

Vorlesungsmitschrift

DIFF II

Prof. Dr. Dorothea Bahns

Henry Ruben Fischer

Auf dem Stand vom 27. Mai 2020

Disclaimer

Nicht von Professor Bahns durchgesehene Mitschrift, keine Garantie auf Richtigkeit ihrerseits.

Inhaltsverzeichnis

1	Metrische Räume	4
2	Normierte Vektorräume	36
3	Differenzierbarkeit in \mathbb{R}^n	53

Kapitel 1

Metrische Räume

Vorlesung 1

Mo 20.04. 10:15

Ziel. Konvergenz, Stetigkeit ... sollten in einem allgemeineren Rahme konzeptualisiert werden.

Erinnerung (DIFF I). Eine Folge $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ konvergiert gegen den Grenzwert a

$$\iff \forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \text{ s. d. } |a_n - a| < \varepsilon \quad \forall n \geq N$$

$(a - \varepsilon, a + \varepsilon)$ wird auch ε -Umgebung von a in \mathbb{R} genannt. Somit lautet die obige Definition in Worten: In jeder noch so kleinen ε -Umgebung von a befinden sich alle bis auf endlich viele Folgenglieder.

Man benötigt für die Formulierung der Definition also lediglich einen Begriff von „(kleine) Umgebung“. Diesen Begriff möchten wir nun verallgemeinern.

Definition 1.1. Sei X eine Menge. Ein System \mathcal{T} von Teilmengen von X heißt Topologie auf X falls gilt:

- a) $\emptyset, X \in \mathcal{T}$.
- b) Sind U und $V \in \mathcal{T}$, so gilt $U \cap V \in \mathcal{T}$.
- c) Ist I eine Indexmenge und $U_i \in \mathcal{T}$ für alle $i \in I$, so gilt auch $\bigcup_{i \in I} U_i \in \mathcal{T}$.

Notation. Ein topologischer Raum ist ein Tupel (X, \mathcal{T}) , wobei X Menge ist und \mathcal{T} eine Topologie auf X .

Eine Teilmenge $U \subset X$ heißt *offen*, falls gilt $U \in \mathcal{T}$. Eine Teilmenge $A \subset X$ heißt *abgeschlossen* falls ihr Komplement $X \setminus A$ offen ist.

Beispiele 1.2. i) $X =$ beliebige Menge. $\mathcal{T} = \{ \emptyset, X \}$.

Beweis. 1.1.a) klar

$$1.1.b) \quad \emptyset \cap X = \emptyset \in \mathcal{T}, \quad X \cap X = X \in \mathcal{T}, \quad \emptyset \cap \emptyset = \emptyset \in \mathcal{T}$$

$$1.1.c) \quad \bigcup_{i \in I} U_i = \begin{cases} X & \text{falls eins der } U_i = X \text{ ist} \\ \emptyset & \text{falls nicht} \end{cases} \quad \square$$

„Klumpentopologie“

ii) $X = \mathbb{R}$

\mathcal{T} = alle Teilmengen $U \subset \mathbb{R}$ mit der Eigenschaft:

$$\forall x \in U \quad \exists \varepsilon > 0 \text{ s. d. } (x - \varepsilon, x + \varepsilon) \subset U$$

Beweis von 1.1.a), 1.1.b) und 1.1.c) als HA (etwas allgemeiner). Hier stellen wir fest, dass insbesondere die offenen Intervalle (a, b) in diesem Sinne offen (also $\in \mathcal{T}$) sind, halb-abgeschlossene und abgeschlossene dagegen nicht.

Beweis. **1. Beh** Zu $x \in [a, b]$ wähle $\varepsilon = \min(\{|x - a|, |x - b|\})$

2. Beh Zu $x = a \in [a, b]$ kann man kein $\varepsilon > 0$ finden s. d. $(a - \varepsilon, a + \varepsilon) \subset [a, b]$, denn $a - \varepsilon/2 \in (a - \varepsilon, a + \varepsilon)$ aber $a - \varepsilon/2 < a$, also $\notin [a, b]$. \square

Abgeschlossene Intervalle sind in diesem Sinn abgeschlossen, denn $\mathbb{R} \setminus [a, b]$ ist nach Definition von \mathcal{T} und Eigenschaft 1.1.c) offen.

Diese Topologie heißt Standard-Topologie auf \mathbb{R} . Wird nichts anderes gesagt, sehen wir \mathbb{R} als mit der Standard-Topologie versehen an.

Definition 1.3. Sei (X, \mathcal{T}) topologischer Raum. Sei $x \in X$. Eine Teilmenge $V \subset X$ heißt *Umgebung von x* , falls es eine offenen Menge $U \subset X$ gibt mit $x \in U$ und $U \subset V$.

Beispiele. i) $V = (a, b)$ ist eine Umgebung für jedes $x \in (a, b)$, aber *nicht* für $x = a$.



ii) $(a - \varepsilon, a + \varepsilon)$, $\varepsilon > 0$, ist eine Umgebung von x .

Lemma 1.4. Eine Teilmenge $V \subset X$ eines topologischen Raumes (X, \mathcal{T}) ist offen gdw für alle $x \in V$ gilt: V ist Umgebung von x .

Beweis. „ \implies “ Ist V offen, so erfüllt $U = V$ für jedes x die Bedingung $x \in U$ und $U \subset V \implies V$ ist Umgebung.

„ \impliedby “ Zu $x \in V$ wähle U_x s.d. $x \in U_x$, $U \subset V$. Dann gilt $V = \bigcup_{x \in U} U_x$ und das ist offen (nach 1.1.c)). \square

Definition 1.5 (Konvergenz in topologischen Räumen). Sei (X, \mathcal{T}) topologischer Raum. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in X . Dann ist $(x_n)_n$ *konvergent mit Grenzwert x* , $x_n \rightarrow x$ in (X, \mathcal{T}) , falls es in jeder Umgebung V von x ein $N \in \mathbb{N}$ gibt, s.d. $x_n \in V \quad \forall n \geq N$.

Beispiele. i) In der Klumpentopologie konvergieren alle Folgen gegen jedes $x \in X$.

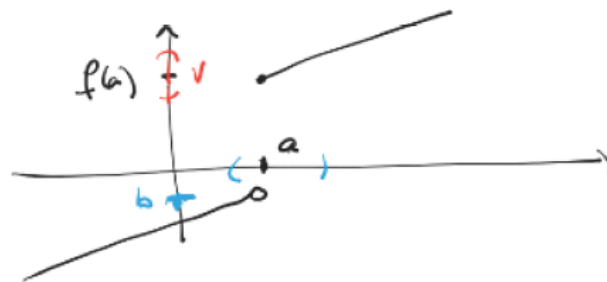
ii) Mit unseren obigen Überlegungen folgern wir, dass Konvergenz in \mathbb{R} im Sinn von Definition 1.5 mit Konvergenz, wie wir sie in der DIFF I

Lemma 1.6. Sei (X, \mathcal{T}) topologischer Raum. Ist (X, \mathcal{T}) ein *Hausdorff-Raum*, gibt es also zu je zwei Punkten $x, y \in X$ mit $x \neq y$ Umgebungen U von x und V von y mit $U \cap V = \emptyset$, so ist der Grenzwert einer konvergenten Folge eindeutig.

Beweis. Seien x und y Grenzwert einer Folge $(x_n)_n$. Angenommen $x \neq y$, so wähle U Umgebung von x , V Umgebung von y mit $U \cap V = \emptyset$. Dann gibt es (wegen der Konvergenz) $N \in \mathbb{N}$ s.d. $x_n \in U \quad \forall n \geq N$ und $M \in \mathbb{N}$ s.d. $x_n \in V \quad \forall n \geq M$. Widerspruch zu $U \cap V = \emptyset$. \square

Definition 1.7. Seien (X, \mathcal{T}) und $(Y, \tilde{\mathcal{T}})$ topologische Räume. Sei $f: X \rightarrow Y$ eine Abbildung. Dann heißt f *stetig in $a \in X$* , falls es zu jeder Umgebung V von $f(a) \in Y$ eine Umgebung U von a gibt, s.d. $f(U) \subset V$. f heißt *stetig* (auf X), falls f stetig in allen $a \in X$ ist.

Bemerkung. Wir werden später sehen, dass diese Definition für $f: \mathbb{R} \rightarrow \mathbb{R}$ mit unserer Definition aus der DIFF I übereinstimmt (ε - δ -Kriterium).



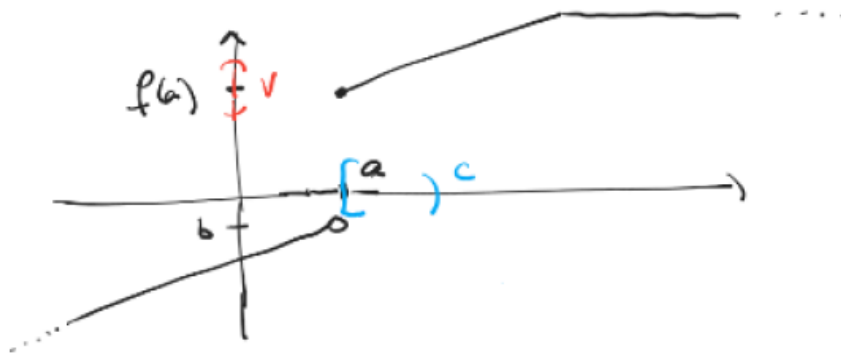
Für jede Umgebung U von a gilt: $f(U)$ enthält auch Punkte $< b$, also außerhalb V

Satz 1.8. Sei $f: X \rightarrow Y$ Abbildung zwischen topologischen Räumen. Dann ist f stetig auf X gdw für jede offene Teilmenge $V \subset Y$ das Urbild $f^{-1}(V)$, also $\{x \in X \mid f(x) \in V\}$ offen in X ist.

Beweis. „ \Rightarrow “ Sei f stetig vorausgesetzt. Sei V offen in Y . Ist das Urbild $f^{-1}(V)$ leer, sind wir fertig.

Sei also $a \in f^{-1}(V)$. Dann gibt es nach Voraussetzung eine Umgebung U von a s. d. $f(U) \subset V$. Also gilt $U \subset f^{-1}(V)$. Somit besitzt also jeder Punkt $a \in f^{-1}(V)$ eine Umgebung U mit $U \subset f^{-1}(V)$ und somit ist $f^{-1}(V)$ selbst Umgebung jedes seiner Elemente $\xrightarrow{1.4} f^{-1}(V)$ ist offen.

„ \Leftarrow “ Sei $a \in X$ beliebig. Sei V eine Umgebung von $f(a)$. Dann gibt es \tilde{V} offen mit $f(a) \in \tilde{V}$ und $\tilde{V} \subset V$. Nach Voraussetzung ist das Urbild $U := f^{-1}(\tilde{V})$ offen. U enthält a , ist also Umgebung von a und es gilt $f(U) = \tilde{V} \subset V \Rightarrow f$ ist stetig in a . \square



$f^{-1}(V) = [a, c)$ ist nicht offen in \mathbb{R}

Bemerkung. Äquivalent: f ist genau dann stetig, wenn das Bild jeder abgeschlossen Menge abgeschlossen ist.

Vorsicht:

Es ist immer Offenheit in X (bzw. Y) gemeint!

Zur Veranschaulichung:

Betrachtet man im Beispiel oben als Definitionsbereich $X = [a, \infty)$, so ist die Funktion stetig! Dies ist konsistent, da $[a, c)$ in $X = [a, \infty)$ versehen mit der Standard-Topologie tatsächlich offen ist:

Definition / Satz 1.9. Sei (X, \mathcal{T}) topologischer Raum. Sei $\tilde{X} \subset X$ eine Teilmenge. Dann induziert \mathcal{T} auf \tilde{X} eine Topologie, die sogenannte *Teilraum-Topologie* vermöge

$$T_{\tilde{X}} := \left\{ U \cap \tilde{X} \mid U \in \mathcal{T} \right\}.$$

Den (einfachen) Beweis, dass dies in der Tat eine Topologie definiert, lassen wir weg.

In unserem Beispiel ist $X = \mathbb{R}$, $\tilde{X} = [a, \infty)$ und da $(a - \varepsilon, c)$ offen in \mathbb{R} ist ($\varepsilon > 0$), ist nach Definitionsbereich $[a, c) = (a - \varepsilon, c) \cap [a, \infty)$ offen in $[a, \infty)$.

Dies ist der tiefere Grund, weshalb man bei Funktionen den Raum, in dem sie ihre Werte annehmen (im Beispiel oben $Y = \mathbb{R}$) angeben sollte, nicht ihr Bild.

Denn in $Y = (-\infty, b) \cup [f(a), \infty)$ wäre das Bild von $[a - \varepsilon, c]$ $\forall \varepsilon > 0$ in der Tat abgeschlossen, denn sein Komplement

$$Y \setminus ([b - \delta, b) \cup [f(a), f(c))) = -(-\infty, b - \delta) \cup (f(c), \infty)$$

wäre offen.

Dagegen ist

$$\mathbb{R} \setminus ([b - \delta, b) \cup [f(a), f(c))) = -(-\infty, b - \delta) \cup [b, f(a)) \cup (f(c), \infty)$$

für kein $\delta > 0$ offen.

Definition / Satz 1.10. Seien (X, \mathcal{T}_X) und (Y, \mathcal{T}_Y) topologische Räume. Betrachte das *kartesische Produkt* $X \times Y = \{ (x, y) \mid x \in X, y \in Y \}$. Dann nennt man das System

$$T := \left\{ U \subset X \times Y \mid U = \text{beliebige Vereinigung von Mengen der Form } V \times W, V \in \mathcal{T}_X, W \in \mathcal{T}_Y \right\}$$

Produkttopologie. Und dies definiert in der Tat eine Topologie auf $X \times Y$.

Beweis. 1.1.a) klar

1.1.b)

$$\begin{aligned} U &= \bigcup_{\alpha} U_{\alpha} \times W_{\alpha} \\ V &= \bigcup_{\beta} \tilde{V}_{\beta} \times \tilde{W}_{\beta} \\ U \cap V &= \bigcup_{\alpha, \beta} \underbrace{(U_{\alpha} \cap \tilde{V}_{\beta})}_{\text{offen in } X} \times \underbrace{(W_{\alpha} \cap \tilde{W}_{\beta})}_{\text{offen in } Y}. \end{aligned}$$

1.1.c)

$$\bigcup_{\rho} \left(\bigcup_{\alpha} V_{\alpha}^{(\rho)} \times W_{\alpha}^{(\rho)} \right) = \bigcup_{\rho, \alpha} V_{\alpha}^{(\rho)} \times W_{\alpha}^{(\rho)}. \quad \square$$

Wir kommen nun zu einer wichtigen Beispiel-Klasse für Topologien:

Definition 1.11. Sei X eine Menge. Eine *Metrik* auf X ist eine Abbildung

$$d: X \times X \rightarrow \mathbb{R}$$

mit den Eigenschaften

- a) $d(x, y) = 0 \iff x = y$ „ d ist nicht ausgeartet.“
- b) $d(x, y) = d(y, x) \quad \forall x, y \in X$ „ d ist symmetrisch.“
- c) $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$ „Es gilt die Dreiecksungleichung.“

Ein *metrischer Raum* ist ein Tupel (X, d) , wobei X eine Menge ist und d eine Metrik auf X . Mist schreibt man nur X , weil Missverständnisse ausgeschlossen sind.

Bemerkung. Aus den Axiomen folgt auch

$$d(x, y) \geq 0 \quad \forall x, y \in X,$$

denn

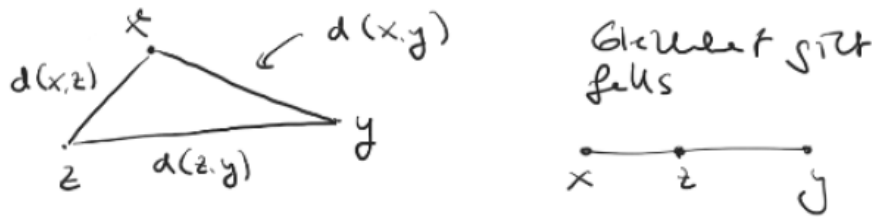
$$0 = \underset{\substack{\uparrow \\ 1.11.a)}}{d(x, x)} \leq \underset{\substack{\uparrow \\ \Delta\text{-Ungl.}}}{d(x, y)} + \underset{\substack{\uparrow \\ \text{Symm.}}}{d(y, x)} = 2d(x, y).$$

Beispiele. i) $\mathbb{R}, d(x, y) = |x - y|$.

ii) X Menge, $d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$, „triviale“ oder „diskrete Metrik“.

iii) (aus AGLA I) \mathbb{R}^n , $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$, „Euklidische Metrik“.

Eine Metrik misst den *Abstand* zwischen zwei Punkten. Im zweiten Beispiel sind alle verschiedenen Punkte gleich weit von einander entfernt. Für $n = 1$ stimmt iii) mit i) überein. Mit iii) wird auch der Name der Dreiecksungleichung klar:



Definition 1.12. Sei (X, d) ein metrischer Raum. Seien $x \in X$, $\varepsilon > 0$. Dann nennt man

$$B_\varepsilon(x) := \{ y \in X \mid d(x, y) < \varepsilon \}$$

den (offenen) ε -Ball um x .

Beispiele. i) $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$.

$$\text{ii) } B_\varepsilon(x) = \begin{cases} x & \varepsilon \leq 1 \\ X & \varepsilon > 1 \end{cases}$$

$$\text{iii) } B_\varepsilon(x) = \text{Diagram of a circle with a point inside and blue diagonal lines inside the circle.}$$

Satz 1.13. Sei (X, d) ein metrischer Raum. Dann wird durch

$$\mathcal{T}_d := \{ U \subset X \mid \forall x \in U \ \exists \varepsilon > 0 \text{ s. d. } B_\varepsilon(x) \subset U \}$$

eine Topologie definiert.

Beweis. Als Hausaufgabe. □

Bemerkungen 1.14. i) 1.2.ii) ist ein Spezialfall dieser Aussage

ii) Die „offenen“ ε -Bälle sind tatsächlich offen: Zu $y \in B_\varepsilon(x)$ wähl $\tilde{\varepsilon} := \varepsilon - d(x, y) > 0$.



Dann ist $B_{\tilde{\varepsilon}}(y) = \{ z \mid d(y, z) < \tilde{\varepsilon} \} \subset B_{\varepsilon}(x)$. Denn für alle $z \in B_{\tilde{\varepsilon}}(y)$ ist

$$\begin{aligned} d(x, z) &\leq d(x, y) + d(y, z) < d(x, y) + \tilde{\varepsilon} \\ &= d(x, y) + \varepsilon - d(x, y) = \varepsilon \end{aligned}$$

- iii) Bezüglich der diskreten Metrik ist *jede* Teilmenge offen.
- iv) Die Klumpentopologie wird nicht von einer Metrik erzeugt (wenn X mehr als 1 Element enthält).

Beweis. Seien $x, y \in X, x \neq y$. Angenommen \exists Metrik d .

$$\implies d(x, y) \neq 0 \implies d(x, y) = c > 0$$

$$\implies B_c(x) \text{ ist offen.}$$

$$\implies_{\text{VOR}} B_c(x) = \emptyset \text{ oder } = X$$

$$\implies B_c(x) = X$$

□

⚡, da $y \notin B_c(x)$.

- v) Ein metrischer Raum ist hausdorffsch. \rightarrow HA.

Wir formulieren nun Konvergenz und Stetigkeit für metrische Räume:

Bemerkungen 1.15. Sei (X, d) metrischer Raum.

- i) [Definition 1.3] $V \subset X$ heißt Umgebung von $x \in X$, falls es $\varepsilon > 0$ gibt s.d. $B_{\varepsilon}(x) \subset V$.
- ii) [Definition 1.5] $(x_n)_n$ konvergiert mit Grenzwert x , falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt s.d. $x_n \in B_{\varepsilon}(x) \quad \forall n \geq N$.

- iii) [Definition 1.7] Sei (Y, \tilde{d}) weiterer metrischer Raum, $f: X \rightarrow Y$ eine Abbildung. Dann ist f stetig a gdw :

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ s. d. } f(B_\delta(a)) \subset B_\varepsilon(f(a)).$$

Bemerkungen. i) 1.15.iii) ist das ε - δ -Kriterium.

- ii) Die Einschränkung auf ε -Bälle in 1.15.ii) und 1.15.iii) (statt allgemeiner Umgebungen) ist keine echte Einschränkung: Gilt etwas für all Umgebungen, so speziell auch für ε -Bälle.

Und gilt eine Inklusion für alle ε -Bälle (etwa $x_n \in B_\varepsilon(x) \quad \forall n \geq N(\varepsilon)$), so auch für beliebige Umgebungen U von x , da es immer einen ε -Ball $B_\varepsilon(x)$ gibt, der ganz in U enthalten ist.

Beispiele 1.16. i) \mathbb{R}^m mit der Euklidischen Metrik. $(x_n)_{n \geq 1}$ Folge in \mathbb{R}^m , also $n \mapsto x_n = (x_n^{(1)}, \dots, x_n^{(m)}) \in \mathbb{R}^m$.

- ii) $x_n = \left(\frac{1}{n} \cos(n), \frac{1}{n} \sin(n), a, \dots, a\right)$

Behauptung. $x_n \rightarrow (0, 0, a, \dots, a) =: x$.

Beweis. Sei $\varepsilon > 0$. Es gilt

$$\begin{aligned} d(x_n, x)^2 &= \sum_{i=1}^m (x_n^{(i)} - x^{(i)})^2 \\ &= \left(\frac{1}{n} \cos(n) - 0\right)^2 + \left(\frac{1}{n} \sin(n) - 0\right)^2 + (a - a)^2 + \dots + (a - a)^2 \\ &= \frac{1}{n^2} (\cos(n)^2 + \sin(n)^2) = \frac{1}{n^2} \\ \implies d(x_n, x) &= \frac{1}{n} \\ \implies d(x_n, x) < \varepsilon \quad \forall n \geq N \text{ mit } N > \frac{1}{\varepsilon} \\ \implies x_n &\in B_\varepsilon(x) \quad \forall n \geq N. \end{aligned}$$

□

- iii) $X = C([a, b])$, $d(f, g) := \|f - g\|_\infty$ mit $\|f - g\|_\infty = \sup_{x \in [a, b]} |f(x) - g(x)|$.

1. Beh d ist eine Metrik auf X .

Beweis. 1.11.a):

$$\begin{aligned} \sup_{x \in [a, b]} |f(x) - g(x)| &= 0 \\ \iff |f(x) - g(x)| &= 0 \quad \forall x \\ \iff f(x) &= g(x) \quad \forall x. \end{aligned}$$

1.11.b):

$$|f(x) - g(x)| = |g(x) - f(x)| \quad \forall x \\ \implies d(f, g) = d(g, f).$$

1.11.c):

$$|f(x) - g(x)| = |f(x) - h(x) + h(x) - g(x)| \\ \leq |f(x) - h(x)| + |h(x) - g(x)| \\ \quad \quad \quad \uparrow \\ \quad \quad \quad \triangle\text{-Ungl. für } |\cdot| \text{ auf } \mathbb{R} \\ \implies \triangle\text{-Ungl. für } d. \quad \square$$

2. Beh $(f_n)_n \subset C([0, 1])$, $f_n(x) = x^n$, konvergiert nicht (vgl. DIFF I).

Beweis. Wir wissen aus der DIFF I, dass wenn Konvergenz vorliegt, der Grenzwert gleich dem punktweisen Grenzwert ist. Dieser ist

$$f(x) = \begin{cases} 1 & x = 1 \\ 0 & \text{sonst} \end{cases}.$$

Aber

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| = \sup_{x \in [0, 1]} |x^n| = 1. \quad \square$$

iv) $X = C([0, 1])$, $d(f, g) = \int_0^1 |f(x) - g(x)| dx$.

1. Beh d ist eine Metrik auf $C([0, 1])$.

Beweis. HA. \square

2. Beh $(f_n)_n \subset C([0, 1])$, $f_n(x) = x^n$ konvergiert, und zwar gegen $f(x) = 0 \quad \forall x$.

Beweis.

$$\int_0^1 |f_n(x) - 0| dx \implies d(f_n, f) = \frac{1}{n+1} < \varepsilon \quad \forall n \geq N \text{ mit } N \geq \frac{1}{\varepsilon}. \quad \square$$

Vorlesung 2

Do 23.04. 10:15

Bevor wir uns mit offenen und abgeschlossenen Mengen und sogenannten vollständigen metrischen Räumen näher befassen, beweisen wir noch zwei nützliche Lemmata zu Konvergenz und Stetigkeit:

Lemma 1.17. (X, d) sei metrischer Raum.

Eine Folge $(x_n)_n$ in X konvergiert in X gegen $a \in X$

$$\iff (d(x_n, a))_n \text{ ist Nullfolge (in } \mathbb{R}\text{)}.$$

Beweis.

$$d(x_n, a) = |d(x_n, a) - 0|.$$

Also ist

$$d(x_n, a) < \varepsilon \iff |d(x_n, a) - 0| < \varepsilon. \quad \square$$

Lemma 1.18. Seien (X, d_x) und (Y, d_y) metrische Räume, $f: X \rightarrow Y$ eine Abbildung. Dann gilt:

f ist in $a \in X$ stetig \iff für jede Folge $(a_n)_n$ mit $a_n \rightarrow a$ in X gilt

$$\lim_{n \rightarrow \infty} f(a_n) = f(\underbrace{\lim_{n \rightarrow \infty} a_n}_{=a}).$$

Notation.

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Beweis. „ \implies “ Sei das ε - δ -Kriterium erfüllt (1.15.iii). Sei $(x_n)_n$ Folge in X mit $x_n \rightarrow a$ in X . Sei $\varepsilon > 0$. Dann $\exists \delta > 0$ s. d.

$$d_Y(f(x), f(a)) < \varepsilon \quad \forall x \in B_\delta(a) \subset X.$$

Wegen der Konvergenz $\exists N = N(\delta)$ s. d.

$$\begin{aligned} x_n &\in B_\delta(a) \quad \forall n \geq N \\ \implies f(x_n) &\in B_\varepsilon(f(a)) \subset Y \quad \forall n \geq N. \end{aligned}$$

Also gilt $f(x_n) \rightarrow f(a)$.

„ \Leftarrow “ Gelte $\lim_{x \rightarrow a}(x) = f(a)$.

Angenommen, das ε - δ -Kriterium wäre verletzt. Dann gäbe es $\varepsilon > 0$ s. d. für *alle* $\delta > 0$ ein $x \in X$ existierte s. d.

$$\begin{aligned} x \in B_\delta(a) \text{ aber } f(x) \notin B_\varepsilon(f(a)) \\ \text{also } d_y(f(x), f(a)) \geq \varepsilon. \end{aligned}$$

Insbesondere gäbe es zu $\delta = \frac{1}{n}$ ein solches x , nennen wir es x_n . Dann gilt für alle n : $d(x_n, a) < \frac{1}{n}$, aber $d_y(f(x_n), f(a)) \geq \varepsilon$, somit $x_n \rightarrow a$ aber $f(x_n) \not\rightarrow f(a)$ (wegen 1.17). \square

Charakterisierung topologischer Grundbegriffe in metrischen Räumen

Lemma 1.19. Sei (X, d) metrischer Raum. Dann ist $A \subset X$ abgeschlossen \iff für jede Folge $(a_n)_n$, $a_n \in A$, die in X konvergiert, gilt:

$$\lim_{n \rightarrow \infty} a_n \in A.$$

Beweis. O.B.d.A. $\emptyset \neq A \neq X$.

„ \implies “ Sei $(a_n)_n$, $a_n \in A$, konvergent in X . Sei $a = \lim a_n$. Angenommen $a \notin A$. Nach Voraussetzung ist $X \setminus A$ offen, also ist $X \setminus A$ Umgebung von $a \implies \exists N$ s. d.

$$a_n \in X \setminus A \quad \forall n \geq N \quad (\text{wegen Konvergenz})$$

\nrightarrow zu $a_n \in A$.

„ \Leftarrow “ Wir zeigen, dass $X \setminus A$ offen ist. Sei also $b \in X \setminus A$. Es gibt $\varepsilon > 0$ s. d. $B_\varepsilon(b) \cap A = \emptyset$, also $B_\varepsilon(b) \subset X \setminus A$.

Denn angenommen es gibt kein solches ε . Dann gilt für *alle* $\varepsilon > 0$: $B_\varepsilon(b) \cap A \neq \emptyset$, also kann man zu jedem $k \geq 1$ ein $x_k \in A$ finden mit $d(x_k, b) < \frac{1}{k} = \varepsilon$.

$$\implies x_k = b \underset{\text{VOR}}{\implies} b \in A.$$

\nrightarrow Widerspruch zu $b \in X \setminus A$.

Also gibt es ein solches $\varepsilon > 0$, also ist $X \setminus A$ offen. \square

Definition 1.20. Sei (X, d) metrischer Raum, $M \subset X$. Ein Punkt $y \in X$ heißt *Randpunkt* von M , falls in jeder Umgebung von y sowohl Punkte von M als auch $X \setminus M$ liegen.

Notation. $\partial(M) = \{ \text{Randpunkte von } M \}$

Beispiel $(\mathbb{R}^n, d_{\text{Eukl.}})$. Kugel im \mathbb{R}^n :

$$K^n := \{ x \in \mathbb{R}^n \mid \|x - 0\|_{\text{E}} \leq R \} \subset \mathbb{R}^n$$

Sphäre:

$$\partial(K)^n = \{ x \in \mathbb{R}^n \mid \|x\|_{\text{E}} = R \} = S^{n-1}$$

Beispiel. $\mathbb{Q} \subset \mathbb{R}$. $\partial(\mathbb{Q}) = \mathbb{R}$.

Satz 1.21. Sei (X, d) metrischer Raum. Sei $M \subset X$. Dann gilt

- i) $M \setminus \partial(M)$ ist offen.
- ii) $M \cup \partial(M)$ ist abgeschlossen.
- iii) $\partial(M)$ ist abgeschlossen.

Beweis. 1.21.i): $a \in M \setminus \partial(M) \implies \exists \varepsilon > 0$ s.d. $B_\varepsilon(a) \cap X \setminus M = \emptyset$. Für dieses gilt auch $B_\varepsilon \cap \partial(M) = \emptyset$ (denn angenommen $\exists y \in B_\varepsilon(a) \cap \partial(M)$, dann wäre (da $y \in \partial(M)$ und $B_\varepsilon(a)$ eine Umgebung von y) $B_\varepsilon(a) \cap (X \setminus M) \neq \emptyset$ ∇ VOR).

