Numerisches Programmieren (IN0019)

Frank R. Schmidt

Winter Semester 2016/2017

Iterationsverfahren

Oft können wir keine direkte Lösung einer Gleichung angeben.

Wenn wir eine Lösung (mathematisch) angeben können, lässt sich diese nicht immer mit Hilfe von endlich vielen Grundoperationen angeben, z.B. $\sqrt{2} \notin \mathbb{Q}$.

Anstelle einer direkten Lösung kann man aber oft ein **Iterationsverfahren** angeben, d.h. wir starten mit einer (schlechten) Approximation und verbessern sie in jedem Schritt.

Wir beginnen also mit einem **Startwert** x_0 und definieren die Folge

$$(x_n)_{n\in\mathbb{N}} \qquad x_{n+1} := \Phi(x_n),$$

wobei Φ eine wohldefinierte **stetige** Rechenvorschrift ist.
Die Generierung dieser Folge heißt **Iterationsverfahren**.

Heron-Verfahren

Eins der ersten Iterationsverfahren ist das **Heron-Verfahren** zur Berechnung der Quadratwurzel $\sqrt{a}$. Das Verfahren wurde von Heron etwa um 100 n. Chr. beschrieben, war aber schon in Babylon (ca. 1750 v. Chr.) bekannt.

Nach der Wahl eines Startwerts $x_0 > 0$ berechnet man

$$x_{n+1} = \Phi(x_n) \qquad \Phi(x) = \frac{x + \frac{a}{x}}{2}$$

Wählt man $x_0 = \frac{a+1}{2}$, so ergeben sich bei der Berechnung von $\sqrt{2}$, $\sqrt{3}$ bzw. $\sqrt{4}$ folgende Reihen

a	x_0	x_1	x_2	x_3
2	1.50000 ...	1.41667 ...	1.41422 ...	1.41421 ...
3	2.00000 ...	1.75000 ...	1.73214 ...	1.73205 ...
4	2.50000 ...	2.05000 ...	2.00061 ...	2.00000 ...

Fixpunktgleichung

Wir sind nur an solchen **Iterationsfolgen** $(x_n)_{n\in\mathbb{N}}$ interessiert, die konvergieren. Dann gilt aber für den Grenzwert $x^* := \lim_{n\to\infty} x_n$ gerade

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \Phi(x_n) = \Phi\left(\lim_{n\to\infty} x_n\right) = \Phi(x^*)$$

Für eine Funktion $\Phi\colon X \to X$, nennen wir $x^* \in X$ mit $\Phi(x^*) = x^*$ einen **Fixpunkt** von Φ. Wir haben gesehen, dass jede konvergente Iterationsfolge gegen einen Fixpunkt konvergiert.

Beim Heron-Verfahren gilt für die Fixpunkte x^* gerade

$$0 = x^* - \Phi(x^*) = x^* - \frac{x^* + \frac{a}{x^*}}{2} = \frac{{x^*}^2 - a}{2x^*},$$

d.h. $\sqrt{a}$ und $-\sqrt{a}$ sind die einzigen Fixpunkte.

Konvergenzverhalten

Das Konvergenzverhalten hängt im Allgemeinen vom Startwert ab.
Betrachten wir hierzu folgende Iterationsvorschrift

$$\begin{aligned} \Phi : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$

mit den Fixpunkten 0 und 1.

x_0	x_n	$\lim_{n\to\infty} x_n$
0	0	0
± 1	1	1
$\lvert x_0\rvert < 1$	x_0^{2n}	0
$\lvert x_0\rvert > 1$	x_0^{2n}	∞

Da der Fixpunkt 1 nur vorkommt wenn man mit ± 1 initialisiert, nennt man 1 einen **abstossenden Fixpunkt**.

