4. NICHTLINEARE GLEICHUNGSSYSTEME

In diesem Kapitel suchen wir zu einer gegebenen Funktion $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ ein $x \in D$, mit

$$f(x) = 0 \quad \Longleftrightarrow \quad \left\{ \begin{array}{rcl} f_1(x_1, \dots, x_n) & = & 0 \\ & \vdots & \\ f_n(x_1, \dots, x_n) & = & 0 \end{array} \right. \tag{4.1}$$

Ein nichtlineares Gleichungssystem hat im Allgemeinen keine eindeutige Lösung. So hat im Fall $n = 1$ zum Beispiel (4.1) für

$$f(x) = e^x$$

keine Lösung, für

$$f(x) = x^2 - 1$$

mehrere (aber endlich viele) Lösungen und für

$$f(x) = x \sin \frac{1}{x}$$

unendlich viele Lösungen.

4.1 Fixpunktiteration

Wir betrachten zunächst ein eng verwandtes Problem, nämlich das, den **Fixpunkt** x einer Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ zu bestimmen, d. h. wir suchen $x \in \mathbb{R}^n$ mit

$$x = \Phi(x). \tag{4.2}$$

Das obige Nullstellenproblem läßt sich leicht in ein Fixpunktproblem umschreiben, etwa durch die Definition $\Phi(x) = x - f(x)$ oder $\Phi(x) = x - Bf(x)$ mit einer geeigneten Matrix $B \in \mathbb{R}^{n,n}$.

Zu einer gegebenen Anfangsnäherung $x_0 \in \mathbb{R}^n$ betrachtet man die Fixpunktiteration

$$x_{k+1} = \Phi(x_k), \qquad k = 0, 1, \dots \tag{4.3}$$

Konvergiert die Folge x_k gegen ein $x \in \mathbb{R}^n$ und ist Φ stetig, dann ist x Fixpunkt von Φ. Was kann man über den Fehler sagen? Zunächst definieren wir allgemein die Konvergenzordnung eines Verfahrens bzw. einer Folge.

Definition 4.1. *Sei $(e_k)_{k \geq 0}$ eine Folge mit $e_k > 0$ und $e_k \to 0$ für $k \to \infty$. Die Folge* **konvergiert** *(mindestens)* **mit Ordnung** *$p \geq 1$, wenn ein $C > 0$ und ein $k_0 \in \mathbb{N}_0$ existiert, so dass*

$$e_{k+1} \leq C e_k^p, \qquad \text{für alle } k \geq k_0$$

gilt. Für $p = 1$ fordert man zusätzlich $C < 1$.

Bemerkung. Für gewisse Konvergenzordnungen findet man in der Literatur auch folgende Bezeichnungen:

Konvergenzordnung	Bezeichnung
$p = 1$	linear
$p > 1$	superlinear
$p = 2$	quadratisch
$p = 3$	kubisch

Bei nichtlinearen Problemen, bei denen in der Regel keine eindeutige Lösung existiert, wird man von einem Iterationsverfahren nicht erwarten können, dass es für beliebige Startwerte gegen eine bestimmte Lösung konvergiert. Hierzu unterscheidet man lokale und globale Konvergenz:

Definition 4.2. *Ein Iterationsverfahren, welches eine Folge von Iterierten $(x_k)_{k\geq 0}$ im $\mathbb{R}^n$ liefert, heißt* **lokal konvergent** *gegen $x \in \mathbb{R}^n$, falls eine Umgebung $U \subset \mathbb{R}^n$ um x existiert, so dass*

$$x_k \to x \qquad \text{für alle } x_0 \in U$$

gilt. Das Verfahren ist **global konvergent**, *falls $U = \mathbb{R}^n$.*

Wir zeigen jetzt, dass die Fixpunktiteration (4.3) unter geeigneten Voraussetzungen linear konvergiert.

Satz 4.3. *(Banach'scher Fixpunktsatz)*
Sei $\Phi : \Omega \to \Omega$, eine kontrahierende Selbstabbildung der abgeschlossenen Teilmenge $\Omega \subset \mathbb{R}^n$ bezüglich $\|\cdot\|$, d. h.