Also gilt $B_\varepsilon(a) \subset M \setminus \partial(M) \implies$ Beh.

1.21.ii): Es gilt $\partial(M) = \partial((X \setminus M) \setminus \partial((X \setminus M)))$ (nach Definition), $(X \setminus M) \setminus \partial((X \setminus M))$ ist offen nach 1.21.i) \implies

$$X \setminus ((X \setminus M) \setminus \partial((X \setminus M))) \overset{\substack{\uparrow \\ \text{Manipulation mit Mengen}}}{=} (X \setminus (X \setminus M)) \cup \partial((X \setminus M)) = M \cup \partial(M)$$

ist offen.



1.21.iii):

$$\begin{aligned} \partial(M) &= (M \cup \partial(M)) \setminus (M \setminus \partial(M)) \\ \implies X \setminus \partial(M) &= X \setminus \underbrace{\left(\underbrace{M \cup \partial(M)}_{\substack{\text{(abgeschl. nach 1.21.ii)}}} \right)}_{\text{offen}} \cup \underbrace{\left(M \setminus \partial(M) \right)}_{\text{offen nach 1.21.i)}. \end{aligned} \quad \square$$

Notation. Sei $M \subset X$.

$M^\circ := M \setminus \partial(M)$ heißt das *Innere* von M .

$\overline{M} := M \cup \partial(M)$ heißt der *Abschluss* von M .

Nach 1.19 können wir \overline{M} konstruieren, indem wir zu M noch alle Grenzwerte von Folgen $(x_n)_n$, $x_n \in M$, die in X konvergieren, hinzunehmen.

Beispiel. $M = [a, b)$, $\overline{M} = [a, b]$.

Bemerkung (als Hausaufgabe).

$$M \subset X \text{ offen} \iff M \cap \partial(M) = \emptyset.$$

$$M \subset X \text{ abgeschlossen} \iff \partial(M) \subset M.$$

Vollständigkeit

Definition 1.22. Sei (X, d) ein metrischer Raum. Eine Folge $(y_n)_n \subset X$ heißt *Cauchy-Folge*, falls gilt

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \text{ s. d. } d(y_n, y_m) < \varepsilon \quad \forall n, m \geq N.$$

Lemma 1.23. Sei (X, d) ein metrischer Raum. Eine konvergente Folge in X ist eine *Cauchy-Folge*.

Beweis. Sei $(y_n)_n$ konvergente Folge mit Grenzwert y (eindeutig wegen 1.14.v) und 1.6). Sei $\varepsilon > 0$.

Dann gibt es $N \in \mathbb{N}$ s. d. $d(y_m, y) < \varepsilon \quad \forall m \geq N$.

$$\implies d(y_n, y_m) \underset{\Delta}{\leq} d(y_n, y) + d(y, y_m) < \varepsilon \quad \forall n, m \geq N. \quad \square$$

Bemerkung. Nicht jede Cauchy-Folge konvergiert:

Beispiel $((\mathbb{Q}, |\cdot|))$. $y_{n+1} = \frac{1}{2}y_n + \frac{1}{y_n}$, $y_0 = 1$.

Check. Es gilt für $n \geq 1$

$$\left[\frac{1}{y_{n+1}}, y_{n+1} \right] \subset \left[\frac{1}{y_n}, y_n \right] \quad (*)$$

und für $l_n := y_n - \frac{1}{y_n}$

$$\begin{aligned} l_{n+1} &\leq \frac{1}{4y_{n+1}} l_n^2 \leq \frac{1}{4} l_n^2 & (**) \\ \implies d(y_n, y_m) = |y_n - y_m| &\leq \left| y_n - \frac{1}{y_n} \right| = l_n \xrightarrow{\text{wg. } (**)} 0. \\ &\quad \text{O.B.d.A. } m \geq n \quad \uparrow \\ &\implies y_m \in \left[\frac{1}{y_n}, y_n \right] \text{ wg. } (*) \end{aligned}$$

$\mathbb{Q} \subset \mathbb{R}$ und somit $(y_n)_n \subset \mathbb{R}$. In \mathbb{R} konvergiert jede Cauchy-Folge. Nennen wir den Grenzwert $a \in \mathbb{R}$. Es gilt dann

$$\underbrace{y_{n+1}}_{\rightarrow a} = \underbrace{\frac{1}{2}y_n}_{\frac{1}{2}a} + \underbrace{\frac{1}{y_n}}_{\frac{1}{a}},$$

also $a^2 = 2$. Aber $\sqrt{2} \notin \mathbb{Q}$.

Definition 1.24. Ein metrischer Raum, in dem jede Cauchy-Folge konvergiert heißt *vollständig*.

Beispiele 1.25. i) $\mathbb{R}, |\cdot|$ ist vollständig (DIFF I).

ii) $(C([a, b], \mathbb{R}), d_{L^1})$ mit $d_{L^1}(f, g) = \int_a^b |f(t) - g(t)| dt$ (vgl. HA Blatt 1, A1) ist *nicht* vollständig.

iii) $(C([a, b], \mathbb{R}), d_{\text{sup}})$, mit

$$d_{\text{sup}} = \|f - g\|_{\infty} = \sup_{t \in [a, b]} |f(t) - g(t)|,$$

ist vollständig. Den Beweis führen wir später allgemeiner.

Zunächst einige

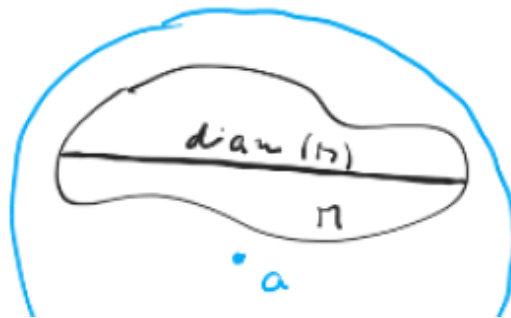
Betrachtungen in vollständigen metrischen Räumen

Definition 1.26. Sei (X, d) metrischer Raum, $M \subset X$,

$$\text{diam}(M) := \sup_{x, y \in M} d(x, y) \text{ „Durchmesser“ (englisch „diameter“).}$$

M heißt *beschränkt*, falls $\text{diam}(M) < \infty$.

Bemerkung. M ist beschränkt $\iff \exists R \geq 0$ und $a \in X$ s. d. $M \subset B_R(a)$



Beispiel. $\text{diam}([a, b]) = b - a$

Satz 1.27 (Schachtelungsprinzip). Sei (X, d) ein *vollständiger* metrischer Raum und sei $A_0 \subset A_1 \subset A_2 \subset \dots$. Eine Familie nicht-leerer abgeschlossener Teilmengen von X mit

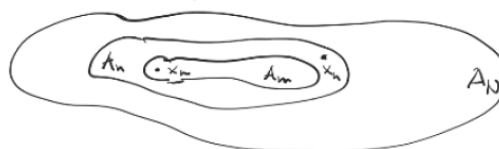
$$\text{diam}(A_k) \rightarrow 0 \text{ (in } \mathbb{R}) \text{ für } k \rightarrow \infty.$$

Dann gibt es genau einen Punkt $a \in X$ der in *allen* A_k liegt.

Beweis. Eindeutigkeit: Angenommen $\exists x \neq y$ mit $x \in A_k \ \forall k$ und $y \in A_k \ \forall k$. Dann kann $\text{diam}(A_k)$ keine Nullfolge sein, da $d(x, y) \neq 0$.

Existenz: Wähle $x_n \in A_n$. Dann ist $(x_n)_n$ eine Cauchy-Folge, denn

$$d(x_n, x_m) \leq \text{diam } A_N \text{ für } n, m \geq N$$



$$d(x_n, x_m) \leq \text{diam } A_N \text{ für}$$

$$\begin{array}{c} \implies x_n \rightarrow x \text{ in } X, \\ \uparrow \\ \text{Vollständigkeit} \end{array}$$

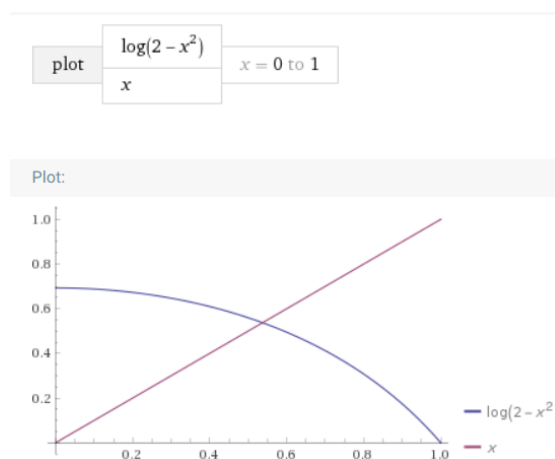
Da $x_n \in A_k \quad \forall n \geq k$, folgt mit 1.19: $x \in A_k \quad \forall k$. □

Ein sehr wichtiger Satz, der viele Anwendungen hat ist der folgende:

Satz 1.28 (Banach'scher Fixpunktsatz). Sei (X, d_X) ein *vollständiger* metrischer Raum. Sei $M \subset X$ eine *abgeschlossene* Teilmenge und $\Phi: M \rightarrow X$ eine Abbildung mit $\Phi(M \subset M)$ und es gebe $0 \leq L < 1$ s. d.

$$d_X(\Phi(x), \Phi(y)) \leq L d_X(x, y) \quad \forall x, y \in M \quad (, \Phi \text{ ist Kontraktion}).$$

Dann gibt es genau ein t_* s. d. $\Phi(t_*) = t_*$. Ein solches t_* heißt *Fixpunkt* von Φ .



$X = \mathbb{R}, M = [0, 1], \log(2 - x^2), (\text{WolframAlpha})$

Beispiel.

Beweis. Eindeutigkeit: Seien $\Phi(t_*) = t_*, \Phi(\tilde{t}_*) = \tilde{t}_*$. Dann gilt

$$\begin{aligned} d(t_*, \tilde{t}_*) &= d(\Phi(t_*), \Phi(\tilde{t}_*)) \\ &\leq d(t_*, \tilde{t}_*) \end{aligned}$$

Da $L > 1$ ist, folgt $d(t_*, \tilde{t}_*) = 0$, also $t_* = \tilde{t}_*$.

Existenz: Wir betrachten die Folge $x_0 \in M$ beliebig, $x_n := \Phi(x_{n-1})$ für $n \geq 1$.

Behauptung. $(x_n)_n$ konvergiert in M und zwar gegen den Fixpunkt.

Beweis. • $(x_n)_n$ ist Cauchy-Folge:

$$\begin{array}{c} d(x_{n+1}, x_n) \leq L d(x_n, x_{n-1}) \quad \forall n \geq 1. \\ \parallel \\ \Phi(x_n) \quad \parallel \\ \Phi(x_{n-1}) \end{array}$$

Iteration liefert

$$d(x_{n+1}, x_n) \leq L^2 d(x_{n-1}, x_{n-2}) \leq \dots \leq L^n d(x_1, x_0).$$

Zudem gilt

$$\begin{aligned} d(x_{n+k}, x_n) &\leq d(x_{n+k}, x_{n+k-1}) \\ &\quad + d(x_{n+k-1}, x_{n+k-2}) \\ &\quad \vdots \\ &\quad + d(x_{n+1}, x_n) \\ &\leq \underbrace{(L^{n+k-1} + L^{n+k-2} + \dots + L^n)}_{=L^n \sum_{r=0}^{k-1} L^r \leq L^n \sum_{r=0}^{\infty} L^r = \frac{L^n}{1-L}} d(x_1, x_0) \\ &\quad \text{geom. Reihe } (L < 1) \end{aligned}$$

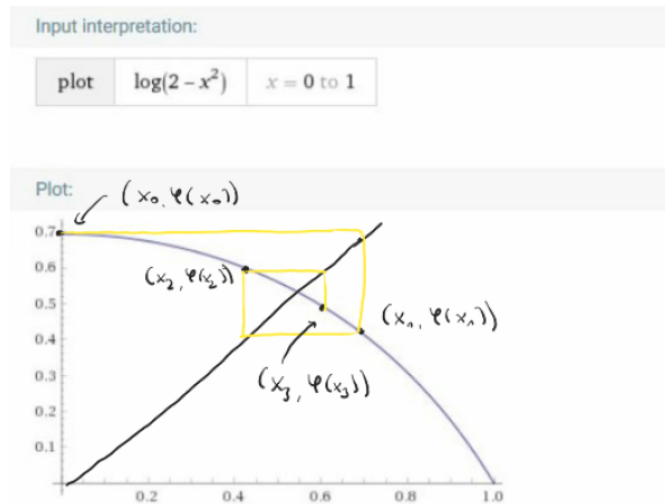
\implies (wegen $L < 1$) Beh.

- Da X vollständig ist, konvergiert $(x_n)_n$. Setze $t_* = \lim_{n \rightarrow \infty} x_n$.
- Da M abgeschlossen ist, ist $t_* \in M$ nach 1.19.
- t_* ist der gesuchte Fixpunkt:

$$t_* = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \Phi(x_{n-1}) = \Phi(t_*)$$

\uparrow
Kontraktionen sind stetig und 1.18

□



$$x_0 = 0, x_1 = \ln(2 - (\ln 2)^2) \approx 0,42, x_3 \approx 0,60, x_4 \approx 0,49$$

Bemerkung. Kontraktionen sind stetig: Zu $\varepsilon > 0$ wähle $\delta = \varepsilon/L$.

Bemerkung. Die Konvergenz ist recht schnell:

$$d(x_n, t_*) \leq \frac{L^n}{1 - L} d(x_1, x_0) \quad (L < 1).$$

Alle Voraussetzungen sind notwendig, gilt eine nicht, so gibt es nicht unbedingt einen Fixpunkt (oder keinen eindeutigen).



Vorlesung 3

Mo 27.04. 10:15

Lemma 1.29 (Cauchy-Kriterium für gleichmäßige Konvergenz).

- a) Sei (X, d_x) metrischer Raum, sei (Y, d_Y) ein *vollständiger* metrischer Raum. Sei $f_n: X \rightarrow Y$ Folge von Funktionen. Dann konvergiert f_n gegen f bezüglich

$$d_{\text{sup}}(h, g) := \sup_{t \in X} \underbrace{d_Y(f(t), g(t))}_{\in \mathbb{R}} \quad h, g: X \rightarrow Y$$

$$\iff \forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N} \text{ s. d.}$$

$$d_Y(f_n(t), f_m(t)) < \varepsilon \quad \forall t \in X \quad n, m \geq N(\varepsilon). \quad (*)$$

Notation. Man spricht von *gleichmäßiger Konvergenz*.

Beachte:

Der wesentliche Punkt in $(*)$ ist, dass N unabhängig von t gewählt werden kann.

Beweis. Wir stellen zunächst fest, dass d_{sup} auf

$$\mathcal{F} := \{ f: X \rightarrow Y \mid \text{Für je zwei Funktionen gilt: } d_{\text{sup}}(f_1, f_2) < \infty \}$$

eine Metrik definiert (auch wenn Y nicht vollständig ist).

1.11.a)

$$\begin{aligned} d_{\text{sup}}(f, g) &= 0 \\ \iff d_Y(f(t), g(t)) &= 0 \quad \forall t \in X \\ \iff f(t) &= g(t) \quad \forall t \in X \\ &\quad \uparrow \\ &\text{\scriptsize } d_Y \text{ ist Metrik} \end{aligned}$$

1.11.b)

$$d_{\text{sup}}(f, g) = \sup d_Y(f(t), g(t)) = \sup d_Y(g(t), f(t)) = d_{\text{sup}}(g, f)$$

1.11.c)

$$\begin{aligned} d_{\text{sup}}(f, g) &= \sup \underbrace{d_Y(f(t), g(t))}_{\leq d_Y(f(t), h(t)) + d_Y(h(t), g(t))} \\ &\leq \sup d_Y(f(t), h(t)) + \sup d_Y(h(t), g(t)) \\ &= d_{\text{sup}}(f, h) + d_{\text{sup}}(h, g) \end{aligned}$$

Zum Beweis der Behauptung:„ \implies “

$$\sup_t d_Y(f_n(t), f(t)) < \varepsilon \quad \forall n \geq N(\varepsilon)$$

impliziert

$$d_Y(f_n(t), f(t)) < \varepsilon \quad \forall t \in X \quad \forall n \geq N(\varepsilon),$$

somit für *alle* $t \in X$

$$\begin{aligned} d_Y(f_n(t), f_m(t)) &\leq_{\Delta(d_Y)} d_Y(f_n(t), f(t)) + d_Y(f(t), f_m(t)) \\ &< 2\varepsilon \quad \forall n, m \geq N(\varepsilon) \end{aligned}$$

„ \Leftarrow “ Gelte (*). Dann ist für jedes $t \in X$, dass $(f_n(t))_n$ ist Cauchy-Folge in Y .Vollständigkeit von $Y \implies (f_n(t))_n$ konvergiert. Setze $f(t) := \lim_{n \rightarrow \infty} f_n(t)$.Wir zeigen f_n konvergiert bezüglich d_{sup} gegen f . Sei also $\varepsilon > 0$. Wähle in (*) $m \geq N(\varepsilon)$ fest. Dann gilt für *alle* t :

$$\begin{aligned} \varepsilon &\geq \lim_{n \rightarrow \infty} d_Y(f_n(t), f_m(t)) \\ &= d_Y(f(t), f_m(t)). \end{aligned}$$

\uparrow
 d_Y ist stetig

Das gilt für alle $m \geq N(\varepsilon)$, $\forall t$, also auch für das Supremum \implies Beh. \square

- b) Seien X, Y metrische Räume, $(f_n)_n$ eine Folge *stetiger* Funktionen $f_n: X \rightarrow Y$, die gleichmäßig konvergiere. Dann ist die Grenzfunktion $f: X \rightarrow Y$.

Beweis. Sei $a \in X$. Sei $\varepsilon > 0$. Gleichmäßige Konvergenz $\implies \exists N = N(\varepsilon) \in \mathbb{N}$ s. d.

$$d_Y(f(t), f_n(t)) < \varepsilon \quad \forall t \in X \quad \forall n \geq N$$

 f_n stetig in $a \implies \exists \delta > 0$ s. d.

$$\begin{aligned} d_Y(f_N(t), f_N(a)) &< \varepsilon \quad \forall t \text{ mit } d_X(t, a) < \delta \\ \implies d_Y(f(t), f(a)) &\leq_{\Delta} d_Y(f(t), f_N(t)) + d_Y(f_N(t), f_N(a)) + d_Y(f_N(a), f(a)) \quad \square \\ &< 3\varepsilon \quad \forall t \text{ mit } d_X(t, a) < \delta. \end{aligned}$$

Folgerung.

$$(C([a, b], \mathbb{R}), d_{\sup}),$$

Stellt man diese Bedingung, ist automatisch garantiert, dass $d_{\sup}(f_1, f_2) = \sup_{t \in [a, b]} |f_1(t) - f_2(t)| < \infty$

ist vollständig.

Beweis. Sei $(f_n)_n$ Cauchy-Folge in $C(D, \mathbb{R})$ bezüglich d_{\sup} , d. h. zu $\varepsilon > 0 \exists N = N(\varepsilon)$ s. d.

$$\begin{aligned} d_{\sup}(f_n, f_m) &< \varepsilon \quad \forall n, m \geq N(\varepsilon) \\ \implies d_Y(f_n(t), f_m(t)) &= |f_n(t) - f_m(t)| < \varepsilon \quad \forall n, m \geq N(\varepsilon) \quad \forall t \in D. \end{aligned}$$

\mathbb{R} ist vollständig

$\xrightarrow{1.29} (f_n)_n$ konvergiert bezüglich d_{\sup} gegen seinen punktweisen Grenzwert

$$f(t) := \lim_{n \rightarrow \infty} f_n(t) \quad (\text{Konvergenz in } \mathbb{R})$$

$\xrightarrow{1.29.b)} t \mapsto f(t)$ ist stetig.

□

Stetige Abbildungen auf metrischen Räumen

Lemma 1.30. Seien X, Y, Z metrische Räume, $f: X \rightarrow Y$, $g: Y \rightarrow Z$, $f(X) \subset Y$. Ist f stetig in $a \in X$ und g stetig in $b = f(a) \in \tilde{Y}$, so ist $g \circ f: X \rightarrow Z$ stetig in a .

Beweis. (Über Folgenstetigkeit, Lemma 1.18) Sei $x_n \rightarrow a \implies \lim f(x_n) = (a) = b$ und $\lim g(f(x_n)) = g(b) = g(f(a)) \implies \lim g \circ f(x_n) = g \circ f(a)$.

Definition 1.31. Auf dem \mathbb{R}^n ist durch

$$d_{\max}(x, y) := \max_{i \in \{1, \dots, n\}} |x_i - y_i|.$$

eine Metrik definiert.

Bemerkungen. i) $d_{\max}(x, y) = d_{\sup}(x, y)$, fasst man x und y als Abbildungen

$$x: \{1, \dots, n\} \rightarrow \mathbb{R}$$

auf, $x(i) = x_i$.

- ii) Eine Folge $(x_m)_m \subset \mathbb{R}^n$, $x_m = (x_m^1, \dots, x_m^n)$ konvergiert bezüglich $d_{\max} \iff$ Alle Komponentenfolgen $(x_m^i)_m$ ($1 \leq i \leq n$) konvergieren in \mathbb{R} .

Beweis. „ \implies “ Zu $\varepsilon > 0 \exists N$ s. d. $\max |x_m^i - a_m^i| < \varepsilon \quad \forall m \geq N$.

„ \impliedby “ Zu $\varepsilon > 0 \exists N_i$ s. d. $|x_m^i - a_m^i| < \varepsilon \quad \forall m \geq N_i$

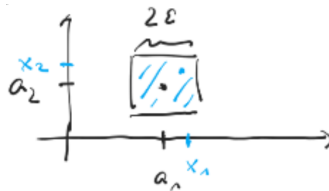
$$\implies \max_i |x_m^i - a_m^i| < \varepsilon \quad \forall m \geq N = \max(\{N_1, \dots, N_n\}). \quad \square$$

- iii) Es folgt: (\mathbb{R}^n, d_{\max}) ist *vollständig*.

- iv) $B_\varepsilon(a)$ bezüglich dieser Metrik:

$$\left\{ x \in \mathbb{R}^n \mid \max_{i \in \{1, \dots, n\}} |x_i - a_i| < \varepsilon \right\},$$

Würfel mit Seitenlängen 2ε um a .



Lemma 1.32. Sei (X, d) metrischer Raum. Sei \mathbb{R}^n mit d_{\max} versehen. Eine Abbildung $f: X \rightarrow \mathbb{R}^n$, $f = (f_1, \dots, f_n)^T$,

$$f(y) = (f_1(y), \dots, f_n(y))^T \in \mathbb{R}^n, y \in X.$$

$f_i: X \rightarrow \mathbb{R}, i \in \{1, \dots, n\}$, „Komponenten-Funktionen“, ist genau dann stetig in $a \in X$, falls alle f_i stetig in a sind.

Beweis. Mit Folgenstetigkeit direkt aus Bemerkung 1.31.ii). Hier nochmals mit ε - δ -Kriterium.

Notation. $\underline{n} = \{1, \dots, n\}$.

„ \implies “ Sei also $f: X \rightarrow \mathbb{R}^n$ stetig in a . Sei $\varepsilon > 0$. Dann $\exists \delta > 0$ s. d.

$$\begin{aligned} \max_{i \in \underline{n}} |f_i(y) - f_i(a)| &< \varepsilon \quad \forall y \in B_\delta(a) \\ &\quad \uparrow \\ &\text{bezüglich } d \\ \implies |f_i(y) - f_i(a)| &< \varepsilon \quad \forall y \in B_\delta(a) \quad \forall i \in \underline{n} \\ \implies f_i &\text{ sind stetig in } a. \end{aligned}$$

„ \Leftarrow “ Seien also die $f_i: X \rightarrow \mathbb{R}$, $i \in \underline{n}$, stetig in a . Sei $\varepsilon > 0$. Dann $\exists \delta_i > 0$ s. d.

$$|f_i(y) - f_i(a)| < \varepsilon \quad \forall y \in B_{\delta_i}(a) \subset X.$$

Wähle $d := \min(\{\delta_1, \dots, \delta_n\})$. Dann ist

$$\max_{i \in \underline{n}} |f_i(y) - f_i(a)| < \varepsilon \quad \forall y \in B_d(a).$$

□

Lemma 1.33. Folgende Abbildungen sind stetig:

$$\begin{aligned} \text{add}: \mathbb{R}^2 &\rightarrow \mathbb{R}, \quad \text{add}(x, y) = x + y \\ \text{mult}: \mathbb{R}^2 &\rightarrow \mathbb{R}, \quad \text{mult}(x, y) = x \cdot y \\ \text{quot}: \mathbb{R} \times \mathbb{R}^* &\rightarrow \mathbb{R}, \quad \text{quot}(x, y) = x/y. \\ &\quad \parallel \\ &\quad \mathbb{R} \setminus \{0\} \end{aligned}$$

Hierbei sei \mathbb{R}^2 , $\mathbb{R} \times \mathbb{R}^*$, mit d_{\max} versehen.

Beweis. Sei $((x_m, y_m))_m \subset \mathbb{R}^2$ mit $(x_m, y_m) \rightarrow (x, y)$ (bezüglich d_{\max})

$$\xRightarrow{\text{Bem 1.31.ii}} x_m \rightarrow x \text{ und } y_m \rightarrow y \text{ in } \mathbb{R}$$

$$\implies \lim(x_m + y_m) = x + y$$

$$\lim(x_m \cdot y_m) = x \cdot y$$

$$\lim(x_m/y_m) = x/y \quad \text{falls } y_m \neq 0, y \neq 0.$$

□

Folgerung. Sei (X, d) metrischer Raum. Seien $f, g: X \rightarrow \mathbb{R}$ stetig. Dann sind auch

$$\begin{aligned} f + g: X &\rightarrow \mathbb{R}, \quad (f + g)(x) = f(x) + g(x) \text{ und} \\ f \cdot g: X &\rightarrow \mathbb{R}, \quad (f \cdot g)(x) = f(x) \cdot g(x) \end{aligned}$$

stetig. Gilt $g(x) \neq 0 \quad \forall x \in X$, so ist auch

$$f/g: X \rightarrow \mathbb{R}, (f/g)(x) = f(x)/g(x)$$

stetig.

Beweis.

$$1.32 \implies \begin{pmatrix} f \\ g \end{pmatrix}: X \rightarrow \mathbb{R}^2, \begin{pmatrix} f \\ g \end{pmatrix}(x) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$$

ist stetig.

Es ist

$$\begin{aligned} f + g &= \text{add} \circ \begin{pmatrix} f \\ g \end{pmatrix} \\ f \cdot g &= \text{mult} \circ \begin{pmatrix} f \\ g \end{pmatrix} \\ f/g &= \text{quot} \circ \begin{pmatrix} f \\ g \end{pmatrix} \end{aligned}$$

Mit 1.33 und 1.30 folgt die Behauptung. \square

Beispiel. Polynomische Funktionen $\mathbb{R}^n \rightarrow \mathbb{R}$

$$x \mapsto \sum_{0 \leq k_i \leq r} c_{\underbrace{k_1 \dots k_n}_{\in \mathbb{R}}} x_1^{k_1} \dots x_n^{k_n}$$

sind stetig.

Bemerkung 1.34. Wir werden später sehen, dass die Aussage in 1.33 auch gilt, wenn man den \mathbb{R}^2 z. B. mit dem Euklidischen Abstand versieht.

Kompaktheit

Definition 1.35. Sei (X, d) metrischer Raum, $M \subset X$. Eine *offene Überdeckung* von M ist eine Familie $(U_i)_{i \in I}$ von offenen Teilmengen $U_i \subset X$ mit $M \subset \bigcup_{i \in I} U_i$ (I eine beliebige Indexmenge).

Definition 1.36. $M \subset X$ heißt *kompakt*, falls es zu *jeder* offenen Überdeckung von $\bigcup_{i \in I} U_i$ von M endlich viele Indizes i_1, \dots, i_N gibt s. d.

$$M \subset U_{i_1} \cup \dots \cup U_{i_N}.$$

Achtung. Ein nicht-kompakter Raum kann eine endliche Überdeckung $U_1 \cup \dots \cup U_N$ besitzen. Die Aussage der Definition ist, dass man aus *jeder* offenen Überdeckung endlich viele offene Mengen wählen kann, die M noch ganz überdecken!

Beispiele 1.37. i) $[a, b]$ ist kompakt (Beweis später).

ii) (a, b) ist nicht kompakt (obwohl etwa (a, b) eine endliche offene Überdeckung ist!)

Beweis.

$$U_j = \left(a + \frac{1}{j}, b\right), \quad j \geq 1$$

$$\bigcup_j U_j = (a, b)$$

aber es gibt *kein* N s. d. $\bigcup_{j=1}^N U_j \supset (a, b)$, denn z. B. $a + \frac{1}{N+1} \notin \bigcup_{j=1}^N U_j$. \square

iii) Sei $(x_n)_n \subset X$ gegen a konvergente Folge. Dann ist $M = \{x_n \mid n \in \mathbb{N}\} \cup \{a\}$ kompakt.

Beweis. Sei $(U_j)_j$ eine offene Überdeckung von M

$$a \in M \implies \exists j_0 \text{ s. d. } a \in U_{j_0}$$

U_{j_0} ist offen, also eine Umgebung von a .

$$\implies \exists N \text{ s. d. } x_n \in U_{j_0} \quad \forall n \geq N. \quad \square$$

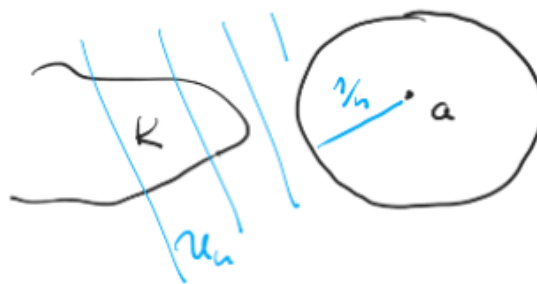
iv) Sei $(X_i, d_{\text{discrete}})$. Dann sind genau die endlichen Mengen kompakt.