Brouwerscher Fixpunktsatz

Der **Brouwersche Fixpunktsatz** besagt, dass eine stetige Funktion $\Phi\colon B_r(x) \to B_r(x)$ immer (mindestens) einen Fixpunkt besitzt, wobei $B_r(x)$ den folgenden Ball beschreibt

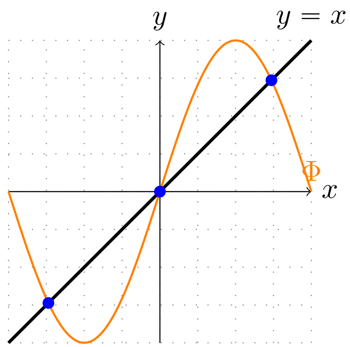
$$B_r(x) = \{y \in \mathbb{R}^n \mid \|y - x\| \leqslant r\}$$

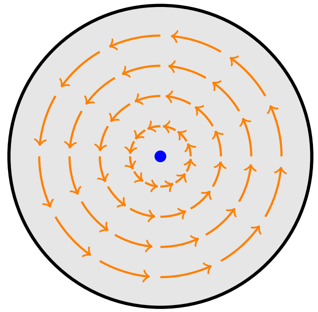
Da der allgemeine Beweis recht kompliziert ist, beschäftigen wir uns nur mit dem Fall $n = 1$. Die Aussage des Satzes ist dann, dass jede stetige Funktion $\Phi\colon [a, b] \to [a, b]$ einen Fixpunkt besitzt.

1. Fall: $\Phi(a) = a$ oder $\Phi(b) = b$.
Dann ist nichts zu zeigen und der Satz ist bewiesen.

2. Fall: $\Phi(a) > a$ und $\Phi(b) < b$.
Für die Funktion $f(x) = \Phi(x) - x$ gilt $f(a) > 0$ und $f(b) < 0$.
Aus Stetigkeitsgründen (Zwischenwertsatz) muss ein $x \in]a; b[$ existieren, so dass $f(x) = 0$ gilt und x ist ein Fixpunkt ($\Phi(x) - x = 0$).

Brouwersche Fixpunkte (Bsp)





Der Brouwersche Fixpunktsatz sagt etwas darüber aus, **ob ein Fixpunkt existiert**, aber nicht wie er zu finden ist.

Im Folgenden wollen wir solche Fixpunkte betrachten, die durch **iteriertes Anwenden** der Funktion Φ gefunden werden können.

Anziehender Fixpunkt

Wir wissen nun, dass jede stetige Funktion $\Phi\colon [a,b] \to [a,b]$ einen Fixpunkt x^* besitzt. Damit wir Φ für ein Iterationsverfahren nutzen können, muss x^* ein **anziehender Fixpunkt** sein, d.h. es sollte mindestens gelten

$$|\Phi(x) - x^*| < |x - x^*|$$

Wäre die Funktion differenzierbar, gäbe es nach dem Mittelwertsatz ein $\xi \in [x, x^*]$ mit

$$\left|\Phi'(\xi)\right| = \left|\frac{\Phi(x) - \Phi(x^*)}{x - x^*}\right| < 1$$

Man könnte sich also alleine auf solche Funktionen Φ beschränken, die die Ungleichung $-1 < \Phi'(x) < 1$ erfüllen. Um weiter mit stetigen Abbildungen arbeiten zu können, betrachten wir stattdessen **kontrahierende Abbildungen**.

Kontrahierende Abbildungen

Sei eine Abbildung $\Phi\colon B \to B$ gegeben, wobei $B \subset \mathbb{R}^n$ ein n-dimensionaler Ball ist. Wir nennen Φ eine **kontrahierende Abbildung**, wenn es ein Konstante $L < 1$ gibt, so dass Folgendes gilt:

$$\|\Phi(x) - \Phi(y)\| \leqslant L\,\|x - y\| \qquad \forall x, y \in B$$

Kontrahierende Abbildungen besitzen **maximal einen Fixpunkt**, denn für zwei unterschiedliche Fixpunkte x^* und $\hat{x}$ gilt

$$\|x^* - \hat{x}\| = \|\Phi(x^*) - \Phi(\hat{x})\| < \|x^* - \hat{x}\|\,,$$

was zu einem Widerspruch führt.

Zusammen mit dem Brouwerschen Fixpunktsatz wissen wir also, dass eine kontrahierende Abbildung einen **eindeutigen Fixpunkt** besitzt.

