2.4 Inverse Abbildungen

Fragestellung: Gegeben sei eine differenzierbare Funktion

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

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 Veränderlicher in Erinnerung:

Satz (Ana I): Ist $f : \mathbb{R} \supset I \longrightarrow \mathbb{R}$ stetig d'bar mit $f'(x) \neq 0 \; \forall x \in I$, so existiert eine Umkehrfunktion

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

die ebenfalls stetig d'bar ist. Für ihre Ableitung gilt $\quad (f^{-1})'(f(x)) = \dfrac{1}{f'(x)}$.

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

Bsp. 1: $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $x \longmapsto Ax + b$ ist invertierbar genau dann, wenn die $n \times n$-Matrix A invertierbar ist, d.h. wenn $\det A \neq 0$ gilt. In diesem Fall ist die Umkehrabbildung gegeben durch

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

Aus den beiden Spezialfällen – $n=1$ und die affin-linearen Abbildungen – ergibt sich die

Vermutung: Die Voraussetzung $f'(x) \neq 0$ ist im höherdimensionalen durch "$Df(x)$ ist invertierbar" bzw. $\det Df(x) \neq 0$ zu ersetzen.

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

Bsp. 2: $f: \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y)^T \mapsto \exp(x)\,(\cos(y), \sin(y))^T$
(Das ist die komplexe e-Fkt, in reeller Schreibweise!)
Für die Jacobi-Matrix finden wir

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

mit $\det Df(x,y) = e^{2x} > 0 \quad \forall (x,y) \in \mathbb{R}^2$.

Aber: $f(x,y) = f(x, y+2\pi k)$, $k \in \mathbb{Z}$. f ist also nicht injektiv, also kann keine globale Umkehrabbildung existieren.

Ausweg: Man schränkt den Definitionsbereich der Funktion soweit ein, dass sie bijektiv wird. z.B. könnte man die Funktion f aus Bsp. 2 einschränken auf

$$\Omega := \mathbb{R} \times \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right) \left(\hookrightarrow \mathbb{R}^2\right)$$

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

und der Umkehrfunktion

$$f^{-1}(u,v) = \left(\ln\left(\sqrt{u^2+v^2}\right), \arctan\left(\tfrac{v}{u}\right)\right)$$

Dies würde man etwa als eine "lokale Umkehrfunktion" bezeichnen. Deren Existenz ist Gegenstand des folgenden Theorems:

<u>Satz</u> über inverse Funktionen: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f: \Omega \to \mathbb{R}^n$ stetig diff'bar. 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

$$f|_U : U \longrightarrow V, \quad x \longmapsto f(x)$$

bijektiv ist. Deren Umkehrfunktion

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

ist ebenfalls stetig d'bar mit

$$Df^{-1}(f(x)) = (Df(x))^{-1} \quad \text{für alle } x \in U.$$

(vgl.: Formel für die Ableitung der Umkehrfunktion!)

Der Beweis dieses Satzes ist deutlich schwieriger als im eindimensionalen Fall und benötigt einige Vorbereitungen. Zunächst ein Lemma über die Invertierbarkeit linearer Abbildungen:

Lemma 1: Gegeben seien zwei lineare Abbildungen $A, B : \mathbb{R}^n \to \mathbb{R}^n$, A sei zusätzlich bijektiv. Wenn für die Operatornormen

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

gilt, ist auch B invertierbar.

Bew.: Ann. nicht, dann ist $\operatorname{Ker}(B) \neq \{0\}$. Für $x \in \operatorname{Ker}(B)$ mit $|x| = 1$ ist dann

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

$\square$

Ein wesentliches Hilfsmittel für den Beweis des Satzes über inverse Abbildungen ist der:

Banachscher Fixpunktsatz (Kontraktionssatz, CMP): Es sei (X, d) ein vollständiger metrischer Raum und $f : X \to X$ eine Kontraktion, d.h. es gebe ein $q \in (0,1)$, so dass für alle $x, y \in X$ gilt

$$d(f(x), f(y)) \leq q \, d(x,y).$$

Dann besitzt f in X genau einen Fixpunkt.

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

<u>Fixpunktsatz von Weissinger:</u> 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

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

Dann besitzt f genau einen Fixpunkt.