Beweis. Betrachte $\bigcup_{x \in M} \{x\}$. \square

Satz 1.38. Sei (X, d) metrischer Raum, $K \subset X$ kompakt. Dann ist K abgeschlossen und beschränkt.

Beweis. Abgeschlossen: Sei $a \in X \setminus K$. Setze zu $n \geq 1$

$$U_n := \left\{ y \in X \mid d(y, a) > \frac{1}{n} \right\}$$



U_n ist offen (denn $X \setminus U_n = \overline{B_{1/n}(a)}$) und $\bigcup_{n=1}^{\infty} U_n = X \setminus \{a\} \supset K$. K kompakt $\implies \exists U_{n_1}, \dots, U_{n_L}$ s. d. $K \subset U_{n_1} \cup \dots \cup U_{n_L}$. Setze $N := \max(\{n_1, \dots, n_L\})$. Dann ist $B_{\frac{1}{N}}(a) \subset X \setminus K \implies X \setminus K$ ist offen \implies Beh.

Beschränktheit: Sei $a \in X$. Dann ist $X = \bigcup_{n=1}^{\infty} B_n(a)$ und somit $(B_n(a))_n$ eine offene Überdeckung von K .

$$\begin{aligned} &\implies \exists n_1, \dots, n_k \text{ s. d. } K \subset B_{n_1}(a) \cup \dots \cup B_{n_k}(a) \\ &\implies K \subset B_N(a) \text{ für } N = \max(\{n_1, \dots, n_k\}) \\ &\implies \text{diam}(K) \leq 2N. \end{aligned}$$

□

Folgerung. Konvergente Folgen sind beschränkt.

Bemerkung. Die Umkehrung von 1.38 gilt im Allgemeinen nicht!

(X, d_{discrete}) , X habe unendlich viele Elemente. Jede Teilmenge ist abgeschlossen (da jede offen ist) und beschränkt (durch 1), aber nur die *endlichen* sind kompakt.

Lemma 1.39. Ist $K \subset X$ kompakt und $A \subset K$ ist abgeschlossen, so ist A kompakt.

Beweis. Sei $(U_j)_j$ offene Überdeckung von A . Es ist

$$\begin{aligned} &\underbrace{(X \setminus A) \cup \bigcup U_j}_{\text{offen (VOR)}} = X \supset K \\ &\implies \exists j_1, \dots, j_L \text{ s. d. } K \subset (X \setminus A) \cup U_{j_1} \cup \dots \cup U_{j_L} \\ &\implies A \subset U_{j_1} \cup \dots \cup U_{j_L}. \end{aligned}$$

□

Satz 1.40. Seien X, Y metrische Räume und $f: X \rightarrow Y$ stetig. Ist $K \subset X$ kompakt, so ist auch $f(K) \subset Y$ kompakt.

Beweis. Sei $(U_j)_j$ offene Überdeckung von $f(K)$. f stetig $\xRightarrow{1.8}$ Die Urbilder $V_j := f^{-1}(U_j)$ sind offen.

Und nach Definition ist $K \subset \bigcup_j V_j$.

$$\begin{aligned} &\xRightarrow{\text{VOR}} \exists j_1, \dots, j_N \text{ s. d. } K \subset V_{j_1} \cup \dots \cup V_{j_N} \\ &\implies f(K) \subset U_{j_1} \cup \dots \cup U_{j_N}. \end{aligned}$$

□

Satz 1.41. Sei \mathfrak{X} kompakter metrischer Raum, $f: \mathfrak{X} \rightarrow \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt ihr Maximum und Minimum an, d. h. $\exists a, b \in \mathfrak{X}$

$$f(a) = \sup \{ f(x) \mid x \in \mathfrak{X} \}, \quad f(b) = \inf \{ f(x) \mid x \in \mathfrak{X} \}.$$

Beweis. 1.40 $\implies f(\mathfrak{X})$ ist kompakt. Mit 1.38 folgt: $f(\mathfrak{X})$ ist beschränkt (somit ist f beschränkt) und abgeschlossen.

Also sind $\sup f(\mathfrak{X})$ und $\inf(\mathfrak{X})$ endlich. Zudem gibt es

$$(y_k)_k \subset f(\mathfrak{X}), \quad y_k \rightarrow \sup(f(\mathfrak{X}))$$

$$(z_k)_k \subset f(\mathfrak{X}), \quad z_k \rightarrow \inf(\mathfrak{X}),$$

somit (Abgeschlossenheit!)

$$\sup(f(\mathfrak{X})) \in f(\mathfrak{X})$$

$$\inf(f(\mathfrak{X})) \in f(\mathfrak{X})$$

\implies Beh. □

Beispiel. Sei (\mathfrak{X}, d) metrischer Raum. $M \subset \mathfrak{X}$. Sei $x \in \mathfrak{X}$. Der *Abstand* von x zu M ist definiert als

$$d(x, M) := \inf \{ d(x, y) \mid y \in M \}.$$

Behauptung. $x \mapsto d(x, M)$ ist stetig auf \mathfrak{X} .

Beweis. Sei $\varepsilon > 0$. Dann ist

$$|d(x, M) - d(\tilde{x}, M)| \leq d(x, \tilde{x}) < \varepsilon \quad \text{falls } d(x, \tilde{x}) < \varepsilon,$$

denn

$$d(x, M) \underset{\Delta}{\leq} d(x, \tilde{x}) + d(\tilde{x}, M) \quad \forall x, \tilde{x} \in \mathfrak{X}. \quad \square$$

Definiere zu $K \subset \mathfrak{X}$

$$d(K, M) := \inf \{ d(x, M) \mid x \in K \}.$$

Behauptung. Ist M abgeschlossen, K kompakt und ist $M \cap K = \emptyset$, so gilt $d(M, K) > 0$.

Beweis. $x \mapsto d(x, M)$ ist stetig auf \mathfrak{X} , somit erst recht auf K . K ist kompakt $\xrightarrow{1.41} \exists a \in K$ s.d. $d(a, M) = d(K, M)$. M abgeschlossen $\implies \exists \varepsilon > 0$ s.d. $B_\varepsilon(a) \subset X \setminus M \implies d(a, M) \geq \varepsilon$.

Achtung. i) Betrachte

$$M = \{ (x, y) \mid xy = 0 \} \subset \mathbb{R}^2, N = \{ (x, y) \mid xy = 1 \} \subset \mathbb{R}^2$$

$$d(M, N) = 0.$$



ii) Betrachte $B_{1/2}(1), B_{1/2}(2) \subset \mathbb{R}^2$, $d_{\text{Euklidisch}}$. Distanz ist 0.



Satz / Definition 1.42. Seien \mathfrak{X}, Y metrische Räume, \mathfrak{X} kompakt. Dann ist jede stetig Abbildung $f: \mathfrak{X} \rightarrow Y$ sogar *gleichmäßig stetig* d. h. im ε - δ -Kriterium kann δ unabhängig von x gewählt werden:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ s. d. } d_Y(f(x), f(\tilde{x})) < \varepsilon \quad \forall x, \tilde{x}, d_{\mathfrak{X}}(x, \tilde{x}) < \delta.$$

Beweis. Sei $\varepsilon > 0$. Dann gibt es zu $a \in \mathfrak{X}$ ein $\delta(a) > 0$ s. d.

$$d_Y(f(a), f(y)) < \varepsilon \quad \forall y \in B_{\delta(a)}(a).$$

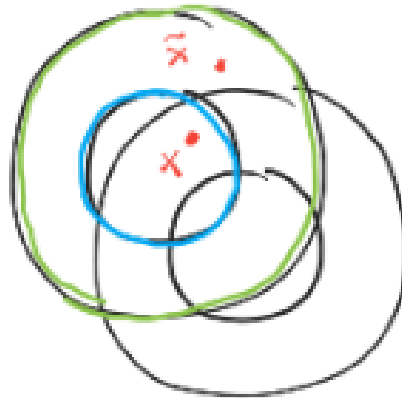
Es gilt $\bigcup_{a \in X} B_{\frac{\delta(a)}{2}}(a) = \mathfrak{X}$.

\mathfrak{X} ist kompakt $\implies \exists a_1, \dots, a_N$ s. d. $X = \bigcup_{j=1}^N B_{\delta(a_j)/2}(a_j)$. Setze

$$\delta := \frac{1}{2} \min(\{ \delta(a_1), \dots, \delta(a_N) \}).$$

Seien jetzt x, \tilde{x} beliebig aus \mathfrak{X} mit $d_{\mathfrak{X}}(x, \tilde{x}) < \delta$. Dann gibt es ein $j \in \{1, \dots, N\}$ s.d. $x \in B_{\delta(a_j)/2}$ und somit $\tilde{x} \in B_{\delta(a_j)}(a_j)$

$$\begin{aligned} \implies d_Y(f(x), f(a_j)) &< \varepsilon & d_Y(f(\tilde{x}), f(a_j)) &< \varepsilon \\ \implies d_Y(f(x), f(\tilde{x})) &< 2\varepsilon & \forall x, \tilde{x}, d_{\mathfrak{X}}(x, \tilde{x}) &< \delta. \end{aligned} \quad \square$$



Satz 1.43 (Bolzano-Weierstraß). Sei (X, d) metrischer Raum. Sei $K \subset X$ kompakt. Dann besitzt jede Folge $(x_n)_n$ in K eine Teilfolge $(x_{n_k})_k$, die gegen einen Punkt $x \in K$ konvergiert.

Beweis. Angenommen, \nexists Teilfolge, die gegen einen Punkt von K konvergiert. Dann besitzt jedes $x \in K$ eine offene Umgebung U_x , in der nur endlich viele Folgenglieder liegen (sonst könnte man eine gegen x konvergente Teilfolge konstruieren). Es gilt: $\bigcup_{x \in K} U_x \supset K$

$$\implies \exists x_1, \dots, x_N \text{ s.d. } \bigcup_{j=1}^N U_{x_j} \subset K \quad \square$$

Aber dann liegen nur endlich viele x_k in K , \nless zur Definition.

Vorlesung 4

Do 30.04. 10:15

Äquivalenz von Metriken

Wir haben gesehen, dass die Eigenschaften derselben Menge sehr verschieden sein können, je nachdem mit welcher Topologie man sie versieht.

Beispiel. \mathbb{R} mit der Standardtopologie $|x - y|$:

- $(a, b]$ ist nicht offen, $[a, b]$ ist kompakt.

\mathbb{R} mit der diskreten Metrik d_{disk}

- Alle Teilmengen sind offen.
- Nur endliche Teilmengen sind kompakt.
- Konvergiert $x_n \rightarrow a$ (bezüglich d_{disk}), so muss gelten $\exists N$ s. d. $x_n = a \quad \forall n \geq N$ (denn $\{a\}$ ist Umgebung von a , oder anders gesagt: damit $d(x_n, a) < \varepsilon < 1$ wird, muss gelten $x_n = a$).
- Alle Abbildungen $f: (X, d_{\text{disc}}) \rightarrow (Y, d)$ sind stetig. (Beweis am einfachsten über Folgenstetigkeit).

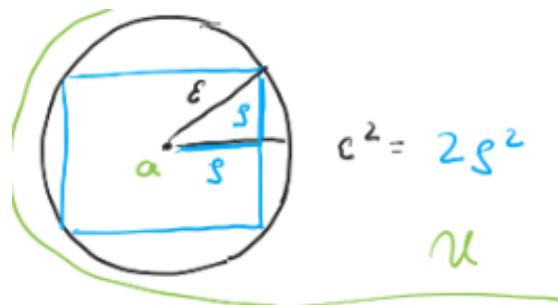
Andererseits gilt:

$U \subset \mathbb{R}^2$ ist offen in $(\mathbb{R}^2, d_{\text{Eukl}}) \iff U$ ist offen in $(\mathbb{R}^2, d_{\text{max}})$.

Beweis. „ \implies “ Sei $a \in U \xrightarrow{\text{VOR}} \exists \varepsilon > 0$ s. d.

$$B_{\varepsilon}^{d_E}(a) := \left\{ x \in \mathbb{R}^2 \mid d_{\text{Eukl}}(x, a) = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2} < \varepsilon \right\} \subset U$$

Da $B_{\rho}^{d_{\text{max}}}(a) \subset B_{\varepsilon}^{d_E}(a)$ für $\rho = \frac{\varepsilon}{\sqrt{2}}$, ist U auch offen $(\mathbb{R}^2, d_{\text{max}})$.



„ \Leftarrow “ Sei $a \in U \xrightarrow{VOR} \exists \varepsilon > 0$ s. d.

$$B_\varepsilon^{d_{\max}}(a) = \{x \mid d_{\max}(x, a) < \varepsilon\} \subset U.$$

□

Es gilt $B_\rho^{d_E}(a) \subset B_\varepsilon^{d_{\max}}(a)$ für $\rho = \varepsilon$, also ist U offen in $(\mathbb{R}^2, d_{\text{Eukl}})$.



Definition 1.44. Sei X eine Menge, seien d und \tilde{d} Metriken auf X . Dann nennt man d *stärker* (feiner) als \tilde{d} , falls jede bezüglich \tilde{d} offene Menge auch offen bezüglich d ist, und *schwächer* (gröber), falls \tilde{d} stärker ist als d . Ist d sowohl stärker als auch schwächer als \tilde{d} , so nennt man d und \tilde{d} *äquivalent*.

Beispiel. d_{\max} ist äquivalent zu d_{Eukl} . d_{disk} ist stärker als d_{\max} und nicht schwächer.

Bemerkungen 1.45. Sei d stärker als \tilde{d} . Dann gilt:

- i) Konvergiert eine Folge bezüglich der stärkeren Metrik, so auch bezüglich der schwächeren.

(denn: Konvergiere $x_n \rightarrow a$ (bezüglich d). Sei $\varepsilon > 0$. Betrachte $B_\varepsilon^{\tilde{d}}(a) = U$ offen bezüglich $\tilde{d} \implies U$ Umgebung von a (bezüglich \tilde{d}) $\implies U$ Umgebung von a (bezüglich d) $\implies \exists N$ s. d. $x_n \in U \quad \forall n \geq N$.)

- ii) Ist eine Funktion $f: (X, \tilde{d}) \rightarrow (Y, d_Y)$ stetig, so auch $f: (X, d) \rightarrow (Y, d_Y)$.

- iii) Ist eine Funktion $f: (Y, d_Y) \rightarrow (X, d)$ stetig, so auch $f: (Y, d_Y) \rightarrow (X, \tilde{d})$.

Beweis. f stetig \iff Urbilder offener Mengen sind offen.

1. $U \subset Y \implies f^{-1}(U)$ offen bezüglich $\tilde{d} \implies f^{-1}(U)$ offen bezüglich d .

2. Sei $U \subset X$ offen bezüglich \tilde{d} , also auch offen bezüglich $d \implies f^{-1}(U)$ offen in Y . □

Bemerkung. Sind d und \tilde{d} äquivalent, sind die selben Folgen konvergent, die selben Mengen offen, kompakt, die selben Funktionen stetig etc.

Kapitel 2

Normierte Vektorräume

Definition 2.1. Sei V ein reeller Vektorraum. Eine *Norm* auf V ist eine Abbildung $\|\cdot\|: V \rightarrow \mathbb{R}$ mit

a)

$$\|x\| = 0 \iff x = 0$$

b)

$$\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \mathbb{R}, x \in V$$

c)

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$$

Dreiecksungleichung.

Ein *normierter VR* $(V, \|\cdot\|)$ ist ein VR mit einer Norm.

Beispiele. • \mathbb{R}^n mit $\|\cdot\|_E$ Euklidische Norm auf \mathbb{R}^n

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2} \quad x = (x_1, \dots, x_n).$$

• \mathbb{R}^n mit $\|\cdot\|_\infty = \|\cdot\|_{\max}$,

$$\|x\|_\infty = \max(\{|x_1|, \dots, |x_n|\}).$$

• \mathbb{R}^n mit $\|\cdot\|_p$ „ p -Norm“, $p \geq 1, p \in \mathbb{R}$.

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \rightarrow \text{Saalübung.}$$

- $C([a, b])$ mit $\|f\|_{L^1} = \int_a^b |f(t)| dt$.
- $C([a, b])$ mit $\|f\|_{\infty} = \sup_{t \in [a, b]} |f(t)|$.

Lemma 2.2. Sei $(V, \|\cdot\|)$ normierter VR. Dann wird durch $d(x, y) := \|x - y\|$ eine Metrik auf V definiert („induziert“).

Beweis. 2.1.a) (Norm) \implies 1.11.a) (Metrik). 1.11.b) (Symmetrie der Metrik): folgt aus $\|x - y\| = \|y - x\|$. \square

Notation. Wir schreiben $(V, \|\cdot\|)$ für den *metrischen* Raum, dessen Metrik von $\|\cdot\|$, dessen Metrik von $\|\cdot\|$ induziert wird.

Bemerkung. Nicht jede Metrik auf einem Vektorraum wird von einer Norm induziert, denn induzierte Metriken erfüllen $d(\lambda x, \lambda y) = |\lambda| d(x, y)$. Die diskrete Metrik erfüllt das nicht.

Lemma 2.3. Seien d_1 und d_2 auf V von Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ induziert. Dann ist d_2 stärker als d_1 genau dann, wenn es eine positive Zahl $C > 0$ gibt s. d.

$$\|x\|_1 \leq C \|x\|_2 \quad \forall x \in V.$$

Beweis. Bezeichne $B_r^j(0)$, $r > 0$, die offenen Kugeln bezüglich d_j .

„ \implies “ Nach VOR ist insbesondere $B_1^1(0)$ offen bezüglich $d_2 \implies \exists \varepsilon > 0$ s. d. $B_\varepsilon^2(0) \subset B_1^1(0) \implies$ für $x \in X$, $x \neq 0$ gilt

$$\begin{aligned} \left\| \frac{\varepsilon x}{2\|x\|_2} \right\|_2 &= \frac{\varepsilon}{2} < \varepsilon \\ \implies \left\| \frac{\varepsilon x}{2\|x\|_2} \right\|_1 &< 1 \\ \implies \|x\|_1 &< \frac{2}{\varepsilon} \|x\|_2. \end{aligned}$$

„ \Leftarrow “ Existiere C wie oben. $B_r^2(x) \subset B_{cr}^1(x) \quad \forall x \in X, r > 0$. Denn $r > \|x - y\|_2 \geq \frac{1}{C}$. Sei U offen bezüglich d_1

$$\begin{aligned} \implies \forall x \in U \quad \exists \varepsilon > 0 \text{ s. d. } B_\varepsilon^1(x) &\subset U \\ \implies B_{\frac{\varepsilon}{C}}^2(x) \subset B_\varepsilon^1(x) &\subset U. \end{aligned} \quad \square$$

Folgerung. d_2 ist äquivalent zu d_1

$$\iff \exists C, \tilde{C} \text{ s. d. } \tilde{C} \|x_2\| \leq \|x_1\| \leq C \|x_2\|,$$

(„die Normen sind äquivalent“).

Bemerkung. Äquivalenz von Normen ist eine Äquivalenz-Relation (reflexiv, symmetrisch, transitiv).

Bemerkung. Für allgemeine Metriken kann man nicht so eine einfache Abschätzung angeben. Etwa ist $\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto \|x - y\|_{\max}$ unbeschränkt, aber $d_{\text{discrete}}(x, y) = 1 \quad \forall x \neq y$. Die Beweisrichtung „ \Leftarrow “ gilt noch (also gibt es ein C wie in 2.3, so ist d_2 stärker als d_1). Die Beweisrichtung „ \Rightarrow “ verwendet die Skalar-Multiplikation des zugrunde liegenden Vektorraums und vor allem die „Betrags-Homogenität“ der Norm

$$\left\| \frac{x}{\|x\|_2} \right\| = \frac{1}{\|x\|_2} \|x\|_1.$$

Satz 2.4. Auf \mathbb{R}^n sind alle Normen äquivalent.

Beweis. Aufgrund der Transitivität genügt es die Äquivalenz einer beliebigen Norm $\|\cdot\|$ zu $\|\cdot\|_{\infty}$ zu beweisen.

1. $\|\cdot\|_{\infty}$ ist stärker als $\|\cdot\|$: Denn sei $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$, $e_j = (0, \dots, \underset{j\text{-te}}{1}, \dots, 0)$. Dann ist

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \underset{\substack{\uparrow \\ \Delta \text{ und Homog.}}}{\leq} |x_j| \cdot \|e_j\| \leq \|x\|_{\infty} \underbrace{\sum_{j=1}^n \|e_j\|}_{=C}$$

2. $\|\cdot\|_{\infty}$ ist schwächer als $\|\cdot\|$:

Betrachte $M := \{x \in \mathbb{R}^n \mid \|x\|_{\infty} = 1\}$ (Einheits„sphäre“ bezüglich $\|\cdot\|_{\infty}$, also Rand des Einheitswürfels ).

Behauptung. $f: M \rightarrow \mathbb{R}$, $x \mapsto \|x\|$ ist stetig bezüglich $\|\cdot\|_{\infty}$.

Beweis.

$$|f(x) - f(y)| = |\|x\| - \|y\|| \underset{(*)}{\leq} \|x - y\| \leq C \|x - y\|_{\infty}$$

(*) umgekehrte Dreiecksungleichung:

$$\begin{aligned} \|x\| - \|y\| &= \|x + y - y\| - \|y\| \stackrel{\Delta}{\leq} \|x + y\| \\ \|y\| - \|x\| &= \|y + x - x\| - \|x\| \leq \|x + y\|. \end{aligned}$$

□

M ist abgeschlossen bezüglich $\|\cdot\|_{\infty}$ (denn $\mathbb{R}^n \setminus M = \text{Urbild der offenen Menge } \mathbb{R} \setminus \{1\}$ unter der stetigen Abbildung $x \mapsto \|x\|_{\infty}$). $M \subset$ abgeschlossenen *Quader* und dieser ist *kompakt* in $(\mathbb{R}^n, \|\cdot\|_{\infty})$ (Lemma 2.5) $\Rightarrow M$ ist kompakt (1.39).

Es folgt: f nimmt sein Minimum b an und (da $f > 0$) somit ist $b > 0$. Nach Definition ist $\|y\| \geq b \quad \forall y \in M$. Für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt $\frac{x}{\|x\|_\infty} \in M$, also ist $\|\frac{x}{\|x\|_\infty}\| \geq b$, also $\|x\| \geq b\|x\|_\infty$ und für $x = 0$ gilt dies ohnehin. \square

Lemma 2.5. Der Quader $Q = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid a_j \leq x_j \leq b_j\}$ ist kompakt in $\mathbb{R}^n, \|\cdot\|_\infty$ ($a_j \leq b_j$).

Beweis. Sei $(U_j)_j$ eine offene Überdeckung von Q . Angenommen, Q kann nicht durch endlich viele U_j 's überdeckt werden.

Wir konstruieren induktiv eine Folge von abgeschlossenen Teilquadern

$$Q_0 \supset Q_1 \supset Q_2 \supset \dots$$

mit

- a) Q_n kann *nicht* durch endlich viele U_j 's überdeckt werden
- b) $\text{diam } Q_m = 2^{-m} \text{diam } Q$.

Beachte:

$\text{diam } Q$ = Länge der längsten Seite bezüglich $\|\cdot\|_\infty$.

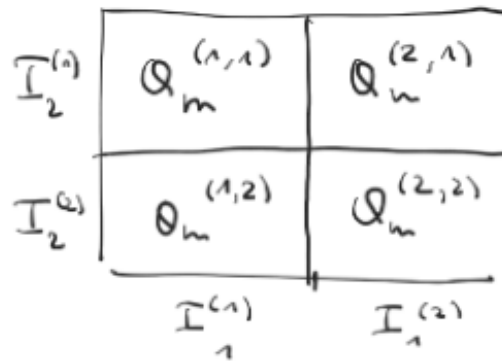


Setze $Q_0 = Q$. Sei Q_m konstruiert. Schreibe $Q_m = I_1 \times \dots \times I_n$, I_j abgeschlossene Intervalle. Zerlege $I_j^{(1)} \cup I_j^{(2)}$ in zwei abgeschlossene Intervalle der halben Länge und setze

$$Q^{(s_1, \dots, s_n)} := I_1^{(s_1)} \times \dots \times I_n^{(s_n)}, \quad s_j \in \{1, 2\}.$$

Das ergibt 2^n Quader mit

$$\bigcup_{s_j \in \{1, 2\}} Q_m^{(s_1, \dots, s_n)} = Q_m$$



Es gibt mindestens einen Quader $Q_m^{(s_1, \dots, s_n)}$, der nicht durch endlich viele U_j 's überdeckt werden kann. Einen solchen wählen wir als Q_{m+1} . Es gilt per Konstruktion

$$\text{diam}(Q_{m+1}) = \frac{1}{2} \text{diam}(Q_m) = \frac{1}{2^{m+1}} \text{diam}(Q).$$

Nach dem Schachtelungsprinzip $\exists a \in Q_m \forall m$. Da $(U_j)_j$ Q überdeckt $\exists U_{j_0}$ s. d. $a \in U_{j_0}$. U_{j_0} offen $\implies \exists \varepsilon > 0$ s. d. $B_\varepsilon^{\|\cdot\|_\infty}(a) \subset U_{j_0}$. Wähle m so groß, dass $\text{diam}(Q_m) < \varepsilon$. $a \in Q_m \implies Q_m \subset B_\varepsilon^{\|\cdot\|_\infty}(a) \subset U_{j_0} \nrightarrow$ Widerspruch Konstruktion der Q_m . \square

Bemerkung 2.6. Aus 2.4 folgt: Q ist bezüglich jeder Norm kompakt. Bolzano-Weierstraß (1.43) \implies In $(\mathbb{R}^n, \|\cdot\|)$ hat jede beschränkte Folge eine konvergente Teilfolge.

Bemerkungen 2.7. Wir haben bereits gesehen:

- i) Auf nicht endlich-dimensionalen Vektor-Räumen sind nicht alle Normen äquivalent: $(C([a, b]), \|\cdot\|_\infty)$ ist vollständig, $(C([a, b]), \|\cdot\|_{L^1})$ nicht.
- ii) Auf dem \mathbb{R}^n sind nicht alle Metriken äquivalent: d_{disc} ist stärker als jede Norm (und nicht schwächer).

Satz 2.8 (Heine-Borel). Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist. (\mathbb{R}^n hier und im Folgenden als normierter VR).

Beweis. „ \implies “ Hatten wir letztes Mal (1.38) für Kompakte in metrischen Räumen bewiesen.

„ \impliedby “ Ist A beschränkt so ist A in einem Quader enthalten (denn $\|x - y\|_\infty \leq \|x - y\|$ somit $\text{diam}_{\|\cdot\|_\infty}(A) \leq C \text{diam}_{\|\cdot\|}(A) < \infty$). Q ist kompakt (bezüglich $\|\cdot\|_\infty$ somit bezüglich $\|\cdot\|$). A abgeschlossen $\implies A$ kompakt (1.39). \square

Bemerkung. 2.8 gilt nicht in unendlich-dimensionalen Vektorräumen:

Betrachte in ℓ_1 , $\|\cdot\|_{\ell_1} = \sum_{k=0}^{\infty} |x_k|$ die Folge $(x^n)_n$ wobei $x^n = (x_k^n)_k$ sei mit $x_k^n = 0$ für $n \neq k$ und $(x^n)_n = 1$. Dann gilt $\|x^n\|_{\ell_1} = 1$ und

$$\|x^n - x^m\|_{\ell_1} = 2 \quad \forall m \in \{0, 1, \dots, n-1\}.$$

\implies Die Folge besitzt keine konvergente Teilfolge, kann also (Bolzano-Weierstrass) nicht kompakt sein, obwohl $\{x^n \mid n \in \mathbb{N}_0\}$ beschränkt und abgeschlossen in $(\ell_1, \|\cdot\|_{\ell_1})$ ist.

Vorlesung 5

Mo 04.05. 10:15

Stetige Abbildungen in normierten Vektorräumen

Lineare Abbildungen

Satz 2.9. Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Vektorräume. Sei $A: V \rightarrow W$ linear. Dann sind folgende Aussagen äquivalent:

- a) A ist stetig
- b) A ist stetig in 0
- c) $\|A(x)\|_W \leq C\|x\|_V$.

Beweis. 2.9.a) \implies 2.9.b) \checkmark

2.9.b) \implies 2.9.c) A stetig in 0 \implies zu $\varepsilon = 1 \exists \delta > 0$ s. d.

$$\|A(y) - A(0)\|_W \stackrel{\text{Lin}}{=} \|A(y)\|_W < 1 \quad \forall y \in V \text{ mit } \|y - 0\|_V = \|y\|_V < \delta.$$

Setze $C := 2/\delta$. Sei $x \in V \setminus \{0\}$ beliebig (für $x = 0$ gilt die Ungleichung 2.9.c) ohnehin).

Setze $\lambda := 1/C\|x\|_V$ und $y := \lambda x$.

Dann ist $\|y\|_V = \frac{1}{C\|x\|_V}\|x\|_V = \delta/2 < \delta$, also $\|A(y)\|_W < 1$.

$$A(y) = A(\lambda x) = \frac{1}{C\|x\|_V} A(x) \implies Beh.$$

2.9.c) \implies 2.9.a) Es gebe $C > 0$ s. d.

$$\|A(x)\|_W \leq C\|x\|_V \quad \forall x \in V.$$

Dann gilt insbesondere für $x = y - a$.

$$\|A(x)\|_W \stackrel{\substack{\uparrow \\ \text{Linearität}}}{=} \|A(y) - A(a)\| \leq C\|y - a\|_V.$$

Sei $\varepsilon > 0$. Dann ist also

$$\|A(y) - A(a)\|_W < \varepsilon \quad \forall y, a \text{ mit } \|y - a\|_V < \frac{\varepsilon}{C}$$

und somit ist A sogar gleichmäßig stetig. □

Beispiele. i) $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$.

$$I: C([a, b]) \rightarrow \mathbb{R}, \quad I(f) := \int_a^b f(t) dt.$$

I ist linear und es gilt

$$\|I(f)\| \leq (b - a)\|f\|_\infty$$

$\Rightarrow I$ ist stetig.

ii) $D: (C^1([a, b]), \|\cdot\|_\infty) \rightarrow (C([a, b]), \|\cdot\|_\infty)$, $D: f \mapsto f'$.

Behauptung. D ist nicht stetig.

