a,b]$, $x_0 = a$, $x_1 = b$ und $f(x_0) \cdot f(x_1) < 0$. Sei weiterhin $f'(x) \neq 0$, $f''(x) \neq 0$ für alle $x \in [a,b]$. Die Konvergenzordnung der Methode der Regula falsi ist dann lokal mindestens Eins.

Bemerkung:

Die Bedingung $f'(x) \neq 0$ in $[a,\, b]$ sichert die Monotonie von f. Insbesondere ist dann x^* eine ***einfache Nullstelle*** von f. Falls $f'(x^*) \neq 0$ und $f''(x^*) \neq 0$, dann existiert immer ein Intervall $[\tilde{a}, \tilde{b}] \subset U_\varepsilon(x^*)$ mit $f'(x) \neq 0$, $f''(x) \neq 0$ für alle $x \in [\tilde{a}, \tilde{b}]$, da f' und f'' stetig sind.

Beweis:

Es gilt

$$\begin{aligned}
& x_{k+1} \;=\; \frac{x_{k-1} \cdot f(x_k) - x_k \cdot f(x_{k-1})}{f(x_k) - f(x_{k-1})} \\
\Longrightarrow\quad & x_{k+1} - x^* \;=\; \frac{(x_{k-1} - x^*) \cdot f(x_k) - (x_k - x^*) \cdot f(x_{k-1})}{f(x_k) - f(x_{k-1})} \\
\Longrightarrow\quad & \frac{(x_{k+1} - x^*) \cdot (f(x_k) - f(x_{k-1}))}{x_k - x_{k-1}} = \frac{(x_{k-1} - x^*) \cdot f(x_k) - (x_k - x^*) \cdot f(x_{k-1})}{x_k - x_{k-1}} \\
& \overset{f(x^*)=0}{=} \left(\frac{\dfrac{f(x_{k-1}) - f(x^*)}{x_{k-1} - x^*} - \dfrac{f(x^*) - f(x_k)}{x^* - x_k}}{x_{k-1} - x_k} \right) \cdot (x^* - x_k) \cdot (x^* - x_{k-1}). \\
\Longrightarrow\quad & (x_{k+1} - x^*) \cdot \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \;=\; f[x_{k-1}, x^*, x_k] \cdot (x^* - x_k) \cdot (x^* - x_{k-1}).
\end{aligned}$$

Nach dem Mittelwertsatz und Lemma 6.7 gilt weiter

$$(x_{k+1} - x^*) \cdot f'(\eta) \;=\; \frac{1}{2} \cdot f''(\xi) \cdot (x^* - x_k) \cdot (x^* - x_{k-1})$$

mit $\eta, \xi \in (x_{k-1}, x_k)$. Damit folgt durch Umformen

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} = \frac{1}{2} \cdot \frac{|f''(\xi)|}{|f'(\eta)|} \cdot |x^* - x_{k-1}|.$$

Aufgrund der Stetigkeit von f' und f'' im Intervall $[a, b]$ existiert eine Konstante $M > 0$ mit

$$\frac{|f''(\xi)|}{|f'(\eta)|} < M \qquad \forall \quad \xi, \eta \in [x_{k-1}, x_k],$$

das heißt, es folgt

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} \leq \frac{1}{2} \cdot M \cdot |x^* - x_{k-1}|.$$

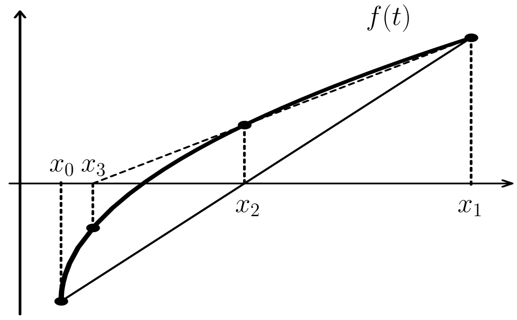
Wegen $\lim\limits_{k\to\infty} x_k = x^*$ ergibt sich lineare Konvergenz. □

Sekantenmethode

Die Sekantenmethode ist eine Modifikation der Methode der Regula falsi. Gegeben seien hier zwei Näherungswerte x_0, x_1, die die Nullstelle ***nicht unbedingt einschließen müssen.***

Idee:

Man berechne den Schnittpunkt der Sekante (der Geraden durch die Punkte $(x_0, f(x_0))$ und $(x_1, f(x_1))$) mit der x-Achse und erhält so x_2. Im Gegensatz zur Methode der Regula falsi wird jetzt die Iteration ***ohne Umbenennen*** der Punkte fortgesetzt, das heißt, man führt den nächsten Iterationsschritt mit den Punkten x_1 und x_2 durch. Es werden immer die letzten beiden Punkte zur Berechnung der neuen Sekante herangezogen. Man erhält zur



Geometrische Darstellung der Sekantenmethode

Berechnung des ″nächsten″ Punktes die Iterationsvorschrift

$$x_{k+1} = x_k - f(x_k) \cdot \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right) = \frac{x_{k-1} \cdot f(x_k) - x_k \cdot f(x_{k-1})}{f(x_k) - f(x_{k-1})},$$

das heißt formal dieselbe Vorschrift wie bei der Methode der Regula falsi. Voraussetzung ist allerdings, dass

$$f(x_0) \neq f(x_1)$$

gilt, da sonst der Nenner des auftretenden Bruches nicht definiert ist.

