

2.4 Das Newton–Verfahren für nichtlineare Gleichungssysteme

Problem: Wir suchen die Nullstellen einer Funktion $f : D \rightarrow \mathbb{R}^n$ mit $D \subset \mathbb{R}^n$:

$$f(\mathbf{x}) = \mathbf{0}$$

Wir kennen bereits die **Fixpunktiteration**

$$\mathbf{x}^{k+1} := \Phi(\mathbf{x}^k)$$

mit Startwert \mathbf{x}^0 und Iterationsvorschrift $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Konvergenzaussagen liefert der **Banachsche Fixpunktsatz**.

Vorteil: Verfahren ist **ableitungsfrei**

Nachteile:

- Numerisches Verfahren konvergiert langsam (linear),
- Keine eindeutige Iterationsvorschrift.

100

Alternative: Newton–Verfahren für vektorwertige Funktionen

Gegeben sei eine \mathcal{C}^1 –Funktion $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen. Wir suchen eine Nullstelle, d.h. ein $\mathbf{x}^* \in D$ mit

$$f(\mathbf{x}^*) = \mathbf{0}$$

Taylor–Entwicklung von $f(\mathbf{x})$ um einen Startwert \mathbf{x}^0 lautet

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \mathbf{J}f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0) + o(\|\mathbf{x} - \mathbf{x}^0\|)$$

Setzen wir $\mathbf{x} = \mathbf{x}^*$, so folgt

$$\mathbf{J}f(\mathbf{x}^0)(\mathbf{x}^* - \mathbf{x}^0) \approx -f(\mathbf{x}^0)$$

Eine Näherungslösung für \mathbf{x}^* ist dann \mathbf{x}^1 , die Lösung des LGS

$$\mathbf{J}f(\mathbf{x}^0)(\mathbf{x}^1 - \mathbf{x}^0) = -f(\mathbf{x}^0)$$

101

Das **Newton–Verfahren** lautet somit

für $k = 0, 1, 2, \dots$

$$\mathbf{Jf}(\mathbf{x}^k) \cdot \Delta \mathbf{x}^k = -f(\mathbf{x}^k)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta \mathbf{x}^k$$

Man muss also im Newton–Verfahren in jedem Iterationsschritt ein LGS lösen. Die Lösung $\Delta \mathbf{x}^k$ heißt **Newton–Korrektur**.

Das Newton–Verfahren ist **skalierungsinvariant**:

Satz: Das Newton–Verfahren ist invariant unter allen Transformationen (Skalierungen) der Form

$$\mathbf{f}(\mathbf{x}) \rightarrow \mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{f}(\mathbf{x})$$

mit einer regulären Matrix $\mathbf{A} \in \mathbb{R}^{(n,n)}$, d.h. die Iterierten für \mathbf{f} und \mathbf{g} sind identisch.

102

Satz: (Konvergenzsatz)

Sei $\mathbf{f} : D \rightarrow \mathbb{R}^n$ eine \mathcal{C}^1 –Funktion, $D \subset \mathbb{R}^n$ offen und konvex.

Sei $\mathbf{x}^* \in D$ mit $\mathbf{f}(\mathbf{x}^*) = 0$. Die Jacobi–Matrix $\mathbf{Jf}(\mathbf{x})$ sei regulär für $\mathbf{x} \in D$, und es gelte eine Lipschitz–Bedingung

$$\|(\mathbf{Jf}(\mathbf{x}))^{-1}(\mathbf{Jf}(\mathbf{y}) - \mathbf{Jf}(\mathbf{x}))\| \leq L\|\mathbf{y} - \mathbf{x}\|$$

für alle $\mathbf{x}, \mathbf{y} \in D$ mit $L > 0$.

Für alle Startwerte $\mathbf{x}^0 \in D$ mit

$$\|\mathbf{x}^0 - \mathbf{x}^*\| < \frac{2}{L} =: r \quad \text{und} \quad K_r(\mathbf{x}^*) \subset D$$

ist dann das Newton–Verfahren wohldefiniert mit $\mathbf{x}^k \in K_r(\mathbf{x}^*)$, $k = 0, 1, 2, \dots$, und die Newton–Iterierten \mathbf{x}^k konvergieren **quadratisch** gegen \mathbf{x}^* :

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \leq \frac{L}{2} \|\mathbf{x}^k - \mathbf{x}^*\|^2$$

Ferner ist \mathbf{x}^* die einzige Nullstelle von $\mathbf{f}(\mathbf{x})$ innerhalb $K_r(\mathbf{x}^*)$.

103

Das gedämpfte Newton–Verfahren

Das Newton–Verfahren konvergiert zwar quadratisch, aber nur **lokal**.