Konvergente Iterationsfolge

Sei $\Phi\colon B \to B$ eine kontrahierende Abbildung, die einen Fixpunkt $x^* \in B$ besitzt, dann konvergiert die Iterationsfolge $(x_n)_{n\in\mathbb{N}}$ mit $x_{n+1} = \Phi(x_n)$ gegen diesen Fixpunkt, denn es gilt

$$\|x_n - x^*\| = \|\Phi(x_{n-1}) - \Phi(x^*)\| \leqslant L\,\|x_{n-1} - x^*\| \leqslant \ldots \leqslant L^n\,\|x_0 - x^*\|\,,$$

d.h. $\|x_n - x^*\|$ ist eine Nullfolge und somit ist $\lim_{n\to\infty} x_n = x^*$.

Bis jetzt haben wir nur solche Abbildungen betrachtet, die bereits einen Fixpunkt besitzen. Betrachten wir nun eine allgemeine Abbildung $\Phi\colon X \to X$, so wollen wir zeigen, dass die Iterationsfolge eine **Cauchy-Folge** ist.

Analog zum oben Bewiesenen haben wir

$$\|x_{n+1} - x_n\| = \|\Phi(x_n) - \Phi(x_{n-1})\| \leqslant L\,\|x_n - x_{n-1}\| \leqslant \ldots \leqslant L^n\,\|x_1 - x_0\|\,.$$

Banachscher Fixpunktsatz

Sei $n \in \mathbb{N}$ beliebig, dann gilt für alle $m > n$ gerade

$$\|x_m - x_n\| \leqslant \sum_{k=n}^{m-1} \|x_{k+1} - x_k\| \leqslant \sum_{k=n}^{m-1} L^k \|x_1 - x_0\| < L^n \cdot \frac{\|x_1 - x_0\|}{1 - L},$$

d.h. $(x_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge.

Damit erhalten wir den **Banachschen Fixpunktsatz**

Theorem 1. *Sei $X \subset \mathbb{R}^n$ eine abgeschlossene, nicht-leere Menge und $\Phi\colon X \to X$ eine kontrahierende Abbildung. Dann besitzt Φ einen eindeutigen Fixpunkt x^* und die Folge*

$$(x_n)_{n\in\mathbb{N}} \qquad\qquad x_{n+1} = \Phi(x_n)$$

konvergiert für jeden Startwert $x_0 \in X$ gegen diesen Fixpunkt.

Klassifikation von Fixpunkten

Sei $\Phi\colon X \to X$ eine C^1-Abbildung und x^* ein Fixpunkt von Φ.

x^* ist ein **anziehender Fixpunkt**, wenn $|\Phi'(x^*)| < 1$ ist.
Dann ist $|\Phi'|$ auch in einer Umgebung von x^* durch ein Zahl $L < 1$ nach oben beschränkt und wir können in dieser Umgebung den Banachschen Fixpunkt anwenden.

x^* ist ein **abstossender Fixpunkt**, wenn $|\Phi'(x^*)| > 1$ ist.
Dann ist $|\Phi'|$ auch in einer Umgebung von x^* durch ein Zahl $L > 1$ nach unten beschränkt und jeder Startwert entfernt sich durch die Anwendung von Φ weiter von x^*.

Der Fall $|\Phi'(x^*)| = 1$ ist oft nicht eindeutig.
Betrachten wir z.B. $\Phi(x) = e^x - 1$ mit $x^* = 0$ und $\Phi'(x^*) = 1$.
Iterationsfolgen mit Startwerten $x_0 < 0$ konvergieren zu x^*.
Iterationsfolgen mit Startwerten $x_0 > 0$ divergieren.

Konvergenzgeschwindigkeit

Aus praktischer Sicht ist nicht nur die Konvergenz einer Folge sondern auch die **Geschwindigkeit** wichtig mit der die Konvergenz stattfindet.

Beide Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ mit

$$a_n = \frac{1}{n} \qquad\qquad b_n = \frac{1}{2^n}$$

sind Nullfolgen, allerdings unterschreitet man eine Fehlerschranke ε bei der Folge (b_n) wesentlich schneller als bei (a_n).

Wir sagen, dass eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \to x^*$ mit der **Ordnung** $p \in \mathbb{N}$ konvergiert falls $c > 0$ existiert, so dass

$$\|x_{n+1} - x^*\| \leqslant c\,\|x_n - x^*\|^p$$

Für $p = 1$ spricht man von **linearer Konvergenz**.
Für $p = 2$ spricht man von **quadratischer Konvergenz**.