Satz 6.9:

Sei $f \in C^2[a,b]$, $x_0, x_1 \in [a,b]$ und sei $x_k \in [a,b]$ für $k = 2, 3, \ldots$. Außerdem gelte $f(a) \cdot f(b) < 0$. Ist weiterhin $f'(x) \neq 0$, $f''(x) \neq 0$ für alle $x \in [a,b]$, dann ist die lokale Konvergenzordnung des Sekantenverfahrens $\frac{1+\sqrt{5}}{2} > 1$ (superlineare Konvergenz).

Beweis:

Aus der Iterationsvorschrift

$$x_{k+1} \;=\; x_k - f(x_k) \cdot \left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right)$$

folgt wie im Beweis von Satz 6.8, dass

$$(x_{k+1} - x^*) \cdot \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \;=\; \pm f[x_{k-1}, x^*, x_k] \cdot (x^* - x_k) \cdot (x^* - x_{k-1})$$

gilt, dabei ist zu beachten, dass (anders als im Beweis zu Satz 6.8) ***nicht unbedingt*** $x_{k-1} < x^* < x_k$ gilt (deshalb entsteht auf der rechten Seite unter Umständen ein anderes Vorzeichen als im letzten Beweis). Mit dem Mittelwertsatz und Lemma 6.7 erhält man eine Darstellung der Form

$$(x_{k+1} - x^*) \cdot f'(\eta_k) \;=\; \pm\frac{1}{2} \cdot f''(\xi_k) \cdot (x^* - x_k) \cdot (x^* - x_{k-1}) \qquad (*)$$

mit $\eta_k, \xi_k \in I_k = \,[\min\{x_{k-1}, x^*, x_k\}, \max\{x_{k-1}, x^*, x_k\}]$. Sei nun

$$M_k := \frac{1}{2}\frac{|f''(\xi_k)|}{|f'(\eta_k)|} \;<\; M := \frac{1}{2} \max_{\eta,\xi \in [a,b]} \frac{|f''(\xi)|}{|f'(\eta)|} < \infty.$$

Zunächst sehen wir ein, dass x_k lokal linear gegen x^* konvergiert, denn aus $(*)$ folgt

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

Zur Abkürzung sei jetzt

$$d_k \;:=\; |x_k - x^*|, \qquad\qquad k = 0, 1, 2, \ldots$$

Dann ergibt sich aus $(*)$

$$d_{k+1} = M_k \, d_k \, d_{k-1}.$$

Es ist nun zu zeigen, dass für $k \to \infty$

$$d_{k+1} \leq c \cdot d_k^{(1+\sqrt{5})/2} \qquad (k \geq 0)$$

mit einer Konstante $0 < c < \infty$ gilt.
Mit dem Ansatz

$$d_k = \gamma_k d_{k-1}^s$$

erhalten wir aus $d_{k+1} = M_k \, d_k \, d_{k-1}$

$$\begin{aligned}
& \gamma_{k+1} \, d_k^s = M_k \, d_k \, d_{k-1} \\
\iff \quad & \gamma_{k+1} \, d_k^{s-1} = M_k \, d_{k-1} \\
\iff \quad & \gamma_{k+1} \, (\gamma_k \, d_{k-1}^s)^{s-1} = M_k \, d_{k-1} \\
\iff \quad & (\gamma_{k+1} \, \gamma_k^{s-1}) \, d_{k-1}^{s^2-s-1} = M_k.
\end{aligned}$$

Für $s = \dfrac{1}{2}(\sqrt{5} + 1)$ folgt $s^2 - s - 1 = 0$ und daher

$$\gamma_{k+1} \, \gamma_k^{s-1} = M_k < M,$$

d.h., die Folge γ_k ist beschränkt. (Da $(d_k)_{k=0}^{\infty}$ eine Nullfolge ist, ist für größere s die Beschränktheit der Folge $(\gamma_k)_{k=0}^{\infty}$ nicht mehr gesichert.) Also ist

$$\lim_{k\to\infty} \frac{d_{k+1}}{d_k^s} = \lim_{k\to\infty} \gamma_{k+1} \le c < \infty.$$

□

Bemerkung:

1) Die Bedingung $x_k \in [a,b]$ für $k = 2, 3, ...$ in Satz 6.9 ist erfüllt, falls ein $d < 1$ so existiert, dass

$$|x_0 - x^*| \cdot M \le d < 1$$

und

$$|x_1 - x^*| \cdot M \le d < 1$$

mit

$$M = \frac{1}{2} \max_{\eta,\xi\in[a,b]} \frac{|f''(\xi)|}{|f'(\eta)|}$$

gilt, denn aus

$$\begin{aligned} & |x_{k+1} - x^*| \le M \cdot |x_k - x^*| \cdot |x_{k-1} - x^*| \\ \Longleftrightarrow \quad & M \cdot |x_{k+1} - x^*| \le M^2 \cdot |x_k - x^*| \cdot |x_{k-1} - x^*| \end{aligned}$$

(vgl. Beweis zu Satz 6.9) folgt

$$M \cdot |x_2 - x^*| \le d^2, \qquad M \cdot |x_3 - x^*| \le d^3, \qquad \text{u.s.w.}$$