

Kapitel 8

Banach-Iteration

Aus Zeitgründen ist dieses Kapitel nicht „offizieller“ Bestandteil der Vorlesung (sprich: ist nicht prüfungsrelevant). Es wird jedoch wärmstens empfohlen, dieses Kapitel in Heimarbeit zu lesen und zu verstehen. Speziell gibt es kein Physikerdasein ohne Kenntnis des Newton-Verfahrens (Abschnitt 8.2).

8.1 Der Banachsche Fixpunktsatz

Wir definieren den Begriff einer kontrahierenden Abbildung:

Definition 8.1:

Eine Abbildung $\Phi : A \subset \mathbb{C} \rightarrow \mathbb{C}$ heißt „**Kontraktion in A** “, wenn eine „**Kontraktionskonstante**“ $k \in [0, 1)$ existiert, so dass für alle $x, y \in A$ gilt:

$$|\Phi(x) - \Phi(y)| \leq k \cdot |x - y|.$$

Zur Namensgebung: der Abstand zweier Bildpunkte $|\Phi(x) - \Phi(y)|$ einer Kontraktion ist stets kleiner als der Abstand der Urbildpunkte $|x - y|$.

Bemerkung 8.2: Kontraktionen sind automatisch stetig: Für jede konvergierende Folge (z_n) konvergiert $\Phi(z_n)$, und es gilt

$$\lim_{n \rightarrow \infty} \Phi(z_n) = \Phi\left(\lim_{n \rightarrow \infty} z_n\right).$$

Dies ist leicht einzusehen. Sei z^* der Grenzwert von (z_n) . Die Kontraktionseigenschaft liefert

$$0 \leq |\Phi(z_n) - \Phi(z^*)| < |z_n - z^*|.$$

Wegen der Konvergenz $z_n \rightarrow z^*$ ist die rechte Seite eine reelle Nullfolge. Mit Satz 2.17 folgt, dass $|\Phi(z_n) - \Phi(z^*)|$ eine Nullfolge ist, was die Konvergenz $\Phi(z_n) \rightarrow \Phi(z^*)$ bedeutet.

Der folgende wichtige Fixpunktsatz für kontrahierende Abbildungen geht auf Stefan Banach (polnischer Mathematiker, 1892 – 1945) zurück. Er setzt neben der Kontraktionseigenschaft als wichtige Annahme voraus, dass die Kontraktion ihren Definitionsbereich in sich selbst abbildet:

Satz 8.3: („BFS“: der Banachsche Fixpunktsatz)

Sei $\Phi : A \rightarrow A$ eine Kontraktion in einer abgeschlossenen Menge $A \subset \mathbb{C}$ mit einer Kontraktionskonstanten $k < 1$. Dann

- a) existiert ein eindeutig bestimmter Fixpunkt $z^* = \Phi(z^*) \in A$,
- b) konvergiert jede Folge $z_{n+1} = \Phi(z_n)$ mit beliebigem Startwert $z_0 \in A$ gegen z^* ,
- c) gelten für jede solche Folge die Abschätzungen

$$|z_n - z^*| \leq \underbrace{\frac{k}{1-k} \cdot |z_n - z_{n-1}|}_{\text{„a posteriori“}} \leq \underbrace{\frac{k^n}{1-k} \cdot |z_1 - z_0|}_{\text{„a priori“}}.$$

Vom (wiederum sehr technischen) Beweis braucht man eigentlich nur zu wissen, dass man zeigen kann, dass die per $z_{n+1} = \Phi(z_n)$ konstruierten Folgen Cauchy-Folgen sind. Der Fixpunkt ergibt sich dann über den Grenzwert der Folgen, dessen Existenz mittels Satz 2.28 gesichert ist. Aus der Kontraktionseigenschaft folgt die Eindeutigkeit (alle Folgen konvergieren gegen denselben Grenzwert). Für technisch Interessierte ist der Beweis der Vollständigkeit halber angegeben:

Beweis: Zeige: (z_n) ist Cauchy-Folge.

