Analysis II

Apl. Prof. Dr. Axel Grünrock

Heinrich-Heine-Universität Düsseldorf

Sommersemester 2020

2.4 Inverse Abbildungen

Fragestellung: Gegeben sei eine differenzierbare Funktion

$$f : \mathbb{R}^n \supset \Omega \to f(\Omega) \subset \mathbb{R}^n.$$

Was läßt sich anhand der Jacobi-Matrix $Df(x)$ über die Umkehrbarkeit von f aussagen?

Rufen wir uns dazu den entsprechenden Satz für reelle Funktionen einer reellen Veränderlichen in Erinnerung:

Erinnerung an Analysis I:

Satz: Ist $f : \mathbb{R} \supset I \to \mathbb{R}$ stetig differenzierbar mit $f'(x) \neq 0$ für alle $x \in I$, so existiert eine Umkehrfunktion

$$f^{-1} : f(I) \to I,$$

die ebenfalls stetig differenzierbar ist. Für ihre Ableitung gilt

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}.$$

Können wir eine ähnliche Aussage für $n \geq 2$ erwarten, und welche Bedingung an $Df(x)$ tritt dann an die Stelle von $f'(x) \neq 0$?

Zwei Beispiele sollen uns den Weg weisen:

Beispiel 1

Es sei

$$f : \mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto Ax + b$$

eine affin-lineare Abbildung (mit eine $n \times n$ - Matrix A und einem Vektor $b \in \mathbb{R}^n$). Dann ist f genau dann umkehrbar, wenn die Matrix A invertierbar ist, das heißt wenn $\det A \neq 0$ gilt. In diesem Fall ist die Umkehrabbildung gegeben durch

$$f^{-1} : \mathbb{R}^n \to \mathbb{R}^n,\ y \mapsto f^{-1}(y) = A^{-1}(y - b).$$

Vermutung

Aus den beiden Spezialfällen - $n = 1$ und die affin-linearen Abbildungen - ergibt sich die Vermutung, dass die Voraussetzung $f'(x) \neq 0$ im Höherdimensionalen durch "$Df(x)$ ist invertierbar" bzw. $\det Df(x) \neq 0$ zu ersetzen ist.

Im Eindimensionalen ebenso wie bei den affin-linearen Abbildungen gibt es *eine* Umkehrabbildung für den gesamten Definitions- bzw. Wertebereich. Diese Situation bezeichnet man als *globale Invertierbarkeit*. Dies ist für nichtlineare Abbildungen im Mehrdimensionalen im Allgemeinen nicht zu erwarten, wie das folgende Beispiel zeigt:

Beispiel 2

Betrachten wir die komplexe exp - Funktion in reeller Schreibweise, das ist

$$f : \mathbb{R}^2 \to \mathbb{R}^2, (x, y)^\top \mapsto \exp(x)(\cos(y), \sin(y))^\top.$$

Für die Jacobi-Matrix von f finden wir

$$Df(x) = \begin{pmatrix} e^x \cos(y) & -e^x \sin(y) \\ e^x \sin(y) & e^x \cos(y) \end{pmatrix}$$

mit $\det Df(x, y) = e^{2x} > 0$ für alle $(x, y)^\top \in \mathbb{R}^2$. Aber: f ist *nicht* injektiv, denn für alle $x, y \in \mathbb{R}$ und $k \in \mathbb{Z}$ ist

$$f(x, y) = f(x, y + 2\pi k).$$

Also existiert keine globale Umkehrabbildung.

Ausweg: Einschränkung des Definitionsbereichs,..

.. um die Bijektivität der Funktion zu erzwingen. Z.B. könnte man die Funktion f aus Beispiel 2 einschränken auf

$$\Omega := \mathbb{R} \times (-\frac{\pi}{2}, \frac{\pi}{2})$$

mit dem Bild $f(\Omega) = \{(u, v) \in \mathbb{R}^2 : u > 0\}$ und der Umkehrfunktion

$$f^{-1}(u, v) = (\ln(\sqrt{u^2 + v^2}), \arctan(\frac{v}{u})).$$

(Nachrechnen!) Dies würde man als eine “lokale Umkehrfunktion” von f bezeichnen. Deren Existenz ist der Gegenstand des folgenden Satzes:

Satz über inverse Funktionen

Es sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^n$ stetig differenzierbar. In $x_0 \in \Omega$ gelte $\det Df(x_0) \neq 0$. Dann existieren offene Umgebungen U von x_0 und V von $f(x_0)$, so dass

$$f\big|_U : U \to V,\, x \mapsto f(x)$$

bijektiv ist. Deren Umkehrfunktion

$$f^{-1} : V \to U$$

ist ebenfalls stetig differenzierbar und für alle $x \in U$ gilt

$$Df^{-1}(f(x)) = Df(x)^{-1}.$$

Bemerkung: Das erste $\quad ^{-1}$ in der letzten Zeile bezeichnet die Umkehrfunktion, das zweite die Inverse der Matrix. Man vergleiche mit der Formel für die Ableitung der Umkehrfunktion auf Seite 3.