Denn: D ist linear ✓, aber die Bedingung aus Satz 2.9 ist verletzt: Betrachte $f_n \in C^1([0, 2])$, $f_n = x^n$. Dann ist $\|f_n\|_\infty = 1$, aber $\|Df_n\|_\infty = n \Rightarrow$ es kann kein $C > 0$ geben s. d.

$$n = \|Df_n\|_\infty \leq C\|f_n\| = C \quad \forall n.$$

□

Definition. Seien V und W normierte Vektorräume. Sei $A: V \rightarrow W$ lineare stetige Abbildung. Die *Operatornorm* von A ist definiert als

$$\|A\|_{\text{op}} := \sup_{\substack{x \in V \\ x \neq 0}} \frac{\|Ax\|_W}{\|x\|_V}.$$

Auf dem VR der stetigen linearen Funktionen $V \rightarrow W$ ist $\|\cdot\|_{\text{op}}$ eine Norm. $\|A\|_{\text{op}}$ ist die kleinste Konstante für die noch die Abschätzung aus 2.9 gilt und es folgt

Bemerkung 2.10. Ein linearer Operator ist genau dann stetig, wenn gilt $\|A\|_{\text{op}} < \infty$.

Beispiel. Ist $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear, so gilt

$$A \in \text{Mat}(m \times n, \mathbb{R}) \simeq \mathbb{R}^{m \cdot n}.$$

Daher ist $\|\cdot\|_{\text{op}}$ in diesem Fall äquivalent zu $\|\cdot\|_\infty$, $\|A\|_\infty = \max_{i,j} |A_{ij}| < \infty$, insbesondere also schwächer und somit ist A stetig.

Konkret gilt: Setze $V = (\mathbb{R}^n, \|\cdot\|_V)$, $W = (\mathbb{R}^n, \|\cdot\|_W)$. Sei $y = Ax \implies y_i = \sum_{j=1}^n A_{ij}x_j$ für $i = 1, \dots, m$.

$$\begin{aligned}
 \|y\|_W &\stackrel{\Delta}{\leq} \sum_{i=1}^m \|y_i e_i\|_W \\
 &\stackrel{\Delta}{\leq} \sum_{i,j} |A_{ij}x_j| \|e_i\|_W \\
 &= \sum_{i,j} |A_{ij}| \cdot |x_j| \cdot \|e_i\|_W \\
 &\leq \|A\|_\infty \cdot \|x\|_{\ell^1} \cdot \underbrace{\sum_{i=1}^m \|e_i\|_W}_{=C_W} \\
 \implies \|A\|_{\text{op}} &= \sup_{x \neq 0} \frac{\|Ax\|_W}{\|x\|_V} \leq \|A\|_\infty \cdot C_W \cdot \underbrace{\sup_{x \neq 0} \frac{\|x\|_{\ell^1}}{\|x\|_V}}_{=C_V},
 \end{aligned}$$

wobei C_V eine Konstante ist mit

$$\|x\|_{\ell^1} \leq C_V \cdot \|x\|_V \quad \forall x \in V = \mathbb{R}^n.$$

Bemerkung. Unsere Beschränkung auf den \mathbb{R}^n (statt beliebige endlich-dimensionale Vektorräume zuzulassen), bedeutet also keine Einschränkung, da ein Basiswechsel nach der Überlegung oben stetig ist.

Beispiele 2.11. $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

a) *Kurven* $\gamma: I \rightarrow \mathbb{R}^n$, I Intervall, stetig.

Beispiele.

$\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$, $t \mapsto (r \cos t, r \sin t)$, $r > 0$. Stetigkeit: Wir versehen \mathbb{R}^2 mit $\|\cdot\|_\infty$. Dann folgt die Stetigkeit von γ aus der Stetigkeit der Komponentenfunktionen $I \rightarrow \mathbb{R}$.

$\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$, $t \mapsto (t^2 - 1, t^3 - 1)$ genauso. Spur von $\gamma = \{ \gamma(t) \mid t \in \mathbb{R} \}$



b) Gebrochen rationale Funktionen:

Beispiele.

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}.$$

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

f ist stetig: auf $\mathbb{R}^2 \setminus \{0\}$ sicherlich als Verknüpfung und Produkt stetiger Funktionen:

$$f = \text{Inv} \circ p_1 \cdot p_2 \quad \text{Inv}(t) = \frac{1}{t}, \quad p_1(x, y) = x^2 + y^2, \quad p_2(x, y) = x^2 y.$$

Stetigkeit in 0: Es gilt $(x - y)^2 \geq 0$

$$\implies 2|xy| \leq x^2 + y^2$$

$$\implies \left| \frac{x^2 y}{x^2 + y^2} \right| < \frac{|x|}{2}$$

für $((x_n, y_n))_n$, $(x_n, y_n) \rightarrow (0, 0)$ (bezüglich irgendeiner Norm) gilt insbesondere $x_n \rightarrow 0$

$$\implies |f(x_n, y_n) - 0| = \left| \frac{x_n^2 y_n}{x_n^2 + y_n^2} \right| < \frac{|x_n|}{2} \rightarrow 0 \text{ in } \mathbb{R}.$$

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

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

f ist stetig auf $\mathbb{R}^2 \setminus \{0\}$ (siehe oben). f ist nicht stetig in 0: Betrachte etwa $(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n^2}\right)$, $n \geq 1$. Dann gilt

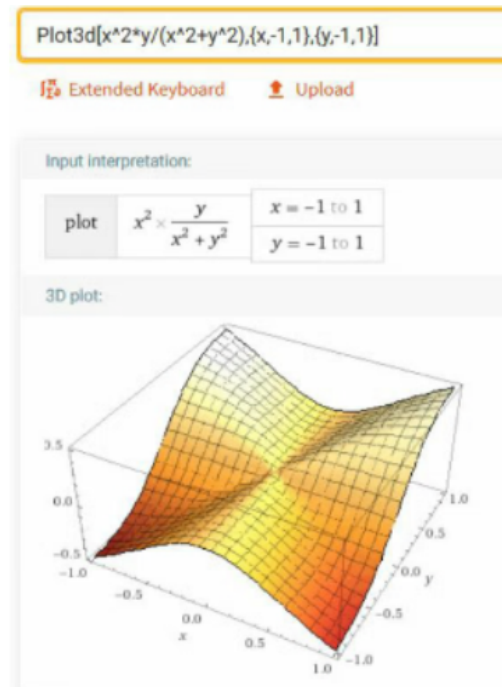
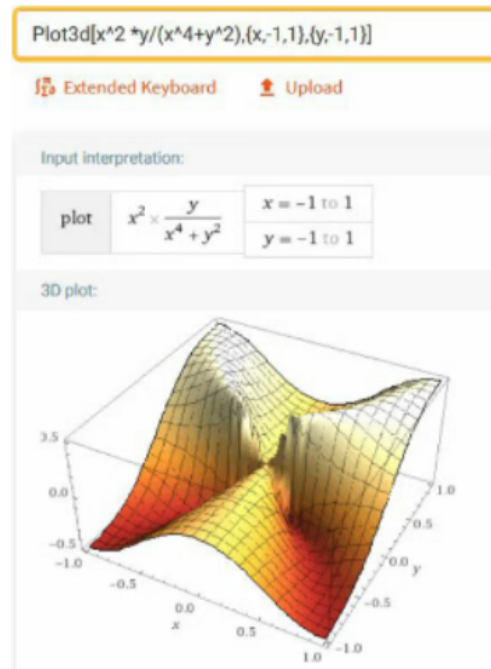
$$f(x_n, y_n) = \frac{1}{n^2 n^2} \left(\frac{n^4}{2} \right) = \frac{1}{2} \not\rightarrow 0.$$

Achtung:

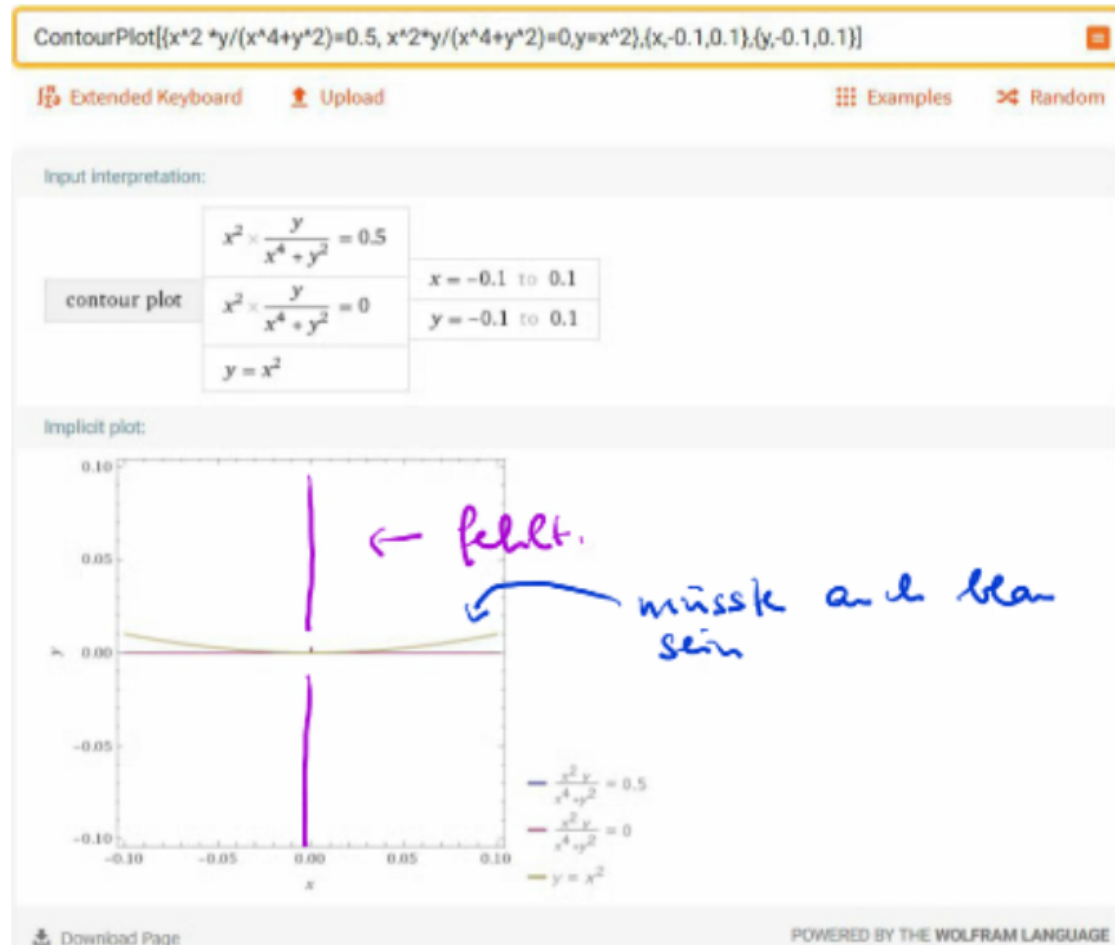
Es gibt durchaus Folgen $(x_n, y_n) \rightarrow 0$ s.d. $f(x_n, y_n) \rightarrow 0$ (für $n \rightarrow \infty$), z.B. $(x_n, y_n) = \left(0, \frac{1}{n}\right)$, wo $f\left(0, \frac{1}{n}\right) = 0 \quad \forall n$ oder $(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n}\right)$ wo

$$f(x_n, y_n) = \frac{1}{n^2} \left(\frac{n^2}{1 + 1/n^2} \right) \rightarrow 0.$$

Daher muss man, wenn man Stetigkeit zeigen will, in Argument finden, dass für *alle* Folgen funktioniert.



Contour-Plot: Eingezeichnet werden alle (x, y) , die die gegebene Gleichung erfüllen. Von Wolfram Alpha.



Vektorräume mit Skalarprodukt

Eine spezielle Klasse von Normen sind solche, die von einem sogenannten Skalarprodukt induziert werden.

Definition 2.12. Sei V ein Vektorraum über \mathbb{R} . Ein *Skalarprodukt* auf V ist eine Abbildung $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ mit

a) Linear:

$$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle \quad \forall x, y, z \in V, \lambda, \mu \in \mathbb{R}$$

b) Symmetrisch:

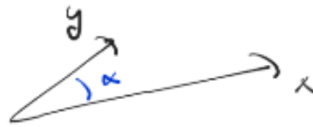
$$\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in V$$

c) Positiv definit:

$$\langle x, x \rangle \geq 0 \text{ und } \langle x, x \rangle = 0 \iff x = 0.$$

Bemerkung. Mit 2.12.b) folgt auch die Linearität im zweiten Argument.

Beispiele. • \mathbb{R}^n , $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$: Euklidisches Skalarprodukt.



interpretation

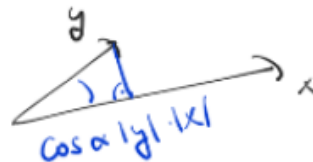
Y geht durch Drehstreckung aus $x \neq 0$ hervor:

$$y = \|y\|_E \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \frac{x}{\|x\|_E}.$$

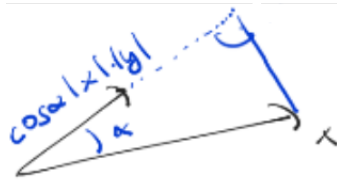
Dann gilt:

$$\begin{aligned} \langle x, y \rangle &= \frac{\|y\|_E}{\|x\|_E} \langle x, \begin{pmatrix} \cos \alpha x_1 - \sin \alpha x_2 \\ \sin \alpha x_1 + \cos \alpha x_2 \end{pmatrix} \rangle \\ &= \frac{\|y\|_E}{\|x\|_E} (x_1^2 \cos \alpha - \cancel{x_1 x_2 \sin \alpha} + \cancel{x_1 x_2 \sin \alpha} + \cancel{x_2^2 \cos \alpha}) \\ &= \|y\|_E \cdot \|x\|_E \cdot \cos \alpha. \end{aligned}$$

Das Skalarprodukt misst die Projektion von y auf x



und umgekehrt



- \mathbb{R}^n mit $\langle x, y \rangle_W = \sum_{i=1}^n w_i x_i y_i$, $w = (w_1, \dots, w_n)$ Gewichtsvektor, $w_i > 0$.
- \mathbb{R}^2 mit $\langle x, y \rangle := 2x_1 y_1 - x_1 y_2 - x_2 y_1 + 2x_2 y_2$ (zu überprüfen ist die Positive Definitheit).
- *Kein* Skalarprodukt ist das Minkowski-Produkt: \mathbb{R}^{n+1} mit $((x, y)) := x_0 y_0 - \sum_{i=1}^n x_i y_i$.
Denn $((x, x)) = 0 \iff x_0 = \pm \|x\|_E$, $x = (x_1, \dots, x_n)$.
- $C([a, b])$ mit $\langle f, g \rangle = \int_a^b f(t)g(t) dt$.

Lemma 2.13. Sei V VR mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann ist durch $\|x\| := \sqrt{\langle x, x \rangle}$ eine Norm auf V definiert.

Beweis. 2.1.a) $\|x = 0\| \implies \langle x, x \rangle = 0 \xrightarrow{2.12.c)} \|0\| = 0 \checkmark$.

2.1.b) $\|\lambda x\| = \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle} \forall \lambda \in \mathbb{R}, x \in V$.

2.1.c)

$$\begin{aligned}
 \|x + y\|^2 &= \langle x + y, x + y \rangle \\
 &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \\
 &\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\
 &\leq (\|x\| + \|y\|)^2 \implies \Delta, \\
 &\quad \uparrow \\
 &\text{siehe } (*) \text{ unten}
 \end{aligned}$$

denn die Wurzel ist monoton wachsend.

Es gilt die *Cauchy-Schwarzsche* Ungleichung:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|. \quad (*)$$

Beweis.

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2 \quad \forall x, y \in V \quad \lambda \in \mathbb{R},$$

also speziell für $y \neq 0$ (für $y = 0$ gilt die Ungleichung sowieso) und $\lambda = \frac{\langle x, y \rangle}{\|y\|^2}$:

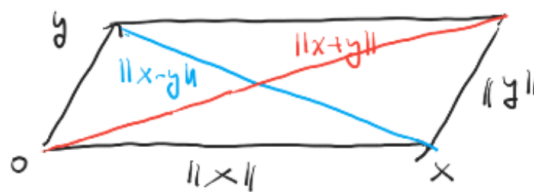
$$0 \leq \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2}. \quad \square$$

Einen Vektorraum mit Skalarprodukt betrachten wir immer als mit der von Skalarprodukt induzierten Norm, also Metrik, also Topologie.

Nicht jede Norm wird von einem Skalarprodukt induziert. Es gilt

Lemma 2.14. Sei $(V, \|\cdot\|)$ normierter VR. Dann wird $\|\cdot\|$ von einem Skalarprodukt induziert genau dann, wenn die Parallelogramm-Gleichung gilt:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall x, y \in V.$$



Erklärung für den Namen. \mathbb{R}^2 , $\|\cdot\| = \|\cdot\|_{\mathbb{E}}$.

Beweis. „ \Rightarrow “ Sei $\|x\| = \sqrt{\langle x, x \rangle}$. Dann gilt

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= 2\|x\|^2 + 0 + 2\|y\|^2. \end{aligned}$$

„ \Leftarrow “ Erfülle $\|\cdot\|$ die Parallelogramm-Gleichung.

Behauptung. Durch „Polarisation“, also

$$\langle x, y \rangle := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

ist ein Skalarprodukt definiert mit $\|x\| = \sqrt{\langle x, x \rangle}$.

Beweis. 2. Beh.: $\langle x, x \rangle = \frac{1}{4} \|2x\|^2 \checkmark$.

1. Beh.: • Aus der 2. Beh. folgt die positive Definitheit aus der Nichtausgeartetheit und Positivität der Norm.

- Die Symmetrie folgt sofort aus der Definition.
- Linearität. Wir zeigen zunächst Additivität:

$$1) \quad \langle x + u, y \rangle + \langle x - u, y \rangle = \langle x, y \rangle$$

denn:

$$\begin{aligned} \text{linke Seite} &= \frac{1}{4} (\|x + u + y\|^2 - \|x + u - y\|^2 + \|x - u + y\|^2 - \|x - u - y\|^2) \quad \square \\ &= \frac{1}{2} (\|x + y\|^2 + \|u\|^2 - \|x - y\|^2 - \|u\|^2) \\ &\quad \uparrow \\ &\quad \text{Parallelogramm-Gleichung} \\ &= \frac{1}{2} (\langle x + y, x + y \rangle - \langle x - y, x - y \rangle) \\ &= 2\langle x, y \rangle. \end{aligned}$$

Damit auch gleich gezeigt:

$$2) \quad \langle 2x, y \rangle = 2\langle x, y \rangle \text{ (setze } u = x \text{) und mit } x = u + v, y = u - v \text{ folgt}$$

3) Additivität:

$$\begin{aligned} \langle x, y \rangle + \langle y, z \rangle &= \langle u + v, z \rangle + \langle u - v, z \rangle \\ &= 2\langle u, z \rangle \\ &= \langle 2u, z \rangle \\ &\quad 2) \\ &= \langle x + y, z \rangle \end{aligned}$$

4) per Induktion $\langle nx, y \rangle = n\langle x, y \rangle \quad \forall n \in \mathbb{N}$, denn

$$\begin{aligned} \langle (n + 1)x, y \rangle &= \langle nx + x, y \rangle \\ &\stackrel{3)}{=} \langle nx, y \rangle + \langle x, y \rangle \\ &\stackrel{\text{IV}}{=} n\langle x, y \rangle + \langle x, y \rangle \\ &= (n + 1)\langle x, y \rangle. \end{aligned}$$

5) Für $\lambda \in -\mathbb{N}_0$ gilt

$$\begin{aligned}\lambda \langle x, y \rangle - \langle \lambda x, y \rangle &= \lambda \langle x, y \rangle - \langle |\lambda|(-x), y \rangle \\ &= \lambda \langle x, y \rangle - |\lambda| \langle -x, y \rangle \\ &= \lambda (\langle x, y \rangle + \langle -x, y \rangle) \\ &= 0\end{aligned}$$

6) Für $\lambda \in \mathbb{Q}$, $\lambda = m/n$, $m, n \in \mathbb{Z}$:

$$n \langle \frac{m}{n} x, y \rangle \stackrel{4), 5)}{=} \langle mx, y \rangle = m \langle x, y \rangle.$$

7) Für $\lambda \in \mathbb{R}$ existiert $(\lambda_n)_n \subset \mathbb{Q}$, $\lambda_n \rightarrow \lambda$. Da $\|\cdot\|$ stetig ist, so auch $\langle \cdot, \cdot \rangle$

$$\begin{aligned}\langle \lambda x, y \rangle &= \langle \lim \lambda_n x, y \rangle \\ &= \lim \langle \lambda_n x, y \rangle \\ &= \lim \lambda_n \langle x, y \rangle \\ &= \lambda \langle x, y \rangle.\end{aligned}$$

Symmetrie \implies es genügt, das erste Argument zu untersuchen. □

Beispiel. $\|\cdot\|_{\max}$ wird nicht von einem Skalarprodukt induziert: Sei $x = e_1$, $y = e_2$. Dann gilt:

$$\|e_1 + e_2\|_{\max}^2 + \|e_1 - e_2\|_{\max}^2 = 1 + 1 = 2,$$

aber

$$2(\|e_1\|_{\max}^2 + \|e_2\|_{\max}^2) = 4.$$

Kapitel 3

Differenzierbarkeit in \mathbb{R}^n

Vorlesung 6

Do 07.05. 10:15

Erinnerung. Approximation einer Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, die in $a \in \mathbb{R}$ differenzierbar ist, durch eine (affin) lineare Funktion

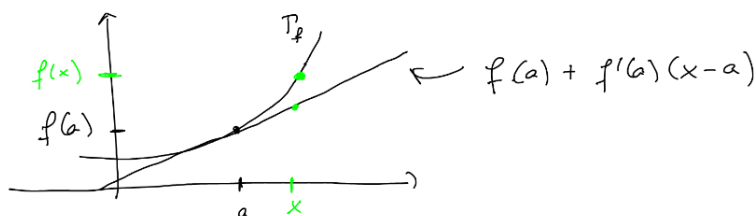
$$f(x) = f(a) + m_a(x - a) + R_a(x)$$

mit $R_a: \mathbb{R} \rightarrow \mathbb{R}$ und $\lim_{x \rightarrow a} \frac{R_a(x)}{x-a} = 0$.

Gibt es ein solches R_a , so ist m_a eindeutig festgelegt und es gilt

$$m_a = \lim_{x \rightarrow a} \underbrace{\frac{f(x) - f(a)}{x - a}}_{\text{Differenzenquotient}}.$$

$f'(a) := m_a$ heißt *Ableitung* von f an der Stelle a .



Für Abbildungen $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ kann man analog definieren:

Definition 3.1. Die *Ableitung* einer Funktion $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, an der Stelle $a \in U$, ist, wenn sie existiert, eine Matrix $Df(a) \in \text{Mat}(m \times n)$, die eine lineare Approximation von f ergibt:

$$f(x) = f(a) + \underbrace{Df(a)}_{\text{Matrix-Multiplikation}} \cdot (x - a) + R_a(x) \quad (*)$$

Sie existiert genau dann, wenn

$$\begin{array}{c} \text{in } \mathbb{R}^m \\ \downarrow \\ \lim_{x \rightarrow a} \frac{R_a(x)}{\|x - a\|} = 0 \\ \uparrow \\ \text{in } \mathbb{R}^n \quad \uparrow \\ \text{eine Norm in } \mathbb{R}^n \end{array}$$

und ist in diesem Fall eindeutig durch $(*)$ bestimmt und man sagt, f ist *differenzierbar*.

Bemerkungen. i) Für $n = m = 1$ stimmt die Definition mit der Üblichen überein, da $\frac{R_a(x)}{x-a} \rightarrow 0 \iff \frac{R_a(x)}{|x-a|} \rightarrow 0$.

ii) Eindeutigkeit: Sei A s. d.

$$\begin{aligned} f(x) &= f(a) + A \cdot (x - a) + \tilde{R}_a(x) \\ &= f(a) + Df(a) \cdot (x - a) + R_a(x) \end{aligned}$$

Dann folgt:

$$\lim_{x \rightarrow a} \underbrace{((A - Df(a)) \cdot (x - a) \cdot 1 / \|x - a\|)}_{= \frac{1}{\|x-a\|} (R_a(x) - \tilde{R}_a(x))} = 0.$$

$\implies A - Df(a) = 0$ (Nullmatrix), da wegen der Stetigkeit linearer Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ der Grenzwert gleich $(A - Df(a)) \cdot \lim (x - a) / \|x - a\|$ ist und $\|\lim (x - a) / \|x - a\|\| = 1 \neq 0$.

iii) Wie in der DIFF I ist es oft zweckmäßig f an der Stelle $a + h$ mit f an der Stelle a zu vergleichen ($x = a + h$).

$$f(a + h) = f(a) + Df(a) \cdot h + \underline{R}_h(a)$$

mit $\lim_{h \rightarrow 0} (\underline{R}_h(x)) / \|h\| = 0$.

Beispiele. i) $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f(x) = A \cdot x + b$, $A \in \text{Mat}(m \times n, \mathbb{R})$, $b \in \mathbb{R}^m$.

$$f(a + h) = f(a) + A \cdot h \implies Df(a) = A.$$

Insbesondere verschwindet die Ableitung einer konstanten Funktion.

ii) $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \langle x, B \cdot x \rangle$, $\langle \cdot, \cdot \rangle$ euklidisches Skalarprodukt, $B \in \text{Mat}(n \times n, \mathbb{R})$.

$$\begin{aligned} f(a + h) &= \langle a + h, B \cdot (a + h) \rangle \\ &= \langle a, B \cdot a \rangle + \langle h, B \cdot a \rangle + \langle a, B \cdot h \rangle + \langle h, B \cdot h \rangle \\ &= \langle a, B \cdot a \rangle + \langle (B + B^T) \cdot a, h \rangle + \langle h, B h \rangle. \\ &\quad \uparrow \\ &\text{CHECK!} \end{aligned}$$

Wegen (wähle $\|\cdot\| = \|\cdot\|_{\mathbb{E}}$)

$$\frac{\langle a, Bh \rangle}{\|h\|} \underset{\text{C-S}}{\leq} \frac{\|h\| \cdot \|Bh\|}{\|h\|} \leq \|B\|_{\text{op}} \cdot \|h\| \rightarrow 0 \text{ für } h \rightarrow 0$$

folgt: f ist in allen $a \in \mathbb{R}^n$ differenzierbar und

$$Df(a) \cdot h = \langle (B + B^T) \cdot a, h \rangle = (b_1, \dots, b_n) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \\ \implies Df(a) = b = ((B + B^T) \cdot a)^T \in \text{Mat}(1 \times n, \mathbb{R}) \quad \forall h \in \mathbb{R}^n.$$

Aus der Definition folgt sofort

Satz 3.2. Sei $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, in $a \in U$ differenzierbar. Dann ist f in a stetig.

