

Nichtlineare Gleichungen

Ein wichtiges Problem in der Praxis ist die Bestimmung einer *Lösung* ξ der Gleichung

$$f(x) = 0, \quad (1)$$

d.h. das Aufsuchen einer *Nullstelle* ξ einer (nicht notwendig linearen) Funktion f . Ist $f(x)$ ein Polynom, so heißt die Gleichung (1) *algebraisch*, ansonsten *transzendent*. Da man nur selten eine Lösung von (1) in endlich vielen Schritten explizit berechnen kann, ist man im allg. auf Näherungsmethoden (*Einschlußverfahren*, *Iterationsverfahren*) angewiesen.

Das Bisektionsverfahren (Intervallhalbierung)

Ist f stetig auf $[a, b]$ und haben $f(a)$ und $f(b)$ entgegengesetzte Vorzeichen ($f(a)f(b) < 0$), so existiert nach dem Zwischenwertsatz mindestens ein $\xi \in (a, b)$ mit $f(\xi) = 0$. Man halbiert nun das Intervall $[a, b]$ und stellt über das Vorzeichen von f am Mittelpunkt fest, ob die gesuchte Nullstelle in der linken oder in der rechten Hälfte des Ausgangsintervalls zu finden ist. Man behält jene Hälfte, auf der das Vorzeichen von f wechselt. Dieses Verfahren wird fortgesetzt, bis eine Nullstelle im Inneren eines hinreichend kleinen Intervalls liegt:

1. $x_{i+1} := \frac{a_i + b_i}{2} \quad : \quad i = 0, 1, 2, \dots \quad \text{und} \quad a_0 = a, b_0 = b$
2. falls $f(a_i)f(x_{i+1}) < 0 \quad : \quad \text{setze } a_{i+1} = a_i \quad \text{und} \quad b_{i+1} = x_{i+1}$
 andernfalls $\quad : \quad \text{setze } a_{i+1} = x_{i+1} \quad \text{und} \quad b_{i+1} = b_i$
3. falls $|b_{i+1} - a_{i+1}| \leq \text{tol} \quad : \quad \text{Abbruch}$
 andernfalls $\quad : \quad \text{gehe zu Schritt 1}$

Nach n Halbierungen liegt eine Nullstelle ξ sicher in einem Intervall der Länge $(b - a)/2^n$, und man erhält die Fehlerabschätzung

$$\epsilon_n := |x_n - \xi| \leq (b - a)/2^n .$$

Damit ist $\epsilon_{n+1} \approx \frac{1}{2}\epsilon_n$. Pro Iterationsschritt verkleinert sich der Fehler um einen konstanten Faktor. Es werden also stets etwa gleich viele Iterationsschritte benötigt, um den Fehler auf einen bestimmten Bruchteil zu reduzieren: Man sagt, das Verfahren sei *linear konvergent*. Wegen $1/2^{10} = 1/1024 \approx 10^{-3}$ gilt die Faustregel, daß man mit je 10 Halbierungsschritten etwa 3 Dezimalstellen an Genauigkeit für die Lösung gewinnt.

- Vorteile $\quad : \quad$ Konvergenz ist garantiert
 nur 1 Funktionsauswertung pro Iteration
- Nachteile $\quad : \quad$ relativ langsame Konvergenz
 funktioniert nur für 1D-Probleme