Globale Konvergenz kann durch einen Dämpfungsterm verbessert werden:

für $k = 0, 1, 2, \dots$

$$\mathbf{J}f(\mathbf{x}^k) \cdot \Delta \mathbf{x}^k = -f(\mathbf{x}^k)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda_k \Delta \mathbf{x}^k$$

Gedämpfte Newton–Verfahren mit **Dämpfungsparameter** $\lambda_k \in (0, 1]$.

Frage: Wie wählt man die Dämpfungsfaktoren λ_k ?

Verwende eine **Testfunktion** $T(\mathbf{x})$ mit

$$T(\mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in D$$

$$T(\mathbf{x}) = 0 \Leftrightarrow f(\mathbf{x}) = \mathbf{0}$$

sodass die Folge $T(\mathbf{x}^k)$ monoton fällt.

105

Bemerkung:

Befindet man sich beim gedämpften Newton–Verfahren *in der Nähe* der gesuchten Lösung \mathbf{x}^* , so sollte $\lambda_k = 1$ gewählt werden. Nur das sichert die (lokale) quadratische Konvergenz.

Beispiel:

Bei praktischen Anwendungen verwendet man häufig die Testfunktion

$$T(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\|_2^2$$

Für diesen Fall gilt der folgende Satz.

Satz:

Sei f eine C^1 –Funktion auf der offenen und konvexen Menge $D \subset \mathbb{R}^n$.

Für $\mathbf{x}^k \in D$ mit $\mathbf{f}(\mathbf{x}^k) \neq \mathbf{0}$ gilt dann:

$$\exists \mu_k : \forall \lambda \in (0, \mu_k) : \|\mathbf{f}(\mathbf{x}^k + \lambda \Delta \mathbf{x}^k)\|_2^2 < \|\mathbf{f}(\mathbf{x}^k)\|_2^2$$

106

Dämpfungsstrategie:

$k = 0$: Man wähle $\lambda_0 \in \{1, \frac{1}{2}, \frac{1}{4}, \dots, \lambda_{min}\}$ möglichst groß,
so dass gilt:

$$T(\mathbf{x}^0 + \lambda_0 \Delta \mathbf{x}^0) < T(\mathbf{x}^0)$$

$k > 0$: $\lambda_k := \lambda_{k-1}$

falls $T(\mathbf{x}^k + \lambda_k \Delta \mathbf{x}^k) < T(\mathbf{x}^k)$:

$$\mathbf{x}^{k+1} := \mathbf{x}^k + \lambda_k \Delta \mathbf{x}^k$$

$$\lambda_k := \begin{cases} \lambda_k & : \text{ falls } \lambda_k = 1 \\ 2\lambda_k & : \text{ sonst} \end{cases}$$

107

Dämpfungsstrategie: (Fortsetzung)

sonst:

$$\mu = \max \left\{ \frac{\lambda_k}{2}, \frac{\lambda_k}{4}, \dots, \lambda_{min} \right\}$$

so dass $T(\mathbf{x}^k + \mu \Delta \mathbf{x}^k) < T(\mathbf{x}^k)$

$$\lambda_k := \mu$$

Bemerkung: Die Testfunktion ist **nicht** skalierungsinvariant.

Daher ist der sogenannte **natürliche Monotonietest** sinnvoll:

$$\|\mathbf{Jf}(\mathbf{x}^k)^{-1} \mathbf{f}(\mathbf{x}^{k+1})\| < \|\mathbf{Jf}(\mathbf{x}^k)^{-1} \mathbf{f}(\mathbf{x}^k)\|$$

Es gilt dabei:

$$\|\Delta \mathbf{x}^k\| = \|\mathbf{Jf}(\mathbf{x}^k)^{-1} \mathbf{f}(\mathbf{x}^k)\|$$

108

Kapitel 3: Integralrechnung mehrerer Variablen

3.1 Bereichsintegrale

Gegeben sei eine Funktion $f : D \rightarrow \mathbb{R}$ mit Definitionsbereich $D \subset \mathbb{R}^n$.

Ziel: Berechnung des Volumens unterhalb des Graphen von $f(\mathbf{x})$:

$$V = \int_D f(\mathbf{x}) d\mathbf{x}$$

Erinnerung Analysis II:

Riemann-Integral einer Funktion $f(x)$ über dem Intervall $[a, b]$:

$$I = \int_a^b f(x) dx$$

Das Integral war als Grenzwert von Ober- und Untersumme, für den Fall, dass die Feinheit der Zerlegung gegen Null strebt und ein gemeinsamer Grenzwert existiert, definiert.

109

Konstruktionsprinzip des Integrals mehrerer Veränderlichen analog, **aber** der Definitionsbereich D ist komplizierter.