Beweis.

$$\begin{aligned} \lim_{h \rightarrow 0} f(a+h) &= \lim_{h \rightarrow 0} (f(a) + Df(a) \cdot h + \underline{R}_a(h)) && \square \\ &= f(a) + 0 + 0 \\ &\quad \uparrow \quad \uparrow \quad \uparrow \\ &\quad \text{Es gilt sogar } \underline{R}_a(h) / \|h\| \rightarrow 0 \\ &\quad \|Df(a) \cdot h\| \leq \|Df(a)\|_{\text{op}} \cdot \|h\| \\ &\quad \uparrow \quad \uparrow \\ &\quad \text{Norm in } \mathbb{R}^m \quad \text{Norm in } \mathbb{R}^n \end{aligned}$$

Satz 3.3 (Kettenregel). Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen, $g: U \rightarrow \mathbb{R}^m$, $f: V \rightarrow \mathbb{R}^k$, $g(U) \subset V$. Ist g in $a \in U$ differenzierbar und f in $b = g(a)$, so ist die Verkettung $f \circ g: U \rightarrow \mathbb{R}^k$ in a differenzierbar und es gilt

$$D(f \circ g)(a) = \underbrace{Df(g(a))}_{\substack{\uparrow \\ \in \text{Mat}(k \times m, \mathbb{R})}} \underbrace{Dg(a)}_{\substack{\uparrow \\ \in \text{Mat}(m \times n, \mathbb{R})}}.$$

Beweis.

$$g(a+u) = g(a) + A \cdot u + \underline{R}_a^g(u) \quad A = Dg(a) \quad (1)$$

$$f(b+v) = f(b) + B \cdot v + \underline{R}_b^f(v) \quad B = Df(b) \quad (2)$$

Setze speziell $v := g(a + u) - g(a) \stackrel{(1)}{=} A \cdot u + \underline{R}_a^g(u)$.

$$\begin{aligned} \implies f \circ g(a + u) &= f(g(a + u)) = f(g(a) + v) && \text{Def } v \\ &= f(g(a)) + B \cdot v + \underline{R}_b^f(v) && b = g(a) \\ &= f(g(a)) + B \cdot A \cdot u + \underbrace{B \cdot \underline{R}_a^g(u) + \underline{R}_b^f(A \cdot u + \underline{R}_a^g(u))}_{\substack{\uparrow \\ \text{zu zeigen: } \frac{\dot{u}}{\|u\|} \rightarrow 0 \text{ f\"ur } u \rightarrow 0}}. \\ &\quad \uparrow \\ &\quad \text{Def } v \text{ und (1)} \end{aligned}$$

- $\frac{\underline{R}_a^g(u)}{\|u\|} \rightarrow 0 \implies \exists C > 0$ s. d. $\|\underline{R}_a^g(u)\| \leq C\|u\|$.
- $\frac{\underline{R}_b^f(v)}{\|v\|} \rightarrow 0 \implies \exists \underline{r}_b^f$ s. d. $\underline{R}_b^f(v) = \|v\|\underline{r}_b^f(v)$ mit $\underline{r}_b^f(v) \rightarrow 0$ ($v \rightarrow 0$).

$$\begin{aligned} \implies \underline{R}_b^f(A \cdot u + \underline{R}_a^g(u)) &\leq \overbrace{A \cdot u + \underline{R}_a^g(u)}^{\leq (\|A\|_{\text{op}} + C)\|u\|} \cdot \underbrace{\|\underline{r}_b^f(Au + \underline{R}_a^g(u))\|}_{\rightarrow 0 \text{ f\"ur } u \rightarrow 0} \\ &\quad \uparrow \quad \quad \quad \uparrow \\ &\quad \text{Norm in } \mathbb{R}^k \quad \quad \text{Norm in } \mathbb{R}^k \\ \implies \frac{\underline{R}_b^f(A \cdot u + \underline{R}_a^g(u))}{\|u\|} &\rightarrow 0 \quad (u \rightarrow 0). \quad \square \end{aligned}$$

Satz 3.4 (Produktregel, Quotientenregel). Seien $f, g: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, differenzierbar in $a \in U$. Dann gilt

i) $f \cdot g$ ist differenzierbar in a und es gilt

$$D(f \cdot g)(a) = Df(a) \cdot g(a) + f(a) \cdot Dg(a).$$

ii) Ist $g(a) \neq 0$ so gilt: f/g ist auf einer Umgebung von a definiert und differenzierbar in a und es gilt

$$D(f/g)(a) = Df(a) \cdot \frac{1}{g(a)} - f(a) \cdot \frac{1}{g(a)^2} \cdot Dg(a).$$

Beweis. 3.4.i)

$$\begin{aligned} f(a + h) &= f(a) + Df(a) \cdot h + \underline{R}_a^f(h) \\ g(a + h) &= g(a) + Dg(a) \cdot h + \underline{R}_a^g(h) \\ f \cdot g(a + h) &= (f(a) + Df(a) \cdot h + \underline{R}_a^f(h)) (g(a) + Dg(a) \cdot h + \underline{R}_a^g(h)) \\ &= f(a) \cdot g(a) + \underbrace{(Df(a) \cdot g(a))}_{\in \text{Mat}(1 \times n, \mathbb{R})} \cdot h + \underbrace{(f(a) \cdot Dg(a))}_{\in \text{Mat}(1 \times n, \mathbb{R})} \cdot h \\ &\quad + \underbrace{(Df(a) \cdot h + \underline{R}_a^f(h))(Dg(a) \cdot h + \underline{R}_a^g(h))}_{\frac{\dot{h}}{\|h\|} \rightarrow 0 \text{ f\"ur } h \rightarrow 0}. \end{aligned}$$

3.4.ii) g ist in a stetig $\implies \exists$ Umgebung von a s.d. $g(x) \neq 0 \forall x \in U$ (wie in der DIFF I: Sei o.B.d.A. $g(a) > 0$. Sei $\varepsilon := g(a)/2$. Sei $\|\cdot\|$ irgendeine Norm auf \mathbb{R}^n . Dann gibt es ein $\delta > 0$ s.d.

$$\begin{aligned} |g(x) - g(a)| &< \varepsilon \quad \forall x \in B_\delta^{\|\cdot\|}(a) \\ \implies -\frac{g(a)}{2} &< g(x) - g(a) < \frac{g(a)}{2} \quad \forall x \in B_\delta^{\|\cdot\|}(a), \end{aligned}$$

also $0 < \frac{g(a)}{2} < g(x) < \frac{3}{2} \frac{g(a)}{2}$.) \implies Auf U ist f/g wohldefiniert. Die Berechnung der Ableitung ist analog zu 3.4.i), nachdem man sich überlegt hat, dass

$$\frac{1}{g} = \text{Inv} \circ g, \quad \text{Inv}(t) = \frac{1}{t}$$

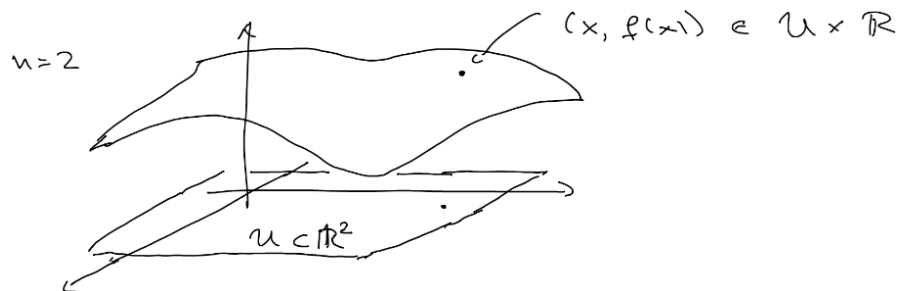
und somit nach der Kettenregel

$$D\left(\frac{1}{g}\right)(a) = D\text{Inv}(g(a)) \cdot Dg(a)$$

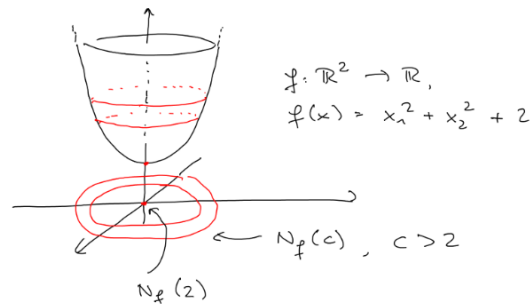
und $D\text{Inv}(t) = -\frac{1}{t^2}$ (DIFF I). □

Geometrische Anschauung, partielle Ableitung

DIFF I: Ableitung beschreibt Rate der Veränderung. Höher-dimensional: Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$. Betrachte den Graph $\Gamma_f = \{(x, f(x)) \mid x \in U\}$ (vgl. die Diskussion bei 2.11.b))



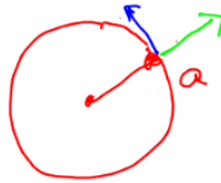
Definition (Niveau-Mengen). Zu $c \in \mathbb{R}$ setze $N_f(c) := \{x \in U \mid f(x) = c\}$.



Es gilt $N_f(c) = \emptyset$ für $c < 2$.

Beispiel. Entlang der Niveauflächen ist f konstant. Wie wird das in der Ableitung sichtbar?

Im Beispiel oben ist $Df(a) = (2a_1, 2a_2)$ (check!). Sei $a = (r \cos \phi, r \sin \phi)$, $r > 0$. Betrachte $f(a + h) = f(a) + Df(a) \cdot h + \underline{R}_a(h)$. Ist $h = (-\varepsilon \sin \phi, \varepsilon \cos \phi)$, $\varepsilon > 0$ (in „Richtung“ der Niveaumenge), so ist $Df(a) \cdot h = 0$.



Ist dagegen $h = (\varepsilon \cos \phi, \varepsilon \sin \phi)$ (von der Niveaumenge weg), so ist $Df(a) \cdot h = 2r\varepsilon > 0$. Das wollen wir im Folgenden systematisch studieren.

Definition. Sei $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, gegeben. Für $a \in U$ und $v \in \mathbb{R}^n$ heißt der Grenzwert (falls er existiert)

$$\partial_v f(a) := \lim_{\substack{t \rightarrow 0 \\ \uparrow \\ \text{in } \mathbb{R}^m}} \frac{f(a + tv) - f(a)}{t}$$

die *Richtungsableitung* von f in a in Richtung v .



Bemerkungen / Beispiele 3.5. i) $\partial_v f(a) = \frac{d}{dt}|_{t=0} g_{a,v}(t) = g'_{a,v}(0)$, $g_{a,v}(t) = f(a + tv)$.

ii) $f = (f_1, \dots, f_m)^T$, so gilt

$$\partial_v f(a) = (\partial_v f_1(a), \dots, \partial_v f_m(a))^T$$

iii) $f: (x_1, x_2) \mapsto x_1^2 + x_2^2 + 2$. Sei $a = (r \cos \phi, r \sin \phi)$. Ist $v = (\varepsilon \cos \phi, \varepsilon \sin \phi)$, so ist

iv)

$$g_{a,v}(t) = ((r + \varepsilon t) \cos \phi)^2 + ((r + \varepsilon t) \sin \phi)^2 + 2$$

und somit $\partial_v f(a) = g'_{a,v}(0) = 2(r + \varepsilon \cdot 0)\varepsilon = 2r\varepsilon$.

Ist $v = (-\varepsilon \sin \phi, \varepsilon \cos \phi)$, so ist

$$g_{a,v}(t) = (r \cos \phi - \varepsilon t \sin \phi)^2 + (r \sin \phi + \varepsilon t \cos \phi)^2$$

und somit

$$\partial_v f(a) = g'_{a,v}(0) = 2(r \cos \phi - 0)(-\varepsilon \sin \phi) + 2(r \sin \phi + 0)(\varepsilon \cos \phi) = 0.$$

Das ist kein Zufall, denn es gilt

Satz 3.6. Sei $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, in $a \in U$ differenzierbar. Dann besitzt f die Richtungsableitungen

$$\partial_v f(a) = Df(a) \cdot v \quad \forall v \in \mathbb{R}^n$$

und $Df(a)$ hat bezüglich der Standardbasis e_1, \dots, e_n die Matrix Darstellung

$$(\partial_1 f \dots \partial_n f) = \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \dots & \partial_n f_2(a) \\ \vdots & & \vdots \\ \partial_1 f_m(a) & \dots & \partial_n f_m(a) \end{pmatrix} \in \text{Mat}(m \times n, \mathbb{R}),$$

„Jacobi-Matrix“, wobei $\partial_j f(a) = \partial_{e_j} f(a) = Df(a) \cdot e_j = j$ -te Spalte.

Beweis. Für $v = 0$ sind beide Seiten 0 ✓.

Für $v \neq 0$ betrachte

$$\begin{aligned} & \| (f(a + tv) - f(a)) / t - Df(a) \cdot v \| \sim_{(\text{Norm auf } \mathbb{R}^m)} \\ &= \frac{1}{|t|} \| f(a + tv) - (f(a) + Df(a) \cdot (tv)) \| \sim \\ &= \frac{1}{\underset{\substack{\uparrow \\ \text{Homogenität} \quad \|\cdot\| \\ \text{Norm auf } \mathbb{R}^n}}{\|t \cdot v\|}} \| f(a + tv) - (f(a) + Df(a) \cdot (tv)) \| \sim \cdot \|v\|. \end{aligned}$$

Differenzierbarkeit \implies strebt gegen 0 für $tv \rightarrow 0$, also für $t \rightarrow 0 \implies$ Die Richtungsableitungen existieren und

$$\partial_v f(a) = Df(a) \cdot v$$

und bezüglich der kanonischen Basen gilt

$$\begin{aligned} \partial_j f(a) &= Df(a) \cdot e_k = j\text{-te Spalte von } Df(a). \\ &\parallel \\ &\begin{pmatrix} \partial_j f_1(a) \\ \vdots \\ \partial_j f_m(a) \end{pmatrix} \end{aligned} \quad \square$$

Achtung. Umgekehrt genügt die Existenz der Richtungsableitungen $\partial_1 f, \dots, \partial_n f$ nicht, um Differenzierbarkeit zu garantieren!

Beispiel.

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

$\partial_1 f(x, y) = \frac{d}{dt} \big|_{t=0} f(x + t, y)$, $\partial_2 f(x, y) = \frac{d}{dt} \big|_{t=0} f(x, y + t)$. Wegen $f(x, 0) = 0 \quad \forall x$, $f(0, y) = 0 \quad \forall y$, ist $\partial_1 f(0, 0) = 0 = \partial_2 f(0, 0)$. Aber f ist in 0 nicht stetig (betrachte etwa $(1/n, 1/n)$), also nicht differenzierbar.

Vorlesung 7

Mo 11.05. 10:15

Beispiele und Erläuterungen

Wir hatten letztes Mal gesehen, dass, wenn $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, in $a \in U$ differenzierbar ist, dass dann die Ableitung mit Hilfe der *partiellen Ableitungen*, also der Richtungsableitungen in Richtung der kanonischen Basis geschrieben werden kann,

$$Df(a) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix},$$

dass aber die Existenz der partiellen Ableitungen nicht unbedingt Differenzierbarkeit garantiert.

Konkret kann man also so vorgehen: Man bestimmt die partiellen Ableitungen und überprüft dann, ob Differenzierbarkeit vorliegt.

Bemerkung 3.7. Berechnung von partiellen Ableitungen. Es gilt $\partial_j f(a) = g'_{a \circ e_j}(0)$, wobei

$$g_{a \circ e_j}(t) = f(a + te_j) = f(a_1, \dots, a_j + t, \dots, a_n) \quad t \in (-\nu, \varepsilon)$$

s. d. $a + te_j$ noch im Definitionsbereich von f liegt.

\Rightarrow Um $\partial_j f(a)$ zu berechnen, kann man die gewöhnlichen Ableitungen bezüglich der j -ten Koordinate bestimmen (und steht stellt sich die Übrigen als Konstanten vor).

Beispiele 3.8. i) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x_1^2 + x_2^2 + 2$.

$$\partial_1 f(a) = 2a_1 \quad \partial_2 f(a) = 2a_2$$

f ist in der Tat differenzierbar in allen $a \in \mathbb{R}^2$, denn

$$\begin{aligned} f(a+h) &= (a_1 + h_1)^2 + (a_2 + h_2)^2 + 2 \\ &= f(a) + A \cdot h + \underbrace{\|h\|_{\mathbb{E}}^2}_{R_a(h)}. \end{aligned}$$

mit $A = \begin{pmatrix} 2a_1 & 2a_2 \end{pmatrix}$ und

$$\frac{R_a(h)}{\|h\|} \leq C\|h\| \rightarrow 0.$$

ii) $f: \mathbb{R}^n \rightarrow \mathbb{R}, n \geq 2, f(x) = \|x\|_{\mathbb{E}} = \sqrt{\sum_{i=1}^n x_i^2}$.

$a \neq 0$ Mittels Kettenregel aus DIFF I:

$$\partial_j f(a) = \frac{1}{2} \frac{1}{\|x\|_{\mathbb{E}}} \cdot 2x_j|_{x=a} = \frac{a_j}{\|a\|_{\mathbb{E}}}$$

$a = 0$

$$\partial_j f(0) = \frac{d}{dt} f(0, \dots, \underset{j\text{-te}}{t}, \dots, 0) = \frac{d}{dt} |_{t=0} |t|.$$

\implies Die partiellen Ableitungen in $a = 0$ existieren nicht (DIFF I: Für $h \in \mathbb{R}$ ist $|h| - 0 / h = \pm 1 \not\rightarrow 0$ für $h \searrow 0$ bzw. $h \nearrow$). $\implies f$ ist *nicht* differenzierbar in $a = 0$.

Für $a \neq 0$ gilt jedoch: $\sqrt{\cdot}: \mathbb{R}_{>0} \rightarrow \mathbb{R}$ ist differenzierbar (DIFF I) und ebenso die polynomiale Funktion $x \mapsto \sum_{i=1}^n x_i^2 \implies f$ ist differenzierbar auf $\mathbb{R}^n \setminus \{0\}$.

Wir hätten $Df(a), a \neq 0$ auch mit der höherdimensionalen Kettenregel bestimmen können:

$$Df(a) = Dw(p(a)) \cdot Dp(a) = \frac{1}{2\sqrt{p(a)}} \cdot (2a_1 \dots \dots 2a_n)$$

$$w: \mathbb{R}_{>0} \rightarrow \mathbb{R}, w(t) = \sqrt{t}, p(a) = \sum_{i=1}^n a_i^2.$$

iii) Eine weitere Anwendung der Kettenregel. Betrachte die „Polarkoordinaten“

$$g: \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \mathbb{R}^2$$

$$g(r, \phi) = (r \cos \phi, r \sin \phi)^T$$



und $f: \mathbb{R}^2 \rightarrow \mathbb{R}$.

$$D(f \circ g)(r, \phi) = Df(g(r, \phi)) \cdot Dg(r, \phi)$$

$$Dg(r, \phi) = \begin{pmatrix} \partial_1 g(r, \phi) & \partial_2 g(r, \phi) \end{pmatrix} = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}.$$

Somit

$$\begin{aligned} D(f \circ g)(r, \phi) &= \begin{pmatrix} \partial_1 f(g(r, \phi)) & \partial_2 f(g(r, \phi)) \end{pmatrix} \cdot \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} \\ &= \begin{pmatrix} \partial_1 f(g(r, \phi)) \cdot \cos \phi + \partial_2 f(g(r, \phi)) \cdot \sin \phi \\ -\partial_1 f(g(r, \phi)) \cdot r \sin \phi + \partial_2 f(g(r, \phi)) \cdot r \cos \phi \end{pmatrix}^T. \end{aligned}$$

Man schreibt dafür manchmal

$$\begin{aligned} \partial_r &= \cos \phi \partial_x + \sin \phi \partial_y \\ \partial_\phi &= -r \sin \phi \partial_x + r \cos \phi \partial_y. \end{aligned}$$

Aber oft ist es *viel* übersichtlicher, die partiellen Ableitungen *nicht* nach den Namen der Variablen zu benennen, sondern durchzunummerieren!

Beispiel. $f(x, y) = \frac{xy}{x^2+y^2}$ auf $U = \mathbb{R}^2 \setminus \{0\}$.

$$\begin{aligned} Df(x, y) &= \begin{pmatrix} \frac{y(y^2-x^2)}{(x^2+y^2)^2} \\ \frac{x(x^2-y^2)}{(x^2+y^2)^2} \end{pmatrix} \\ D(f \circ g)(r, \phi) &= \frac{1}{r} (\cos^2 \phi - \sin^2 \phi) \begin{pmatrix} \overbrace{-\sin \phi \cos \phi + \cos \phi \sin \phi}^{=0} \\ \underbrace{r \sin^2 \phi + r \cos^2 \phi}_{=r} \end{pmatrix} \\ &= (0, \cos^2 \phi - \sin^2 \phi). \end{aligned}$$

In diesem Fall rechnet man allerdings schneller direkt:

$$D(f \circ g)(r, \phi) = D(\tilde{f})(r, \phi) = \cos \phi \cdot \sin \phi.$$

Bemerkung / Definition 3.9. Wir hatten allgemeiner gesehen:

Ist $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, differenzierbar in a , so gilt

$$\partial_v f(a) \cdot Df(a) \cdot v \quad \forall v \in \mathbb{R}^n.$$

Ist speziell $m = 1$, so definiert man

$$\text{grad } f(a) := Df(a)^T = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \in \mathbb{R}^n$$

„Gradient von f in a “, und schreibt $\partial_v f(a) = \langle \text{grad } f(a), v \rangle$.

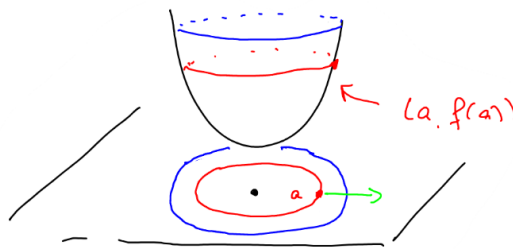
Ist $\text{grad } f(a) \neq 0$ und $\|v\|_E = 1$, so ist

$$\langle \text{grad } f(a), v \rangle = \|\text{grad } f(a)\|_E \cdot \cos \alpha,$$

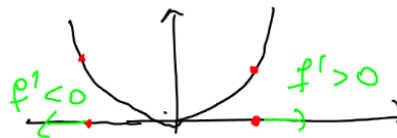
wobei α der zwischen $\text{grad } f(a)$ und v in \mathbb{R}^n eingeschlossene Winkel ist (in der durch die beiden Vektoren aufgespannten Ebene).

Es folgt: $\partial_v f(a)$ ist dann am größten, wenn v in die selbe Richtung zeigt wie $\text{grad } f(a)$
 \implies der Gradient gibt die Richtung stärksten Anstiegs von f in a an.

Beispiel. $f(x) = x_1^2 + x_2^2 + 2$.



Beispiel. $f(x) = x^2$.



Ist f lediglich *partiell differenzierbar*, d. h. die partiellen Ableitungen existieren auf U , so definiert man dennoch

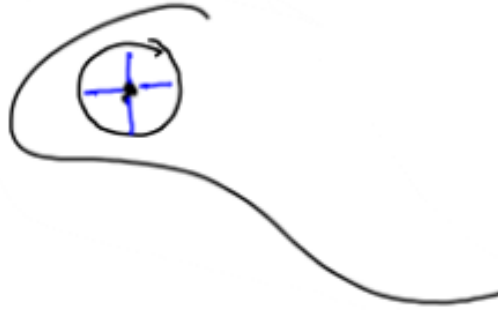
$$\text{grad } f(x) = (\partial_1 f(x), \dots, \partial_n f(x))^T$$

als den Vektor der partiellen Ableitungen bei x .

Satz 3.10. Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, *partiell differenzierbar* auf U . Sei $a \in U$ ein *lokales Maximum* (oder *Minimum*) von f , d. h. \exists Umgebung V von a s. d. $f(x) \leq f(a)$ (oder $f(x) \geq f(a)$ für alle $x \in V$). Dann gilt

$$\boxed{\text{grad } f(a) = 0}.$$

Beweis. Betrachte $g_i(t) := f(a + te_i)$, $i = 1, \dots, n$, mit $t \in (-\varepsilon, \varepsilon)$ s. d. $B_\varepsilon^{\|\cdot\|}(a) \subset U$.



Ist a lokales Extremum (also lokales Maximum oder Minimum) von f , so ist $t = 0$ lokales Extremum von g_i . Die g_i sind in $t = 0$ differenzierbar, denn

$$g'_i(0) = \partial_i f(a) \quad (\text{Definition Richtungsableitung})$$

DIFF I $\implies g'_i(0) = 0 \implies$ Beh. □

Beispiele. i) $f(x) = x_1^2 + x_2^2 + 2$. Außer in $a = 0$ kann kein Extremum vorliegen.

ii) $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$ $\partial_1 f(0) = 0 = \partial_2 f(0)$, also könnte 0 ein Extremum sein. Ist es aber nicht, da für $x = y = \varepsilon > 0$ gilt $f(\varepsilon, \varepsilon) = \frac{1}{2} > 0$ und für $-x = y = \varepsilon > 0$. $f(-\varepsilon, \varepsilon) = -\frac{1}{2} < 0$.

Bemerkung. Hinreichende Kriterien für das Vorliegen lokaler Extremstellen werden wir erst später kennen lernen. Wie in der DIFF I benötigen wir dafür die 2. Ableitung.

Satz 3.11. Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen. Existieren alle partiellen Ableitungen $\partial_j f(x)$ für alle $x \in U$, und sind sie stetig in $a \in U$, so ist f in a differenzierbar.

Beweis. Wir wählen ein Norm $\|\cdot\|$ auf \mathbb{R}^n . U offen $\implies \exists \delta > 0$ s. d. $B_\delta^{\|\cdot\|}(a) \subset U$. Sei $h \in B_\delta^{\|\cdot\|}(0)$, also $a + h \in B_\delta^{\|\cdot\|}(a)$. Setze

$$x^{(j)} := a + \sum_{i=1}^j h_i e_i \in \mathbb{R}^n \quad j = 0, \dots, n,$$

also $x^{(0)} = a$, $x^{(1)} = (a_1 + h_1, a_2, \dots, a_n), \dots, x^{(n)} = a + h$. Es ist $x^{(j)} - x^{(j-1)} = (0, \dots, 0, h_j, 0, \dots, 0) \implies (\text{MWS, DIFF I}) \exists \eta_j \in [0, 1]$ s. d.

$$\begin{aligned} f(x) &= \partial_j f(\underbrace{x_1^{(j-1)}, \dots, x_{j-1}^{(j-1)}, x_j^{(j-1)} + \eta_j h_j, a_{j+1}, \dots, a_n}_{x^{j-1} + \eta_j h_j e_j =: y^{(j)}}) \cdot h_j \\ \implies f(a + h) - f(a) &= \sum_{j=1}^n (f(x^{(j)}) - f(x^{(j-1)})) \\ &= \sum_{j=1}^n \partial_j f(y^{(j)}) h_j \\ &\stackrel{!}{=} \left(\partial_1 f(a) \cdot \dots \cdot \partial_n f(a) \right) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \underline{R}_a(h) \end{aligned}$$

also ist

$$\begin{aligned} \underline{R}_a(h) &= \sum_{j=1}^n (\partial_j f(y^{(j)}) - A_j) h_j \quad A = \partial_j f(a) \\ \xrightarrow{\text{CS}} \frac{|\underline{R}_a(h)|}{\|h\|} &\leq C \|(\partial_1 f(y^{(1)}) - A_1, \dots, \partial_n f(y^{(n)}) - A_n)\|_{\mathbb{E}} \rightarrow 0 \quad h \rightarrow 0, \end{aligned}$$

denn $\lim_{h \rightarrow 0} y^{(j)} = \lim_{h \rightarrow 0} x^{(j-1)} + \eta_j h_j e_j = a$ (in \mathbb{R}^n) und die $\partial_j f$ sind in a stetig nach Voraussetzung, s. d.

$$\lim_{h \rightarrow 0} \partial_j f(y^{(j)}) = (\partial_j f)(\lim y^{(j)}) = \partial_j f(a). \quad \square$$

Bemerkungen 3.12. i) Man sieht: Sind die ∂_j auf U stetig, so ist auch die Ableitung $x \mapsto Df(x)$ eine stetige Abbildung $U \rightarrow \text{Mat}(m \times n, \mathbb{R})$. Man sagt in dem Fall: f ist *stetig differenzierbar*.

ii) Die Einschränkung auf reellwertige Funktionen ist keine, denn: $f: U \rightarrow \mathbb{R}^m$ ist genau dann in $a \in U \subset \mathbb{R}^n$ differenzierbar, wenn alle Komponentenfunktionen $f_j: U \rightarrow \mathbb{R}$ differenzierbar sind ($j = 1, \dots, m$)-

Beweis. 3.12.ii) $f(a + h) = f(a) + A \cdot h + \underline{R}(h)$, $A_{ji} = \partial_j f_i(a)$

$$\iff f_j(a + h) = f_j(a) + \left(\partial_1 f_j(a) \cdot \dots \cdot \partial_n f_j(a) \right) \cdot h + \underline{R}_j(h) \quad \forall j = 1, \dots, m$$

und

$$\frac{\underline{R}(h)}{\|h\|} \rightarrow 0 \text{ in } \mathbb{R}^m \iff \frac{\|\underline{R}(h)\|_{\max}}{\|h\|} \rightarrow 0 \iff \frac{|\underline{R}_j(h)|}{\|h\|} \rightarrow 0 \quad j = 1, \dots, m.$$

3.12.i) Dito: $U \rightarrow \text{Mat}(m \times n, \mathbb{R})$ ist stetig \iff alle Komponentenfunktionen $A_{ij}: U \rightarrow \mathbb{R}$ sind stetig. \square

Bemerkung. Stetig differenzierbar \implies differenzierbar \implies partiell differenzierbar.

Die Umkehrungen sind im Allgemeinen falsch.

Bemerkung (Erinnerung (DIFF I) MWS). $f: I \rightarrow \mathbb{R}$ differenzierbar, I Intervall, $\exists \epsilon \in [0, 1]$ s. d.

$$f(a+h) - f(a) = f'(a + \eta h) \cdot h.$$

Ist f stetig differenzierbar, folgt aus dem Hauptsatz der Differenzial- und Integralrechnung eine andere Variante:

$$f(a+h) - f(a) = \int_a^{a+h} f'(u) du = \int_0^1 f'(a+th) dt \cdot h.$$

Eine analoge Aussage wollen wir nun im \mathbb{R}^n beweisen.

Definition. Sei $A: I \rightarrow \mathbb{R}^k, I \subset \mathbb{R}$ Intervall, stetig. Dann ist das *Integral von A* über $[a, b] \in I$ definiert als

$$\int_a^b A(t) dt = \begin{pmatrix} \int_a^b A_1(t) dt \\ \vdots \\ \int_a^b A_k(t) dt \end{pmatrix}.$$

Insbesondere ist das Integral einer matrixwertigen, stetigen Funktion $A: I \rightarrow \text{Mat}(m \times n, \mathbb{R})$ die Matrix, deren Einträge gleich den Integralen der Komponenten von $A(t)$ ist, also

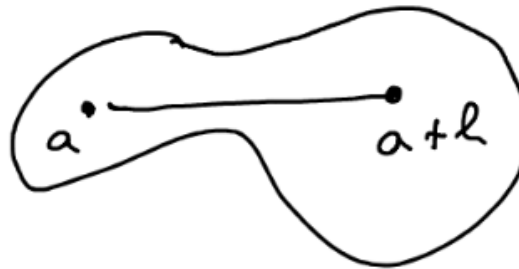
$$\int_a^b A(t) dt = \left(\int_a^b A_{ij}(t) dt \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Satz 3.13. Sei $f: U \rightarrow \mathbb{R}^m, U \subset \mathbb{R}^n$, stetig differenzierbar. Sei $a \in U$ und $h \in \mathbb{R}^n$ s. d. die Verbindungsstrecke

$$\{ a + th \mid t \in [0, 1] \} \subset U.$$

Dann gilt:

$$f(a+h) - f(a) = \left(\int_0^1 Df(a+th) dt \right) \cdot h.$$



Beweis. Wir arbeiten zeilenweise, betrachten also die Komponentenfunktionen. Setze $g_i(t) := f_i(a + th)$. Dann ist g_i stetig differenzierbar, denn mit der Kettenregel gilt $g'_i(t) = Df_i(a + th) \cdot h$.

$$\begin{aligned}
 f_i(a + h) - f_i(a) &= g_i(1) - g_i(0) \\
 &= \int_0^1 g'_i(t) \, dt \\
 &= \int_0^1 Df_i(a + th) \cdot h \, dt \\
 &= \underbrace{\left(\int_0^1 Df_i(a + th) \, dt \right)}_{\substack{\uparrow \text{ linear} \\ (1 \times n)\text{-Matrix}}} \cdot h \quad 1 \leq i \leq m.
 \end{aligned}$$

Dies sind die Zeilen der Matrix $\int_0^1 Df(a + th) \, dt$. □

Folgerung 3.14. Unter den Voraussetzungen von 3.13 gilt $\|f(x + h) - f(x)\|_E \leq C\|h\|$ mit

$$C = \sup_{t \in [0,1]} \|Df(x + th)\|_{\text{op}}.$$

Der Beweis benötigt noch ein Lemma:

Lemma 3.15. Sei $v: [a, b] \rightarrow \mathbb{R}^m$ stetig. Dann gilt

$$\left\| \int_a^b v(t) \, dt \right\|_E \leq \int_a^b \|v(t)\|_E \, dt.$$

Beweis. Sei $\mathbb{R}^m \ni u = \int_a^b v(t) dt$. Dann gilt

$$\begin{aligned}
 \|u\|_E^2 &= \langle u, u \rangle \\
 &= \left\langle \int_a^b v(t) dt, u \right\rangle \\
 &= \int_a^b \langle v(t), u \rangle dt \\
 &\quad \uparrow \\
 &\quad \text{Linearität} \\
 &\leq \int_a^b \|v(t)\|_E \|u\|_E dt \\
 &\quad \uparrow \\
 &\quad \text{Monotonie und C-S} \\
 &= \|u\|_E \int_a^b \|v(t)\|_E dt.
 \end{aligned}$$

□

Bemerkungen. i) Wegen der Äquivalenz aller Normen auf \mathbb{R}^n gilt diese Abschätzung ebenso wie Folgerung 3.14 auch für beliebige Normen auf \mathbb{R}^n :

$$\begin{aligned}
 \left\| \int_a^b v(t) dt \right\| &\leq C_1 \left\| \int_a^b v(t) dt \right\|_E \\
 &\leq C_1 \int_a^b \|v(t)\|_E dt \\
 &\leq C_1 C_2 \int_a^b \|v(t)\| dt.
 \end{aligned}$$

ii) Es folgt, dass für

$$X = (C([a, b], \mathbb{R}^k), \|\cdot\|_\infty),$$

\uparrow
 $\sup \|v(t)\| < \infty$, da $[a, b]$ kompakt und v stetig.