(i) $\Phi(\Omega) \subset \Omega$ *(Selbstabbildung)*

(ii) $\|\Phi(x) - \Phi(y)\| \leq L\|x - y\|$ *für alle $x, y \in \Omega$ (Lipschitz-stetig mit Lipschitz-Konstante $L < 1$).*

Dann hat die Fixpunktgleichung (4.2) genau eine Lösung $\widehat{x} \in \Omega$ und die Fixpunktiteration (4.3) mit beliebigem $x_0 \in \Omega$ konvergiert linear gegen $\widehat{x}$. Genauer gilt für den Fehler $e_k = x_k - \widehat{x}$

(a) $\|e_{k+1}\| \leq L\|e_k\|$ *(Monotonie)*

(b) $\|e_k\| \leq \dfrac{L^k}{1-L}\|x_1 - x_0\|$ *(a-priori-Schranke)*

(c) $\|e_k\| \leq \dfrac{L}{1-L}\|x_k - x_{k-1}\|$ *(a-posteriori-Schranke)*

Beweis. Aus (i) folgt sofort $x_k \in \Omega$ für alle $k \geq 0$. Wegen (ii) gilt

$$\begin{aligned}\|x_{k+1} - x_k\| = \|\Phi(x_k) - \Phi(x_{k-1})\| &\leq L\|x_k - x_{k-1}\| \\ &\leq L^2\|x_{k-1} - x_{k-2}\| \leq \ldots \leq L^k\|x_1 - x_0\|.\end{aligned}$$

Ferner ist (x_k) eine Cauchy-Folge, denn für $m > n$ ist mit der Dreiecksungleichung und dem eben Gezeigten

$$\begin{aligned}\|x_m - x_n\| &\le \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \ldots + \|x_{n+1} - x_n\| \\ &\le \big(L^{m-1} + L^{m-2} + \ldots + L^n\big)\|x_1 - x_0\| \\ &\le L^n\big(1 + L + L^2 + \ldots\big)\|x_1 - x_0\| \\ &= \frac{L^n}{1-L}\|x_1 - x_0\|. \end{aligned} \tag{4.4}$$

Für $n \to \infty$ geht $\|x_m - x_n\| \to 0$, da $L < 1$. Die Folge (x_k) konvergiert wegen der Abgeschlossenheit von Ω gegen ein $\widehat{x} \in \Omega$. $\widehat{x}$ erfüllt nach Iterationsvorschrift $\widehat{x} = \Phi(\widehat{x})$, ist also ein Fixpunkt. Ist x ein weiterer Fixpunkt, so folgt wiederum aus der Kontraktionseigenschaft (ii)

$$\|x - \widehat{x}\| = \|\Phi(x) - \Phi(\widehat{x})\| \le L\|x - \widehat{x}\|,$$

was für $L < 1$ nur möglich ist, wenn $x = \widehat{x}$ gilt. Es gibt also nur den einen Fixpunkt $\widehat{x} \in \Omega$.

Es bleibt noch der Nachweis der Ungleichungen. (a) folgt wie gehabt aus der Kontraktionseigenschaft. (b) folgt aus (4.4) durch Grenzübergang $m \to \infty$. Schließlich erhält man aus

$$\begin{aligned}L\|x_k - x_{k-1}\| \ge \|x_{k+1} - x_k\| &\ge \|x_k - \widehat{x}\| - \|x_{k+1} - \widehat{x}\| \\ &\ge \|x_k - \widehat{x}\| - L\|x_k - \widehat{x}\| \\ &= (1 - L)\|x_k - \widehat{x}\|\end{aligned}$$

die letzte Ungleichung. □

Der Banach'sche Fixpunktsatz gilt allgemein in vollständigen metrischen Räumen. Er liefert Existenz, Eindeutigkeit und Konvergenz, jedoch lassen sich in praktischen Anwendungen die Voraussetzungen wenn überhaupt nur mit großem Aufwand überprüfen. In vielen Fällen ist folgende Variante hilfreich:

Satz 4.4. *Es sei $\Phi : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ zweimal stetig differenzierbar und $\widehat{x}$ sei eine Lösung von (4.2). Ist $\|\Phi'(\widehat{x})\| < 1$ in einer beliebigen (von einer Vektornorm induzierten) Matrixnorm, dann konvergiert die Fixpunktiteration lokal linear und für den Fehler $e_k = x_k - \widehat{x}$ gilt*

$$e_{k+1} = \Phi'(\widehat{x})e_k + O(\|e_k\|^2).$$

Beweis. Wegen $\widehat{x} = \Phi(\widehat{x})$ ist

$$e_{k+1} = x_{k+1} - \widehat{x} = \Phi(x_k) - \Phi(\widehat{x}) = \Phi'(\widehat{x})e_k + O(\|e_k\|^2)$$

nach dem Satz von Taylor. □

Häufig verzichtet man jedoch auf eine vorherige Überprüfung der Voraussetzungen und versucht, wie in Algorithmus 4.1, während des Algorithmus geeignete Konvergenz- und Abbruchkriterien einzubauen.