Bew.: 1. Existenz: Wir betrachten die Folge $(f^n(x_0))_n$ der Iterierten

$$f(x_0),\ f^2(x_0) = f(f(x_0)),\ f^3(x_0) = f(f(f(x_0))), \ldots$$

mit einem beliebigen Startpunkt $x_0 \in X$. Hierfür erhalten wir mit der Δ's-Ungleichung $(m > n)$

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

$$\text{Vor.} \longrightarrow \leq \sum_{k=n}^{m-1} \alpha_k \, d(f(x_0), x_0) \longrightarrow 0 \ (m, n \to \infty).$$

Also ist $(f^n(x_0))_n$ eine Cauchy-Folge in (X, d) und aufgrund der Vollständigkeitsvoraussetzung konvergent. Sei $x^* := \lim_{n\to\infty} f^n(x_0)$. Dann gilt

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

denn wg. $d(f(x), f(y)) \le \alpha_1 d(x,y)$ ist f stetig.

Also ist x^* ein Fixpunkt von f.

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

$$d(x,y) = d(f^n(x), f^n(y)) \le \underbrace{\alpha_n}_{<1\,!} d(x,y)$$

$$\Rightarrow d(x,y) = 0 \Rightarrow x = y. \qquad \square$$

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

(2) Fehlerabschätzung: $d(x^*, f^n(x_0)) = d(\lim_{m\to\infty} f^m(x_0), f^n(x_0))$

$$= \lim_{m\to\infty} d(f^m(x_0), f^n(x_0)) \underset{\text{s.o.}}{\le} \left(\sum_{k=n}^{\infty} \alpha_k\right) d(f(x_0), x_0)$$

Im Fall $\alpha_k = q^k$: $\sum_{k=n}^{\infty} q^k = q^n \cdot \sum_{k=0}^{\infty} q^k = \frac{q^n}{1-q}$.

(3) Die Vollständigkeitsvoraussetzung ist wesentlich:

Bsp. $(X, d) = ([0,\infty)\cap\mathbb{Q}, |\cdot|)$,

$f(x) = \frac{1}{2+x}$. Dann ist $f(x) = x \Leftrightarrow x^2+2x-1 = 0 \Leftrightarrow x = \sqrt{2}-1 \notin \mathbb{Q}\,!$

Aber $|f(x) - f(y)| = \frac{|x-y|}{(2+x)(2+y)} \le \frac{|x-y|}{4}$. Auch führt die Anwendung von f nicht aus dem Bereich der rationalen Zahlen hinaus.

Beweis des Satzes über inverse Abbildungen:

① Zu $y \in \mathbb{R}^n$ (zunächst beliebig) definieren wir eine Abbildung

$$\varphi_y : \Omega \to \mathbb{R}^n, \quad x \mapsto \varphi_y(x) := x + A^{-1}(y - f(x)), \quad A = Df(x_0).$$

Dann gilt $y \in f(x) \iff x = \varphi_y(x)$.

(D.h. wir suchen eine Lösung x von $f(x) = y$ als Fixpunkt von φ_y. Um den Banachschen Fixpunktsatz anwenden zu können, schränken wir den Definitionsbereich von φ_y so-weit ein, dass φ_y eine Kontraktion wird.)

$$D\varphi_y(x) = I_n - A^{-1}Df(x) = A^{-1}(Df(x_0) - Df(x))$$

und daher

$$\| D\varphi_y(x) \| \leq \| A^{-1} \| \; \| Df(x_0) - Df(x) \|$$

Da f stetig d'bar, existiert ein $\delta > 0$, so daß

$$\| Df(x_0) - Df(x) \| \leq \frac{1}{2\| A^{-1} \|} \quad \forall x \in B_\delta(x_0) =: U$$

Für solche x_1, x_2 erhalten wir

$$| \varphi_y(x_1) - \varphi_y(x_2) | \leq \sup \{ \| D\varphi_y(\xi) \| : \xi \in B_\delta(x_0) \} \, |x_1 - x_2| \leq \tfrac{1}{2} |x_1 - x_2|$$

D.h.: $\varphi_y : U \to \mathbb{R}^n$ ist kontrahierend, besitzt also höchstens einen Fixpunkt. Es folgt, dass

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

injektiv ist. Ferner ergibt Lemma 1, daß $Df(x)$ für alle $x \in U$ invertierbar ist. ($\rightsquigarrow$ step ③)

② Zu $U = B_\delta(x_0)$ wie in ① bestimmt setzen wir (D60)

$V := f(U)$, so daß

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

bijektiv wird. Wir zeigen, dass V offen ist.