$\int_a^b: X \rightarrow \mathbb{R}^k$ stetig ist, denn \int_a^b ist linear und

$$\begin{aligned}
 \left\| \int_a^b \right\|_{\text{op}} &= \sup_{0 \neq v \in C} \left\| \int_a^b v(t) dt \right\| / \|v\|_\infty \\
 &\leq C \sup_{v \neq 0} \left\| \int_a^b v(t) dt \right\|_E / \|v\|_\infty \\
 &\leq C \sup_{v \neq 0} \int_a^b \|v(t)\|_E dt / \|v\|_\infty \\
 &\leq \tilde{C}(b-a),
 \end{aligned}$$

denn $\|v(t)\|_E \leq \sup_t \|v(t)\|_E \leq C_0 \sup_t \|v(t)\| = C_0 \|v\|_\infty$.

Beweis von Folgerung 3.14.

$$\begin{aligned}\|f(x+h) - f(x)\|_{\mathbb{E}} &\leq \int_0^1 \|Df(x+th) \cdot h\|_{\mathbb{E}} dt \\ &\leq \int_0^1 \|Df(x+th)\|_{\text{op}} \cdot \|h\|_{\mathbb{E}} dt.\end{aligned}$$

Das Supremum wird angenommen, da $t \mapsto \|Df(x+th)\|_{\text{op}}$ stetig ist und $[0, 1]$ kompakt. \square

Implizite Funktionen

Eine Anwendung der Kettenregel

Manchmal ist es einfacher, eine 1-dimensionale Ableitung über eine Ableitung einer höherdimensionalen Funktion zu berechnen. Betrachte etwa zwei differenzierbare Funktionen $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^2$ offen, und $g: I \rightarrow \mathbb{R}$. Sei $\Gamma_g \subset U$ und es gelte

$$\underbrace{f(t, g(t))}_{\in \Gamma_g} = 0 \quad \forall t \in I.$$

Dann gilt (Kettenregel!) (mit $\text{Id}(t) = t$):

$$\begin{aligned}0 &= D * f \circ \begin{pmatrix} \text{Id} \\ g \end{pmatrix}(t) \\ &= Df((t, g(t))) \cdot D \begin{pmatrix} \text{Id} \\ g \end{pmatrix}(t) \\ &= \begin{pmatrix} \partial_1 f(t, g(t)) & \partial_2 f(t, g(t)) \end{pmatrix} \cdot \begin{pmatrix} 1 \\ g'(t) \end{pmatrix}\end{aligned}$$

\implies Ist $\partial_2 f(t, g(t)) \neq 0$, so gilt

$$g'(t) = -\frac{\partial_1 f(t, g(t))}{\partial_2 f(t, g(t))}.$$

Beispiel. $g: (0, 1) \rightarrow \mathbb{R}$, $g(t) = \arcsin \sqrt{1-t^3}$, also $\sin(g(t)) = \sqrt{1-t^3}$, also $(\sin(g(t)))^2 =$

$1 - t^3$. Also gilt $f(t, g(t)) = 0$ für

$$f(x, y) = (\sin y)^2 - 1 + x^3$$

$$Df(x, y) = (3x^2, 2 \sin y \cos y)$$

$$\implies g'(t) = -\frac{3t^2}{2 \sin(g(t)) \underbrace{\cos(g(t))}_{>0}}$$

$$\text{da } t \in (0, 1) \text{ und } \sqrt{1 - t^3} \in (0, 1), \text{ somit } g(t) \in (0, \pi/2) \implies \cos g(t) = + \sqrt{1 - \underbrace{\sin^2 g(t)}_{1-t^3}} = \sqrt{t^3}$$

$$= -\frac{3t^2}{2\sqrt{1-t^3}\sqrt{t^3}}.$$

Vorlesung 8

Do 14.05. 10:15

Satz 3.16. Sei $f: B_{r_1}^{\|\cdot\|}(a) \times B_{r_2}^{\|\cdot\|}(b) \rightarrow \mathbb{R}^m$ eine Abbildung mit $f(a, b) = 0$, die in (a, b) differenzierbar sei. Sei zudem $D_2f(a, b)$, definiert über

$$Df(a, b) = \underbrace{(\partial_1 f(a, b) \cdots \partial_k f(a, b))}_{=: D_1 f(a, b) \in \text{Mat}(m \times k)} \underbrace{(\partial_{k+1} f(a, b) \cdots \partial_{k+m} f(a, b))}_{=: D_2 f(a, b) \in \text{Mat}(m \times m)},$$

invertierbar. Sei zudem $g: B_{r_1}(a) \rightarrow \mathbb{R}^m$ stetig und gelte $g(B_{r_1}(a)) \subset B_{r_2}(b)$ und $g(a) = b$ und $f(x, g(x)) = 0 \forall x \in B_{r_1}(a)$. Dann ist g in a differenzierbar und es gilt

$$Dg(a) = -(D_2f(a, b))^{-1} D_1f(a, b) \in \text{Mat}(m \times k, \mathbb{R}).$$

Beweis. Es gilt für $h = (h_1, h_2) \in B_{r_1}(0) \times B_{r_2}(0)$

$$f(\underbrace{(a, b) + h}_{=(a+h_1, b+h_2)}) = 0 + D_1f(a, b) \cdot h_1 + D_2f(a, b) \cdot h_2 + \underline{R}_{(a,b)}(h)$$

mit $\frac{\underline{R}_{(a,b)}(h)}{\|h\|} \rightarrow 0 \forall x \in B_{r_1}(a)$ folgt

$$\begin{aligned} 0 &= f(\underbrace{a + h_1}_{\in B_{r_1}(a)}, g(a + h_1)) = D_1f(a, b) \cdot h_1 \\ &\quad + D_2f(a, b) \cdot (g(a + h_1) - b) \\ &\quad + \underline{R}_{(a,b)}(h_1, g(a + h_1) - b) \quad \forall h_1 \in B_{r_1}(0). \end{aligned}$$

Also

$$\begin{aligned} g(a + h_1) &= g(a) - D_2f(a, b)^{-1} \cdot D_1f(a, b) \cdot h_1 \\ &\quad - \underbrace{D_2f(a, b)^{-1} \cdot \underline{R}_{(a,b)}(h_1, g(a + h_1) - b)}_{=: \tilde{R}_{a,b}^g(h_1)} \quad \forall h_1 \in B_{r_1}(0). \end{aligned}$$

Wir sind fertig, wenn wir zeigen können, dass

$$\frac{\tilde{R}_{a,b}^g(h_1)}{\|h_1\|} \rightarrow 0 \text{ für } h_1 \rightarrow 0.$$

Wegen der Differenzierbarkeit von f in $(a, b) \exists \tilde{C}$ und $\delta_1, \delta_2 > 0, \delta_i < r_i$, s. d.

$$\|\underline{R}_{(a,b)}(h_1, h_2)\| \leq \tilde{C}(\|h_1\| + \|h_2\|) \quad \forall h_1 \in B_{\delta_1}(0), h_2 \in B_{\delta_2}(0),$$

also

$$\|\underline{R}_{(a,b)}(h_1, g(a+h_1) - b)\| \leq \tilde{C}(\|h_1\| + \|g(a+h_1) - b\|)$$

für alle h_1 s. d. $\|h_1\| < \delta_1$ und $\|g(a+h_1) - b\| < \delta_2$. Wegen der Stetigkeit von g in a gibt es $\delta > 0, \delta < \delta_1$, $\|g(a+h_1) - \underbrace{g(a)}_{=b}\| < \delta_2$ für alle $h_1 \in B_\delta(0)$. \square

Bemerkungen. i) Ist f, g wie in Satz 3.16 und f überall *stetig* differenzierbar. Ist dann $D_2f(a, b)$ invertierbar. So gibt es eine Umgebung $V_1 \times V_2$ von (a, b) s. d. $D_2f(x, y)$ invertierbar ist für alle $(x, y) \in V_1 \times V_2$ (denn $(x, y) \mapsto \det D_2f(x, y)$ ist stetig, da es ein Polynom stetiger Funktionen ist und $\det D_2f(a, b) \neq 0$).

ii) Das Beispiel vom letzten Mal ist genau von diesem Typ.

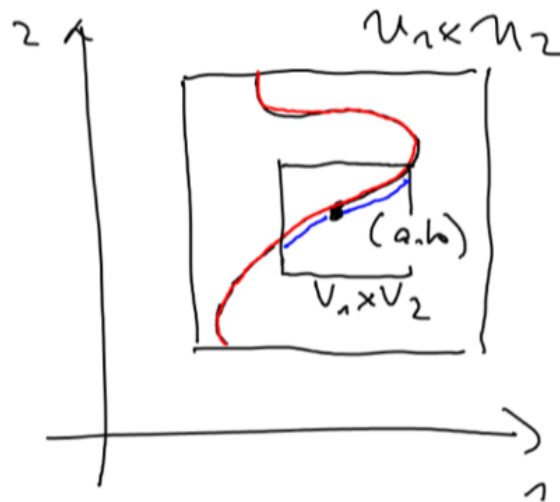
Umgekehrt kann man eine Funktion g wie oben durch „Auflösen der Gleichung $f(x, g(x)) = 0$ “ bestimmen (unter gewissen Voraussetzungen):

Satz 3.17 (Satz von der impliziten Funktion). Seien $U_1 \subset \mathbb{R}^k, U_2 \subset \mathbb{R}^m$ offen und $f: U_1 \times U_2 \rightarrow \mathbb{R}^m$ stetig differenzierbar. Sei $(a, b) \in U_1 \times U_2$ s. d. $f(a, b) = 0$ und $D_2f(a, b) \in \text{Mat}(m \times m, \mathbb{R})$ invertierbar. Dann gibt es offene Umgebungen $V_1 \subset U_1, V_2 \subset U_2$ von a bzw. b und eine stetige Funktion $g: V_1 \rightarrow \mathbb{R}^m, g(V_1) \subset V_2$, s. d. $f(x, g(x)) = 0 \forall x \in V_1$. Ist $(x, y) \in V_1 \times V_2$ s. d. $f(x, y) = 0$, so ist $y = g(x)$.

Bemerkungen. i) Aus 3.16 und der folgenden Bemerkung folgt, dass g in einer eventuell verkleinerten Umgebung $\tilde{V}_1 \subset V_1$ von a sogar stetig differenzierbar ist und gilt

$$Dg(x) = -D_2f(x, g(x))^{-1} \cdot D_1f(x, g(x)) \quad \forall x \in \tilde{V}_1$$

ii) Für den Satz ist wichtig, dass $U_1 \times U_2$ gegebenenfalls verkleinert wird:

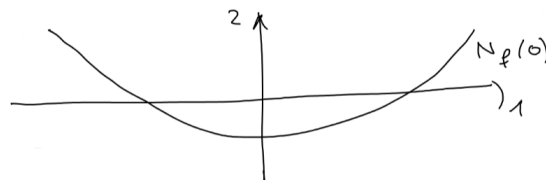


$$\Gamma_g = \{ (x, g(x)) \mid x \in V_1 \},$$

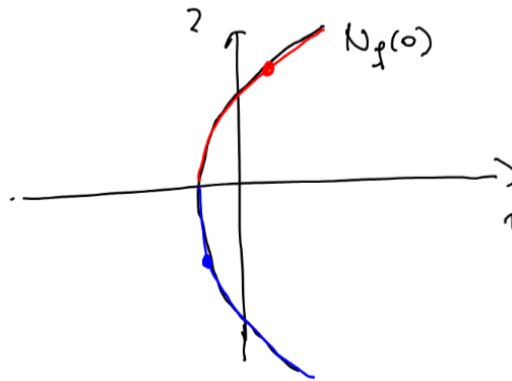
Betrachtete man auch den oberen Teil der Kurve, könnte man x nicht ein eindeutiges y zuordnen.

- iii) Die Einschränkung auf Definitionsbereiche der Form $U_1 \times U_2$ ist keine, sie vereinfacht nur die Notation. Ist f auf $U \subset \mathbb{R}^{k+m}$ offen, findet man stets $U_1 \subset \mathbb{R}^k$, $U_2 \subset \mathbb{R}^m$ s. d. $U_1 \times U_2 \subset U$.
- iv) Die Einschränkung auf $N_f(0)$ ist keine: Will man etwa die Gleichung $f(x, y) = c$ auflösen, wendet man den Satz auf \tilde{f} an mit $\tilde{f}(x, y) = f(x, y) - c$.
- v) Durch Umnummerierung kann man auch andere $m \times m$ -Untermatrizen von Df betrachten als die letzten m .
- vi) Unter den Voraussetzungen von 3.17 sagt man: g ist durch $f(x, y) = 0$ *implizit* gegeben und man *löst* $f(x, y) = 0$ *nach* y *auf*.

Beispiele. i) $f(x, y) = 3y - x^2 + 1$ auf \mathbb{R}^2 . $Df(a, b) = \begin{pmatrix} -2a & 3 \end{pmatrix}$, $3 \neq 0$ ist invertierbar, 3.17 $\implies \exists g: I \rightarrow \mathbb{R}$ s. d. $f(x, g(x)) = 0 \forall x \in I$. In diesem Fall sogar $I = \mathbb{R}$: $g(x) = \frac{1}{3}(x^2 - 1)$.



- ii) $f(x, y) = 3x - y^2 + 1$ auf \mathbb{R}^2 . $Df(a, b) = \begin{pmatrix} 3 & -2b \end{pmatrix}$. 3.17 \implies Zu $b \neq 0$ gibt es $g: I \rightarrow \mathbb{R}$. $b > 0$: $g(x) = +\sqrt{3x+1}$, $x > -\frac{1}{3}$. $b < 0$: $g(x) = -\sqrt{3x+1}$.



$$f(x, g(x)) = 0$$

Beweis von Satz 3.17. Setze $B := D_2f(a, b)$ und definiere eine Abbildung $h: U_1 \times U_2 \rightarrow \mathbb{R}^m$ vermöge

$$h(x, y) = y - b - B^{-1}f(x, y).$$

Dann gilt

$$D_2h(x, y) = \mathbb{1} - B^{-1}D_2f(x, y).$$

$\implies D_2h(a, b) = 0 \implies$ (da alle Ableitungen stetig sind) $\exists W_1 \subset U_1, W_2 \subset U_2$ offene Umgebungen von a bzw. b s. d.

$$\|D_2h(x, y)\| \leq \frac{1}{2} \quad \forall x \in W_1, y \in W_2. \quad (*)$$

Wähle $r > 0$ s. d. $V_2 := B_r^{\|\cdot\|}(b) \subset W_2$. Es ist $h(a, b) = 0 \implies$ (da h differenzierbar ist und somit auch stetig) \exists offene Umgebung $V_1 \subset W_1$ von a s. d.

$$\varepsilon := \sup_{x \in V_1} \|h(x, b)\| < \frac{r}{2} \quad (**)$$

(auf einem Kompaktum $\subset W_1$ um a ist $x \mapsto h(x, b)$ beschränkt und wird auf einem hinreichend kleinen Kompaktum beliebig klein. Um V_1 offen zu erhalten, nehmen wir das Innere eines solchen Kompaktums).

Wir zeigen jetzt: Zu jedem $x \in V_1$ gibt es höchstens ein $y \in V_2$ s. d. $f(x, y) = 0$ also s. d. $h(x, y) = y - b$.

Sei also $x \in V_1$ und seien y_1 und y_2 s. d. $h(x, y_1) = y_1 - b$ und $h(x, y_2) = y_2 - b$.

$$\begin{aligned}
 &\implies y_1 - y_2 = h(x, y_1) - h(x, y_2) \\
 &\stackrel{\text{MWS und } (*)}{\implies} \|y_1 - y_2\| = \|h(x, y_1) - h(x, y_2)\| \\
 &\quad = \|D_2 h(x, \zeta)\| \cdot \|y_1 - y_2\| \\
 &\quad \quad \quad \uparrow \\
 &\quad \zeta \text{ auf der Verbindungsstrecke zw. } y_1 \text{ und } y_2 \text{ (liegt in } V_2 = B_r(b)) \\
 &\quad \leq \frac{1}{2} \|y_1 - y_2\| \\
 &\implies \|y_1 - y_2\| = 0 \implies y_1 = y_2.
 \end{aligned}$$

1) 2) Wir zeigen nun die Existenz einer Funktion g wie im Satz behauptet. Setze dazu $g_0(x) = b$ und definiere rekursiv für $x \in V_1$:

$$g_{j+1}(x) := b + h(x, g_j(x)).$$

a) Es gilt

$$\|g_{j+1} - g_j\|_{\infty, V_1} \leq 2^{-j} \varepsilon.$$

\uparrow
aus (**)

Induktionsanfang:

$$\|g_1 - g_0\|_{\infty, V_1} = \|h(x, b)\|_{\infty, V_1} = \varepsilon.$$

Induktionsschritt: Sei die Behauptung für $i \leq n$ bewiesen.

$$\begin{aligned}
 &g_{n+2}(x) - g_{n+1}(x) = h(x, g_{n+1}(x)) - h(x, g_n(x)). \\
 &\stackrel{\text{MWS und } (*)}{\implies} \|g_{n+2} - g_{n+1}\|_{\infty, V_1} \leq \frac{1}{2} \|g_{n+1} - g_n\|_{\infty, V_1}.
 \end{aligned}$$

Bemerkung. Der MWS darf tatsächlich angewendet werden. $g_{n+1}(x), g_n(x)$ und somit auch die Verbindungsstrecke zwischen ihnen liegen in V_2 , denn nach Induktionsvoraussetzung gilt für alle $j \leq n$

$$\|g_{j+1} - b\|_{\infty, V_1} \leq \sum_{i=0}^j \|g_{i+1} - g_i\| \leq 2\varepsilon < r$$

(da $g_{j+1} - b = \sum_{i=0}^j (g_{i+1} - g_i)$ ist). Somit darf der MWS auf

$$h(x, g_{n+1}(x)) - h(x, g_n(x))$$

angewendet werden.

b) Es folgt $\|g_n - b\|_{\infty, V_1} < r$ und somit $g_n(V_1) \subset V_2$. Denn

$$g_n = \sum_{j=0}^{n-1} (g_{j+1} - g_j) + b \quad (\text{Teleskopsumme}) \implies \|g_n - b\|_{\infty, V_1} \leq \sum_{j=0}^{n-1} 2^{-j} \varepsilon \leq 2\varepsilon < r.$$

\uparrow
geom. Reihe

c) Zudem gilt: $\|\sum_{j=0}^{\infty} (g_{j+1} - g_j)\|_{\infty, V_1}$ hat die Majorante $\sum_{j=0}^{\infty} 2^{-j} \varepsilon$. \implies Die Reihe konvergiert gleichmäßig auf $V_1 \implies$ (DIFF I)

$$g := \lim_{n \rightarrow \infty} g_n = \sum_{j=0}^{\infty} (g_{j+1} - g_j) + b$$

ist stetig auf V_1 und

$$\|g - b\|_{\infty, V_1} \leq 2\varepsilon < r,$$

also $g(V_1) \subset V_2$.

Aus der Definition folgt durch Grenzübergang auf beiden Seiten (h ist stetig)

$$g(x) = b + h(x, g(x)) \quad \forall x \in V_1,$$

also $g(x) = g(x) - B^{-1}f(x, g(x))$, also

$$f(x, g(x)) = 0 \quad \forall x \in V_1. \quad \square$$

Folgerung 3.18. Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^2$ offen, stetig differenzierbar und sei $(a, b) \in U$, $f(a, b) = c$ und $\text{grad } f(a, b) \neq 0$. Dann kann man ein Stück der Höhenlinie $N_f(c)$ als Graph einer Funktion beschreiben. Denn sei $\partial_2 f(a, b) \neq 0$, Satz 3.17, angewandt auf $\tilde{f}(x, y) = f(x, y) - c$, impliziert:

\exists Intervalle I_1, I_2 , $a \in I_1$, $b \in I_2$, $I_1 \times I_2 \subset U$ und eine stetig differenzierbare Funktion (Bemerkung 3.17.i)) $g: I_1 \rightarrow \mathbb{R}$, mit $g(I_1) \subset I_2$ und

$$N_f(c) \cap I_1 \times I_2 = \left\{ (x, y) \in I_1 \times I_2 \mid \tilde{f}(x, y) = 0 \right\} = \left\{ (x, g(x)) \mid x \in I_1 \right\} = \Gamma_g.$$

Dito für den Fall, dass $\partial_1 f(a, b) \neq 0$ ist (mit vertauschten Rollen für x, y), also

$$N_f(c) \cap I_1 \times I_2 = \left\{ (g(x), x) \mid x \in I_2 \right\}.$$

Vorlesung 9

Mo 17.05. 10:15

Beispiele (Weitere zum Satz über implizite Funktionen). i) $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $f(x, y, z) = \begin{pmatrix} x^2 - y^2 + 1 \\ x^2 - z^2 - 2 \end{pmatrix}$ ist stetig differenzierbar.

$$Df(x, y, z) = \begin{pmatrix} 2x & -2y & 0 \\ 2x & 0 & -2z \end{pmatrix}.$$

Ist $y_0 \cdot z_0 \neq 0$ (also beide $\neq 0$), so ist $-2 \begin{pmatrix} y_0 & 0 \\ 0 & z_0 \end{pmatrix}$ invertierbar.

Es gibt also ein offenes Intervall I , $x_0 \in I$ und eine offene Umgebung V_2 von $\begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \in \mathbb{R}^2 \setminus \text{Achsen}$ und eine stetig differenzierbare Funktion $g: I \rightarrow \mathbb{R}$ mit $g(I) \subset V_2$ s. d.

$$f(x, y, z) = C = f(x_0, y_0, z_0) \iff \begin{pmatrix} y \\ z \end{pmatrix} = g(x)$$

und es gilt

$$g'(x) = \frac{1}{2} \begin{pmatrix} \frac{1}{y} & \\ & \frac{1}{z} \end{pmatrix} \begin{pmatrix} 2x \\ 2x \end{pmatrix} = \begin{pmatrix} x/g_1(x) & \\ & x/g_2(x) \end{pmatrix},$$

also $g'_1(x)g_1(x) = x$ und $g'_2(x)g_2(x) = x$. Hauptsatz der Differential- und Integralrechnung

$$\implies \int_{x_0}^x g'_j(t)g_j(t) dt = \frac{x^2}{2} - \frac{x_0^2}{2} \quad j = 1, 2$$

$$\text{(Substitution)} \quad \int_{g_j(x_0)}^{g_j(x)} u du \stackrel{||}{=} \frac{1}{2}(g_j^2(x) - g_j^2(x_0))$$

$$\implies x^2 > x_0^2 - y_0^2: g_1(x) = \operatorname{sgn}(y_0) \sqrt{x^2 - x_0^2 + y_0^2}$$

$$x^2 > x_0^2 - y_0^2: g_2(x) = \operatorname{sgn}(z_0) \sqrt{x^2 - x_0^2 + z_0^2}.$$

Beispiel. Für $(x_0, y_0, z_0) = (3, 2, -1)$

$$g: (2\sqrt{2}, \infty) \rightarrow \mathbb{R}^2 \quad g(x) = \begin{pmatrix} \sqrt{x^2 - 5} \\ -\sqrt{x^2 - 8} \end{pmatrix}$$



- ii) $f: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$, $f(x, t) = t^n + \sum_{j=1}^n x_j t^{n-j}$, $f(x, t) = 0 \implies t$ Nullstelle des Polynoms $p_x(t) = f(x, t)$.

Sei t_0 eine einfache Nullstelle. Dann ist $D_2 f(x, t_0) \neq 0$, da $p_x(t) = (t - t_0)q_x(t)$ und $q_x(t_0) \neq 0$ und somit $D_2 f(x, t_0) = q_x(t_0) + \underbrace{(t_0 - t_0)q'_x(t_0)}_{=0}$.

$\implies \exists$ Umgebung U von $x_0 \in \mathbb{R}^n$ und eine stetig differenzierbare Funktion $g: U \rightarrow \mathbb{R}$ mit $f(x, g(x)) = 0$ und $g(x_0) = t_0 \implies$ Die einfachen Nullstellen eines Polynoms hängen in stetig differenzierbarer Art von den Koeffizienten ab, d. h. insbesondere hängen sie stetig davon ab, d. h. zu $\varepsilon > 0$, $\exists \delta > 0$ s. d. $|t - t_0| = |g(x) - g(x_0)| < \varepsilon$ für $x \in B_\delta(x_0) \subset U$.

Beispiel. $3t^2 - t$ Einfache Nullstelle $t_0 = \frac{1}{3}$.

$\frac{10}{3}t^2 - t$ Einfache Nullstelle $t_0 = \frac{3}{10}$.

- iii) Die Bedingung im Satz über implizite Funktionen ist *nicht* notwendig. $f(x, y) = y^2$ erfüllt in $(1, 0)$ die Bedingung $D_2 f(1, 0)$ invertierbar *nicht*, $f(x, y) = 0$ besitzt aber die Auflösung $g(x) = 0 \quad \forall 0$.

Der Satz von der Umkehrabbildung

Sei $f: U \rightarrow \mathbb{R}^n$ stetig differenzierbar, $U \subset \mathbb{R}^n$. Gibt es einfach zu prüfende Kriterien, die garantieren, dass f umkehrbar ist mit stetig differenzierbarer Umkehrfunktion $g: f(U) \rightarrow \mathbb{R}^n$?

Notwendiges Kriterium ist, dass $Df(x)$ invertierbar ist, da mit der Kettenregel aus $g \circ f = \text{Id}_U$ folgt $Dg(f(x)) \cdot Df(x) = \mathbb{1}$. Tatsächlich ist sogar folgendes hinreichend:

Satz 3.19. Sei $U \subset \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}^n$, stetig differenzierbar. Sei $a \in U$ s. d. $Df(a)$ invertierbar ist. Dann gibt es eine offene Umgebung $V \subset U$ von a und eine offene

Umgebung $W \in \mathbb{R}^n$ von $b := f(a)$, s.d. $f: V \rightarrow \mathbb{R}^n$, V bijektiv auf W abgebildet und die Umkehrabbildung $g: W \rightarrow \mathbb{R}^n$ stetig differenzierbar ist. Es gilt dann

$$Dg(f(x)) = Df(x)^{-1} \quad \forall x \in V.$$

Bemerkung. Die letzte Identität, die wir oben schon bewiesen hatten, leuchtet unmittelbar ein, wenn man sich überlegt, dass die Ableitung eine differenzierbare Funktion linear approximiert, und dass die Umkehrabbildung einer invertierbaren linearen Abbildung die inverse Matrix ist.

Beweis von 3. Betrachte $F: \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $d(z, x) = z - f(x)$. Dann ist

$$Df(z, x) = \begin{pmatrix} \underbrace{1_{n \times n}}_{D_1 F(z, x)} & - \underbrace{Df(x)}_{D_2 F(z, x)} \end{pmatrix}.$$

Wegen $F(b, a) = 0$ und $D_2 F(b, a) = -Df(a)$ invertierbar können wir den Satz von der impliziten Funktion 3.17 auf F anwenden und F nach x auflösen. Es gibt also offene Umgebungen $V_1 \subset \mathbb{R}^n$ von b und $V_2 \subset U$ von a und eine stetig differenzierbare Funktion (vgl. Bemerkung 3.17.i) nach 3.17) $g: V_1 \rightarrow \mathbb{R}^n$ mit $g(V_1) \subset V_2$ und $F(z, g(z)) = 0 \quad \forall z \in V_1$, also $z - f(g(z)) = 0 \quad \forall z \in V_1$ und für all $(z, x) \in V_1 \times V_2$ mit $F(z, x) = 0$ gilt $x = g(z)$.