Wir diskutieren zunächst das Abbruchkriterium. Nach dem Banach'schen Fixpunktsatz wissen wir, dass falls Konvergenz eintritt, diese (mindestens) linear ist, d. h.

$$L_k := \frac{\|x_{k+1} - x_k\|}{\|x_k - x_{k-1}\|} < 1$$

Algorithmus 4.1 Fixpunktiteration

while noch nicht konvergent {Konvergenzkriterium} **do**
  $x_{k+1} = \Phi(x_k)$
  if keine Konvergenz möglich **then**
    Stop {Abbruchkriterium}
  end if
end while

gilt. Ist also $L_k \geq 1$, brechen wir ab.

Das optimale Konvergenzkriterium wäre $\|e_{k+1}\| \leq \texttt{tol}$, d. h. in einer vorgegebenen Norm ist der Fehler kleiner als eine gewünschte Toleranz. Leider verwendet dieses Kriterium die unbekannte Lösung $\widehat{x}$, ist also in der Praxis unbrauchbar. Satz 4.3 liefert die Abschätzung

$$\|e_{k+1}\| \leq \frac{L}{1-L}\|x_{k+1} - x_k\|,$$

in der wir die unbekannte Lipschitz-Konstante L durch die berechenbare Kontraktionszahl L_k ersetzen. Das führt auf das Konvergenzkriterium

$$\frac{L_k}{1-L_k}\|x_{k+1} - x_k\| \leq \texttt{tol}.$$

4.2 Newton-Verfahren

Ist die Funktion f mindestens einmal stetig differenzierbar, so kann man mit einer Startnäherung $x_0 \in \mathbb{R}^n$ sukzessive die gesuchte Lösung von (4.1) approximieren, indem man linearisiert (im skalaren Fall ersetzt man f durch eine Gerade). Im k-ten Schritt f approximiert man

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k),$$

d. h. man vernachlässigt in der Taylor-Entwicklung die Terme der Ordnung mindestens zwei. Als nächste Näherung sucht man dann eine Lösung x_{k+1} der Linearisierung

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0.$$

Ist $f'(x_k)$ nicht singulär, so erhält man

$$\Delta x_k := x_{k+1} - x_k = -f'(x_k)^{-1} f(x_k). \tag{4.5}$$

Mit dieser Näherung fährt man dann genauso fort, vgl. Algorithmus 4.2. In jedem Schritt muss ein lineares Gleichungssystem mit der Jacobi-Matrix $f'(x_k)$ gelöst werden (zum Beispiel mit Gauß-Elimination oder mit QR-Zerlegung).

Es stellen sich folgende Fragen: Unter welchen Voraussetzungen an f und x_0 konvergiert das Newton-Verfahren? Wie schnell ist die Konvergenz?

Satz 4.5. *Es sei f dreimal stetig differenzierbar, $f(\widehat{x}) = 0$ und $f'(x)$ sei invertierbar in einer Umgebung von $\widehat{x}$. Ist die Fehlernorm $\|e_k\|$, $e_k = x_k - \widehat{x}$ hinreichend klein, dann gilt für die Folge $(x_k)_{k\geq 0}$ aus dem Newton-Verfahren*

$$e_{k+1} = \frac{1}{2} f'(x_k)^{-1} f''(x_k)[e_k, e_k] + O(\|e_k\|^3).$$

Algorithmus 4.2 Newton-Verfahren

```
while noch nicht konvergent {Konvergenzkriterium} do
  Berechne f(x_k) und f'(x_k)
  Löse das linearisierte Problem: f'(x_k)Δx_k = −f(x_k)
  Setze x_{k+1} = x_k + Δx_k
  if keine Konvergenz möglich then
    Stop {Abbruchkriterium}
  end if
end while
```

Bemerkung. Für $f : \mathbb{R}^n \to \mathbb{R}^n$ ist $f' \in \mathbb{R}^{n,n}$ eine quadratische Matrix und f'' ein Tensor zweiter Stufe. Die j-te Komponente von $f''(x)[y, y]$ ist durch

$$\big(f''(x)[y, y]\big)_j = \sum_{k,l} \frac{\partial^2 f_j}{\partial x_k \partial x_l} y_k y_l$$

gegeben. Das Newton-Verfahren konvergiert damit lokal quadratisch.

Beweis. Der Satz von Taylor liefert

$$0 = f(\widehat{x}) = f(x_k - e_k) = f(x_k) - f'(x_k)e_k + \frac{1}{2} f''(x_k)[e_k, e_k] + O(\|e_k\|^3).$$

Aus (4.5) folgt

$$f(x_k) = -f'(x_k)(x_{k+1} - x_k) = -f'(x_k)(e_{k+1} - e_k).$$

Zusammen mit der obigen Gleichung ergibt sich aus

$$0 = -f'(x_k)e_{k+1} + \frac{1}{2} f''(x_k)[e_k, e_k] + O(\|e_k\|^3)$$

die Behauptung. □

Unter stärkeren Voraussetzungen kann man auch globale Konvergenz des Newton-Verfahrens zeigen.