$$\begin{aligned} |z_{n+m} - z_n| &= |z_{n+m} - z_{n+m-1} + z_{n+m-1} - \dots - z_{n+1} + z_{n+1} - z_n| \\ &\leq |z_{n+m} - z_{n+m-1}| + |z_{n+m-1} - z_{n+m-2}| + \dots + |z_{n+1} - z_n| \end{aligned}$$

für jedes $n, m \geq 0$. Aus $|z_j - z_{j-1}| = |\Phi(z_{j-1}) - \Phi(z_{j-2})| \leq k \cdot |z_{j-1} - z_{j-2}|$, d.h.,

$$|z_j - z_{j-1}| \leq k \cdot |z_{j-1} - z_{j-2}| \leq k^2 \cdot |z_{j-2} - z_{j-3}| \leq \dots$$

folgt

$$\begin{aligned} |z_{n+m} - z_n| &\leq k^{m-1} \cdot |z_{n+1} - z_n| + k^{m-2} \cdot |z_{n+1} - z_n| + \dots + |z_{n+1} - z_n| \\ &= (1 + k + k^2 + \dots + k^{m-1}) \cdot |z_{n+1} - z_n| = \frac{1 - k^m}{1 - k} \cdot |z_{n+1} - z_n| \\ &\leq \frac{|z_{n+1} - z_n|}{1 - k} \stackrel{(\#)}{\leq} \frac{k |z_n - z_{n-1}|}{1 - k} \leq \frac{k^2 |z_{n-1} - z_{n-2}|}{1 - k} \leq \dots \stackrel{(\#\#)}{\leq} \frac{k^n |z_1 - z_0|}{1 - k}. \end{aligned}$$

Mit $k^n \rightarrow 0$ folgt die Cauchy-Eigenschaft. Es existiert somit ein Grenzwert z^* . Mit $z_0 \in A$ und $\Phi : A \rightarrow A$ folgt $z_n \in A \Rightarrow z^*$ ist Häufungspunkt von $A \Rightarrow z^* \in A$ (abgeschlossen). Da Φ im Sinne von Bemerkung 8.2 stetig ist:

$$z^* = \lim_{n \rightarrow \infty} z_{n+1} = \lim_{i \rightarrow \infty} \Phi(z_n) = \Phi(\lim_{n \rightarrow \infty} z_n) = \Phi(z^*).$$

Eindeutigkeit: für einen weiteren Fixpunkt $z^{**} \neq z^*$ folgt der Widerspruch

$$|z^* - z^{**}| = |\Phi(z^*) - \Phi(z^{**})| \leq k \cdot |z^* - z^{**}| < |z^* - z^{**}|.$$

Die Abschätzungen c) ergeben sich aus (#) und (##):

$$\lim_{m \rightarrow \infty} |z_{n+m} - z_n| = |z^* - z_n| \stackrel{(\#)}{\leq} \frac{k}{1-k} |z_n - z_{n-1}| \stackrel{(\#\#)}{\leq} \frac{k^n}{1-k} |z_1 - z_0|.$$

Q.E.D.

Interpretation und Anwendung 8.4:

Eine Gleichung $f(z) = 0$ sei zu lösen. Der BFS gibt ein Rezept, wie man Näherungen für die Lösung konstruieren kann:

- 1) Formuliere die Gleichung $f(z) = 0$ in ein äquivalentes Fixpunktproblem $z = \Phi(z)$ um.
- 2) Ist Φ kontrahierend auf einer Umgebung A der Lösung und bildet Φ diese Menge A auf sich selbst ab, so läßt sich der BFS anwenden: wähle einen beliebigen Punkt $z_0 \in A$ und iteriere $z_{n+1} = \Phi(z_n)$. Diese Folge konvergiert gegen eine Lösung des Fixpunktproblems und damit gegen eine Lösung des Ausgangsproblems $f(z) = 0$.
- 3) Hat man durch Abschätzungen eine Kontraktionskonstante k für die Menge A gefunden, kann man mit den a-priori- bzw. a-posteriori-Abschätzungen bestimmen, wie weit man noch von der Lösung entfernt ist und abbrechen, sobald eine vorgegebene Zielgenauigkeit erreicht ist.

Mit der a-priori-Abschätzung kann man aus dem Startpunkt z_0 und dem nächsten Punkt z_1 sofort ermitteln, wie oft man höchstens iterieren muß, um die Zielgenauigkeit zu erreichen (die Iterationswerte werden dafür nicht benötigt). Nachdem die Iteration durchgeführt worden ist und Zahlenwerte für z_n vorliegen, kann man a-posteriori abschätzen, welche Approximationsgenauigkeit nun wirklich erreicht ist (die a-posteriori-Abschätzung ist prinzipiell genauer als die a-priori-Abschätzung).