Beweisvorbereitungen

Der Beweis dieses Satzes ist deutlich aufwändiger als im Eindimensionalen und bedarf einiger Vorbereitungen. Zunächst ein Lemma über die Invertierbarkeit linearer Abbildungen:

Lemma 1: Es seien $A, B : \mathbb{R}^n \to \mathbb{R}^n$ lineare Abbildungen, davon A bijektiv. Wenn für die Operatornormen gilt, dass

$$\|B - A\| \|A^{-1}\| < 1,$$

dann ist auch B invertierbar.

Beweisvorbereitungen

Der Beweis dieses Satzes ist deutlich aufwändiger als im Eindimensionalen und bedarf einiger Vorbereitungen. Zunächst ein Lemma über die Invertierbarkeit linearer Abbildungen:

Lemma 1: Es seien $A, B : \mathbb{R}^n \to \mathbb{R}^n$ lineare Abbildungen, davon A bijektiv. Wenn für die Operatornormen gilt, dass

$$\|B - A\| \|A^{-1}\| < 1,$$

dann ist auch B invertierbar.

Beweis: Nehmen wir an, dies sei nicht der Fall, so existiert ein $x \neq 0$ im Kern von B. Hierfür können wir $\|x\| = 1$ annehmen, denn der Kern ist ein linearer Raum. Dann folgt

$$1 = |A^{-1}Ax| = |A^{-1}(A - B)x| \leq \|A^{-1}\| \|A - B\| |x| < 1.$$

$\square$

Der Banachsche Fixpunktsatz

Ein wesentliches Hilfsmittel für den Beweis des Satzes über inverse Abbildungen ist der *Banachsche Fixpunktsatz*, auch Kontraktionssatz genannt, im Englischen *Contraction mapping principle*, was die Sache eigentlich besser trifft, denn es handelt sich hierbei nicht um einen gewöhnlichen Satz, sondern vielmehr um ein Beweisprinzip, um einen Satz, mit dem man eine Reihe anderer Sätze beweisen kann. Er lautet:

Es sei (X, d) ein vollständiger metrischer Raum und $f : X \to X$ eine Kontraktion, d.h. eine Lipschitz-stetige Abbildung mit einer Lipschitz-Konstanten $q \in (0, 1)$. Dann besitzt f in X genau einen Fixpunkt.

Wir werden gleich eine leichte Verallgemeinerung dieses Satzes beweisen, die uns später bei den gewöhnlichen Differenzialgleichungen von Nutzen sein wird, nämlich den

Fixpunktsatz von Weissinger

Es sei (X, d) ein vollständiger metrischer Raum, $(\alpha_n)_n$ eine Folge positiver reeller Zahlen, so dass $\sum_{n=1}^{\infty} \alpha_n < \infty$ ist und $f : X \to X$ eine Funktion mit

$$d(f^n(x), f^n(y)) \leq \alpha_n d(x, y)$$

für alle $x, y \in X$ und $n \in \mathbb{N}$. Dann besitzt f genau einen Fixpunkt.

Bemerkung: Ein Fixpunkt von f ist eine Lösung der Gleichung $f(x) = x$. - Mit f^n ist die n-te Iterierte der Abbildung f bezeichnet, also $f^0(x) = x$, $f^1(x) = f(x)$, $f^2(x) = f(f(x))$, $f^3(x) = f(f(f(x)))$ usw..

Beweis: 1. Existenz

Wir betrachten die Folge $(f^n(x_0))_n$ der Iterierten mit einem beliebigen Startpunkt $x_0 \in X$. Ist $m > n$, erhalten wir mit der Dreiecksungleichung hierfür

$$d(f^m(x_0), f^n(x_0) \leq \sum_{k=n}^{m-1} d(f^{k+1}(x_0), f^k(x_0)) \leq \sum_{k=n}^{m-1} \alpha_k d(f(x_0), x_0) \to 0,$$

letzteres, wenn $n, m \to \infty$. Also ist $(f^n(x_0))_n$ eine Cauchy-Folge in (X, d) und aufgrund der Vollständigkeitsvoraussetzung konvergent. Der Grenzwert sei mit x^* bezeichnet. Dann ist

$$f(x^*) = f(\lim_{n\to\infty} f^n(x_0)) = \lim_{n\to\infty} f^{n+1}(x_0) = x^*,$$

denn wegen $d(f(x), f(y)) \leq \alpha_1 d(x, y)$ ist f stetig. Also ist x^* ein Fixpunkt von f.