Satz 4.6. *(Newton-Mysovskii)*
Sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \to \mathbb{R}^n$ stetig differenzierbar und $f'(x)$ sei invertierbar für jedes $x \in \Omega$. Für beliebiges $x_0 \in \Omega$ gebe es $\alpha, \omega \geq 0$ mit

(a) $\|\Delta x_0\| \leq \alpha$, $\Delta x_0 = -f'(x_0)^{-1} f(x_0)$

(b) $\|f'(x)^{-1}\big(f'(y) - f'(z)\big)(y - z)\| \leq \omega \|y - z\|^2$ *für alle* $x, y, z \in \Omega$, *für die* $y \in [x, z]$ *(y liegt auf der Strecke von x nach z).*

(c) $\gamma := \alpha\omega/2 < 1$

(d) *Für* $\rho = \alpha/(1 - \gamma)$ *sei* $B(x_0, \rho) = \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq \rho\} \subset \Omega$.

Dann bleibt die Newton-Folge (x_k) in $B(x_0, \rho)$ und konvergiert gegen eine Lösung $\widehat{x}$ von $f(x) = 0$. Es gilt die Abschätzung

$$\|x_{k+1} - x_k\| \leq \frac{\omega}{2} \|x_k - x_{k-1}\|^2. \tag{4.6}$$

Beweis. Wir zeigen zuerst die Fehlerabschätzung (4.6). Für $x_{k-1}, x_k \in B(x_0, \rho)$ gilt nach (4.5)

$$\begin{aligned}\|x_{k+1} - x_k\| &= \|f'(x_k)^{-1} f(x_k)\| \\ &= \|f'(x_k)^{-1}\big[f(x_k) - \big(f(x_{k-1}) + f'(x_{k-1})\Delta x_{k-1}\big)\big]\|.\end{aligned}$$

Verwendet man

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt \qquad \text{mit} \qquad \varphi(t) = f(x_{k-1} + t\Delta x_{k-1}),$$

so ergibt sich

$$f(x_k) - f(x_{k-1}) = \int_0^1 f'(x_{k-1} + t\Delta x_{k-1})\Delta x_{k-1} dt.$$

Damit ist nach (b) mit $x = x_k$, $y = x_{k-1} + t\Delta x_{k-1}$ und $z = x_{k-1}$ wegen $y - z = t\Delta x_{k-1}$

$$\begin{aligned}\|x_{k+1} - x_k\| &= \|\int_0^1 f'(x_k)^{-1}\big[f'(x_{k-1} + t\Delta x_{k-1}) - f'(x_{k-1})\big]\Delta x_{k-1} dt\| \\ &\le \int_0^1 \|f'(x_k)^{-1}\big[f'(x_{k-1} + t\Delta x_{k-1}) - f'(x_{k-1})\big]\Delta x_{k-1} t\| t^{-1} dt \\ &\le \int_0^1 \omega \|\Delta x_{k-1}\|^2 t\, dt \\ &= \frac{\omega}{2}\|\Delta x_{k-1}\|^2\end{aligned}$$

Im zweiten Teil des Beweises nehmen wir an, dass für ein k bereits $x_j \in B(x_0, \rho)$ für $j \le k$ gezeigt ist. Dies ist gerechtfertigt, da die Behauptung für $k = 0$ wegen $\rho > \alpha$ stimmt. Dann gilt nach dem gerade Gezeigten und mit der Formel für die endliche geometrische Reihe

$$\begin{aligned}\|\Delta x_k\| = \|x_{k+1} - x_k\| &\le \frac{\omega}{2}\|\Delta x_{k-1}\|^2 \\ &\le \frac{\omega}{2}\left(\frac{\omega}{2}\right)^2 \|\Delta x_{k-2}\|^4 \le \ldots \le \left(\frac{\omega}{2}\right)^{1+2+4+\ldots+2^{k-1}} \|\Delta x_0\|^{2^k} \\ &= \frac{2}{\omega}\left(\frac{\omega}{2}\|\Delta x_0\|\right)^{2^k} \le \frac{2}{\omega}\gamma^{2^k}.\end{aligned}$$

Für $0 \le l \le k$ folgt daraus

$$\begin{aligned}\|x_{k+1} - x_l\| &\le \|x_{k+1} - x_k\| + \ldots + \|x_{l+1} - x_l\| \\ &\le \frac{2}{\omega}\gamma^{2^l}\left[\ldots + \gamma^{15\cdot 2^l} + \gamma^{7\cdot 2^l} + \gamma^{3\cdot 2^l} + \gamma^{2^l} + 1\right] \\ &\le \frac{2}{\omega}\frac{\gamma^{2^l}}{1 - \gamma^{2^l}}.\end{aligned} \tag{4.7}$$

Für $l = 0$ ist

$$\|x_{k+1} - x_0\| \le \frac{2}{\omega}\frac{\gamma}{1 - \gamma} = \rho.$$

Mit Induktion liegt daher die gesamte Folge in $B(x_0, \rho)$.