Konvergenzgeschwindigkeit (Heron-Verfahren)

Eine Iterationsfolge konvergiert also linear, wenn die Iterationsvorschrift Φ die Voraussetzungen des Banachschen Fixpunktsatzes erfüllt.

Nehmen wir an, dass wir beim Heron-Verfahren ein Folgenglied $x_n = \sqrt{a}(1+\varepsilon)$ haben. Dann gilt $|x_n - x^*| = \sqrt{a}\,|\varepsilon|$.

Weiter haben wir

$$x_{n+1} = \frac{x_n + \frac{a}{x_n}}{2} = \frac{x_n^2 + a}{2x_n} = \frac{a(2+2\varepsilon+\varepsilon^2)}{2\sqrt{a}(1+\varepsilon)} \leqslant \sqrt{a}\left(1 + \frac{\varepsilon^2}{2}\right).$$

Somit gilt für den Fehler

$$|x_{n+1} - x^*| \leqslant \frac{\sqrt{a}}{2}\varepsilon^2 = \frac{1}{2\sqrt{a}}\,|x_n - x^*|^2,$$

was die **quadratische Konvergenz des Heron-Verfahrens** für $a \neq 0$ beweist.

Newton-Verfahren

Das Ziel des Newton-Verfahrens ist es, eine Nullstelle einer Funktion $f \colon \mathbb{R} \to \mathbb{R}$ zu finden. Hierzu ersetzen wir f durch seine lineare Taylorapproximation, dessen Nullstelle einfach zu berechnen ist.

Sei also ein Startwert $x_0 \in \mathbb{R}$ gegeben. Dann gilt für die Nullstelle x gerade

$$0 = f(x) = f(x_0 + (x - x_0)) \approx f(x_0) + f'(x_0)(x - x_0)$$

Die Nullstelle der rechten Seite ist

$$x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Somit erhalten wir die **Newton-Iteration**

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

Newton-Verfahren zur Wurzelberechnung

Wollen wir $\sqrt{a}$ mit Hilfe des **Newton-Verfahrens** lösen, müssen wir eine Nullstelle von

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

berechnen.

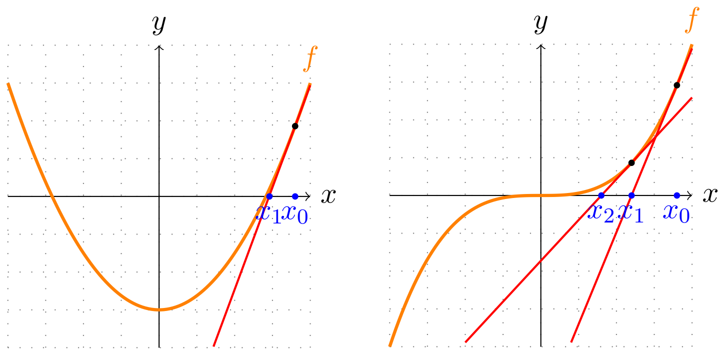
Die Newton-Iteration ist dann

$$\Phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x^2 - a}{2x} = \frac{x^2 + a}{2x} = \frac{x + \frac{a}{x}}{2}.$$

Dies entspricht exakt dem **Heron-Verfahren**.

Wir können also das Heron-Verfahren als einen Spezialfall des Newton-Verfahrens verstehen.

Newton-Verfahren



Geometrisch wird eine **Tangente** an den Graphen von f bei $(x_n, f(x_n))$ gelegt und der eindeutige **Schnittpunkt** mit der x-Achse berechnet. Dieser Schnittpunkt x_{n+1} wird bei der nächsten Iteration benutzt.

Fixpunkte der Newton-Iteration

Wenn die Newton-Iteration konvergiert, konvergiert sie zu einem Fixpunkt von Φ, d.h. es muss gelten

$$0 = x^* - \Phi(x^*) = \frac{f(x^*)}{f'(x^*)}$$

Somit kann eine Reihe die auf der Newton-Iteration Φ beruht nur gegen eine Nullstelle von f konvergieren.

Gilt für die Nullstelle x^* außerdem $f'(x^*) \neq 0$, so ist

$$\Phi'(x^*) = \frac{\mathrm{d}}{\mathrm{dx}}\left[x - \frac{f(x^*)}{f'(x^*)}\right] = \frac{f(x^*)f''(x^*)}{f'(x^*)^2} = 0,$$

d.h. die Nullstelle x^* ist ein **anziehender Fixpunkt** der Newton-Iteration.