8.2 Das Newton-Verfahren

Hier eine für die Praxis besonders wichtige Anwendung des BFS. Es geht darum, numerische Approximationen von Nullstellen x^* einer Funktion zu konstruieren. Die prinzipielle Idee ist, eine Zahlenfolge (x_n) zu konstruieren, die gegen einen Grenzwert x^* mit $f(x^*) = 0$ konvergiert. Für hinreichend großes n bricht man die Iteration ab und akzeptiert x_n als eine Approximation von x^* . Zunächst hat man die Aufgabe, ein Nullstellenproblem $f(x^*) = 0$ in ein Fixpunktproblem $x^* = \Phi(x^*)$ umzuformulieren, um den BFS anwenden zu können.

Bemerkung 8.5: *Es gibt viele Wege, eine gegebene Gleichung $f(x) = 0$ in eine Fixpunktgleichung $x = \Phi(x)$ umzuformen, z.B.*

$$\Phi(x) = x - g(x) \cdot f(x)$$

mit einer (praktisch beliebig wählbaren) Funktion $g(x)$. Ist $f(x)$ differenzierbar und ist die Lösung im Sinne von Definition 1.24 eine einfache Nullstelle, so ist $g(x) = 1/f'(x)$ eine ausgezeichnete Wahl. Die Iteration lautet dann

$$x_{n+1} = \Phi(x_n) = x_n - \frac{f(x_n)}{f'(x_n)} \quad (\text{das „Newton-Verfahren“}).$$

Es gibt folgende einfache **geometrische Interpretation** des Newton-Verfahrens. Sei x_n ein Approximation einer Nullstelle von $f(x)$. Ersetze $f(x)$ in einer Umgebung von x_n durch die Tangente

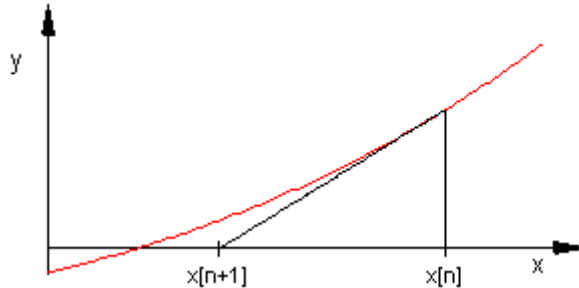
$$t(x) = f(x_n) + (x - x_n) \cdot f'(x_n).$$

Die Nullstelle der Tangente ist: $t(x) = 0 \Rightarrow x = x_n - \frac{f(x_n)}{f'(x_n)}$, also:

Der Newton-Nachfolger $x_{n+1} = x_n - f(x_n)/f'(x_n)$ von x_n ist die Nullstelle der Tangente von f am Punkt x_n .

Hier eine Beispiel-Graphik (mit MuPAD 3.0) zur Demonstration:

```
>> f:= x -> x^4 - 2: x0:= 1.5:
>> plot(plot::Function2d(f(x), x = 1.1..1.6, Color = RGB::Red),
        plot::Function2d(f(x0) + (x - x0)*f'(x0),
                          x = x0 - f(x0)/f'(x0) .. x0,
                          Color = RGB::Black),
        plot::Line2d([x0, 0], [x0, f(x0)], Color = RGB::Black),
        TicksNumber = None,
        XTicksAt = [x0 = "x[n]", x0 - f(x0)/f'(x0) = "x[n+1]"],
        Height = 4*unit::cm, Width = 8*unit::cm)
```



Hier einige numerische Werte bei der Berechnung von $\sqrt[4]{2}$ durch Lösen von $f(x) = x^4 - 2 = 0$:

```
>> f:= x -> float(x^4 - 2): x0:= 1.5:
>> Phi:= x -> x - f(x)/f'(x):
>> x0 := 1.5
                                1.5
>> x1 := Phi(x0)
                                1.273148148
>> x2 := Phi(x1)
                                1.19714982
>> x3 := Phi(x2)
                                1.189285812
>> x4 := Phi(x3)
                                1.189207123
>> x5 := Phi(x4)
                                1.189207115
```

Bemerkung 8.6: Man kann zeigen, dass es immer eine (eventuell kleine) Umgebung A einer Nullstelle von f gibt, auf der Φ eine Kontraktion ist und für die $\Phi(A) \subset A$ gilt. Damit gilt der BFS auf einer (leider oft nicht konkret bekannten) Umgebung einer Lösung, und es gilt:

Für hinreichend genaue Startwerte x_0 dicht bei einer Lösung x^* von $f(x) = 0$ konvergiert die Newton-Folge $x_{n+1} = x_n - f(x_n)/f'(x_n)$ gegen x^* .