Ferner ist $(x_k)_{k\ge 0}$ wegen (4.7) eine Cauchy-Folge, konvergiert also gegen ein $\widehat{x}$. Aus (4.5) folgt aus der Stetigkeit von f und f', dass $f(\widehat{x}) = 0$. □

In Algorithmus 4.2 haben wir noch offen gelassen, wann wir die Iteration beenden. Ein vernünftiges Konvergenzkriterium wäre etwa

$$\|\Delta x_k\| \leq \texttt{tol},$$

wobei `tol` eine vorgegebenen Fehlerschranke ist, denn falls das Newton-Verfahren konvergiert, ist die Konvergenz quadratisch, d. h.

$$\|e_{k+1}\| \leq C\|e_k\|^2$$

und es gilt

$$\|e_k\| = \|x_k - \widehat{x}\| \leq \|x_k - x_{k+1}\| + \|e_{k+1}\| \leq \|x_k - x_{k+1}\| + C\|e_k\|^2.$$

Ist $C\|e_k\| \ll 1$, so folgt

$$\|e_k\| \leq \frac{\|\Delta x_k\|}{1 - C\|e_k\|} \lesssim \|\Delta x_k\|$$

Als Norm $\|\cdot\|$ verwendet man eine gegebenenfalls an das Problem angepasste Norm. Denkbar ist eine skalierte Euklid-Norm der Form

$$\|v\| = \|Dv\|_2, \qquad D = \operatorname{diag}(d_1, \ldots, d_n).$$

Um eine Endlosschleife zu vermeiden, sollte man zusätzlich die maximale Anzahl von Iterationen auf $k \leq k_{\max}$ beschränken.

Nicht sinnvoll wäre das Konvergenzkriterium

$$\|f(x_k)\| \leq \texttt{tol},$$

denn wenn wir $f(x) = 0$ durch das äquivalente System $Af(x) = 0$ mit einer invertierbaren Matrix A ersetzen, ändern sich die Nullstellen von f und die Iterierten des Newton-Verfahrens nicht:

$$\begin{aligned} f &\longrightarrow Af \\ f'(x)^{-1}f(x) &\longrightarrow f'(x)^{-1}A^{-1} \cdot Af(x). \end{aligned}$$

Also sollten auch die Konvergenz- und Abbruchkriterien invariant gegenüber dieser Transformation sein.

Eine sinnvolle Forderung an die Newton-Folge wäre die Monotonie der Fehlernormen

$$\|e_{k+1}\| \leq \|e_k\|. \tag{4.8}$$

Leider ist diese Forderung nicht realisierbar, da die unbekannte Lösung benötigt wird. Aus oben genannten Gründen ist auch $\|f(x_{k+1})\| \leq \|f(x_k)\|$ nicht brauchbar. Wir versuchen daher, (4.8) näherungsweise zu erfüllen. Mit dem Satz von Taylor erhalten wir die Näherungen

$$\begin{aligned} f(x_k) &\approx f'(\widehat{x})e_k, \\ f(x_{k+1}) &\approx f'(\widehat{x})e_{k+1}. \end{aligned}$$

Ersetzt man hier die unbekannte Jacobi-Matrix $f'(\widehat{x})$ durch die schon berechnete $f'(x_k)$, so kann man (4.8) durch

$$\|f'(x_k)^{-1}f(x_{k+1})\| \leq \|f'(x_k)^{-1}f(x_k)\| = \|\Delta x_k\|$$

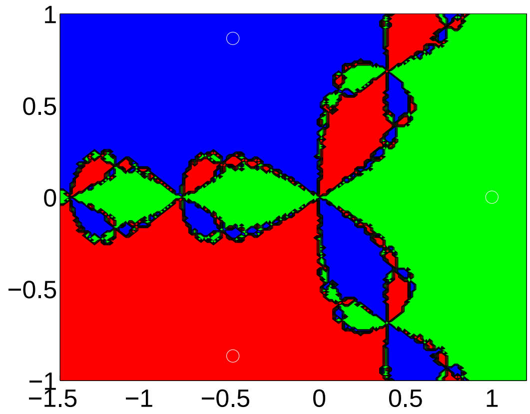

Abb. 4.1: Attraktionsmengen für die Newton-Folgen zu $f(z) = z^3 - 1$

ersetzen. Falls dieses affin-invariante Kriterium nicht erfüllt ist, brechen wir ab. Zudem kann die Funktionsauswertung $f(x_{k+1})$ im nächsten Schritt verwendet werden und das Gleichungssystem

$$f'(x_k)\overline{\Delta x_{k+1}} = -f(x_{k+1}) \tag{4.9}$$

kann mit der schon berechneten LR-Zerlegung von $f'(x_k)$ gelöst werden.