Beweis: 2. Eindeutigkeit

Es seien x und y Fixpunkte von f. Wir wählen n so groß, dass $\alpha_n < 1$ ist. Dann erhalten wir

$$d(x, y) = d(f^n(x), f^n(y)) \leq \alpha_n d(x, y).$$

Da $\alpha_n < 1$ ist, folgt $d(x, y) = 0$, also $x = y$.

$\square$

Bemerkungen

(1) Mit $q^n = \alpha_n$ geht der Fixpunktsatz von Weissinger in den Banachschen über.

Bemerkungen

(1) Mit $q^n = \alpha_n$ geht der Fixpunktsatz von Weissinger in den Banachschen über.

(2) Fehlerabschätzung:

$$d(x^*, f^n(x_0)) = \lim_{m\to\infty} d(f^m(x_0), f^n(x_0)) \le \sum_{k=n}^{\infty} \alpha_k d(f(x_0), x_0),$$

wobei für $\alpha_n = q^n$ gilt $\sum_{k=n}^{\infty} q^k = \frac{q^n}{1-q}$.

Bemerkungen

(1) Mit $q^n = \alpha_n$ geht der Fixpunktsatz von Weissinger in den Banachschen über.

(2) Fehlerabschätzung:

$$d(x^*, f^n(x_0)) = \lim_{m\to\infty} d(f^m(x_0), f^n(x_0)) \leq \sum_{k=n}^{\infty} \alpha_k d(f(x_0), x_0),$$

wobei für $\alpha_n = q^n$ gilt $\displaystyle\sum_{k=n}^{\infty} q^k = \frac{q^n}{1-q}$.

(3) Die Vollständigkeitsvoraussetzung ist wesentlich für den Existenzteil des Satzes. Ein Beispiel dazu finden Sie im Skript. Für die Eindeutigkeitsaussage wird die Vollständigkeit hingegen nicht benötigt.

Beweisskizze zum Satz über inverse Abbildungen

Den vollständigen Beweis finden Sie im Skript auf den Seiten D59 - D61. Er ist aufgeteilt in drei Beweisschritte:

Schritt 1: Für ein (zunächst beliebiges) $y \in \mathbb{R}^n$ definiert man mit $A := Df(x_0)$ die Abbildung

$$\varphi_y : \Omega \to \mathbb{R}^n,\ x \mapsto \varphi_y(x) := x + A^{-1}(y - f(x)).$$

Dann ist $y = f(x)$ genau dann, wenn x ein Fixpunkt der Abbildung φ_y ist. Nun schränkt man den Definitionsbereich von φ_y ein auf eine offene Kugel $B_\delta(x_0) = U$ und rechnet mit Hilfe der Mittelwertungleichung nach, dass für ein hinreichend kleines $\delta > 0$ die Abbildung $\varphi_y : U \to \mathbb{R}^n$ kontrahierend ist, also höchstens einen Fixpunkt besitzt. Daraus folgt, dass

Beweisskizze zum Satz über inverse Abbildungen

$$f\Big|_U : U \to \mathbb{R}^n$$

injektiv ist, und mit Lemma 1 erhält man, dass $Df(x)$ für alle $x \in U$ invertierbar ist (was man im 3. Schritt benötigt).

Schritt 2: Man setzt $V = f(U)$ ($U = B_\delta(x_0)$ wie in Schritt 1. bestimmt), so dass

$$f\Big|_U : U \to V$$

bijektiv ist. Eine nichttriviale Aufgabe ist es dann, die Offenheit von V zu zeigen. Dies geschieht mit Hilfe des Banachschen Fixpunktsatzes.

Beweisskizze zum Satz über inverse Abbildungen

Schritt 3: Nun existiert also eine (lokale) Inverse $f^{-1} : V \to U$, deren stetige Differenzierbarkeit noch zu zeigen ist (, wobei die Ableitungsformel gleich mit abfallen wird). Dazu wählt man $y, y + k \in V$ mit Urbildern $x, x + h \in U$, setzt

$$\psi(k) := f^{-1}(y + k) - f^{-1}(y) - Df(x)^{-1} k$$

und zeigt unter Berücksichtigung zuvor getroffener Wahlen, dass $\lim_{k \to 0} \frac{\psi(k)}{|k|} = 0$.