Zunächst betrachten wir $n = 2$ und einen Definitionsbereich D der Form

$$D = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$$

d.h. D ist ein kompakter Quader (Rechteck).

Weiter sei $f : D \rightarrow \mathbb{R}$ eine beschränkte Funktion.

Definition:

- 1) Man nennt $Z = \{(x_0, x_1, \dots, x_n), (y_0, y_1, \dots, y_m)\}$ eine Zerlegung des Quaders $D = [a_1, b_1] \times [a_2, b_2]$, falls gelten

$$a_1 = x_0 < x_1 < \dots < x_n = b_1$$

$$a_2 = y_0 < y_1 < \dots < y_m = b_2$$

Mit $Z(D)$ wird die **Menge der Zerlegungen** von D bezeichnet.

110

Definition: (Fortsetzung)

2) Die **Feinheit** einer Zerlegung $Z \in \mathbf{Z}(D)$ ist gegeben durch

$$\|Z\| := \max_{i,j} \{|x_{i+1} - x_i|, |y_{j+1} - y_j|\}$$

3) Für eine vorgegebene Zerlegung Z nennt man die Mengen

$$Q_{ij} := [x_i, x_{i+1}] \times [y_j, y_{j+1}]$$

die **Teilquader** der Zerlegung Z . Das **Volumen** des Teilquaders Q_{ij} ist

$$\text{vol}(Q_{ij}) := (x_{i+1} - x_i) \cdot (y_{j+1} - y_j)$$

4) Für beliebige Punkte $x_{ij} \in Q_{ij}$ der jeweiligen Teilquader nennt man

$$R_f(Z) := \sum_{i,j} f(x_{ij}) \cdot \text{vol}(Q_{ij})$$

eine **Riemannsche Summe** zur Zerlegung Z .

111

Definition: (Fortsetzung)

5) Analog zum Integral einer Variablen heißen

$$U_f(Z) := \sum_{i,j} \inf_{x \in Q_{ij}} f(x) \cdot \text{vol}(Q_{ij})$$

$$O_f(Z) := \sum_{i,j} \sup_{x \in Q_{ij}} f(x) \cdot \text{vol}(Q_{ij})$$

die **Riemannsche Unter- bzw. Obersumme** der Funktion $f(x)$ zur Zerlegung Z .

Bemerkung:

Eine Riemannsche Summe zur Zerlegung Z liegt stets zwischen der Unter- und Obersumme dieser Zerlegung, d.h.

$$U_f(Z) \leq R_f(Z) \leq O_f(Z)$$

112

Bemerkung:

Ensteht eine Zerlegung Z_2 aus der Zerlegung Z_1 durch Hinzunahme weiterer Zwischenpunkte x_i und/oder y_j , so gilt

$$U_f(Z_2) \geq U_f(Z_1) \quad O_f(Z_2) \leq O_f(Z_1)$$

Für zwei beliebige Zerlegungen Z_1 und Z_2 gilt stets:

$$U_f(Z_1) \leq O_f(Z_2)$$

Was passiert mit den Unter- und Obersummen im Grenzwert $\|Z\| \rightarrow 0$:

$$U_f := \sup\{U_f(Z) : Z \in \mathbf{Z}(D)\}$$

$$O_f := \inf\{O_f(Z) : Z \in \mathbf{Z}(D)\}$$

Die beiden Werte U_f und O_f existieren, da Unter- und Obersumme geeignete Monotonieeigenschaften besitzen.

113

Definition:

1) **Riemannsches Unter- bzw. Oberintegral** der Funktion $f(x)$ über D :

$$\int_D^* f(x) dx := \sup\{U_f(Z) : Z \in \mathbf{Z}(D)\}$$

$$\int_D^* f(x) dx := \inf\{O_f(Z) : Z \in \mathbf{Z}(D)\}$$

2) Die Funktion $f(x)$ nennt man **Riemann-integrierbar** über D , falls Unter- und Oberintegral übereinstimmen. Das **Riemann-Integral** von $f(x)$ über D ist dann

$$\int_D f(x) dx := \int_D^* f(x) dx = \int_D^* f(x) dx$$

114

Bemerkung:

Wir haben bis jetzt den Fall von **zwei** Variablen betrachtet:

$$f : D \rightarrow \mathbb{R}, \quad D \in \mathbb{R}^2$$

betrachtet.

In höheren Dimensionen $n > 2$ ist die Vorgehensweise analog.

Schreibweise für $n = 2$ und $n = 3$:

$$\int_D f(x, y) dx dy \quad \int_D f(x, y, z) dx dy dz$$

oder auch

$$\iint_D f(x, y) dx dy \quad \iiint_D f(x, y, z) dx dy dz$$