Beispiel. Es sei

$$f(z) = z^3 - 1 \qquad \text{oder} \qquad f(x,y) = \begin{bmatrix} x^3 - 3xy^2 - 1 \\ 3x^2y - y^3 \end{bmatrix}$$

für $z = x + iy$. Das Newton-Verfahren ist durch

$$z_{k+1} = z_k - \frac{z_k^3 - 1}{3z_k^2} = \frac{1}{3}\big(2z_k + \frac{1}{z_k^2}\big)$$

gegeben. Für die drei Nullstellen $1, (-1 \pm i\sqrt{3})/2$ von f definiert man

$$A(a) = \big\{z_0 \in \mathbb{C} \mid \{z_k\} \text{ konvergiert gegen } a\big\}.$$

In Abbildung 4.1 erkennt man, dass die Folge (z_k) nicht notwendigerweise gegen die Nullstelle von f konvergiert, die am nächsten an z_0 liegt. $\diamond$

4.3 Vereinfachtes Newton-Verfahren

Im gewöhnlichen Newton-Verfahren besteht der wesentliche Aufwand darin, in jedem Iterationsschritt die Jacobi-Matrix $f'(x_k)$ und deren LR-Zerlegung zur Lösung des linearen Gleichungssystems zu berechnen. Beim vereinfachten Newton-Verfahren berechnet man hingegen nur einmal eine Approximation $A \approx f'(x_0)$ und deren LR-Zerlegung und verwendet diese auch in den folgenden Schritten, vgl. Algorithmus 4.3.

Den reduzierten Aufwand des vereinfachten Newton-Verfahrens "bezahlt" man mit nur noch lokal linearer Konvergenz:

Algorithmus 4.3 Vereinfachtes Newton-Verfahren

Berechne $A \approx f'(x_0)$ und die LR-Zerlegung von A
for $k = 0, 1, 2, \dots$ **do**
  Berechne $f(x_k)$
  Löse mit Hilfe der LR-Zerlegung $A\Delta x_k = -f(x_k)$
  Setze $x_{k+1} = x_k + \Delta x_k$
end for

Satz 4.7. *Es sei $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ zweimal stetig differenzierbar, $f(\widehat{x}) = 0$ und $A \in \mathbb{R}^{n,n}$ sei invertierbar. Dann gilt für den Fehler $e_k = x_k - \widehat{x}$ der Näherungen x_k des vereinfachten Newton-Verfahrens*

$$e_{k+1} = \big(I - A^{-1}f'(x_k)\big)e_k + O(\|e_k\|^2),$$

falls $\|e_k\|$ hinreichend klein ist.

Beweis. Der Satz von Taylor liefert wieder

$$0 = f(\widehat{x}) = f(x_k - e_k) = f(x_k) - f'(x_k)e_k + O(\|e_k\|^2).$$

Beim vereinfachten Newton-Verfahren gilt jedoch

$$f(x_k) = -A(x_{k+1} - x_k) = -A(e_{k+1} - e_k).$$

Zusammen mit der obigen Gleichung ergibt sich aus

$$Ae_{k+1} = \big(A - f'(x_k)\big)e_k + O(\|e_k\|^2)$$

die Behauptung. □

Eine globale Konvergenzaussage gibt

Satz 4.8. *Es sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \to \mathbb{R}^n$ stetig differenzierbar und $A \in \mathbb{R}^{n,n}$ invertierbar. Ferner sei $x_0 \in \Omega$ gewählt und es gelte*

(a) $\|\Delta x_0\| \le \alpha$,

(b) $\|I - A^{-1}f'(x)\| \le \gamma < 1$ *für alle* $x \in \Omega$,

(c) *für* $\rho = \alpha/(1-\gamma)$ *sei* $B(x_0, \rho) \subset \Omega$.