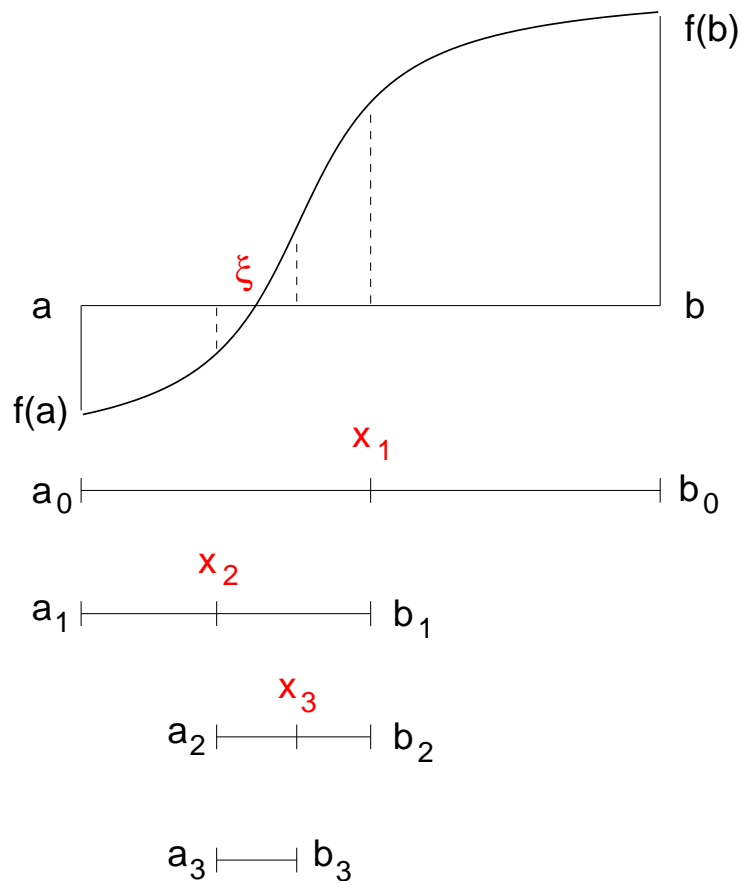


Abbildung 1: Bisektionsverfahren.

Fixpunktiteration (Methode der sukzessiven Approximation)

Hier wird das Nullstellenproblem $f(x) = 0$ in ein äquivalentes *Fixpunktproblem*

$$x = \varphi(x) \quad (2)$$

mit einer geeigneten *Iterationsfunktion* φ umformuliert. Jede Nullstelle ξ von $f(x)$ ist dann ein *Fixpunkt* von $\varphi(x)$, also eine Stelle ξ , welche durch φ auf sich selbst abgebildet wird (*fix* bleibt). Anschaulich beschreibt (2) die Schnittpunkte von $y(x) = x$ mit $\varphi(x)$. Bei der Wahl von φ hat man gewisse Freiheiten, eine kann zum Erfolg, eine andere zum Scheitern führen. Ausgehend von einem geeigneten *Startwert* x_0 wird mit der Iterationsvorschrift

$$\boxed{x_{i+1} := \varphi(x_i)} \quad (3)$$

für $i = 0, 1, 2, \dots$ eine Folge $\{x_i\}$ von Näherungswerten berechnet, von der man hofft, daß sie gegen die gesuchte Lösung ξ konvergiert.

Beispiel: Ist etwa die Gleichung $f(x) := x - \cos x = 0$ zu lösen, so liegt es nahe, die Iterationsvorschrift $x_{i+1} = \cos x_i$, also $\varphi(x) := \cos x$ zu verwenden. Man erhält eine alternierend konvergente Folge $\{x_i\}$ von Näherungswerten an ξ (die Iteration verläuft in einem "Käfig"). Diese Iteration läßt sich leicht auf einem Taschenrechner durch wiederholtes Drücken der **cos**-Taste erzeugen: Egal, mit welchem Wert man beginnt, nach einigen Schritten erscheint immer die gleiche Zahl.

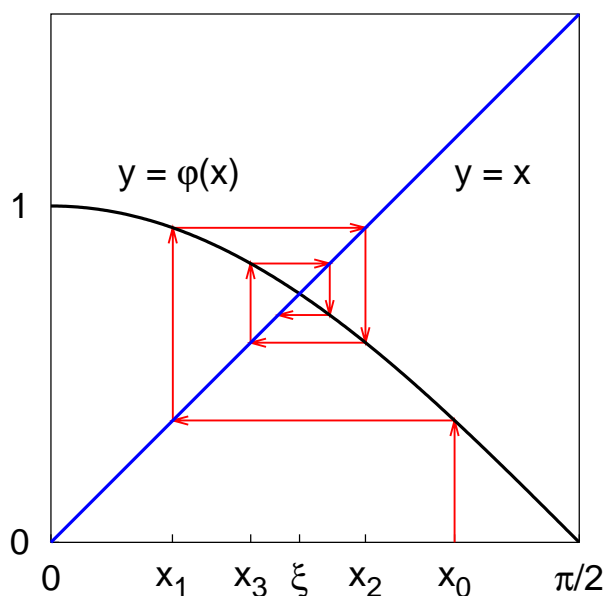


Abbildung 2: Iterationskäfig für $\varphi(x) = \cos x$.

Bezüglich der Konvergenz des Iterationsverfahrens liefert der folgende Kontraktionssatz eine *hinreichende* Konvergenzbedingung:

Kontraktionssatz (Fixpunktsatz von Banach)

Ist $\varphi : I \rightarrow I$ eine kontrahierende Selbstabbildung eines abgeschlossenen Intervalls I in sich mit

$$|\varphi(x) - \varphi(y)| \leq L|x - y| \quad \forall x, y \in I$$

und einer Kontraktionskonstante (Lipschitz-Konstante) $0 \leq L < 1$, so gilt:

- (a) φ besitzt in I genau einen Fixpunkt $\xi = \varphi(\xi)$.
- (b) Die Iterationsfolge $x_{i+1} = \varphi(x_i)$ konvergiert für jeden Startwert $x_0 \in I$ gegen ξ .
- (c) Es gilt die (a priori) Fehlerabschätzung

$$|x_n - \xi| \leq \frac{L^n}{1 - L} |x_1 - x_0| .$$

- (d) Das Verfahren konvergiert zumindest linear, d.h.

$$\epsilon_{n+1} \leq L \epsilon_n .$$

Anmerkungen:

- Da I sehr klein sein kann, hat der Kontraktionssatz oft nur *lokalen* Charakter, d.h. für eine Konvergenz des Verfahrens muß der Startwert x_0 hinreichend nahe bei der Lösung ξ liegen.
- φ ist auf I genau dann eine Kontraktion, wenn $|\varphi'(x)| < 1$ für alle $x \in I$ gilt (Mittelwertsatz der Differentialrechnung):
 - $0 \leq \varphi' < 1$: $\{x_i\}$ konvergiert monoton (Abbildung 3).
 - $-1 < \varphi' < 0$: $\{x_i\}$ konvergiert alternierend (Abbildung 2).
- Das Iterationsverfahren kann leicht auf n -dimensionale Probleme (d.h. auf *Systeme* von nichtlinearen Gleichungen) verallgemeinert werden.
- Rundungsfehler werden beim Iterationsverfahren nicht akkumuliert. Das Verfahren "vergißt" die früheren Rundungsfehler, da jeder Schritt als ein erster Schritt angesehen werden kann.

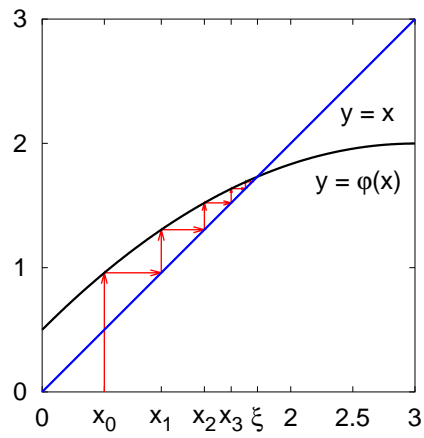


Abbildung 3: Monotone Konvergenz.

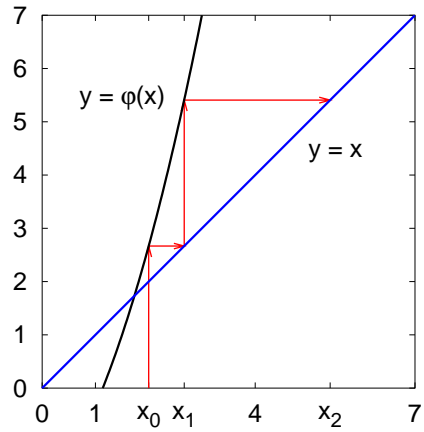


Abbildung 4: Monotone Divergenz.

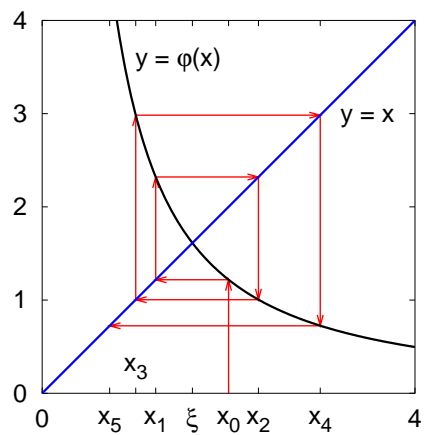


Abbildung 5: Alternierende Divergenz.

Das Newton-Verfahren

Die Funktion f sei genügend oft stetig differenzierbar und besitze im Intervall (a, b) eine *einfache* Nullstelle ξ , d.h. $f(\xi) = 0$ und $f'(\xi) \neq 0$. Sei x_i schon eine gute Näherung für eine Nullstelle ξ . Um zu einer besseren Näherung zu gelangen, ersetzt man die (nichtlineare) Funktion f durch ihre Tangente im Punkt $(x_i, f(x_i))$,

$$t(x) := f(x_i) + f'(x_i)(x - x_i) , \quad (4)$$

und nimmt die Nullstelle von $t(x)$ als neue Näherung x_{i+1} für ξ (Abbildung 6):

$$0 = f(x_i) + f'(x_i)(x_{i+1} - x_i) .$$

Dies führt für $i = 0, 1, 2, \dots$ auf die Iterationsvorschrift (*Newton-Raphson-Verfahren*)