Quadratische Konvergenz

Wenn die Newton-Iteration gegen die Nullstelle x^* mit $f'(x^*) \neq 0$ konvergiert, kann man sogar quadratische Konvergenz zeigen. Hierzu betrachten wir die quadratische Taylor-Reihe, die bei x_n entwickelt wird ($\xi \in [x_n; x^*]$):

$$0 = f(x^*) = f(x_n) + f'(x_n)(x^* - x_n) + \frac{1}{2} f''(\xi)(x^* - x_n)^2$$

Daraus ergibt sich

$$\frac{f(x_n)}{f'(x_n)} + (x^* - x_n) = -\frac{f''(\xi)}{2f'(x_n)}(x^* - x_n)^2$$

In einer Umgebung von x^* lässt sich $\frac{f''(\xi)}{2f'(x_n)}$ durch $M > 0$ beschränken und

$$|x^* - x_{n+1}| \leqslant M \, |x^* - x_n|^2$$

beweist die **quadratische Konvergenz**.

Mehrfache Nullstelle

Wenn x^* eine Nullstelle von $f\colon \mathbb{R} \to \mathbb{R}$ ist, so kann man eine Funktion $g\colon \mathbb{R} \to \mathbb{R}$ finden, so dass $f(x) = g(x) \cdot (x - x^*)$ gilt:

$$g_1(x) = \begin{cases} \frac{f(x)}{x-x^*} & \text{falls } x \neq x^* \\ f'(x^*) & \text{falls } x = x^* \end{cases}$$

Ist $f'(x^*) = 0$ kann man diese Operation noch einmal durchführen und erhält eine Darstellung $f(x) = g_2(x) \cdot (x - x^*)^2$ mit $g_2(x^*) = \frac{f''(x^*)}{2}$.

Gilt nun $0 = f(x^*) = f'(x^*) = \ldots = f^{(m-1)}(x^*)$ und $f^{(m)}(x^*) \neq 0$, so heißt m die **Vielfachheit der Nullstelle** x^*. Außerdem hat man die Darstellung

$$f(x) = g(x) \cdot (x - x^*)^m \qquad g(x^*) = \frac{f^{(m)}(x^*)}{m!} \neq 0$$

Konvergenz bei mehrfacher Nullstelle

Sei x^* eine Nullstelle von der Vielfachheit $m > 1$, dann gilt

$$\begin{aligned} f(x) =& g(x) \cdot (x - x^*)^m \\ f'(x) =& g'(x) \cdot (x - x^*)^m + m \cdot g(x) \cdot (x - x^*)^{m-1} \end{aligned}$$

Für die Newton-Iteration gilt also

$$\Phi(x) = x - \frac{g(x) \cdot (x - x^*)}{g'(x) \cdot (x - x^*) + m \cdot g(x)}$$

und wir haben wegen

$$\Phi'(x^*) = 1 - \frac{g(x^*)}{m \cdot g(x^*)} = 1 - \frac{1}{m}$$

nur **lineare Konvergenz**.

Sekantenverfahren

Das **Sekantenverfahren** ist vom Newton-Verfahren inspiriert.
Anstatt die Funktion durch die Tangente an einem Punkt zu approximieren, approximiert man sie durch eine Sekante, die durch zwei Punkte verläuft.

Das heisst man ersetzt $f(x)$ durch

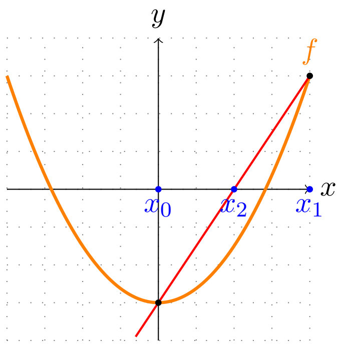
$$s(x) = f(x_{k-1}) + \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x - x_{k-1})$$

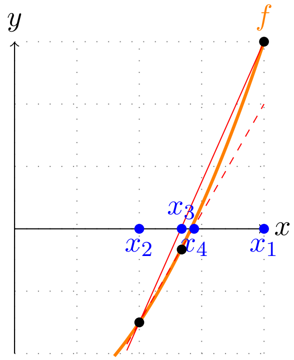
Die Sekante hat ihre Nullstelle bei

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

Man kann zeigen, dass das Sekantenverfahren eine Konvergenzordnung von $1 < p < 2$ hat und in der Nähe der Nullstelle numerisch instabil wird.
Allerdings muss keine Ableitung berechnet werden!