Dann gilt für die Näherungen des vereinfachten Newton-Verfahrens $x_k \in B(x_0, \rho)$. x_k konvergiert gegen eine Nullstelle $\widehat{x}$ von f und es gilt die Abschätzung

$$\|x_{k+1} - x_k\| \le \gamma\|x_k - x_{k-1}\|.$$

Beweis. Das vereinfachte Newton-Verfahren kann als Fixpunktiteration mit $\Phi(x) = x - A^{-1}f(x)$ aufgefasst werden.

Aus (b) schließen wir, dass Φ in $B(x_0, \rho)$ kontraktiv ist, denn für $x, y \in B(x_0, \rho)$ gilt

$$\|\Phi(x) - \Phi(y)\| \le \sup_{z \in [x,y]} \|\Phi'(z)\| \cdot \|x - y\| \le \gamma\|x - y\|, \qquad \gamma < 1.$$

Die Voraussetzung (c) besagt, dass $\Phi : B(x_0, \rho) \to B(x_0, \rho)$ eine Selbstabbildung ist, denn für $x \in B(x_0, \rho)$ ist wegen $x_1 = \Phi(x_0)$

$$\|\Phi(x) - x_0\| \le \|\Phi(x) - \Phi(x_0)\| + \|x_1 - x_0\| \le \gamma\|x - x_0\| + \alpha \le \gamma\rho + \alpha = \rho.$$

Die Behauptung folgt jetzt direkt aus dem Banach'schen Fixpunktsatz 4.3. □

4.4 Das gedämpfte Newton-Verfahren

Beim Newton-Verfahren haben wir das Abbruchkriterium

$$\|\overline{\Delta x_{k+1}}\| > \theta \, \|\Delta x_k\|$$

mit $\theta = 1$ für $\overline{\Delta x_{k+1}}$ definiert in (4.9) verwendet. Die Erfahrung und auch eine theoretische Analyse zeigen, dass die Wahl $\theta = 1/2$ günstiger ist. Eine Alternative zu einem Neustart mit hoffentlich besserem Startwert ist, eine Dämpfung der Newton-Schritte vorzunehmen, also die nächste Iterierte als

$$x_{k+1} = x_k + \lambda_k \Delta x_k, \qquad 0 < \lambda_k \leq 1.$$

zu wählen. Hierbei ist λ_k ein Dämpfungsfaktor, den wir so wählen, dass die gedämpfte Iterierte den Monotonietest in obigem Abbruchkriterium mit $\theta = 1 - \lambda_k/2$ besteht. Hierzu wählt man λ_k aus einer endlichen Folge $\{2^{-j}, j = 0, \ldots, j_{\max}\}$ aus und bricht die Iteration erst dann ab, wenn $\lambda_k < 2^{-j_{\max}}$ gewählt werden müsste. War λ_k erfolgreich, so versucht man den nächsten Schritt mit $\lambda_{k+1} = \min\{1, 2\lambda_k\}$, um asymptotisch die quadratische Konvergenz des Newton-Verfahrens ($\lambda = 1$) zu erreichen.

4.5 Das Sekantenverfahren

Der Hauptaufwand des Newton-Verfahrens besteht in der Auswertung von f und von f'. Besonders die Berechnung von f' kann in Anwendungen sehr schwierig oder sogar unmöglich sein. In vielen Fällen ist f' deutlich komplizierter als f. Für den skalaren Fall $f : \mathbb{R} \to \mathbb{R}$ erhält man ein weniger aufwendiges Verfahren, wenn man f' durch einen Differenzenquotienten ersetzt, etwa

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

Einsetzen dieser Approximation in die Newton-Vorschrift liefert das **Sekantenverfahren**

$$\begin{aligned} x_{k+1} &= x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \, f(x_k) \\ &= \frac{x_{k-1} f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}. \end{aligned}$$

Der Name Sekantenverfahren beruht darauf, dass man f durch die Sekante durch $(x_{k-1}, f(x_{k-1}))$ und $(x_k, f(x_k))$ ersetzt und die Nullstelle dieser Sekante als Näherung verwendet.

Satz 4.9. *Es sei $f : \Omega \to \mathbb{R}$ zweimal stetig differenzierbar im Intervall $\Omega = [a, b]$, $f(\widehat{x}) = 0$ für ein $\widehat{x} \in (a, b)$ mit $f'(\widehat{x}) \neq 0$ und $f''(\widehat{x}) \neq 0$. Dann konvergiert das Sekantenverfahren lokal gegen $\widehat{x}$ mit der exakten Konvergenzordnung*

$$p = \frac{1}{2}(1 + \sqrt{5}) \approx 1.61803\ldots$$

Beweis. (a) Aus der Iterationsvorschrift folgt für den Fehler $e_k = x_k - \widehat{x}$ die Rekursion