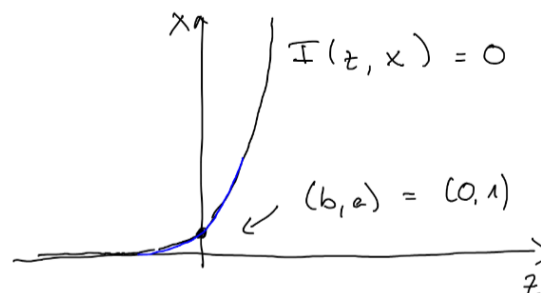
Setze nun

$$\begin{aligned} V &:= V_2 \cap f^{-1}(V_1) \\ &\quad \uparrow \\ &\quad \text{Urbild von } V_1 \\ &= \{ x \in V_2 \mid f(x) \in V_1 \}. \end{aligned}$$

V ist offen (da f stetig und V_1, V_2 offen) und f bildet V bijektiv auf $W := V_1$ ab und die Umkehrabbildung ist g . \square

Man sagt, wenn f wie oben ist: „ f ist bei a lokal umkehrbar.“

Beispiele 3.20. i) $f(x) = \ln x$, $x > 0$, $a = 1$, $b = 0$, $f'(1) = 1$.



$$\Gamma_g = \{ (z, g(z)) \mid z \in V \}$$

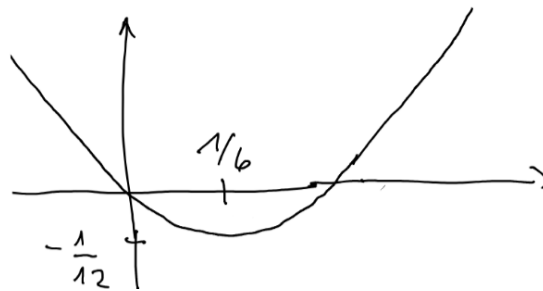
$$g'(\ln x) = \left(\frac{1}{x}\right)^{-1} = x$$

$$g'(z) = g(z)$$

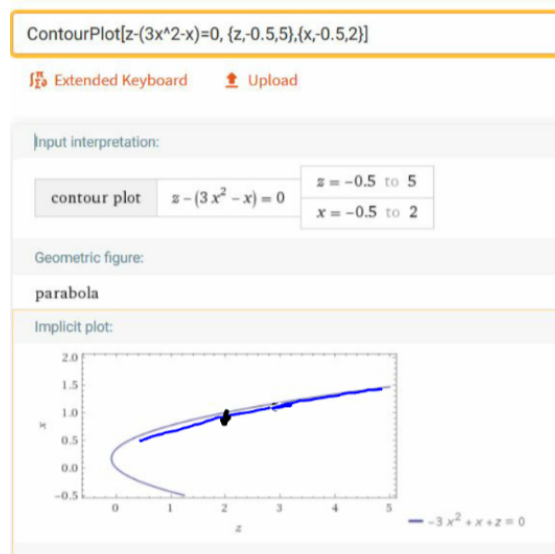
$$\stackrel{\text{(DIFF I)}}{\implies} g(z) = C \exp(z),$$

und wegen $g(0) = 1$ folgt $C = 1$.

ii) $f(x) = 3x^2 - x$, $f(1) = 2$, $Df(1) = 5 \neq 0$.



$$F(z, x) = z - f(x) \quad N_f(0) = \{ (f(x), x) \mid x \in U \}$$



$N_F(0), \Gamma_g = \{ \underbrace{(z, g(z))}_{=(f(x), x)} \mid z \in W \}$, von Wolfram Alpha.

Die Rekursionsformel $g_0(z) = a, g_{j+1}(z) = a + h(z, g_j(z))$,

$$h(z, x) = x - D_2 F(b, a)^{-1} \cdot F(z, x)$$

ist meist nicht sehr nützlich zur Bestimmung von g . Aber die Ableitung von g können wir sofort bestimmen: $Dg(z) = Df(g(z))^{-1}$ oder $Dg(f(x)) = Df(x)^{-1}$.

Im Beispiel: $g'(z) = \frac{1}{6g(z)-1}$ für $g(z) \neq 1/6$. \implies Ansatz $g(z) = \alpha + \sqrt{\beta + \gamma z}$

$$z \stackrel{!}{=} f(g(z)) = 3(\alpha^2 + \beta + \gamma z + 2\alpha\sqrt{\beta + \gamma z}) - \alpha - \sqrt{\beta + \gamma z}$$

$$6\alpha = 1$$

$$\implies 3\alpha^2 + 3\beta - \alpha = 0 \implies g(z) = \frac{1}{6} + \frac{1}{6}\sqrt{+1 + 12z}$$

$$3\gamma = 1$$

auf $(-\frac{1}{12}, \infty)$ definiert \implies es muss $f > -\frac{1}{12}$ sein, also ist das maximale $V = (\frac{1}{12}, z)$.

iii) $f: \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \mathbb{R}^2, (r, \phi) \mapsto (r \cos \phi, r \sin \phi)^T$.

$$\det Df(r, \phi) = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}$$

$$\det Df(r, \phi) = r > 0 \quad \forall (r, \phi).$$

\implies Bei allen r, ϕ ist f lokal invertierbar. Es gilt

$$Df(r, \phi)^{-1} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\frac{\sin \phi}{r} & \frac{\cos \phi}{r} \end{pmatrix}.$$

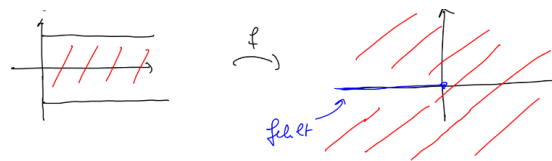
Setze $f(r, \phi) = \begin{pmatrix} x \\ y \end{pmatrix} \implies r = \sqrt{x^2 + y^2}, \frac{x}{r} = \cos \phi, \frac{y}{r} = \sin \phi.$

$$\begin{aligned} \implies Df(r, \phi)^{-1} &= Df(\underbrace{f(r, \phi)}_{\begin{pmatrix} x \\ y \end{pmatrix}}) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

mit g eine lokale Umkehrung.

Beachte. $f(\mathbb{R}_{>0} \times \mathbb{R}) = \mathbb{R}^2 \setminus \{0\}$, aber es gibt *keine* globale Umkehrfunktion $g: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$, denn $f(r, \phi + k2\pi) = f(r, \phi) \quad \forall k \in \mathbb{Z}$, also f nicht injektiv. Man kann maximal Intervalle der Länge 2π (in ϕ) betrachten.

Betrachte etwa $U = \mathbb{R}_{>0} \times (-\pi, \pi)$. Das wird unter f bijektiv auf $W = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$ abgebildet.



„geschlitzte Ebene“

Wie kommt man darauf, dass $f|_U$ auf W abbildet? ϕ ist entweder $\neq 0$, dann ist $r \sin \phi \in (-r, 0) \cup (0, r)$, oder $\phi = 0$, dann ist $r \cos \phi = r > 0$.

$f|_U$ ist injektiv:

Ist $r \cos \phi = r' \cos \phi'$ und $r \sin \phi = r' \sin \phi'$, so ist $r^2 = r'^2 \implies$ (da, $r, r' > 0$) $r = r'$ und aus $\cos \phi = \cos \phi'$ folgt zunächst $\phi = \pm \phi'$, aber aus $\sin \phi = \sin \phi' = \pm \sin \phi'$ folgt $\phi = +\phi'$.

$f|_U$ bildet surjektiv auf W ab:

Sei $\begin{pmatrix} x \\ y \end{pmatrix} \in W$, d.h. es ist $y \neq 0$ oder $x > 0 \implies r = \{x^2 + y^2\} > 0 \implies \left(\frac{x}{r}\right)^2 + \left(\frac{y}{r}\right)^2 = 1 \implies |x/r| \leq 1$. Es ist $x/r \neq -1$, denn sonst wäre $y/r = 0$, aber $(-1, 0) \notin W$. Somit ist $y = 0 \iff x/r = 1$ und wir legen einen Winkel ϕ fest mit $\cos \phi = x/r$. $\phi \in (-\pi, 0)$, falls $y < 0$, oder $(0, \pi)$, falls $y > 0$, oder $\phi = 0$, falls $y = 0$.

Behauptung. Für diesen Winkel gilt $\sin \phi = y/r$.

Beweis.

$$\sin \phi = \operatorname{sgn}(\sin \phi) \sqrt{\sin^2 \phi} = \operatorname{sgn}(y) \sqrt{1 - \cos^2 \phi} = \operatorname{sgn}(y) \sqrt{1 - x^2/r^2} = \operatorname{sgn}(y) |y|/r = y/r. \quad \square$$

Ist man nicht an einer maximalen Umgebung von (r_0, ϕ_0) interessiert, kann man so schneller zum Ziel:

Ist beispielsweise $-\pi/2 < \phi_0 < \pi/2$ betrachte $\phi \in (-\pi/2, \pi/2)$, woraus $x > 0$ folgt, so dass man recht schnell rät, dass $f|_V := \mathbb{R}_{>0} \times (-\pi/2, \pi/2)$ bijektiv auf $W = \mathbb{R}_{>0} \times \mathbb{R}$ abbildet, indem man sich überlegt, dass die Umkehrfunktion g durch $g(x, y) = (\sqrt{x^2 + y^2}, \arctan y/x)^T$ gegeben ist (wohldefiniert für $x > 0$), da

$$g(r \cos \phi, r \sin \phi) = (r, \underbrace{\arctan \sin \phi / \cos \phi}_{=\phi})$$

und es ist klar, dass $g|_W$ bijektiv auf V abbildet, da $\arctan: \mathbb{R} \rightarrow (-\pi/2, \pi/2)$ bijektiv ist, s.d. wenn $\arctan x/y = \arctan \tilde{x}/\tilde{y}$ gilt $y = \tilde{y}x/\tilde{x}$ und dann

$$\sqrt{x^2 + y^2} = \sqrt{x^2 + \frac{\tilde{y}^2 x^2}{\tilde{x}^2}} = \frac{x}{\tilde{x}} \sqrt{\tilde{x}^2 + \tilde{y}^2}$$

nur gleich $\sqrt{\tilde{x}^2 + \tilde{y}^2}$, wenn $x = \tilde{x}$ (und somit $y = \tilde{y}$) und $\mathbb{R}_{>0} \times \mathbb{R} \ni (x, y) \mapsto \sqrt{x^2 + y^2}$ bildet surjektiv auf $\mathbb{R}_{>0}$ ab (da streng monoton wachsend).

Es folgen einige etwas tiefsinnigere Folgen aus dem Umkehrsatz.

Satz 3.21. Sei $f: U \rightarrow \mathbb{R}^n$ stetig differenzierbar, $U \subset \mathbb{R}^n$ offen. Sei $V \subset U$ offen. Ist $Df(x)$ invertierbar für alle $x \in V$, so ist $f|_V$ eine *offene Abbildung*, d.h. offene Teilmengen $\subset V$ werden auf offene Mengen abgebildet.

Anti-Beispiel. $f: (0, 1) \rightarrow \mathbb{R}$, $f(x) = 1 \quad \forall x$. $f'(x) = 0$ nicht invertierbar. $\{1\}$ nicht offen in \mathbb{R} .

Beweis von 3.21. Sei $\tilde{V} \subset U$ offen, dann ist \tilde{V} auch offen U . Sei $b \in f(\tilde{V})$. Wähle ein $a \in \tilde{V}$ s.d. $b = f(a)$. Wende den Satz von der lokalen Umkehrfunktion auf $f|_{\tilde{V}}$ an $\implies \exists$ offene Umgebungen $V_1 \subset \tilde{V}$ und $W \subset \mathbb{R}^n$ von a beziehungsweise b s.d. $f|_{V_1}$ umkehrbar ist mit $g: W \rightarrow \mathbb{R}^n$ und insbesondere $f|_{V_1}$ bijektiv auf W abbildet. Somit gilt $W = f(V_1) \subset f(\tilde{V})$ und ist Umgebung von $b \implies f(\tilde{V})$ ist offen. \square

Satz 3.22 (Minimum- / Maximum-Prinzip). Sei $f: U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^n$, stetig differenzierbar. Sei $V \subset U$ offen und sei $Df(x)$ invertierbar für alle $x \in V$. Dann gilt für die Funktion

$$V \ni x \mapsto \|f(x)\|$$

- i) Sie besitzt kein Maximum.
- ii) Ist $f(x) \neq 0 \quad \forall x \in V$ besitzt sie kein Minimum.

Beachte. Man kann den Beweis dieses Satzes auf 3.10 zurückführen, obwohl $\|\cdot\|$ nur auf $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar ist. Wir wollen hier dennoch anders vorgehen:

Beweis. 3.22.i) Angenommen a ist Maximalstelle von $\|f\|$. Dann ist $f(a) \neq 0$, da sonst $\|f(x)\| = 0 \quad \forall x \in V$ und somit $f|_V = 0$ und $Df(x) = 0 \quad \forall x \in V$. 3.21 $\implies \exists \varepsilon > 0$ s.d. $B_\varepsilon(f(a)) \subset f(V)$. Setze $b := f(a) + \rho f(a)$ mit $\rho = \varepsilon / (2\|f(a)\|)$

$$\begin{aligned} \implies \|b - f(a)\| &= \rho \|f(a)\| = \varepsilon / 2 < \varepsilon \\ \implies b \in B_\varepsilon &\implies \exists x_0 \in V \text{ s.d. } b = f(x_0) \\ \implies \|f(x_0)\| &= \|b\| = (1 + \rho)\|f(a)\| > \|f(a)\|. \end{aligned}$$

3.22.ii) Analog mit $b = f(a) - \rho f(a)$. \square

Vorlesung 10

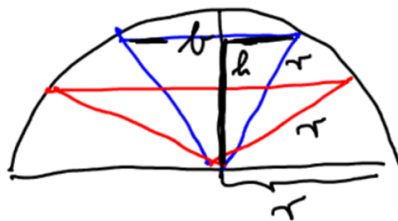
Do 21.05. 10:15

Lokale Extrema unter Nebenbedingungen

Wir betrachten folgende Situation:

Wir suchen Extremstellen einer Funktion f unter gewissen Nebenbedingungen z. B.

- a) (Mikro-Ökonomie) Es soll das Maximum einer Nutzenfunktion $f(x, y) = \sqrt{xy}$, $x > 0, y > 0$, unter der Budgetbedingung $64 = 2x + 8y$ bestimmt werden.
- b) (Geometrie) Welches Dreieck kopfüber im Halbkreis hat den größten Flächeninhalt?
Also $f(b, h) = \frac{1}{2}bh$ unter der Nebenbedingung $\frac{1}{4}b^2 + h^2 = r^2$ (Pythagoras).



- c) Berücksichtige, dass eine Bewegung eingeschränkt ist, z. B. auf dem Innenrand einer Kugelschale (Nebenbedingung: $x^2 + y^2 + z^2 = r^2$).

Mathematisch geht es um folgendes Problem:

Definition 3.23. Sei $U \subset \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$, sei $h: U \rightarrow \mathbb{R}^m$. Dann sagen wir f hat in $a \in U$ ein *lokales Maximum* bzw. *Minimum* (Extremum) *unter der Nebenbedingung* $h(x) = 0$, wenn: $a \in N_h(0)$ und es eine Umgebung U_0 von a gibt s. d. $f(a) \geq f(x)$ bzw. $f(a) \leq f(x) \forall x \in U_0 \cap N_h(0)$.

In einfachen Fällen kann man die Nebenbedingung direkt auflösen und die (3.18) Funktion einsetzen, z. B. in a)

$$y = 8 - \frac{1}{4}x \quad f(x, 8 - \frac{1}{4}x) = \sqrt{x(8 - \frac{1}{4}x)}.$$

Allgemeiner: Sei f, h wie in 3.23 und zudem f differenzierbar und $h: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^{n-m} \times \mathbb{R}^m$, erfülle in $a = (a_1, a_2) \in N_h(0)$ die Bedingungen des Satzes über implizite Funktionen. Also gibt es Umgebungen V_1 von a_1 und V_2 von a_2 und eine stetig differenzierbare Funktion $g: V_1 \rightarrow V_2$ s. d. $N_h(0) \cap V_1 \times V_2 = \{ (u, g(u)) \mid u \in V_1 \}$. Also besitzt

in diesem Fall f eine lokales Extremum in a unter der Nebenbedingung $h = 0$ genau dann, wenn $u \mapsto f(u, g(u))$ ein lokales Extremum in a_1 besitzt.

Dies führt zur sogenannten

Satz 3.24 (Lagrange-Multiplikator-Regel). Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, in $a \in U$ differenzierbar. Sei $h: U \rightarrow \mathbb{R}^m$ stetig differenzierbar, $m < n$. Es sei $\text{rang } Dh(a) = m$.

Es besitze f in a unter der Nebenbedingung $h(x) = 0$ ein lokales Extremum. Dann gibt es $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ (Lagrange-Multiplikatoren), s. d.

$$\text{grad } f(a) = \lambda_1 \text{grad } h_1(a) + \dots + \lambda_m \text{grad } h_m(a).$$

. Sei o.B.d.A. $U = U_1 \times U_2$ mit $U_1 \subset \mathbb{R}^{n-m}$, $U_2 \subset \mathbb{R}^m$, $a = (a_1, a_2) \in U_1 \times U_2$, (Notation $(x, y) \in U_1 \times U_2$) und $Dh(x, y) = \begin{pmatrix} D_1 h(x, y) & \underbrace{D_2 h(x, y)}_{\in \text{Mat}(m \times m, \mathbb{R})} \end{pmatrix}$ mit $D_2 h(a_1, a_2)$

invertierbar.

Satz 3.17 $\implies \exists V_1$ offene Umgebung von a_1 , V_2 offene Umgebung von a_2 , $g: V_1 \rightarrow \mathbb{R}^m$ stetig differenzierbar, $g(V_1) \subset V_2$, $Dg(x) = -D_2 h(x, g(x))^{-1} \cdot D_1 h(x, g(x))$ und

$$N_h(0) \cap V_1 \times V_2 = \{ (x, g(x)) \mid x \in V_1 \}.$$

a ist lokale Extremstelle von f unter der Nebenbedingung $h(x) = 0 \implies a_1$ ist lokale Extremstelle von

$$V_1 \ni u \mapsto f(u, g(u)) = f \circ \begin{pmatrix} \text{Id} \\ g \end{pmatrix} (u)$$

\implies (Kettenregel und 3.10)

$$\begin{aligned} 0 &= D * f \circ \begin{pmatrix} \text{Id} \\ g \end{pmatrix} (a_1) \\ &= Df(a_1, g(a_1)) \cdot \begin{pmatrix} \mathbb{1}_{n-m} \\ Dg \end{pmatrix} (a_1) \\ &= D_1 f(a_1, a_2) - \underbrace{D_2 f(a_1, a_2) \cdot D_2 h(a_1, a_2)^{-1}}_{(1 \times m) \text{ mal } (m \times m) = (1 \times m) =: (\lambda_1, \dots, \lambda_m)} \cdot D_1 h(a_1, a_2) \end{aligned}$$

$$\begin{aligned}
 \implies D_1 f(a) &= (\lambda_1, \dots, \lambda_m) \cdot D_1 h(a) \\
 &= (\lambda_1, \dots, \lambda_m) \cdot \begin{pmatrix} \partial_1 h_1 & \dots & \partial_{n-m} h_1 \\ \vdots & & \vdots \\ \partial_1 h_m & \dots & \partial_{n-m} h_m \end{pmatrix} (a) \\
 &= \left(\sum_{j=1}^m \lambda_j \partial_1 h_j, \sum_{j=1}^m \lambda_j \partial_2 h_j, \dots, \sum_{j=1}^m \lambda_j \partial_{n-m} h_j \right) \\
 &= \sum_{j=1}^m \lambda_j D_1 h_j(a).
 \end{aligned}$$

Gleichzeitig gilt (Definition der λ_j)

$$D_2 f(a_1, a_2) \cdot D_j h(a_1, a_2)^{-1} =: (\lambda_1, \dots, \lambda_m), \quad (3.1)$$

also (wie oben)

$$\begin{aligned}
 D_2 f(a_1, a_2) &= (\lambda_1, \dots, \lambda_m) \cdot D_2 h(a) = \sum_{j=1}^m \lambda_j D_2 h_j(a) \\
 \implies \begin{pmatrix} D_1 f(a) & D_2 f(a) \end{pmatrix} &= \sum_{j=1}^m \lambda_j \begin{pmatrix} D_1 h_j(a) & D_2 h_j(a) \end{pmatrix}. \quad \square
 \end{aligned}$$

Wir nutzt man diesen Satz zur Lösung von Problemen wie oben?

Bemerkung 3.25. Unter Voraussetzungen an f und h wie oben betrachte das Gleichungssystem

$$\left. \begin{aligned} \partial_1 f(a) + \lambda_1 \partial_1 h_1(a) + \dots + \lambda_m \partial_1 h_m(a) &= 0 \\ \partial_2 f(a) + \lambda_1 \partial_2 h_1(a) + \dots + \lambda_m \partial_2 h_m(a) &= 0 \\ &\vdots \\ \partial_n f(a) + \lambda_1 \partial_n h_1(a) + \dots + \lambda_m \partial_n h_m(a) &= 0 \\ h_1(a) &= 0 \\ &\vdots \\ h_m(a) &= 0 \end{aligned} \right\} \quad (*) \quad (**)$$

(*) ist gleichbedeutend mit

$$Df(a) = 0, \quad (**)$$

wenn $F(a, \lambda) = f(a) + \sum_{j=1}^m \lambda_j h_j(a)$, denn $D_1 F = 0$ liefert die ersten n Zeilen von (*), $D_2 F = 0$ liefert die letzten m Zeilen von (*).

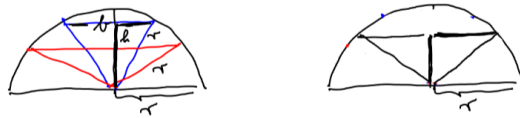
Man versucht dann (*) oder (**) zu lösen.

Beispiele. i) Zunächst b) von oben. $f: \mathbb{R}_{>0} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$, $f(b, h) = \frac{1}{2}bh$, $H(b, h) = \frac{1}{4}b^2 + h^2 - r^2$. $a = (b, h)$. $Df(a) = \left(\frac{1}{2}h, \frac{1}{2}b\right)$, $DH(a) = \left(\frac{1}{2}b, 2h\right)$. $h > 0 \implies$ die Voraussetzungen des Satzes sind erfüllt.

$$\left. \begin{aligned} \partial_1 f(a) + \lambda \partial_1 H(a) &= \frac{1}{2}h + \frac{1}{2}\lambda b \stackrel{!}{=} 0 \\ \partial_2 f(a) + \lambda \partial_2 H(a) &= \frac{1}{2}b + \lambda 2h \stackrel{!}{=} 0 \\ H(a) &= \frac{1}{4}b^2 + h^2 - r^2 = 0 \end{aligned} \right\} \quad (*)$$

$$\implies \lambda = -b/4h \quad h = b^2/4h \implies 4h^2 = b^2$$

$$\stackrel{\text{3. Zeile}}{\implies} \frac{1}{2}b^2 = r^2 \implies b = \sqrt{2} \quad h = \frac{1}{\sqrt{2}r}.$$



Der Flächeninhalt wird maximal $\frac{1}{2}r^2$, denn für $b \rightarrow 0$ (spitzerwerdendes Dreieck), strebt $f \rightarrow 0$, ebenso für $h \rightarrow 0$ (breiterwerdendes Dreieck).

ii) Sei $A \in \text{Mat}(n \times n, \mathbb{R})$ symmetrisch. Sei $f(x) := \langle x, Ax \rangle$ die zugehörige quadratische Form. Gesucht: Extrema von f unter der Nebenbedingung $h(x) = \|x\|_E^2 - 1 = 0$. $Dh(x) = 2x \neq 0 \forall x \in N_h(0)$. Wir können also unseren Satz anwenden

$$\begin{aligned} Df(a) &= 2A \cdot a \\ 2A \cdot a + \lambda 2a &\stackrel{!}{=} 0 \end{aligned} \quad (*)$$

und $\|a\|_E = 1$. Das heißt notwendig für das Vorliegen einer Extremstelle ist, dass a Eigenvektor ist. Da A symmetrisch, ist A diagonalisierbar $\implies \exists$ Eigenvektoren v_1, \dots, v_n . Wegen $f(v_j) = \langle v_j, Av_j \rangle = \lambda_j$ Eigenwert folgt: f wird maximiert von normierten Eigenvektoren zum größten Eigenwert und minimiert von denen zum kleinsten Eigenwert.

iii) Gesucht ist der Punkt in der Ebene

$$E = \left\{ (x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x_1 + x_2 - x_3 = 0 \right\},$$

der von Punkt $(1, 0, 0)^T$ den kleinsten Abstand hat. Da $\{(1, 0, 0)^T\}$ kompakt ist und E abgeschlossen, gibt es einen solchen (vgl. unser Beispiel nach Satz 1.41). Wir wollen also $f(x) = \|x - e_1\|_E^2$ minimieren unter der Neben-Bedingung $h(x) = x_1 + x_2 - x_3 = 0$.

$Dh(a) = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$ hat maximalen Rang (1) d.h. jeder gesuchte Punkt a muss erfüllen

$$2(a_1 - 1) + \lambda = 0$$

$$2a_2 + \lambda = 0$$

$$2a_3 - \lambda = 0$$

$$a_1 + a_2 - a_3 = 0,$$

also ist $a = \begin{pmatrix} 2/3 & -1/3 & 1/3 \end{pmatrix}$ der einzige Kandidat für ein Extremum. Da es ein Minimum gibt, muss a dieses sein.

Höhere Ableitungen, Taylorformel

Ist $f: U \rightarrow \mathbb{R}^m$, $U \subset \mathbb{R}^n$ offen, stetig differenzierbar, kann man sich fragen, ob auch die Ableitung $U \ni x \mapsto Df(x)$ als Funktion $Df: U \rightarrow \text{Mat}(m \times n, \mathbb{R})$ wieder differenzierbar ist (in a , auf ganz U , auf $V \subset U \dots$), also ob es eine *lineare* Abbildung $A: \mathbb{R}^n \rightarrow \text{Mat}(m \times n, \mathbb{R})$ gibt, s. d.

$$Df(a+h) = Df(a) + A(h) * \underline{R}_a^{Df}(h)$$

mit $\frac{\underline{R}_a^{Df}(h)}{\|h\|} \rightarrow 0$ wenn $h \rightarrow 0$. Hier ist $\underline{R}_a^{Df}: B_\delta(0) \rightarrow \text{Mat}(m \times n, \mathbb{R})$.

Benennt man $Df =: g$ sieht man schnell, dass alle Sätze über differenzierbare Abbildungen natürlich auf für g gelten, insbesondere die Eindeutigkeit der Abbildung, wenn sie existiert, Kettenregel etc. Nur kann man, wenn man nicht den Raum $\text{Mat}(m \times n, \mathbb{R})$ mit $\mathbb{R}^{m \cdot n}$, die Ausdrücke im Allgem. einen mehr so leicht hinschreiben, etwa $A(h)$ als Produkt einer Matrix mit dem Vektor h .

Ist $m = 1$ oder betrachtet man nur die Komponentenfunktionen einer \mathbb{R}^m -wertigen Funktion, sieht man jedoch:

$$g := Df: U \rightarrow \text{Mat}(1 \times n, \mathbb{R})$$

nimmt Werte in den $1 \times n$ -Matrizen an, identifiziert man nun $\text{Mat}(1 \times n, \mathbb{R}) \simeq \mathbb{R}^n$ (Spaltenvektor), hat man es mit einer Funktion $g: U \rightarrow \mathbb{R}^n$ zu tun.

Einen *Kandidaten* für $Dg(a)$ findet man wieder über die partiellen Ableitungen

$$\begin{pmatrix} \partial_1 g(a) & \dots & \partial_n g(a) \end{pmatrix} \in \text{Mat}(n \times n, \mathbb{R}) \quad (3.2)$$

und

$$g(x) = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix}$$

(wenn sie existieren). Insgesamt also

$$\begin{pmatrix} \partial_1^2 f(a) & \partial_2 \partial_1 f(a) & \dots & \partial_n \partial_1 f(a) \\ \partial_1 \partial_2 f(a) & \partial_2^2 f(a) & \dots & \partial_n \partial_2 f(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 \partial_n f(a) & \partial_2 \partial_n f(a) & \dots & \partial_n^2 f(a) \end{pmatrix} =: A$$

mit der Notation $\partial_j \partial_i f(a) = \partial_j g_i(a) = \partial_j (\partial_i f)(a)$ und $\partial_j^2 = \partial_j \partial_j$.

Dann berechnet man den Rest wie üblich:

$$g(a+h) = g(a) + A \cdot h + \underline{R}_a^g(h)$$

und überprüft, ob $\frac{\underline{R}_a^g(h)}{\|h\|} \rightarrow 0$ für $h \rightarrow 0$.

Oder man überprüft, ob die partiellen Ableitungen stetig sind. Wenn ja, folgt die Differenzierbarkeit von Df aus Satz 3.11.

Wenn nein, muss man prüfen, ob \underline{R}_a^g das richtige Verhalten bei 0 hat. Wir werden allerdings bei den sogenannten höheren Ableitungen (wie der 2. Ableitung oben) meist stetig differenzierbare Funktionen betrachten.

Definition 3.26. $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, heißt in a k -fach partiell differenzierbar, wenn alle partiellen Ableitungen

$$\partial_{j_1} f(a), \partial_{j_1} \partial_{j_2} f(a), \dots, \partial_{j_1} \partial_{j_2} \dots \partial_{j_k} f(a)$$

(Notation wie oben) $j_i \in \{1, \dots, n\}$ existieren.

Achtung. Beachte, dass man zur Berechnung einer partiellen Ableitung die Funktion in einer Umgebung des Punktes kennen muss. Um also etwa eine k -te partielle Ableitung von f in a zu berechnen, muss man die $(k-1)$ -te partielle Ableitung in einer Umgebung von a kennen.

Ist f auf U k -fach partiell differenzierbar und sind alle partiellen Ableitungen der Ordnung $\leq k$ stetig, so heißt f k -fach stetig partiell differenzierbar.

Beispiel. $f(x_1, \dots, x_4) = x_1^3 + x_2^2 x_3 - 3x_4 x_2$. $\partial_1^2 f(x) = 6x_1$, $\partial_2^2 \partial_4 f(x) = \partial_2^2(-3x_2) = 0$,
 $\partial_2 \partial_4 \partial_3 f(x) = \partial_2 \partial_4(2x_2 x_3 - 3x_4) = \partial_2(-3x_4) = 0$.

Dass die beiden letzt genannten Ableitungen gleich sind, ist kein Zufall:

Satz 3.27 (Satz von Schwarz). Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$, zweimal stetig partiell differenzierbar. Dann gilt für alle $a \in U$ und alle $i, j \in \{1, \dots, n\}$

$$\partial_j \partial_i f(a) = \partial_i \partial_j f(a).$$

Beweis. Sei o.B.d.A. $n = 2$, $a = (a_1, a_2)$. $\exists \delta > 0$ s.d. $B_\delta(a_1) \times B_\delta(a_2) \subset U$. Setze zu

$y \in B_\delta(a_2)$

$$F_y: B_\delta(a_1) \rightarrow \mathbb{R} \quad F_y(x) = f(x, y) - f(x, a_2).$$

MWS $\implies \exists \zeta \in B_\delta(a_1) = (a_1 - \delta, a_1 + \delta)$ und zwar zwischen a_1 und x s.d.

$$F_y(x) - F_y(a_1) = F'_y(\zeta)(x - a_1).$$

Es ist $F'_y(\zeta) = \partial_1 f(\zeta, y) - \partial_1 f(\zeta, a_2)$. Betrachte die Funktion $B_\delta(a_2) \ni y \mapsto \partial_1 f(\zeta, y)$.