Dazu sei $y_1 \in V$ vorgegeben mit Urbild $x_1 \in U$.

Wir wählen $B = B_r(x_1)$ so, daß $\overline{B} \subset U$. Nun sei

$y \in B_\varepsilon(y_1)$, wobei $\varepsilon > 0$ noch zu wählen ist. Dann

gilt für dieses y und für jedes $x \in \overline{B}$:

$$|\varphi_y(x) - x_1| \leq |\varphi_y(x) - \varphi_y(x_1)| + |\varphi_y(x_1) - x_1|$$

$$\leq \sup\{\|D\varphi_y(\xi)\| : \xi \in U\}\,|x - x_1| + |A^{-1}(\underbrace{f(x_1)}_{y_1} - y)|$$

$$\leq \frac{r}{2} + \|A^{-1}\| \cdot \varepsilon \,.$$

Die Wahl $\varepsilon = \dfrac{r}{\|A^{-1}\| \cdot 2}$ ergibt also $|\varphi_y(x) - x_1| \leq r$,

d.h. $\varphi_y(x) \in \overline{B}$ $\forall x \in \overline{B}$. Wie in ① gesehen, ist

φ_y kontrahierend. Damit sind die Voraussetzungen des Fixpunktsatzes erfüllt, also existiert

ein $x \in \overline{B}$ mit $\varphi_y(x) = x$, d.h. mit $y = f(x)$.

Das heißt $y \in f(\overline{B}) \subset f(U) = V$. Dies gilt für jedes

$y \in B_\varepsilon(y_1)$, also $B_\varepsilon(y_1) \subset V$. Da $y_1 \in V$ beliebig vorgegeben war, ist damit die Offenheit von V gezeigt.

③ Bisher gezeigt: Es existieren offene Umgebungen (DG1)

U von x_0 und V von $f(x_0)$, so dass

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

bijektiv ist. $f^{-1}: V \to U$ sei die Umkehrabbildung.

Dann ist noch zu zeigen, dass f^{-1} differenzierbar

ist mit $\quad Df^{-1}(f(x)) = (Df(x))^{-1} \quad \forall x \in U$

(Dann ist, da Df stetig ist, auch die Stetigkeit von

Df^{-1} gezeigt.)

Zu $y, y+k \in V$ wählen wir $x, x+h \in U$ mit $f(x)=y$

und $f(x+h) = y+k$. Dann ist

Existiert, siehe Ende ①

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

$$= x+h-x - Df(x)^{-1} \cdot k = (Df(x))^{-1}(Df(x) \cdot h - k)$$

$$= (Df(x))^{-1}(Df(x) \cdot h - (f(x+h) - f(x)))$$

$$\Rightarrow \frac{|\varphi(k)|}{|k|} \leq \|(Df(x))^{-1}\| \cdot \frac{|h|}{|k|} \cdot \frac{1}{|h|} \, |Df(x) \cdot h - (f(x+h) - f(x))|,$$

Nun ist $\varphi_y(x+h) - \varphi_y(x) = h + A^{-1}(f(x) - f(x+h)) = h - A^{-1}k$

und daher $|h - A^{-1}k| = |\varphi_y(x+h) - \varphi_y(x)| \underset{\text{s.o.}}{\leq} \frac{1}{2}|h|$.

$\Rightarrow |h| \leq |h - A^{-1}k| + |A^{-1}k| \leq \frac{|h|}{2} + \|A^{-1}\||k| \Rightarrow |h| \leq 2\|A^{-1}\||k|$

Also: $|\varphi(k)|/|k| \leq \|(Df(x))^{-1}\| \cdot 2\|A^{-1}\| \cdot \frac{1}{|h|} |f(x+h) - f(x) - Df(x) \cdot h|$

und damit $\lim\limits_{k \to 0} \frac{|\varphi(k)|}{|k|} = 0$, denn für $k \to 0$ geht

auch $h \to 0$. $\square$