Sekanten-Verfahren





Beim Sekantenverfahren kann es vorkommen,
dass eine x_{n+1} nicht im Intervall $[x_{n-1}; x_n]$ liegt.

Das heisst, dass das Sekantenverfahren **keine Intervallschachtelung** von x^* berechnet.

Bisektionsverfahren

Beim **Bisektionsverfahren** geht man davon aus, dass zwei Stellen $a_0, b_0 \in \mathbb{R}$, so dass $f(a_0) \cdot f(b_0) < 0$ gilt.

Ist $c_0 = \frac{a_0+b_0}{2}$ eine Nullstelle, terminiert das Verfahren.

Setze $a_1 = a_0$ und $b_1 = c_0$, wenn $f(a_0)f(c_0) < 0$.
Setze $a_1 = c_0$ und $b_1 = b_0$, wenn $f(b_0)f(c_0) < 0$.

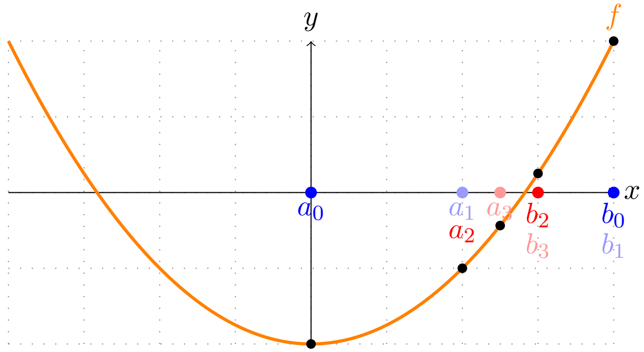
Damit gilt nun ebenfalls $f(a_1)f(b_1) < 0$, aber die Länge des Intervalls wurde halbiert. Nach n Schritten gilt also

$$|a_n - b_n| = \frac{1}{2^n} |a_0 - b_0|$$

Für die Nullstelle x^* gilt insbesondere

$$|a_n - x^*| + |b_n - x^*| = |a_n - b_n| = \frac{1}{2^n} (|a_0 - x^*| + |b_0 - x^*|)$$

Bisektionsverfahren



Beim Bisektionsverfahren wird eine Intervallschachtelung von x^* berechnet.

Das Verfahren benutzt nur die Stetigkeit von f.

Regula Falsi

Beim Bisektionsverfahren konvergiert also weder a_n noch b_n linear, aber die Summe der Fehler sinkt linear.

Das **Regula Falsi-Verfahren** verbindet die Ideen des Sekantenverfahrens mit dem Bisektionsverfahrens.

Es werden wieder die Intervallgrenzen in jedem Schritt verbessert. Allerdings wird bei der Berechnung von c_n die Sekanteniteration benutzt.

Da in jedem Schritt $f(a_n)f(b_n) < 0$ gilt, liegt c_n im Intervall $[a_n, b_n]$.

Damit kann es vorkommen, dass die Länge des Intervalls schneller sinkt.

Allerdings kann es auch vorkommen, dass $|b_n - a_n|$ keine Nullfolge mehr ist.

Praktische Nullstellen-Berechnung

In der Praxis wird oft eine Kombination von **Bisektion** und **Newton** benutzt.

Starte mit sicherer Bisektion, bis der Einzugsbereich der quadratischen Konvergenz des Newtonverfahrens erreicht ist.

Dazu benötigt man Stellen a und b mit $f(a)f(b) < 0$.
Dies kann man durch äquidistante Auswertungen oder mit Hilfe von Zufallszahlen erreichen.

Falls man keine solche Stellen a und b findet, liegt eventuell keine Nullstelle vor oder eine Nullstelle gerader Ordnung.

Nullstellen gerader Ordnung kann man finden, indem man das Newton-Verfahren bzgl. der Ableitung f' anwendet.