Bei einfachen Nullstellen ist die Konvergenz beim Newton-Verfahren sehr schnell, da die Kontraktionskonstanten auf kleinen Umgebungen der Lösung prinzipiell sehr klein sind. Gilt $f'(x^*) \neq 0$ an der Nullstelle x^* , so erhält man durch Taylor-Entwicklung um x^* :

$$f(x) = \underbrace{f(x^*)}_{=0} + f'(x^*) \cdot (x - x^*) + \frac{f''(x^*)}{2} \cdot (x - x^*)^2 + \dots,$$

$$\begin{aligned}
f'(x) &= f'(x^*) + f''(x^*) \cdot (x - x^*) + \dots, \\
\Rightarrow x_{n+1} - x^* &= x_n - x^* - \frac{f'(x^*) \cdot (x_n - x^*) + \frac{f''(x^*)}{2} \cdot (x_n - x^*)^2 + \dots}{f'(x^*) + f''(x^*) \cdot (x_n - x^*) + \dots} \\
&= (x_n - x^*) \cdot \left(1 - \frac{f'(x^*) + \frac{f''(x^*)}{2} \cdot (x_n - x^*) + O((x - x^*)^2)}{f'(x^*) + f''(x^*) \cdot (x_n - x^*) + \dots}\right) \\
&= (x_n - x^*) \cdot \frac{-\frac{f''(x^*)}{2} \cdot (x_n - x^*) + O((x - x^*)^2)}{f'(x^*) + f''(x^*) \cdot (x_n - x^*) + \dots} \\
&= (x_n - x^*)^2 \cdot \frac{-\frac{f''(x^*)}{2} + O(x_n - x^*)}{f'(x^*) + O(x_n - x^*)} \\
&\approx - (x_n - x^*)^2 \cdot \frac{f''(x^*)}{2 \cdot f'(x^*)}.
\end{aligned}$$

Ist der Approximationsfehler $x_n - x^*$ erst einmal klein, so wird er durch Quadrieren sehr schnell sehr klein (man nennt dies „**quadratische Konvergenz**“).

Beispiel 8.7: In einer Softwareumgebung gebe es die Grundarithmetik, aber keine Wurzelfunktion. Um diese zu implementieren, soll die Gleichung $y^2 = b$ für gegebenes positives $b \in \mathbb{R}$ nach y gelöst werden. Mittels Division durch eine geeignete 4er-Potenz kann b auf das Intervall $[1, 4]$ transformiert werden (in Binärdarstellung kostet dies nichts). Sei nun $a = b/4^n \in [1, 4]$. Ist $x \in [1, 2]$ eine Lösung von $x^2 = a$, so ist $y = 2^n \cdot x$ die gesuchte Lösung des Ausgangsproblems $y^2 = b$.

Das verbleibende Problem ist also, ein $x \in [1, 2]$ zu finden, das das Nullstellenproblem $f(x) = x^2 - a = 0$ mit $a \in [1, 4]$ erfüllt. Hierzu soll das Newton-Verfahren benutzt werden. Betrachte also

$$\Phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x^2 - a}{2 \cdot x} = \frac{x^2 + a}{2 \cdot x} = \frac{1}{2} \cdot \left(x + \frac{a}{x}\right).$$

Die entsprechende Iteration lautet also

$$x_{n+1} = \frac{1}{2} \cdot \left(x_n + \frac{a}{x_n}\right).$$

Das Verfahren konvergiert sehr schnell gegen $x^* = \sqrt{a}$. Beispiel: $a = 2$, $x_0 = 1.5$:

$$x_0 = 1.5, \quad x_1 = 1.416666666\dots, \quad x_2 = 1.414215686\dots,$$

$$x_3 = 1.414213562\dots, \quad x_4 = 1.414213562\dots, \quad \dots$$

Die folgende Analyse ist wiederum sehr technisch und an technisch Interessierte adressiert:

Analyse: betrachte das Intervall $A = [\sqrt{a}, \frac{1+a}{2}]$, das die Lösung \sqrt{a} enthält. (Dieses Intervall fällt hier vom Himmel.) Die folgenden Rechnungen zeigen, dass dieses Intervall

in der Tat so ist, dass der BFS angewendet werden kann.