$$\begin{aligned} e_{k+1} &= e_k - \frac{e_k - e_{k-1}}{f(x_k) - f(x_{k-1})} \, f(x_k) \\ &= \frac{e_{k-1} f(x_k) - e_k f(x_{k-1})}{f(x_k) - f(x_{k-1})} \end{aligned}$$

oder

$$\frac{e_{k+1}}{e_k e_{k-1}} = \frac{1}{f(x_k) - f(x_{k-1})}\left(\frac{f(x_k)}{x_k - \widehat{x}} - \frac{f(x_{k-1})}{x_{k-1} - \widehat{x}}\right) = \frac{g(x_k) - g(x_{k-1})}{f(x_k) - f(x_{k-1})}, \tag{4.10}$$

mit

$$g(x) = \frac{f(x)}{x - \widehat{x}}, \qquad g'(x) = \frac{f'(x)(x - \widehat{x}) - f(x)}{(x - \widehat{x})^2}.$$

Nach dem verallgemeinerten Mittelwertsatz gibt es ein ξ_k zwischen x_{k-1} und x_k so, dass

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

gilt. Also folgt aus (4.10)

$$\frac{e_{k+1}}{e_k e_{k-1}} = \frac{g'(\xi_k)}{f'(\xi_k)} = \frac{1}{f'(\xi_k)} \frac{f'(\xi_k)(\xi_k - \widehat{x}) - f(\xi_k)}{(\xi_k - \widehat{x})^2} = \frac{1}{2}\frac{f''(\zeta_k)}{f'(\xi_k)} \tag{4.11}$$

für ζ_k zwischen $\widehat{x}$ und ξ_k. Die letzte Identität folgt aus dem Satz von Taylor mit $f(\widehat{x}) = 0$.

Nach Voraussetzung ist in einer hinreichend kleinen Umgebung um $\widehat{x}$ die rechte Seite betragsmäßig beschränkt, d. h. es gibt $C > 0$ mit

$$|e_{k+1}| \le C|e_{k-1}| \cdot |e_k|, \qquad k \ge 1.$$

Der Fehler des Sekantenverfahrens konvergiert also monoton und mindestens linear gegen Null, wenn $|e_0| \le C^{-1}$ und $|e_1| \le C^{-1}$.

(b) Um die Konvergenzordnung nachzuweisen, setzen wir

$$\epsilon_k = \frac{|e_k|}{|e_{k-1}|^p}, \qquad p = \frac{1}{2}(1 + \sqrt{5}) \tag{4.12}$$

und leiten eine Rekursion für $\log \epsilon_k$ her. Wegen

$$\frac{1}{p} = \frac{2}{\sqrt{5} + 1} = \frac{\sqrt{5} - 1}{2} = p - 1$$

ist wegen (4.11)

$$\epsilon_{k+1} = \frac{|e_{k+1}|}{|e_k|^p} = \frac{|e_{k+1}|}{|e_k|}\,|e_k|^{1-p} = \alpha_k |e_{k-1}|\,|e_k|^{-1/p} = \alpha_k \epsilon_k^{-1/p}, \qquad \alpha_k = \frac{1}{2}\frac{|f''(\zeta_k)|}{|f'(\xi_k)|}.$$

Dies ist äquivalent zu

$$\begin{aligned}
\log \epsilon_{k+1} &= \log \alpha_k - \frac{1}{p}\log \epsilon_k \\
&= \log \alpha_k - \frac{1}{p}\log \alpha_{k-1} + \frac{1}{p^2}\log \epsilon_{k-1} \\
&= \dots \\
&= \sum_{j=1}^{k} (-p)^{j-k} \log \alpha_j + (-p)^{-k} \log \epsilon_1.
\end{aligned}$$

Wir betrachten jetzt eine hinreichend kleine Umgebung von $\widehat{x}$. Diese wählen wir so, dass die Konvergenz monoton ist (also $|e_{k+1}| \leq |e_k|$ gilt) und $|f'| > 0$, $|f''| > 0$ und beschränkt sind. Dann existiert ein $\alpha > 0$, so dass

$$|\log \alpha_j| \leq \alpha < \infty \qquad \forall j \in \mathbb{N}.$$

Für x_0, x_1 in dieser Umgebung folgt aus $p^{-1} < 1$

$$|\log \epsilon_{k+1}| \leq |\log \epsilon_1| + \alpha \sum_{j=0}^{\infty} p^{-j} =: c < \infty,$$

also $\epsilon_k \in (e^{-c}, e^c)$. Aus (4.12) folgt daraus

$$|e_k| \leq e^c |e_{k-1}|^p, \qquad |e_k| \geq e^{-c} |e_{k-1}|^p,$$

das Sekantenverfahren konvergiert genau mit der Ordnung p. □