$$\boxed{x_{i+1} := x_i - \frac{f(x_i)}{f'(x_i)}} . \quad (5)$$

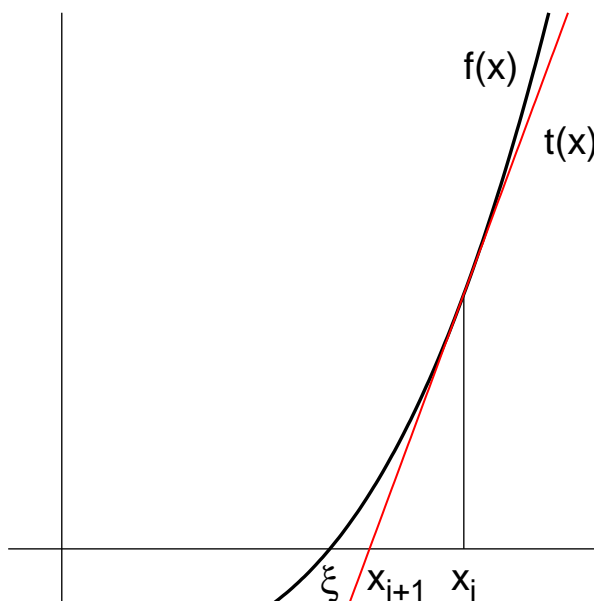


Abbildung 6: Newton-Verfahren.

$t(x)$ in Gleichung (4) ist nichts anderes als der lineare Anteil der Taylor-Entwicklung von f um x_i . Diese *Linearisierung* ist der entscheidende Schritt des Newton-Verfahrens: Die Lösung einer nichtlinearen Gleichung $f(x) = 0$ wird durch eine Folge von Lösungen linearer Gleichungen ersetzt bzw. approximiert.

Formal gehört das Newton-Verfahren zur Klasse der Fixpunktiterationen mit

$$\varphi(x) := x - \frac{f(x)}{f'(x)} . \quad (6)$$

Falls der Startwert x_0 in (5) eine hinreichend gute Anfangsnäherung für ξ ist, so ist das Verfahren (lokal) mindestens *quadratisch konvergent*, d.h.

$$\epsilon_{n+1} \leq C \epsilon_n^2 .$$

Ist der Fehler $\epsilon_n = |x_n - \xi|$ von der Größenordnung 10^{-k} , so ist der nächste Fehler $\epsilon_{n+1} = |x_{n+1} - \xi|$ von der Größenordnung 10^{-2k} (falls C von der Größenordnung 1 ist). Bei jedem Iterationsschritt wird also die Anzahl der mit ξ übereinstimmenden Dezimalstellen ungefähr verdoppelt. Diese rasche Konvergenz ist der Grund für die große Beliebtheit des Newton-Verfahrens. Die Konvergenz des Verfahrens hängt jedoch empfindlich von der Wahl eines *guten Startwertes* x_0 ab, insbesondere reagiert das Verfahren empfindlich auf lokale Extrema von f , da die Iterationsfunktion $\varphi(x)$ (Gleichung (6)) dort singulär wird. Bei *mehrfachen* Nullstellen konvergiert das Newton-Verfahren nur noch *linear*.

- Vorteile : rasche (mind. quadratische) Konvergenz bei einfachen Nullstellen
läßt sich auf *Systeme* nichtlinearer Gleichungen übertragen
- Nachteile : in jedem Schritt muß f und f' berechnet werden
das Verfahren muß nicht konvergieren

Beispiel: Die m -te Wurzel $a^{1/m}$ aus einer positiven reellen Zahl $a > 0$ ist Lösung der Gleichung

$$x^m - a = 0 .$$

Das Newton-Verfahren ergibt mit $f(x) = x^m - a$, $f'(x) = mx^{m-1}$ die Vorschrift ($i = 0, 1, 2, \dots$)

$$x_{i+1} = x_i - \frac{x_i^m - a}{mx_i^{m-1}} = \frac{a + (m-1)x_i^m}{mx_i^{m-1}} ,$$

$$\boxed{x_{i+1} := \frac{1}{m} \left\{ \frac{a}{x_i^{m-1}} + (m-1)x_i \right\}} . \quad (7)$$

Daraus leiten sich folgende Spezialfälle ab:

- (a) $m = 2$: $x_{i+1} = \frac{1}{2} \left(\frac{a}{x_i} + x_i \right)$, Quadratwurzel von a .
- (b) $m = -1$: $x_{i+1} = (2 - ax_i) x_i$, Kehrwert von a .

Die Iteration (a) ist das (schon den Babyloniern bekannte) *Heron-Verfahren* zur Quadratwurzelbestimmung. Die Iteration (b) ermöglichte für die ersten Computer ohne eingebaute Division die Zurückführung dieser Operation auf Multiplikationen und Subtraktionen.