Es ist zunächst zu zeigen, dass $\Phi(A) \subset A$ gilt. In Erinnerung an die Schule berechnen wir dazu

$$\Phi'(x) = \frac{1}{2} \cdot \left(1 - \frac{a}{x^2}\right) = \frac{x^2 - a}{2 \cdot x^2} \geq 0$$

für $x \in [\sqrt{a}, \frac{1+a}{2}]$. Die Funktion ist also in diesem Bereich monoton steigend, und damit gilt

$$\Phi(\sqrt{a}) = \sqrt{a} \leq \Phi(x) \leq \Phi\left(\frac{1+a}{2}\right) = \frac{1+a}{4} + \frac{a}{1+a} \leq \frac{1+a}{2}$$

für alle $x \in [\sqrt{a}, \frac{1+a}{2}]$. Als Kontraktionskonstante auf A schätzt man ab:

$$\begin{aligned} |\Phi(x) - \Phi(y)| &= \frac{1}{2} \cdot \left|x + \frac{a}{x} - y - \frac{a}{y}\right| = \frac{1}{2} \cdot \left|x - y - \frac{a \cdot (x - y)}{x \cdot y}\right| \\ &= \frac{|x - y|}{2} \cdot \left|1 - \frac{a}{x \cdot y}\right| = \frac{|x - y|}{2} \cdot \left(1 - \frac{a}{x \cdot y}\right) \leq \frac{|x - y|}{2} \cdot \left(1 - \frac{a}{\left(\frac{a+1}{2}\right)^2}\right) \\ &= |x - y| \cdot \frac{1}{2} \cdot \left(\frac{a-1}{a+1}\right)^2. \end{aligned}$$

Also, in Abhängigkeit von a ist

$$k = \frac{1}{2} \cdot \left(\frac{a-1}{a+1}\right)^2$$

eine Kontraktionskonstante für Φ über A . Für alle $a \in [1, 4]$ gilt $k \leq 0.18$, d.h., Φ ist auf A in der Tat eine Kontraktion.

Speziell, für $a = 2$ ist $k = 1/18 \approx 0.0555 \dots$. Starten wir mit $x_0 = 3/2$, so ergibt sich $x_1 = 17/12$ und die a-priori-Abschätzung liefert

$$|x_n - \sqrt{2}| \leq \frac{k^n}{1-k} \cdot |x_0 - x_1| = \frac{3}{34 \cdot 18^n}.$$

Nach $n = 5$ Schritten ergibt sich beispielsweise

$$|x_5 - \sqrt{2}| \leq 4.7 \cdot 10^{-8},$$

d.h., x_5 beschreibt garantiert die ersten 7 bis 8 Dezimalstellen von $\sqrt{2}$ korrekt. (In Wirklichkeit ist x_5 schon wesentlich genauer, aber mehr gibt die a-priori-Abschätzung nicht her.)

Bemerkung 8.8: Das letzte Beispiel hat gezeigt, dass das Abschätzen von Kontraktionskonstanten mühselig ist. Es geht aber auch einfacher! Für stetig differenzierbares Φ gilt, dass

$$k = \sup \{|\Phi'(x)|; x \in A\}$$

die bestmögliche (weil kleinste) Kontraktionskonstante über einem Intervall $A \subset \mathbb{R}$ ist. Um dies einzusehen, brauchen wir nur den Mittelwertsatz der Differentialrechnung 6.20 zu betrachten. Es gilt

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

mit ξ im Intervall zwischen x und y . Es folgt

$$|\Phi(x) - \Phi(y)| = |\Phi'(\xi)| \cdot |x - y| \leq k \cdot |x - y|,$$

wobei k das obige Supremum ist.

Bemerkung 8.9: Im Fall skalarer Gleichungen $f(x) = 0$ hat das Newton-Verfahren angenehme lokale Eigenschaften: Der Erfolg ist garantiert, wenn nur der Startpunkt x_0 hinreichend gut ist. Dies gilt sogar bei p -fach entarteten Nullstellen (bei denen für $p > 1$ die quadratische Konvergenz dann allerdings zu linearer Konvergenz mit der asymptotischen Kontraktionskonstanten $k = 1 - 1/p$ wird).

Die globalen Eigenschaften des Newton-Verfahrens sind deutlich weniger glorreich. Gerät die Newton-Folge in die Nähe eines lokalen Extremums, wird die Iteration $x \rightarrow x - f(x)/f'(x)$ in der Nähe einer Nullstelle von f' ausgewertet, und die Iterationspunkte können weit weg landen.