MWS $\implies \exists \eta \in B_\delta(a_2) = (a_2 - \delta, a_2 + \delta)$ und zwar zwischen a_2 und y s.d.

$$\partial_1 f(\zeta, y) - \partial_1 f(\zeta, a_2) = \partial_2(\partial_1 f)(\zeta, \eta)(y - a_2).$$

Insgesamt folgt:

$$f(x, y) - f(x, a_2) - f(a_1, y) + f(a) = \partial_2 \partial_1 f(\zeta, \eta)(y - a_2) \cdot (x - a_1). \quad (*)$$

Dieselbe Konstruktion nehmen wir nun zu festen $x \in B_\delta(a_1)$ für $G_x(y) = f(x, y) - f(a_1, y)$ vor.

MWS:

$\exists \tilde{\eta} \in (a_2 - \delta, a_2 + \delta)$ zwischen a_2 und y s.d. $G_x(y) - G_x(a_2) = G'_x(\tilde{\eta})(y - a_2)$. Es ist $G'_x(\tilde{\eta}) = \partial_2 f(x, \tilde{\eta}) - \partial_2 f(a_1, \tilde{\eta})$.

MWS:

$\exists \tilde{\zeta} \in (a_1 - \delta, a_1 + \delta)$ zwischen a_1 und x s.d. $\partial_2 f(x, \tilde{\eta}) - \partial_2 f(a_1, \tilde{\eta}) = \partial_1(\partial_2 f)(\tilde{\zeta}, \tilde{\eta})(x - a_1)$.

$$f(x, y) - f(a_1, y) - f(x, a_2) + f(a) = \partial_1 \partial_2 f(\tilde{\zeta}, \tilde{\eta})(y - a_2) \cdot (x - a_1). \quad (**)$$

Aus (*) und (**) folgt für $x \neq a_1$, $y \neq a_2$

$$\partial_2 \partial_1 f(\tilde{\zeta}, \tilde{\eta}) = \partial_1 \partial_2 f(\zeta, \eta),$$

mit $\zeta, \tilde{\zeta}$ zwischen a_1 und x und $\eta, \tilde{\eta}$ zwischen a_2 und y .

Grenzübergang $x \rightarrow a_1$ und $y \rightarrow a_2$ liefert wegen der Stetigkeit von $\partial_2 \partial_1 f$ und $\partial_1 \partial_2 f$ die Beh. \square

Anti-Beispiel.

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Es gilt $\partial_2 \partial_1 f(0, 0) \neq \partial_1 \partial_2 f(0, 0)$. Denn partielle Ableitung in 0:

$$\partial_1 f(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0 = \partial_2 f(0, 0).$$

Um die zweiten Ableitungen zu bestimmen, müssen wir aber $\partial_1 f$ auch in der Nähe der 0 kennen. Daher bestimmen wir für $(x, y) \neq (0, 0)$

$$\begin{aligned} \partial_1 f(x, y) &= \frac{x^4 y + 4x^4 y^3 - y^5}{(x^2 + y^2)^2} \\ \partial_2 f(x, y) &= \frac{x^5 - xy^4 - 4x^3 y^4}{(x^2 + y^2)^2} \\ \implies \partial_2 \partial_1 f(0, 0) &= \lim_{h \rightarrow 0} \frac{1}{h} (\partial_1 f(0, h) - \partial_1 f(0, 0)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (-h^5 / h^4) = -1 \\ \partial_1 \partial_2 f(0, 0) &= \lim_{h \rightarrow 0} \frac{1}{h} (\partial_2 f(h, 0) - \partial_2 f(0, 0)) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (h^5 / h^4) = +1. \end{aligned}$$

Der Laplace-Operator

Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, zweimal partiell stetig differenzierbar. Man setzt

$$\Delta f(x) := \partial_1^2 f(x) + \partial_2^2 f(x) + \cdots + \partial_n^2 f(x)$$

und nennt $\Delta = \partial_1^2 + \cdots + \partial_n^2$ den *Laplace-Operator*.

Beispiele. i) $h: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$, $h(x) = f(\|x\|_E)$ mit einer auf $\mathbb{R} \setminus \{0\}$ zweimal stetig partiell differenzierbaren Funktion f .

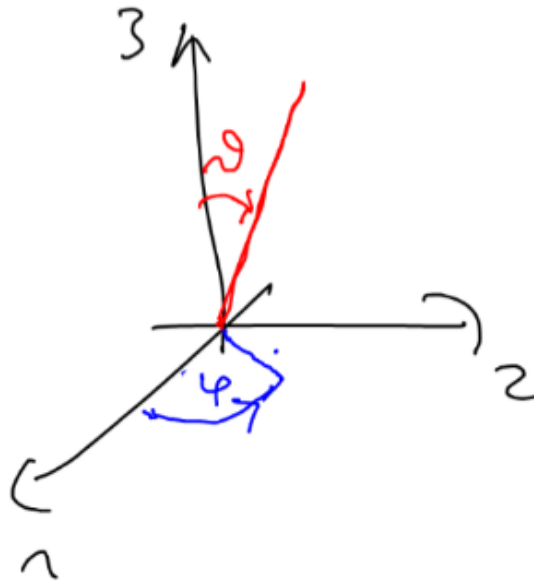
$$\text{grad } f(\|x\|_E) = f'(\|x\|_E) \frac{1}{\|x\|_E} x$$

und

$$\begin{aligned} \partial_i \left(f'(\|x\|_E) \frac{1}{\|x\|_E} x_i \right) &= f''(\|x\|_E) \frac{x_i}{\|x\|_E} \cdot \frac{x_i}{\|x\|_E} + f'(\|x\|_E) \left(-\frac{1}{2} \frac{2x_i}{\sqrt{\sum x_j^2}^3} \cdot x_i + \frac{1}{x_E} \right) \\ \implies \Delta h(x) &= f''(\|x\|_E) + f'(\|x\|_E) \frac{1}{\|x\|_E} (n-1) \end{aligned}$$

ii) In Kugelkoordinaten: Sei $g: \mathbb{R}_{>0} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^3$

$$g(r, \tau, \phi) = \begin{pmatrix} r \cos \phi \sin \theta \\ r \sin \phi \sin \theta \\ r \cos \theta \end{pmatrix}.$$



g ist zweimal stetig partiell differenzierbar.

Sei f zweimal stetig differenzierbar auf $\mathbb{R}^3 \rightarrow \mathbb{R}$. Δf in Kugelkoordinaten bestimmen, also wenn wir $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(r, \theta, \phi)$ schreiben. Zunächst überlegen wir, was es heißt, f in Kugelkoordinaten zu schreiben.

Beispiel. $f(x, y, z) = xyz$.

$$f \circ g(r, \theta, \phi) = f(r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) = r^3 \sin^2 \theta \cos \theta \cos \phi \sin \phi =: F(r, \theta, \phi).$$

Sodann:

$$Dg(r, \theta, \phi) = \begin{pmatrix} \cos \phi \sin \theta & r \cos \phi \cos \theta & -r \sin \phi \sin \theta \\ \sin \phi \sin \theta & r \sin \phi \cos \theta & r \cos \phi \sin \theta \\ \cos \theta & -r \sin \theta & 0 \end{pmatrix}$$

$\det Dg(r, \theta, \phi) = r^2 \sin \theta \implies$ Ist $v \neq k\pi$, $k \in \mathbb{Z}$, so ist g lokal umkehrbar mit stetig differenzierbarer Umkehrfunktion $h: W \rightarrow V$ und in $(x, y, z) \in W$ gilt

$$\begin{aligned}
 f(\underbrace{x, y, z}_{=:g(r, \theta, \phi)}) &= f \circ g \circ h(x, y, z) =: F \circ h(x, y, z) \\
 \implies Df(x, y, z) &= \underbrace{1 \times 3\text{-Matrix}}_{DF(r, \theta, \phi)} \cdot \underbrace{Dh(x, y, z)}_{=:Dg(r, \theta, \phi)^{-1}} =: \tilde{F}(r, \phi, \theta) = \tilde{F} \circ h(x, y, z) \\
 \implies (\partial_1^2 + \partial_2^2 + \partial_3^2)f(x, y, z) &= \partial_1(\tilde{F} \circ h)_1(x, y, z) + \partial_2(\tilde{F} \circ h)_2(x, y, z) + \partial_3(\tilde{F} \circ h)_3(x, y, z) \\
 &= \sum_{j=1}^3 (\partial_j \cdot \tilde{F}_j)(\underbrace{h(x)}_{=: (r, \phi, \theta)}) \cdot \underbrace{\partial_j h_j(x)}_{=: \frac{1}{\partial_j g_j(r, \phi, \theta)}} \quad (\text{Umkehrsatz}) \\
 &= \sum_{j=1}^3 \partial_j \tilde{F}_j(r, \theta, \phi) \cdot \frac{1}{\partial_j g_j(r, \phi, \theta)}.
 \end{aligned}$$

Eine etwas längere Rechnung, bei der man u. a. $Dg(r, \theta, \phi)$ invertieren muss, liefert die explizite Formel für die \tilde{F}_j , etwa

$$\tilde{F}_1 = \begin{pmatrix} \partial_1 F & \partial_2 F & \partial_3 F \end{pmatrix} \cdot \begin{pmatrix} \cos \phi \sin \theta \\ \frac{1}{r} \cos \phi \sin \theta \\ -\frac{1}{r} \frac{\sin \phi}{\sin \theta} \end{pmatrix}.$$

\uparrow
 1. Spalte von $Dg(r, \theta, \phi)^{-1}$

Von diesen berechnet man die jeweilige partielle Ableitung $\partial_j \tilde{F}_j$ und multipliziert mit $\frac{1}{\partial_j g_j}$.

z. B.

$$\begin{aligned}
 \partial_1 \tilde{F}_1 &= \partial_1^2 F \cdot \cos \phi \sin \theta + \partial_1 \partial_2 \cdot \frac{1}{r} \cos \phi \sin \theta - \partial_2 F \cdot \frac{1}{r^2} \cos \phi \sin \theta - \partial_1 \partial_3 \cdot \frac{1}{r} \frac{\sin \phi}{\sin \theta} + \partial_3 F \cdot \frac{1}{r^2} \frac{\sin \phi}{\sin \theta} \\
 1 / \partial_1 g_1 &= \frac{1}{\cos \phi \sin \theta}.
 \end{aligned}$$

Das Ergebnis ist

$$\frac{1}{r^2} \frac{\partial}{\partial_1} (r^2 \frac{\partial}{\partial_1} F) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial_2} (\sin \theta \frac{\partial}{\partial_2} F) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial_3^2} F.$$

Vorlesung 11

Mo 25.05. 10:15

Taylor-Formel, lokale Extrema

Notation 3.28 (Multi-Index-Schreibweise). Sei $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$. Dann definiert man

$$|\alpha| := \alpha_1 + \alpha_2 + \dots + \alpha_n$$

$$\alpha! := \alpha_1! \alpha_2! \dots \alpha_n!.$$

Ist $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, so setzt man

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}.$$

Ist f eine $|\alpha|$ -mal stetig differenzierbare Funktion, so ist

$$\partial^\alpha f := \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n} f$$

Lemma 3.29. Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, eine k -fach stetig differenzierbare Funktion. Sei $a \in U$ und sei $h \in \mathbb{R}^n$ so, dass $\{a + th \mid t \in [0, 1]\}$ ganz in U liegt.



Dann ist die Funktion $g: [0, 1] \rightarrow \mathbb{R}$, $g(t) = f(a + th)$ k -mal stetig differenzierbar und es gilt

$$g^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(a + th) h^\alpha.$$

↑
Summe über alle $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ mit $\alpha_1 + \alpha_2 + \dots + \alpha_n = k$

Beweis. 1. Wir zeigen per Induktion über k , dass gilt

$$g^{(k)}(t) = \sum_{i_1, \dots, i_k \in \{1, \dots, n\}} \partial_{i_1} \dots \partial_{i_k} f(a + th) h_{i_1} \dots h_{i_k}.$$

Induktionsanfang: $k = 1$

$$g'(t) = Df(a + th) \cdot h \quad \text{Kettenregel}$$

$$\sum_{i=1}^n \partial_i f(a + th) h_i.$$

$k-1 \rightarrow k$

$$\frac{d}{dt} \left(\sum_{i_1, \dots, i_{k-1} \in \{1, \dots, n\}} \partial_{i_1} \cdots \partial_{i_{k-1}} f(a+th) h_{i_1} \cdots h_{i_{k-1}} \right) = \sum_{i_1, \dots, i_{k-1} \in \{1, \dots, n\}} \sum_{i=1}^n \partial_i \partial_{i_1} \cdots \partial_{i_{k-1}} f(a+th) h_{i_1} \cdots h_{i_{k-1}} h_i$$

2. Kommt unter den Indizes (i_1, \dots, i_k) der Index 1 α_1 -mal vor, der Index 2 α_2 -mal, \dots , der Index n α_n -mal, so ist

$$\partial_{i_1} \cdots \partial_{i_k} f(a+th) = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f(a+th).$$

Es gibt

$$\frac{k!}{\alpha_1! \cdots \alpha_n!}$$

k -Tupel (i_1, \dots, i_k) von Zahlen $1 \leq i_j \leq n$, in denen 1 genau α_1 -mal, 2 genau α_2 -mal, \dots , n genau α_n -mal vorkommt (ohne Beweis). \implies Beh. \square

Satz 3.30 (Taylorsche Formel). Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, $(k+1)$ -mal stetig differenzierbar. Sei $a \in U$ und $h \in \mathbb{R}^n$ s.d. $\{a + th \mid t \in [0, 1]\} \subset U$. Dann existiert $\theta \in [0, 1]$ s.d.

$$f(a+h) = \sum_{\substack{|\alpha| \leq k \\ \alpha \in \mathbb{N}_0^n}} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha + \sum_{\substack{|\alpha|=k+1 \\ \alpha \in \mathbb{N}_0^n}} \frac{1}{\alpha!} \partial^\alpha f(a+\theta h) h^\alpha.$$

Beweis. Setze $g: [0, 1] \rightarrow \mathbb{R}$, $g(t) := f(a+th)$. Taylorsche Formel aus der DIFF I $\implies \exists \theta \in [0, 1]$ s.d.

$$g(1) = \sum_{j=0}^k \underbrace{\frac{1}{j!} g^{(j)}(0)}_{\parallel \sum_{|\alpha|=j} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha} + \underbrace{\frac{1}{(k+1)!} g^{(k+1)}(\theta)}_{\parallel \frac{1}{\alpha!} \partial^\alpha f(a+\theta h) h^\alpha} \quad \square$$

(3.29)

Folgerung 3.31. Sei $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, k -mal stetig differenzierbar. Sei $a \in U$ und $\delta > 0$ s.d. $B_\delta(a) \subset U$. Dann gilt für alle $h \in B_\delta(0)$:

$$f(a+h) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha + R_{k,a}(h)$$

mit $\frac{R_{k,a}(h)}{\|h\|^k} \rightarrow 0$ für $h \rightarrow 0$.

Äquivalent:

$$f(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \partial^\alpha f(a) (x-a)^\alpha + R_{k,a}(x-a)$$

für $x \in B_\delta(a)$. „Taylor-Entwicklung bis zur Ordnung k mit Entwicklungspunkt a “.

Beweis. $\exists \theta = \alpha_{a,h} \in [0, 1]$ s. d.

$$\begin{aligned} f(a+h) &= \sum_{|\alpha| \leq k-1} \frac{1}{\alpha!} f(a) h^\alpha + \sum_{|\alpha|} \partial^\alpha f(a + \theta h) h^\alpha \\ &= \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} \underbrace{(\partial^\alpha f(a + \theta h) - \partial^\alpha f(a))}_{:= r_{\alpha,a}(h)} h^\alpha. \end{aligned}$$

$\partial^\alpha f$ ist stetig $\implies r_{\alpha,a}(h) \rightarrow r_{\alpha,a}(0) = 0$ für $h \rightarrow 0$ für $h \rightarrow 0$. Setze

$$R_{k,a}(h) := \sum_{|\alpha|=k} \frac{1}{\alpha!} r_{\alpha,a}(h) h^\alpha,$$

dann gilt $\frac{R_{k,a}(h)}{\|h\|^k} \rightarrow 0$, denn

$$\frac{|h^\alpha|}{\|h\|_{\max}^k} = \frac{|h_1|^{\alpha_1} \cdots |h_n|^{\alpha_n}}{(\|h\|_{\max})^k} \leq 1 \quad \square$$

für $|\alpha| = \alpha_1 + \cdots + \alpha_n = k$.

Notation.

$$P_m(h) := P_{m,a}(h) = \sum_{|\alpha|=m} \frac{1}{\alpha!} \partial^\alpha f(a) h^\alpha$$

ist ein homogenes Polynom von Grad m (in h), das sogenannte *Taylorpolynom* zu f zum Entwicklungspunkt a . Es gilt

$$f(a+h) = \sum_{m=0}^k P_{m,a}(h) + R_{k,a}(h).$$

$m=0$ $P_0(h) = f(a)$ konstant.

$m=1$ Die Summe läuft über alle

$$(\alpha_1, \dots, \alpha_n) \in \{ (1, 0, \dots, 0), (0, 1, 0, \dots), \dots, (0, \dots, 0, 1) \}.$$

Also gilt

$$P_1(h) = \sum_{j=1}^n \partial_j f(a) h_j = \left(\partial_1 f(a) \cdots \partial_n f(a) \right) \cdot h.$$

Die Formel aus der Folgerung entspricht also genau der linearen Approximation von f , die durch Ableitung gegeben ist.

$m = 2$ Die Summe läuft über alle

$$\begin{aligned}
 (\alpha_1, \alpha_2, \dots, \alpha_n) \in & \{(1, 1, 0, \dots, 0), (1, 0, 1, 0, \dots, 0), \dots \\
 & , (0, 1, 1, 0, \dots, 0), (0, 1, 0, 1, 0, \dots, 0), \dots \\
 & , \dots \\
 & , (2, 0, \dots, 0), (0, 2, 0, \dots, 0), \dots, (0, \dots, 0, 2)\}.
 \end{aligned}$$

Also ist

$$\begin{aligned}
 P_2(h) &= \sum_{j=1}^n \sum_{i=j+1}^n \partial_i \partial_j f(a) h_i h_j + \sum_{j=1}^n \frac{1}{2} \partial_j^2 f(a) h_j^2 = \frac{1}{2} \sum_{i,j} \partial_j \partial_i f(a) h_i h_j. \\
 &= \underbrace{\sum_{j < i} \partial_i \partial_j f(a) h_i h_j}_{\substack{\uparrow \\ \text{Schwarz}}} = \frac{1}{2} \sum_{i \neq j} \partial_i \partial_j f(a) h_i h_j
 \end{aligned}$$

Man definiert daher die sogenannte *Hesse-Matrix*

$$H_f(a) := (\partial_i \partial_j f(a))_{\substack{i \in \{1, \dots, n\} \\ j \in \{1, \dots, n\}}}$$

und schreibt

$$\boxed{f(a+h) = f(a) + \langle \text{grad}(f)(a), h \rangle + \frac{1}{2} \langle h, H_f(a) h \rangle} + R_{z,a}(h)$$

für $f: U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und $h \in B_\delta(0)$, $\delta > 0$ s. d. $B_\delta(a) \subset U$.

Lokale Extrema

Wir hatten in 3.10 bereits gesehen: Ist $a \in U$ lokale Extremstelle von $f: U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, differenzierbar, so ist $\text{grad}(f)(a) = 0$.

Satz 3.32. Sei $f: U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, $U \subset \mathbb{R}^n$ offen. Sei $a \in U$ s. d. $\text{grad}(f)(a) = 0$. Dann gilt:

- Ist $H_f(a)$ positiv definit (d. h. $\langle h, H_f(a) h \rangle > 0 \forall h \in \mathbb{R}^n, h \neq 0$), so hat f in a ein lokales (isoliertes) Minimum.
- Ist $H_f(a)$ negativ definit (d. h. $\langle h, H_f(a) h \rangle < 0 \forall h \neq 0$), so hat f in a ein lokales (isoliertes) Maximum.
- Ist $H_f(a)$ indefinit (d. h. $\exists h$ s. d. $\langle h, H_f(a) h \rangle > 0$ und $\exists \tilde{h}$ s. d. $\langle \tilde{h}, H_f(a) \tilde{h} \rangle < 0$), so hat f in a kein lokales Extremum.

Bemerkung. Ist $H_f(a)$ positiv oder negativ *semidefinit* s. d. $\forall h \neq 0$ ist $\langle h, H_f(a)h \rangle \geq 0$ (bzw. ≤ 0), so ist keine allgemeine Aussage möglich.

Beispiel. $f_1(x, y) = x^2 + y^4$, $f_2(x, y) = x^2$, $f_3(x, y) = x^2 + y^3$, $H_{f_j}(0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \forall j$ ist positiv semidefinit ($\langle h, H_{f_j}(0)h \rangle = 2h_1^2 \geq 0$, $= 0$ für $h_1 = 0$).

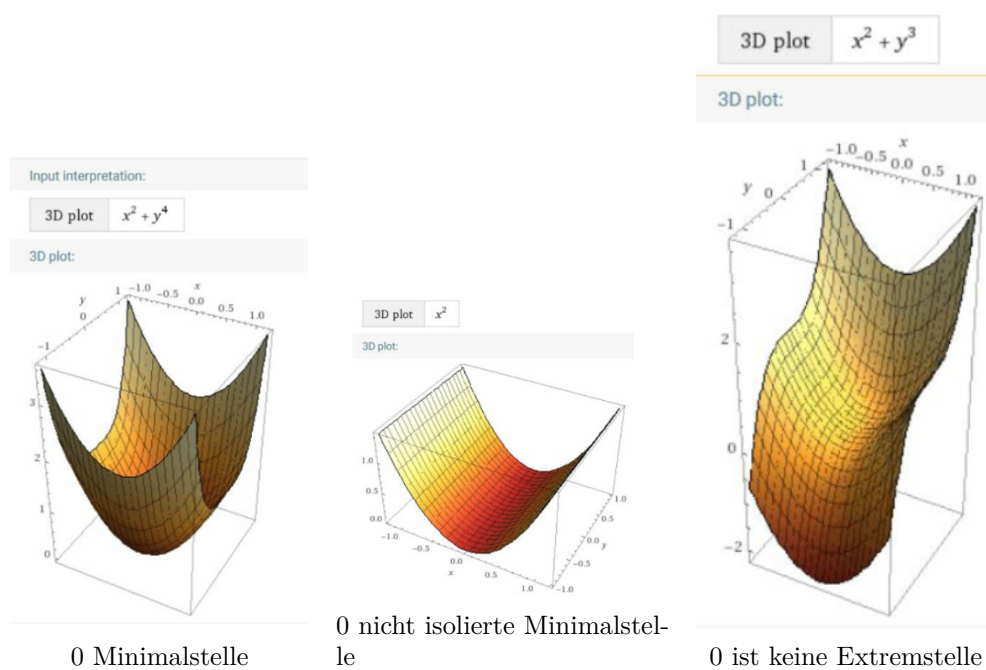


Abbildung 3.1: Wolfram Alpha

Beweis von 3.32. Setze $A := H_f(a)$. $\exists \delta > 0$ s. d. $B_\delta(a) \subset U$. Und somit

$$f(a+h) = f(a) + \underbrace{\langle \text{grad}(f)(a), h \rangle}_{=0} + \frac{1}{2} \langle h, Ah \rangle + R_2(h) \quad \forall h \in B_\delta(0)$$

mit $\frac{R_2(h)}{\|h\|^2} \rightarrow 0$ ($h \rightarrow 0$).

$$\implies \forall \varepsilon > 0 \quad \exists \tilde{\delta} > 0 \quad (\tilde{\delta} < \delta) \text{ s. d. } |R_2(h)| \leq \varepsilon \|h\|^2 \quad \forall h \in B_{\tilde{\delta}}(0)$$

3.32.a) Sei A positiv definit. Betrachte $S = \{h \in \mathbb{R}^n \mid \|h\|_E = 1\}$. S ist abgeschlossen und beschränkt $\subset \mathbb{R}^n \implies S$ ist kompakt $\implies S \ni h \mapsto \langle h, Ah \rangle$ nimmt Maximum und Minimum an (da stetig).

$$\alpha := \inf_{h \in S} \underbrace{\langle h, Ah \rangle}_{>0} > 0.$$

>0 (VOR) ($h \neq 0$ wenn $h \in S$)

Behauptung.

$$\langle h, Ah \rangle \geq \alpha \|h\|_{\mathbb{E}}^2 \quad \forall h \in \mathbb{R}^n.$$

Beweis. $h = 0$: ✓. Sei also $h \neq 0$. Setze $\tilde{h} := \frac{1}{\|h\|_{\mathbb{E}}} h$.

$$\implies \tilde{h} \in S \implies \langle \tilde{h}, A\tilde{h} \rangle \geq \alpha.$$

Die Behauptung folgt wegen

$$\langle \tilde{h}, A\tilde{h} \rangle = \frac{1}{\|h\|_{\mathbb{E}}^2} \langle h, Ah \rangle. \quad \square$$

Wähle nun $\tilde{\delta} > 0$ s. d. $|R_2(h)| \leq \frac{\alpha}{4} \|h\|_{\mathbb{E}}^2$ für $\|h\| < \tilde{\delta}$.

$$\implies f(a+h) \geq f(a) + \underbrace{\frac{1}{2} \langle h, Ah \rangle}_{\geq \alpha \|h\|_{\mathbb{E}}^2} - |R_2(h)| \geq f(a) + \underbrace{\frac{1}{4} \alpha \|h\|_{\mathbb{E}}^2}_{\geq 0} > f(a) \quad \forall \|h\| < \tilde{\delta}, \quad h \neq 0.$$

3.32.b) Ist A negativ definit, betrachte $-f$ und wende 3.32.a) an.

3.32.c) Ist A indefinit, gibt es in jeder Umgebung V von a Punkte $x, x' \in V$ s. d.

$$f(x) < f(a) < f'(x).$$

Es gilt $v \in \mathbb{R}^n$, s. d. $\langle v, Av \rangle =: \alpha > 0$.

$$\implies f(a+tv) = f(a) + \frac{1}{2} t^2 \alpha + R_2(tv)$$

für $|t|$ klein genug (s. d. $tv \in B_{\delta}(0)$). Wähle $|t| > 0$ so klein, dass zudem $R_2(tv) \leq \frac{\alpha}{2} t^2$

$$\implies f(a+sv) \geq f(a) + \frac{1}{2} s^2 \alpha - |R_2(sv)| > f(a)$$

für $0 < s < t$ (o.B.d.A. $t > 0$).

Genauso: Es gibt $v \in \mathbb{R}^n$, s. d. $\langle v, Av \rangle =: \alpha < 0$

$$\implies f(x+tv) = f(x) - \frac{1}{2} t^2 |\alpha| + R_2(tv)$$

für $|t|$ klein genug (s. d. $tv \in B_{\delta}(0)$). Wähle $|t| > 0$ so klein, dass zudem $|R_2(tv)| \leq \frac{\alpha}{4} t^2$

$$\implies f(a+sv) \leq f(a) - \frac{1}{2} s^2 |\alpha| + |R_2(sv)| < f(a) \quad \forall 0 < s < t \quad (\text{o.B.d.A. } t > 0). \quad \square$$

Beispiele. i) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x_1^2 + x_2^2 + 2$, $\text{grad}(f)(a) = (2a_1, 2a_2) = 0$ nur für $a = 0$.

$$H_f(0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

positiv definit \implies isoliertes Minimum in 0.

ii) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = 2 - x_1^2 - x_2^2$, $\text{grad}(f)(a) = -2a = 0$ nur für $a = 0$.

$$H_f(0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

negativ definit \implies isoliertes Maximum in 0.

iii) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = 2 + x_1^2 - x_2^2$, $\text{grad}(f)(a) = 2a_1 - 2a_2 = 0$ nur für $a = 0$.

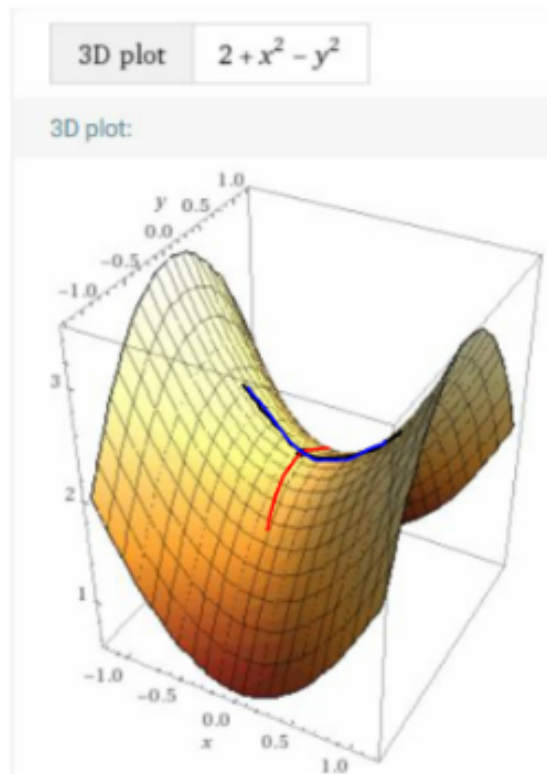
$$H_f(0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

indefinit:

$$\langle e_1, H_f(0)e_1 \rangle = 2 = -\langle e_2, H_f(0)e_2 \rangle$$

$$f(0 + te_1) > f(0) \quad \forall t \neq 0$$

$$f(0 + te_2) < f(0) \quad \forall t \neq 0.$$



Sattelfläche

Bemerkung. Wie untersucht man in komplizierten Fällen $H_f(a)$ auf Definitheit? Kennt man die Eigenwerte (\exists wegen Symmetrie!) ist klar: Sind all positiv (negativ), so ist $H_f(a)$ positiv (negativ) definit, gibt es positive und negative Eigenwerte, so ist sie indefinit. Aber Eigenwerte großer Matrizen sind schwer zu bestimmen. Besser geeignet ist

Lemma 3.33 (aus der AGLA: Kriterium von Hurwitz). Sei $A \in \text{Mat}(n \times n, \mathbb{R})$ symmetrisch. Dann ist A genau dann positiv definit, wenn

$$\det \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix} > 0 \quad \forall k \in \{1, \dots, n\}$$

\uparrow
 Haupt-Minoren von A

und genau dann negativ definit, wenn

$$\det(A_{11}) = A_{11} < 0$$

$$\det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} > 0$$

und dann immer abwechselnd.